



**HAL**  
open science

# Recherche de pas par Majoration-Minoration. Application à la résolution de problèmes inverses.

Emilie Chouzenoux

## ► To cite this version:

Emilie Chouzenoux. Recherche de pas par Majoration-Minoration. Application à la résolution de problèmes inverses.. Traitement du signal et de l'image [eess.SP]. Ecole Centrale de Nantes (ECN), 2010. Français. NNT: . tel-00555643

**HAL Id: tel-00555643**

**<https://theses.hal.science/tel-00555643>**

Submitted on 14 Jan 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ÉCOLE CENTRALE DE NANTES  
ÉCOLE DOCTORALE  
SCIENCES ET TECHNOLOGIES  
DE L'INFORMATION ET DES MATHÉMATIQUES

Année : 2010

## Thèse de Doctorat

Spécialité : AUTOMATIQUE, ROBOTIQUE, TRAITEMENT DU SIGNAL  
ET INFORMATIQUE APPLIQUÉE

*Présentée et soutenue publiquement par :*

**Émilie CHOUZENOUX**

le 8 décembre 2010  
à l'École Centrale de Nantes

### TITRE

**Recherche de pas par Majoration-Minoration  
Application à la résolution de problèmes inverses**

### Jury

Rapporteurs :	Gilles AUBERT Jean-Christophe PESQUET	Professeur à l'Université de Nice Sophia-Antipolis Professeur à l'Université de Marne-la-Vallée (Paris)
Examineurs :	Christian JUTTEN Henri LANTÉRI Saïd MOUSSAOUI Jérôme IDIER	Professeur à l'Université Joseph Fourier (Grenoble) Professeur émérite à l'Université de Nice Sophia-Antipolis Maître de conférences à l'ECN (Nantes) Directeur de recherche CNRS à l'IRCCyN (Nantes)
Invité :	François MARIETTE	Directeur de recherche au CEMAGREF (Rennes)

**Directeur de thèse :** Jérôme IDIER  
Laboratoire / composante : IRCCyN / CNRS  
**Co-encadrant :** Saïd MOUSSAOUI  
Laboratoire / composante : IRCCyN / École Centrale de Nantes



“Voici la première question raisonnable qui m’ait été adressée.  
Entendez-vous, vous autres ? Ha ! ha ! ha !  
Vois-tu, je m’ennuyais. ”  
*F. Dostoïevski (Le Joueur)*

“Il faut avoir encore du chaos en soi  
pour enfanter une étoile dansante.”  
*F. W. Nietzsche*



# Remerciements

Tout d'abord, mes remerciements s'adressent aux personnes qui m'ont proposé le sujet de thèse et qui m'ont encadré tout au long de ces années d'étude : MM. Jérôme Idier et Saïd Moussaoui. Je tiens à les remercier pour leur enthousiasme, leur disponibilité et leurs conseils. Je garderai en mémoire la confiance et la sympathie qu'ils m'ont témoignées au cours de ces trois années de thèse.

Je tiens également à exprimer mes remerciements aux membres du jury, qui ont accepté d'évaluer mon travail de thèse.

Merci à M. Christian Jutten, professeur à l'Université Joseph Fourier, Grenoble, d'avoir accepté de présider le jury de cette thèse, et à MM. les professeurs Jean-Christophe Pesquet de l'Université de Marne-la-Vallée, Paris, et Gilles Aubert de l'Université de Nice Sophia-Antipolis, d'avoir accepté d'être les rapporteurs de ce manuscrit. Leurs remarques et suggestions lors de la lecture de ma thèse m'ont permis d'apporter des améliorations à la qualité de cette dernière. Merci également à M. le professeur Henri Lantéri de l'Université de Nice pour avoir accepté de faire partie de mon jury de thèse.

J'exprime aussi ma reconnaissance à M. François Mariette, directeur de recherche au CE-MAGREF, Rennes, pour sa précieuse collaboration sur la relaxation RMN. Ses propositions et corrections ont aidé à améliorer significativement le dernier chapitre de ce mémoire. Je le remercie vivement d'avoir accepté de participer à mon jury de thèse.

Je remercie M. Michel Malabre, directeur de l'IRCCyN, de m'avoir accueilli dans le laboratoire où j'ai pu réaliser mes travaux de recherche dans des conditions excellentes. Je remercie également Emily Thureau, notre secrétaire, pour sa gentillesse et sa disponibilité.

Je tiens enfin à remercier les amis qui m'ont accompagnés au cours des trois ans de cette thèse : Di, mon collègue de bureau, pour sa sympathie et ses bons plans, Didier pour ses jeux de mots inimitables, ainsi que Pedro et Adrien pour ces inoubliables soirées durant notre dernière année de thèse. Je n'oublie pas non plus mes anciens collègues centraliens parmi lesquels je remercierai plus particulièrement Benoît pour ses conseils avisés, et Sylvain, mon cher binôme.

Un remerciement particulier pour Charles qui m'a tant apporté pendant ses années de thèse.

Finalement, j'adresse un grand merci à toute ma famille qui a toujours été présente lorsque j'en ai eu besoin, en particulier merci à ma grand-mère pour m'avoir transmis sa passion pour les mathématiques.



*A mes sœurs Sandrine et Claire.*



# Table des matières

<b>1</b>	<b>Introduction générale</b>	<b>1</b>
1.1	Position du problème . . . . .	1
1.2	Contributions . . . . .	2
1.3	Organisation du document . . . . .	4
<b>2</b>	<b>Inversion par approche pénalisée en traitement du signal et de l'image</b>	<b>7</b>
2.1	Introduction . . . . .	8
2.2	Problème inverse en traitement du signal . . . . .	8
2.2.1	Définition . . . . .	8
2.2.2	Caractère mal posé . . . . .	9
2.2.3	Inversion généralisée . . . . .	10
2.2.4	Régularisation par contrôle de la dimension . . . . .	11
2.3	Régularisation par minimisation d'un critère pénalisé . . . . .	12
2.3.1	Principe . . . . .	12
2.3.2	Formulation du critère d'adéquation aux données . . . . .	12
2.3.3	Modèle <i>a priori</i> . . . . .	14
2.3.4	Réglage du paramètre de régularisation . . . . .	20
2.4	Conclusion . . . . .	20
<b>3</b>	<b>Minimisation d'un critère pénalisé</b>	<b>23</b>
3.1	Introduction . . . . .	24
3.2	Formulation du problème d'optimisation . . . . .	24
3.2.1	Propriétés du critère . . . . .	24
3.2.2	Conditions d'optimalité . . . . .	27
3.3	Méthode de descente itérative . . . . .	28
3.3.1	Principe . . . . .	28
3.3.2	Méthodes de premier ordre pour la construction de la direction . . . . .	29
3.3.3	Méthodes de second ordre pour la construction de la direction . . . . .	31
3.3.4	Prise en compte de contraintes . . . . .	33
3.3.5	Discussion . . . . .	34
3.4	Méthode itérative de sous-espace . . . . .	34
3.4.1	Principe . . . . .	34
3.4.2	Construction du sous-espace . . . . .	34
3.4.3	Préconditionnement . . . . .	36
3.5	Conclusion . . . . .	37
<b>4</b>	<b>Recherche de pas par Majoration-Minoration quadratique</b>	<b>39</b>
4.1	Introduction . . . . .	40
4.2	Recherche de pas pour l'algorithme de descente itérative . . . . .	40
4.2.1	Règles de pas exactes . . . . .	41
4.2.2	Règles de pas approchées . . . . .	41
4.2.3	Calcul du pas . . . . .	43
4.2.4	Condition de Zoutendijk . . . . .	46

4.3	Recherche de pas pour l'algorithme itératif de sous-espace . . . . .	46
4.3.1	Règle de pas multi-dimensionnelle . . . . .	46
4.3.2	Pondération analytique pour l'algorithme SMG . . . . .	47
4.3.3	Méthode de régions de confiance . . . . .	48
4.4	Algorithme de Majoration-Minoration - Application à la recherche de pas . . . . .	50
4.4.1	Approximation tangente majorante quadratique . . . . .	51
4.4.2	La récurrence de Majoration-Minoration . . . . .	52
4.4.3	Recherche de pas MMQ scalaire . . . . .	54
4.4.4	Recherche de pas MMQ multi-dimensionnelle . . . . .	55
4.5	Propriétés de convergence du pas MM quadratique . . . . .	56
4.5.1	Hypothèses . . . . .	56
4.5.2	Propriétés de la séquence de pas . . . . .	56
4.5.3	Convergence des algorithmes gradient reliés . . . . .	58
4.5.4	Convergence des algorithmes de gradient conjugué non linéaire . . . . .	62
4.6	Conclusion . . . . .	63
<b>5</b>	<b>Application à la déconvolution d'image</b>	<b>65</b>
5.1	Introduction . . . . .	66
5.2	Le problème de déconvolution d'image . . . . .	66
5.3	Algorithmes comparés . . . . .	66
5.3.1	Algorithmes de gradient conjugué non linéaire . . . . .	66
5.3.2	Algorithme à super mémoire de gradient . . . . .	67
5.3.3	Préconditionnement . . . . .	68
5.4	Mise en œuvre des stratégies de pas MMQ . . . . .	68
5.4.1	Construction d'approximations majorantes quadratiques . . . . .	68
5.4.2	Réduction du coût de calcul . . . . .	70
5.5	Interprétation des résultats expérimentaux . . . . .	72
5.5.1	Algorithme GCNL . . . . .	72
5.5.2	Algorithme SMG . . . . .	73
5.5.3	Analyse comparée GCNL / SMG . . . . .	74
5.6	Résultats complémentaires . . . . .	75
5.6.1	Construction du sous-espace . . . . .	75
5.6.2	Reconstruction dans un dictionnaire . . . . .	76
5.7	Conclusion . . . . .	76
<b>6</b>	<b>Minimisation d'un critère présentant une barrière</b>	<b>77</b>
6.1	Introduction . . . . .	78
6.2	Fonction barrière . . . . .	78
6.2.1	Définition . . . . .	78
6.2.2	Fonctions barrières scalaires . . . . .	78
6.2.3	Critères pénalisés à barrière . . . . .	80
6.2.4	Fonction barrière en optimisation contrainte . . . . .	81
6.3	Minimisation d'un critère avec barrière . . . . .	82
6.3.1	Difficultés lors de la résolution du problème . . . . .	83
6.3.2	Restriction du pas . . . . .	84
6.3.3	Interpolation log-quadratique . . . . .	84
6.3.4	Formule de pas pour les fonctions auto-concordantes . . . . .	85
6.3.5	Pas MMQ pour les fonctions barrières . . . . .	87
6.4	Conclusion . . . . .	88
<b>7</b>	<b>Recherche de pas MM pour les fonctions avec barrières</b>	<b>91</b>
7.1	Introduction . . . . .	92
7.2	Description de la stratégie de pas . . . . .	92
7.2.1	Cadre de l'étude . . . . .	92

7.2.2	Fonction majorante log-quadratique . . . . .	92
7.2.3	Construction de la fonction majorante . . . . .	93
7.2.4	Expression du minimiseur . . . . .	98
7.3	Propriétés de convergence de la recherche de pas MMLQ 1D . . . . .	99
7.3.1	Rappel des hypothèses . . . . .	99
7.3.2	Propriétés de la séquence de pas . . . . .	100
7.3.3	Résultats de convergence . . . . .	102
7.4	Conclusion . . . . .	104
<b>8</b>	<b>Résultats expérimentaux sur les algorithmes de points intérieurs</b>	<b>105</b>
8.1	Introduction . . . . .	106
8.2	La programmation quadratique . . . . .	106
8.3	Algorithmes comparés . . . . .	106
8.3.1	Algorithme primal . . . . .	106
8.3.2	Algorithme primal-dual . . . . .	108
8.3.3	Stratégies de recherche de pas . . . . .	109
8.4	Contexte expérimental . . . . .	110
8.4.1	Génération des problèmes quadratiques . . . . .	110
8.4.2	Paramètres des algorithmes . . . . .	111
8.5	Analyse des résultats . . . . .	111
8.5.1	Recherche de pas MM log-quadratique . . . . .	111
8.5.2	Comparaison avec les méthodes de recherche de pas existantes . . . . .	112
8.6	Conclusion . . . . .	113
<b>9</b>	<b>Application à la reconstruction de spectres RMN 2D</b>	<b>117</b>
9.1	Introduction . . . . .	118
9.2	La résonance magnétique nucléaire . . . . .	118
9.2.1	Les phénomènes de relaxation . . . . .	118
9.2.2	Mesure des temps de relaxation . . . . .	121
9.2.3	Cas de systèmes hétérogènes . . . . .	121
9.3	Méthodes d'inversion pénalisée existantes . . . . .	125
9.3.1	Méthode de Tikhonov sous contraintes de positivité . . . . .	125
9.3.2	Reconstruction par maximum d'entropie . . . . .	126
9.4	Algorithme de minimisation ME proposé . . . . .	128
9.4.1	Description de l'algorithme . . . . .	128
9.4.2	Réduction du coût de calcul . . . . .	131
9.5	Application à la reconstruction de données réelles . . . . .	132
9.5.1	Réglage du coefficient de régularisation . . . . .	132
9.5.2	Réglage du paramètre d'angle de bascule . . . . .	134
9.5.3	Application à l'analyse d'échantillons de pomme par RMN . . . . .	135
9.6	Conclusion . . . . .	136
<b>10</b>	<b>Conclusions et perspectives</b>	<b>139</b>
	<b>Annexes</b>	<b>147</b>
<b>A</b>	<b>Démonstrations du chapitre 7</b>	<b>147</b>
A.1	Propriétés de la récurrence MM . . . . .	147
A.1.1	Résultats préliminaires . . . . .	147
A.1.2	Démonstration du lemme 7.5 . . . . .	149
A.1.3	Démonstration du lemme 7.6 . . . . .	149
A.1.4	Démonstration de la propriété 7.1 . . . . .	150
A.1.5	Démonstration de la propriété 7.2 . . . . .	151

A.1.6	Démonstration de la propriété 7.3 . . . . .	152
A.2	Convergence de l'algorithme de conjugué non linéaire . . . . .	152
A.2.1	Résultats préliminaires . . . . .	152
A.2.2	Démonstration du théorème 7.4 . . . . .	153
	<b>Bibliographie</b>	<b>155</b>
	<b>Résumé</b>	<b>172</b>

# Table des figures

2.1	Formation d'un problème inverse linéaire . . . . .	9
2.2	Problème de déconvolution d'image . . . . .	10
2.3	Régularisation de Tikhonov . . . . .	15
2.4	Exemples de fonctions de pondération convexes . . . . .	17
2.5	Choix de la fonction de régularisation . . . . .	21
2.6	Influence du paramètre de régularisation . . . . .	22
3.1	Barrières logarithmique et entropique . . . . .	27
3.2	Direction de descente . . . . .	29
3.3	Principe de la méthode de descente itérative . . . . .	30
4.1	Principe schématique des règles de pas usuelles . . . . .	42
4.2	Algorithme de régions de confiance . . . . .	49
4.3	Sélection de la région de confiance . . . . .	50
4.4	Approximation tangente majorante . . . . .	52
4.5	Algorithme de Majoration-Minoration . . . . .	53
4.6	Recherche de pas par Majoration-Minoration . . . . .	54
5.1	Reconstruction de <b>peppers</b> avec $(\lambda, \delta) = (0.2, 8)$ . . . . .	67
5.2	Principe de construction des majorantes semi-quadratiques . . . . .	70
6.1	Exemples de fonctions barrières scalaires . . . . .	79
6.2	Exemples de $\phi$ -divergences entre $u$ et $a > 0$ . . . . .	80
6.3	Approximation quadratique des fonctions barrières . . . . .	88
7.1	Principe de la recherche de pas MMLQ pour un critère à barrière . . . . .	94
7.2	Forme des termes composants la fonction $b_1$ . . . . .	96
7.3	Forme des termes composants la fonction $b_2$ . . . . .	96
7.4	Caractère gradient Lipschitz d'une fonction barrière . . . . .	100
8.1	Performances de l'algorithme de barrière associé au pas MMLQ 1D . . . . .	112
9.1	Phénomène de résonance magnétique nucléaire . . . . .	119
9.2	Relaxation du moment magnétique $M(t)$ après un basculement de $\Phi = 90^\circ$ . . . . .	120
9.3	Acquisition du temps de relaxation longitudinal . . . . .	122
9.4	Acquisition du temps de relaxation transversal : Echo de spin . . . . .	122
9.5	Acquisition des mesures en RMN 2D . . . . .	123
9.6	Courbe en S pour un jeu de données synthétique . . . . .	133
9.7	Estimation de l'angle de bascule sur un jeu de données synthétiques . . . . .	134
9.8	Influence du paramètre de phase sur la reconstruction de données réelles . . . . .	137
9.9	Estimation du paramètre de phase à partir de données réelles . . . . .	137
9.10	Spectre RMN 2D reconstruit par la méthode ME à partir des données réelles. . . . .	138
9.11	Spectres RMN 1D reconstruits par la méthode ME à partir des données réelles. . . . .	138
9.12	Spectre RMN 2D reconstruit par la méthode TIK <sup>+</sup> à partir des données réelles. . . . .	138
9.13	Spectres RMN 1D reconstruits par la méthode TIK <sup>+</sup> à partir des données réelles. . . . .	138



# Liste des tableaux

3.1	Principaux algorithmes de sous-espace . . . . .	35
4.1	Fonctions interpolantes polynômiales (Sun et Yuan, 2006, Chap.2). . . . .	45
5.1	Expression récursive des sous-espaces du tableau 3.1 . . . . .	72
5.2	Algorithmes GCNL associés au pas de Wolfe . . . . .	73
5.3	Algorithmes GCNL associés au pas MM quadratique . . . . .	73
5.4	Performances de SMG associé à la recherche de pas MM ou exacte . . . . .	74
5.5	Comparaison entre algorithmes MG et GCNL . . . . .	74
6.1	Exemples de fonctions barrières scalaires . . . . .	79
6.2	Exemples de $\phi$ -divergences entre $u$ et $a > 0$ . . . . .	80
6.3	Fonctions interpolantes log-quadratiques (Murray et Wright, 1994) . . . . .	85
8.1	Comparaison des stratégies de pas pour l'algorithme de barrière . . . . .	113
8.2	Résolution des problèmes CUTEr par l'algorithme primal . . . . .	114
8.3	Résolution des problèmes CUTEr par l'algorithme primal-dual . . . . .	115



# Liste des algorithmes

4.1	Algorithme de rebroussement pour la règle d'Armijo . . . . .	44
4.2	Détermination d'un pas satisfaisant la règle de Wolfe . . . . .	44
4.3	Détermination d'un pas satisfaisant la règle de Goldstein . . . . .	45
4.4	Algorithme de régions de confiance ( <a href="#">Conn et al., 1987</a> , Chap.6) . . . . .	49
4.5	Recherche de pas multi-dimensionnelle par régions de confiance ( <a href="#">Shi et Xu, 2009</a> ) .	50
8.1	Algorithme primal de points intérieurs . . . . .	107
8.2	Algorithme primal-dual de points intérieurs . . . . .	108
9.1	Algorithme de gradient conjugué préconditionné (GCP) . . . . .	130
9.2	Algorithme TN pour le problème ME . . . . .	131
9.3	Méthode de la courbe en S pour l'estimation du paramètre de régularisation . . . .	133



# Notations

$\mathbb{R}, \mathbb{R}^+$	: ensemble des réels et des réels positifs
$\mathbb{R}^n$	: ensemble des vecteurs réels à $n$ dimensions
$\mathbb{R}^{m \times n}$	: ensemble des matrices réelles à $m$ lignes et $n$ colonnes
$a$	: les scalaires sont notés par des minuscules
$\mathbf{v}$	: les vecteurs sont notés par des minuscules en gras
$\mathbf{v}^T$	: transposé du vecteur $\mathbf{v}$
$\mathbf{e}$	: vecteur colonne rempli de 1
$\mathbf{e}_i$	: vecteur colonne valant 1 à la ligne $i$ et zéro ailleurs
$\mathbf{M}$	: les matrices sont notées par des majuscules en gras
$M_{ij}$	: valeur de la matrice $\mathbf{M}$ à la ligne $i$ et à la colonne $j$
$v_i$	: valeur du vecteur $\mathbf{v}$ à la ligne $i$
$[\mathbf{M}\mathbf{v}]_i$	: valeur du produit $\mathbf{M}\mathbf{v}$ à la ligne $i$
$x_{i,k}$	: valeur du vecteur $\mathbf{x}_k$ à la ligne $i$
$\mathbf{M}^T$	: transposée de la matrice $\mathbf{M}$
$\mathbf{M}^{-1}$	: matrice inverse de la matrice carrée $\mathbf{M}$
$\text{Diag}(\mathbf{v})$	: matrice diagonale dont la diagonale est le vecteur $\mathbf{v}$
$\text{diag}(\mathbf{M})$	: vecteur contenant la diagonale de la matrice $\mathbf{M}$
$\mathbf{I}$	: matrice identité
$\text{Ker}(\mathbf{M})$	: noyau de la matrice $\mathbf{M}$
$\text{Im}(\mathbf{M})$	: image de la matrice $\mathbf{M}$
$\dot{f}(x)$	: dérivée première de la fonction $f$ par rapport à $x$
$\mathcal{D}_f$	: domaine de définition du critère $F$
$\partial\mathcal{D}_f$	: frontière du domaine $\mathcal{D}_f$
$\nabla F(\mathbf{x})$	: gradient du critère $F$ par rapport à $\mathbf{x}$
$\nabla^2 F(\mathbf{x})$	: Hessien du critère $F$ par rapport à $\mathbf{x}$
$\ \mathbf{v}\ $	: norme euclidienne du vecteur $\mathbf{v}$
$\text{vect}(\mathbf{M})$	: transformation de la matrice $\mathbf{M} \in \mathbb{R}^{m \times n}$ en le vecteur de taille $\mathbb{R}^{mn}$ , dans l'ordre lexicographique, c'est à dire par concaténation des éléments de la matrice $\mathbf{M}$ parcourue ligne par ligne.
$\text{mat}(\mathbf{x})$	: transformation du vecteur $\mathbf{x} \in \mathbb{R}^{mn}$ en la matrice de taille $\mathbb{R}^{m \times n}$ , dans l'ordre lexicographique.
s.c.	: sous les contraintes
$a \triangleq b$	: $a$ est défini comme étant égal à $b$



# Abréviations

MM	:	Majoration-Minoration
MMQ	:	Majoration-Minoration quadratique
MMQ 1D	:	Recherche de pas scalaire par MMQ
MMQ $r$ D	:	Recherche de pas multi-dimensionnelle par MMQ
SVD	:	Décomposition en valeurs singulières
TSVD	:	SVD tronquée
GC	:	Gradient conjugué
GCP	:	Gradient conjugué préconditionné
GCNL	:	Gradient conjugué non linéaire
TN	:	Newton tronqué (Truncated Newton)
MG	:	Mémoire de gradient
SMG	:	Super mémoire de gradient
SMD	:	Super mémoire de descente
GS	:	Sous-espace de gradient (Gradient subspace)
ORTH	:	Sous-espace orthogonal
SESOP	:	Sous-espace séquentiel (Sequential Subspace Optimization)
QNS	:	Sous-espace de Quasi-Newton (Quasi-Newton subspace)
SESOP-TN	:	Sous-espace séquentiel de Newton tronqué
SQ	:	Semi-quadratique
GR	:	Geman et Reynolds
GY	:	Geman et Yang
PRP	:	Polak-Ribière-Polyak
HS	:	Hestenes-Stiefel
LS	:	Liu-Storey
FR	:	Fletcher-Reeves
DY	:	Dan-Yuan
MT	:	Moré et Thuente
MMLQ 1D	:	Recherche de pas scalaire par Majoration-Minoration log-quadratique
LD	:	Lin et Doyle
MW	:	Murray et Wright
RMN	:	Résonance Magnétique Nucléaire
ME	:	Maximum d'Entropie



# Chapitre 1

## Introduction générale

De nombreuses avancées dans les techniques de mesure et d'imagerie mènent au problème numérique complexe de la restitution d'un signal à partir de l'observation de signaux physiquement associés. Ce problème, appelé *problème inverse*, peut être résolu efficacement en mettant en œuvre une approche pénalisée qui consiste à se ramener à l'optimisation d'un critère composite. Un tel critère permet d'obtenir la solution la plus plausible, compte tenu des données et des informations sur l'application considérée. La principale difficulté à surmonter est le volume croissant des données qui entraîne un coût de calcul rédhibitoire lors de la résolution numérique du problème d'optimisation.

Un domaine particulièrement confronté à des données de grande taille est celui de l'imagerie, incluant des problèmes de restauration et de reconstruction d'images. Dans ces applications, le nombre d'inconnues à estimer correspond au nombre de pixels de l'image et est par conséquent très important. De plus, la taille actuelle des problèmes d'imagerie est amenée à augmenter, avec la volonté de développer l'imagerie 3D voire l'imagerie dynamique "2D+t" ou "3D+t". Bien que la puissance des calculateurs augmente régulièrement, traiter de tels volumes de données en un temps raisonnable reste un défi majeur.

Dans cette thèse, nous proposons des méthodes algorithmiques dédiées à l'optimisation des critères pénalisés dans l'objectif de résoudre les problèmes inverses de grande taille du traitement du signal et des images pour un coût de calcul acceptable.

### 1.1 Position du problème

La solution de nombreux problèmes inverses de traitement du signal et de l'image peut être définie comme le minimiseur d'un critère pénalisé  $F : \mathbb{R}^n \rightarrow \mathbb{R}$  qui prend en compte conjointement les données observées et les informations préalables sur la solution. La solution ne peut généralement pas s'exprimer sous une forme analytique et un algorithme de minimisation doit être mis en œuvre pour en fournir une estimée. Bien que cette approche pénalisée fournisse des solutions de qualité satisfaisante, son implémentation actuelle a souvent pour inconvénient d'avoir une charge calculatoire trop importante pour les problèmes de grande taille. Dans le cadre de l'optimisation différentiable, une famille fondamentale d'algorithmes est basée sur la stratégie de *descente itérative* alternant calcul d'une direction de descente et détermination d'un pas scalaire quantifiant l'avancée le long de la direction choisie. Une stratégie pertinente de calcul du pas doit tirer profit de la structure du critère pénalisé, ce qui est le cas des algorithmes de pas par Majoration-Minoration (MM) auxquels nous prêterons une attention particulière tout au long de cette thèse.

### Recherche de pas MM quadratique

Le principe de Majoration-Minoration (MM) apparaît pour la première fois dans l'ouvrage (Ortega et Rheinboldt, 1970). Il consiste à substituer une séquence de problèmes plus simples au pro-

blème initial d'optimisation, en approchant le critère par une *approximation tangente majorante*, c'est-à-dire une fonction tangente au critère en un point et supérieure ou égale au critère en tous les autres points. De nombreuses contributions algorithmiques dans le cadre de l'inversion pénalisée s'identifient avec un schéma MM. C'est le cas par exemple des algorithmes semi-quadratiques de (Geman et Reynolds, 1992; Geman et Yang, 1995) largement utilisés pour la restauration d'images (Allain *et al.*, 2006; Charbonnier *et al.*, 1997; Nikolova et Ng, 2005).

Récemment, la stratégie MM a été utilisée comme principe constructif d'une méthode de recherche de pas (Fessler et Booth, 1999; Labat et Idier, 2008). Le calcul du pas utilise des approximations majorantes quadratiques du critère le long de la direction de descente, construites à partir d'informations sur la structure analytique de celui-ci. Il en résulte une stratégie très simple de détermination du pas, notée MMQ 1D, possédant de nombreuses propriétés de convergence (Labat et Idier, 2008).

## Recherche de pas MM quadratique multi-dimensionnelle

Une accélération significative de l'algorithme de descente itérative est apportée en utilisant une stratégie de sous-espaces (Conn *et al.*, 1994a; Miele et Cantrell, 1969; Narkiss et Zibulevsky, 2005; Yuan, 2007). Cette procédure consiste à explorer le critère non plus le long de directions de descente mais à l'intérieur de sous-espaces engendrés par plusieurs directions. La recherche de pas scalaire est alors remplacée par une recherche de pas multi-dimensionnelle visant à minimiser le critère dans le sous-espace. Néanmoins, les stratégies de calcul du pas multi-dimensionnelles existant dans la littérature de l'optimisation souffrent d'une mise en œuvre complexe ou d'une absence de résultats de convergence.

Comme souligné dans (Labat, 2006), la recherche de pas MM scalaire se généralise simplement au cas multi-dimensionnel, l'intérêt de ce schéma étant sa simplicité de mise en œuvre. **La question se pose de savoir si la recherche de pas MM quadratique multi-dimensionnelle MMQ  $rD$  assure la convergence des algorithmes de sous-espace.** Ce point fait l'objet d'une étude spécifique dans le cadre de ce travail.

## Recherche de pas MM pour les fonctions barrières

L'élaboration des procédures de pas MMQ se basent sur l'hypothèse de l'existence de fonctions quadratiques majorantes du critère à minimiser. Cependant, certains critères pénalisés ne vérifient pas cette hypothèse. C'est le cas en particulier des critères contenant un terme de *barrière*, c'est à dire un terme à gradient non borné. Cette singularité se rencontre fréquemment dans le contexte du traitement du signal par exemple dans les critères de maximum d'entropie (Mohammad-Djafari et Demoment, 1988b), de vraisemblance Poissonienne (Bertero *et al.*, 2009) et est également utilisée comme outil d'optimisation sous contrainte, par exemple dans les algorithmes de points intérieurs (Fiacco et McCormick, 1967).

Lorsqu'un algorithme de descente itérative est utilisé pour la minimisation d'un tel critère, la présence de la barrière rend inefficace les procédures standards de recherche de pas (Murray et Wright, 1994; Wright, 1995). De plus, la forme singulière de la barrière rend impossible l'utilisation de la recherche de pas MM quadratique. **La question de la modification de cette méthode en vue de son utilisation pour l'optimisation des fonctions barrières se pose.** Ce deuxième point fait aussi l'objet d'une étude approfondie dans le cadre de ce travail.

## 1.2 Contributions

Les travaux menés pendant cette thèse concernent à la fois les aspects algorithmiques liés à la minimisation de critères et les aspects expérimentaux à travers l'application de ces approches à un problème de reconstruction de signal en spectroscopie RMN.

## Aspects algorithmiques

Cette thèse se place dans le prolongement des travaux (Labat, 2006; Labat et Idier, 2008) pour les aspects algorithmiques. La procédure de recherche de pas MMQ 1D développée dans la thèse de Christian Labat est une méthode itérative simple qui garantit la convergence de l’algorithme de gradient conjugué non linéaire GCNL. Nous étendons ce résultat de convergence à la famille des algorithmes de descente gradient reliés. Une perspective importante concernait l’extension de la méthode MMQ 1D au calcul du pas multi-dimensionnel dans les algorithmes de sous-espace où la nouvelle itérée est obtenue en explorant le critère dans un sous-espace engendré par un ensemble de directions. Nos contributions consistent à :

- Modifier la recherche de pas afin de pouvoir l’implémenter dans le cadre d’un algorithme de sous-espace. La méthode, notée MMQ  $rD$ , est basée sur des approximations majorantes quadratiques du critère à l’intérieur du sous-espace et conserve la simplicité de son homologue scalaire. Son association avec le sous-espace de mémoire de gradient (MG) définit un algorithme de structure similaire au GCNL qui permet de s’affranchir du choix du coefficient de conjugaison.
- Démontrer la convergence des algorithmes de sous-espaces associés à la technique de pas MMQ  $rD$ , dont l’algorithme MG.
- Illustrer expérimentalement en déconvolution d’image le fait que l’algorithme MG + MMQ 2D est plus efficace que l’algorithme GCNL + MMQ 1D.

Les fonctions barrières ne rentrent pas dans le cadre d’application des techniques de pas MM quadratiques, à cause du caractère non borné de leur gradient. Nos contributions sur les aspects algorithmiques de la minimisation des fonctions barrières consistent à :

- Modifier la recherche de pas afin de pouvoir l’implémenter dans le cadre d’une fonction barrière. Nous proposons pour cela la stratégie MMLQ 1D basée sur une fonction majorante log-quadratique permettant de tenir compte de la singularité de la barrière.
- Démontrer la convergence d’une famille large d’algorithmes de descente itérative associés à MMLQ 1D.
- Illustrer expérimentalement le fait que la recherche de pas MMLQ 1D est plus efficace que les méthodes développées dans la littérature utilisant l’interpolation logarithmique ou cubique.

## Spectroscopie RMN

La résonance magnétique nucléaire (RMN) est une méthode permettant d’accéder à des informations structurales aux échelles microscopiques ou moléculaires de produits biologiques. Dans le domaine agro-alimentaire, cette technique est employée pour déterminer la teneur en eau ou en matière grasse d’un échantillon de matière. Nous nous sommes intéressés à l’estimation d’un spectre 2D à partir de mesures RMN. Les difficultés de l’estimation sont liées au caractère instable de ce problème inverse, à la taille importante des données à traiter et à une incertitude de modèle importante liée à la méconnaissance précise d’un paramètre expérimental. Nous avons développé dans cette thèse une stratégie efficace de résolution basée sur la recherche de la solution ayant la plus grande entropie. Cela conduit à la minimisation d’un critère pénalisé contenant une fonction barrière. Nos contributions consistent à :

- Implémenter la recherche de pas MMLQ 1D sur ce cas concret, en association avec une direction de Newton tronqué (TN)
- Accélérer l’algorithme TN + MMLQ 1D par une stratégie de préconditionnement adaptée à la forme séparable du modèle
- Réduire la complexité de l’algorithme en tenant compte des caractéristiques du modèle RMN
- Développer une stratégie simple de réglage du paramètre de régularisation
- Développer une procédure d’estimation du paramètre d’angle de bascule, intervenant dans le modèle direct
- Illustrer expérimentalement l’efficacité de la méthode sur des données RMN réelles

### 1.3 Organisation du document

**Inversion par approches pénalisées** L'exposé débute par une présentation du cadre méthodologique associé aux problèmes inverses de traitement du signal et des images. Nous montrons dans le chapitre 2 qu'une stratégie efficace de résolution consiste à minimiser un critère pénalisé. Le chapitre 3 traite de la minimisation des critères pénalisés et étudie largement les algorithmes de descente itérative. Les stratégies de calcul de pas développées pour accélérer ces algorithmes font l'objet de cette thèse.

**Recherche de pas par Majoration-Minoration quadratique** Le chapitre 4 débute par un tour d'horizon des procédures de recherche de pas utilisées dans la littérature de l'optimisation. Nous introduisons la recherche de pas par Majoration-Minoration quadratique MMQ 1D, développée par (Labat, 2006) pour la minimisation de critères pénalisés par l'algorithme de descente itérative. Cette approche permet d'assurer la convergence des algorithmes de GCNL (Labat et Idier, 2008). Nous étendons ce résultat aux algorithmes *gradient reliés*, incluant les méthodes standards de Newton, de quasi-Newton et Newton tronqué. Nous montrons également dans ce chapitre que la procédure MM permet de définir une nouvelle méthode de calcul de pas multi-dimensionnelle MMQ  $r$ D qui assure la convergence des algorithmes itératifs de sous-espace, incluant les stratégies à super mémoire de gradient (SMG). Les performances des algorithmes GCNL + MMQ 1D et SMG + MMQ  $r$ D sont comparées dans le chapitre 5 sur un problème de déconvolution d'image de grande taille.

**Recherche de pas par Majoration-Minoration log-quadratique** La Majoration-Minoration quadratique ne s'applique pas lorsque le critère pénalisé contient une *barrière*, c'est à dire une fonction à gradient non borné. L'optimisation des fonctions barrières fait l'objet de la deuxième partie de cette thèse. Le chapitre 6 apporte une définition formelle d'une fonction barrière, accompagnée d'exemples concrets issus du traitement du signal. Nous montrons que les difficultés rencontrées lors de la minimisation d'une fonction barrière s'expliquent majoritairement par une mauvaise stratégie de pas. Nous terminons ce chapitre par une description détaillée des solutions proposées de la littérature pour remédier à ce problème de calcul du pas. Le chapitre 7 décrit une contribution majeure de cette thèse, à savoir le développement d'une nouvelle stratégie de pas MM tenant compte du terme de barrière, basée sur une fonction majorante log-quadratique. La convergence des algorithmes classiques de descente itérative associés à cette recherche de pas MMLQ 1D est démontrée. Pour alléger la lecture, les démonstrations des résultats sont reportées dans l'annexe A. Les performances de la méthode sont illustrées dans le chapitre 8 sur un ensemble d'exemples de programmation quadratique résolu par des méthodes de points intérieurs.

**Résolution d'un problème inverse en résonance Magnétique Nucléaire bi-dimensionnelle**

Le chapitre 9 est consacré à l'application de la méthode de pas MMLQ 1D sur un problème réel de traitement du signal, la reconstruction de spectres issus de la Résonance Magnétique Nucléaire bi-dimensionnelle (RMN 2D). Après avoir souligné les difficultés de ce problème inverse, nous présentons les approches pénalisées proposées dans la littérature pour résoudre le problème. L'algorithme d'inversion que nous proposons est basé sur la pénalisation par maximum d'entropie. La résolution, nécessitant la minimisation d'une fonction barrière, est effectuée par un algorithme de Newton tronqué associé à la recherche de pas MMLQ 1D et à un préconditionneur adapté à la forme du critère. Les performances de la méthode sont illustrées sur des exemples réels.

## Publications

### Articles acceptés

- [1] E. Chouzenoux, S. Moussaoui, J. Idier and F. Mariette. **Efficient Maximum Entropy Reconstruction of Nuclear Magnetic Resonance T1-T2 Spectra**. à paraître, *IEEE Transactions on Signal Processing*, 2010.
- [2] E. Chouzenoux, J. Idier and S. Moussaoui. **A Majorize-Minimize strategy for subspace optimization applied to image restoration**, révision mineure, *IEEE Transactions on Image Processing*, 2010.

### Article en préparation

- [3] E. Chouzenoux, S. Moussaoui and J. Idier. **A Majorize-Minimize line search algorithm for barrier functions**, *Tech. Rep.*, IRCCyN, 2009.  
<http://hal.archives-ouvertes.fr/hal-00362304/fr/>.

### Conférences

- [4] E. Chouzenoux, S. Moussaoui, J. Idier and F. Mariette. **Optimization algorithm for maximum entropy 2d nuclear resonance reconstruction**, in *Proc. 35th IEEE ICASSP 2010*, Dallas, Texas, USA, mars 2010.
- [5] E. Chouzenoux, S. Moussaoui, J. Idier and F. Mariette. **Reconstruction d'un spectre RMN 2D par maximum d'entropie**, in *Actes 22e coll. GRETSI*, Dijon, France, septembre 2009.
- [6] E. Chouzenoux, S. Moussaoui and J. Idier. **A Majorize-Minimize line search algorithm for barrier function optimization**, in *Proc. 17th EUSIPCO*, pages 1379-1383, Glasgow, UK, août 2009. EURASIP Press.



## Chapitre 2

# Inversion par approche pénalisée en traitement du signal et de l'image

### Sommaire

---

2.1	Introduction . . . . .	<b>8</b>
2.2	Problème inverse en traitement du signal . . . . .	<b>8</b>
2.2.1	Définition . . . . .	8
2.2.2	Caractère mal posé . . . . .	9
2.2.3	Inversion généralisée . . . . .	10
2.2.4	Régularisation par contrôle de la dimension . . . . .	11
2.3	Régularisation par minimisation d'un critère pénalisé . . . . .	<b>12</b>
2.3.1	Principe . . . . .	12
2.3.2	Formulation du critère d'adéquation aux données . . . . .	12
2.3.3	Modèle <i>a priori</i> . . . . .	14
2.3.4	Réglage du paramètre de régularisation . . . . .	20
2.4	Conclusion . . . . .	<b>20</b>

---

## 2.1 Introduction

Dans divers domaines d'applications (physique, biologie, médecine, économie, par exemple), la grandeur mesurable est représentée sous forme mathématique par un signal. Le traitement du signal a pour objectif d'entreprendre et de développer des techniques de transformation, d'analyse et d'interprétation des signaux. Dans cette thèse, nous nous intéressons à la résolution de problèmes inverses en traitement du signal. Le but d'un problème inverse est de restituer un signal originel à partir de l'observation de signaux physiquement associés et d'informations *a priori*. Nous nous concentrons sur la résolution des problèmes inverses linéaires au sens où le lien entre le signal d'intérêt et les données observées peut être représenté par un modèle linéaire. Nous présentons tout d'abord en section 2.2 les approches non pénalisées qui tentent de reconstruire le signal originel à partir de la seule connaissance des observations. Cette approche naïve conduit à des difficultés qui peuvent être résolues à l'aide du principe de régularisation. Nous décrivons dans la section 2.3 une stratégie efficace de régularisation basée sur la minimisation d'un critère pénalisé.

## 2.2 Problème inverse en traitement du signal

### 2.2.1 Définition

Il existe plusieurs branches particulières du traitement du signal, en fonction de la nature des signaux considérés. En particulier, on distingue le traitement du son, le traitement d'image et le traitement de la vidéo. Selon la nature du signal traité, le traitement du signal peut avoir différentes finalités : la détection, le codage, la compression et enfin la restauration selon des critères qualitatifs physiologiques (liés par exemple à l'écoute ou la visualisation).

Parmi ces applications, certaines s'expriment comme l'estimation d'une quantité d'intérêt  $x$  inaccessible directement mais physiquement reliées à une grandeur mesurable  $y$ . C'est le cas par exemple de la *restauration d'image* qui consiste à estimer une image originelle à partir d'une version dégradée par un bruit, un flou ou un sous-échantillonnage (Boccacci et Bertero, 2002; Frieden, 1975; Jansson, 1997). Citons également la *reconstruction d'image* qui détermine une image originelle à partir de données indirectement liées à cette image par exemple par une opération de projection (Gordon, 1971; Lewitt et Matej, 2003; Zeng, 2001). Le problème de la restitution de  $x$  à partir de  $y$  constitue le *problème inverse* (Bertero et Boccacci, 1998).

La résolution du problème inverse passe tout d'abord par l'élaboration d'un *modèle direct* liant  $y$  et  $x$  à partir d'équations de la physique. Le modèle idéal n'est généralement jamais satisfait par les mesures en raison des erreurs de modélisation, des incertitudes de mesures et des erreurs de discrétisation. Ces perturbations sont prises en compte par un terme de bruit. De façon générique, le modèle direct s'écrit

$$y = K(x) + \text{bruit}, \quad (2.1)$$

où  $K$  est un opérateur qui modélise la relation directe entre les mesures et les inconnues. En restauration d'image,  $K$  représente la dégradation causée par l'instrument optique de mesure. Il s'agit par exemple d'un opérateur de convolution. En reconstruction d'image,  $K$  modélise la transformation effectuée sur l'image, par exemple une projection par transformée de Radon.

Dans cette thèse, nous nous intéressons au cas particulier du modèle linéaire discret

$$\mathbf{y} = \mathbf{K}\mathbf{x} + \boldsymbol{\epsilon}, \quad \mathbf{y} \in \mathbb{R}^m, \mathbf{x} \in \mathbb{R}^n, \mathbf{K} \in \mathbb{R}^{n \times m}, \quad (2.2)$$

schématisé en figure 2.1. Le *problème inverse* associé à (2.2) consiste à estimer le vecteur  $\mathbf{x}$  à partir de  $\mathbf{y}$  et de  $\mathbf{K}$ . Le problème de déconvolution d'image présenté en figure 2.2 ainsi que les applications traitées dans les chapitres 5 et 9 de cette thèse rentrent dans ce cadre.

Dans certaines applications telle que la déconvolution aveugle (Lam et Goodman, 2000), la matrice  $\mathbf{K}$  est inconnue et doit être estimée conjointement à  $\mathbf{x}$ . Dans cette thèse, nous supposons que la détermination de cette matrice a été préalablement effectuée.

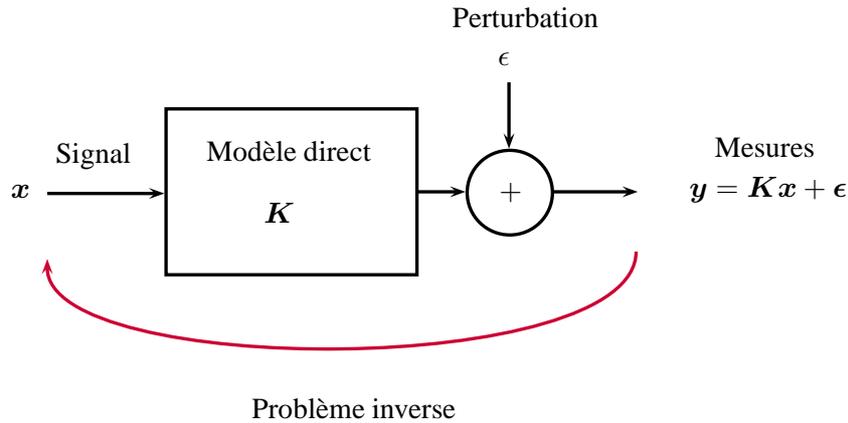


FIGURE 2.1 – Formation d'un problème inverse linéaire

### 2.2.2 Caractère mal posé

La résolution du problème inverse (2.1) est rendue difficile par son caractère *mal posé*. Au début du XX siècle, Hadamard a défini les trois conditions nécessaires pour qu'un problème soit bien posé :

**Définition 2.1.** (*Hadamard, 1902*) La résolution de l'équation (2.1) est dite bien posée si

**Existence :** Pour chaque donnée  $\mathbf{y}$  dans une classe définie  $\mathcal{Y}$ , il existe une solution  $\mathbf{x}$  dans une classe prescrite  $\mathcal{X}$

**Unicité :** La solution est unique dans  $\mathcal{X}$

**Continuité :** La dépendance de  $\mathbf{x}$  par rapport à  $\mathbf{y}$  est continue : lorsque l'erreur sur la donnée  $\mathbf{y}$  tend vers zéro, l'erreur induite sur la solution  $\mathbf{x}$  tend aussi vers zéro

Si la grande majorité des problèmes directs en physique sont bien posés, leurs problèmes inverses associés ne le sont pas. On dit alors qu'ils sont *mal posés*. Les conditions sur l'opérateur  $K(\cdot)$  pour que le problème soit bien posé s'écrivent

**Existence :** L'ensemble des  $\mathbf{y}$  de  $\mathcal{Y}$  sont des images par  $K(\cdot)$  de  $\mathbf{x}$  de  $\mathcal{X}$ , soit  $\mathcal{Y} = \text{Im}(K)$

**Unicité :** Le noyau de  $K(\cdot)$  est réduit au vecteur nul  $\mathbf{x} = \mathbf{0}$ , c'est à dire que l'unique solution de l'équation  $K(\mathbf{x}) = \mathbf{0}$  est  $\text{Ker}(K) = \{\mathbf{0}\}$

**Continuité :** L'image de  $K(\cdot)$  est fermée,  $\text{Im}(K) = \overline{\text{Im}(K)}$ , c'est à dire que toute limite de suites dans  $\text{Im}(K)$  reste dans  $\text{Im}(K)$

Dans le cas linéaire de dimension finie i.e.,  $K(\cdot) = \mathbf{K}$ , la condition de continuité est toujours vérifiée, mais pas forcément celle d'unicité et d'existence. En effet, si le nombre de valeurs singulières  $s$  de  $\mathbf{K} \in \mathbb{R}^{m \times n}$  est inférieur à  $n$ , le noyau de l'opérateur et son image seront respectivement de dimensions  $n - s$  et  $s$ , détruisant la condition d'unicité. De plus, si  $s$  est inférieur à  $n$ , alors l'existence de la solution n'est pas assurée.

Lorsque la matrice d'observation est carrée ( $n = m$ ) et inversible ( $\text{Ker}(\mathbf{K}) = \{\mathbf{0}\}$ ), une inversion directe, au sens de la solution calculée par  $\hat{\mathbf{x}} = \mathbf{K}^{-1}\mathbf{y}$ , peut être envisagée. Dans tous les autres cas, cette inversion directe est inapplicable. La nécessité de définir des solutions bien posées utilisables dans des situations générales est à l'origine des développements de théories alternatives à l'inversion directe, comme l'*inversion généralisée* et surtout la *régularisation* (Tikhonov et Arsenin, 1977; Titterton, 1985b).

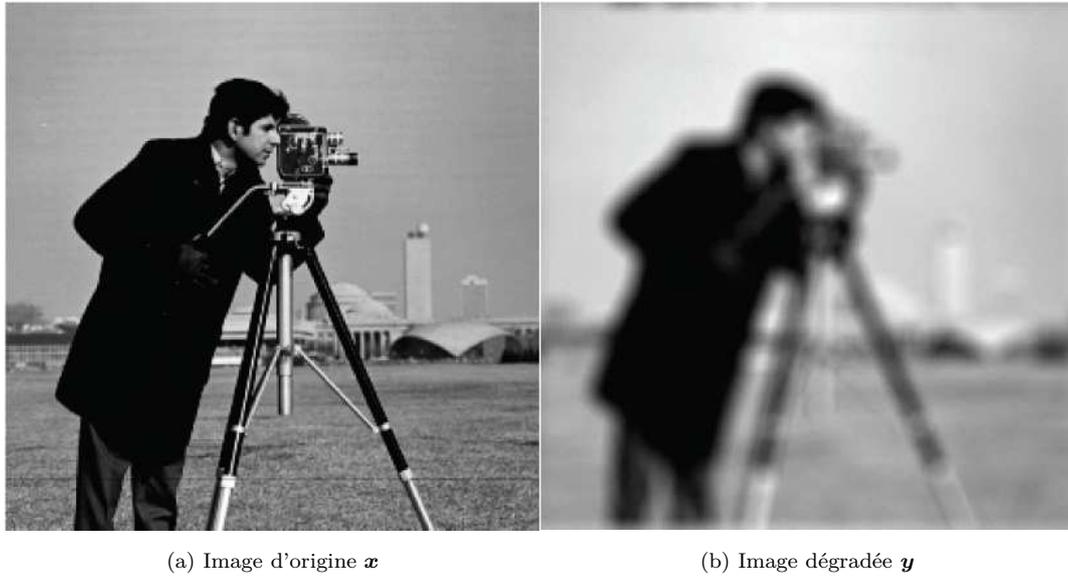


FIGURE 2.2 – Problème de déconvolution de l'image *cameraman* (Soulez, 2008) : L'image d'origine  $\mathbf{x} \in \mathbb{R}^{256 \times 256}$  a été floutée par une opération de convolution circulaire et dégradée par un bruit additif Gaussien blanc centré. Le problème inverse consiste à estimer  $\mathbf{x}$  à partir de l'image  $\mathbf{y}$  et de la matrice de convolution  $\mathbf{K}$ .

## 2.2.3 Inversion généralisée

### 2.2.3.1 Principe

Une approche naturelle pour résoudre le problème inverse linéaire (2.2) consiste à utiliser la méthode des *moindres carrés*, c'est à dire à choisir une solution  $\mathbf{x}^{MC}$  de l'équation normale

$$\mathbf{K}^T \mathbf{K} \mathbf{x} = \mathbf{K}^T \mathbf{y}. \quad (2.3)$$

Si la matrice  $\mathbf{K}^T \mathbf{K}$  n'est pas inversible, alors il n'y a pas unicité de  $\mathbf{x}^{MC}$  et le problème reste mal posé. Pour pallier ce problème, la solution *généralisée*  $\mathbf{x}^\dagger$  est retenue. Elle est définie comme étant la solution de norme minimale de l'ensemble des solutions des moindres carrés. L'existence et l'unicité de  $\mathbf{x}^\dagger$  sont assurées par la convexité de cet ensemble. On peut exprimer  $\mathbf{x}^\dagger$  facilement à l'aide de la décomposition en valeurs singulières (*Singular Value Decomposition* ou SVD) de  $\mathbf{K}$ .

### 2.2.3.2 Expression en fonction de la SVD

La SVD (Golub et Van Loan, 1996) consiste à écrire  $\mathbf{K}$  sous la forme

$$\mathbf{K} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \quad (2.4)$$

avec  $\mathbf{U} = (\mathbf{u}_j)$ ,  $\mathbf{V} = (\mathbf{v}_i)$  des matrices orthogonales et  $\mathbf{\Sigma}$  une matrice diagonale contenant les valeurs singulières de  $\mathbf{K}$  :

$$\Sigma_{ji} = \begin{cases} \sigma_i & \text{si } j = i \leq \min(n, m) \\ 0 & \text{sinon} \end{cases}. \quad (2.5)$$

L'inverse généralisée s'écrit

$$\mathbf{x}^\dagger = \sum_{i=1}^n \alpha_i \mathbf{v}_i \quad \text{avec} \quad \alpha_i = \begin{cases} \mathbf{u}_i^T \mathbf{y} / \sigma_i & \text{si } \sigma_i > 0 \\ 0 & \text{si } \sigma_i = 0 \end{cases} \quad (2.6)$$

L'expression (2.6) définit  $\mathbf{x}^\dagger$  de façon unique. Par conséquent, la solution de l'inverse généralisée satisfait bien le caractère bien posé de Hadamard. Cependant, nous allons voir que cette solution n'est toujours pas satisfaisante.

### 2.2.3.3 Sensibilité numérique de l'inverse généralisée

La condition de continuité est une condition nécessaire, mais en pratique non suffisante, de stabilité numérique de la solution généralisée. En effet, cette solution manque de robustesse au sens où une petite perturbation dans les données peut engendrer une grande perturbation dans la solution. Une perturbation additive  $\delta\mathbf{y}$  sur les observables  $\mathbf{y}$  entraîne une perturbation additive  $\delta\mathbf{x}$  sur la solution  $\mathbf{x}^\dagger$  comme l'illustre le théorème suivant.

**Théorème 2.1.** *Si  $\delta\mathbf{y}$  est une faible perturbation sur l'observation  $\mathbf{y}$  et  $\delta\mathbf{x}^\dagger$  la perturbation induite sur la valeur de  $\mathbf{x}^\dagger$ , alors*

$$\frac{\|\delta\mathbf{x}^\dagger\|}{\|\mathbf{x}^\dagger\|} \leq \kappa(\mathbf{K}) \frac{\|\delta\mathbf{y}\|}{\|\mathbf{y}\|} \quad (2.7)$$

avec  $\kappa(\mathbf{K}) = \frac{\sigma_{\max}}{\sigma_{\min}}$ , fonction des valeurs singulières extrêmes de  $\mathbf{K}$ .

La stabilité de la solution inverse généralisée dépend de la valeur du coefficient  $\kappa(\mathbf{K})$ , appelé *conditionnement* de la matrice  $\mathbf{K}$  (ou conditionnement du problème (2.2)). L'influence de l'erreur sur l'observation est d'autant plus marquée sur la solution de l'inverse généralisé que la valeur de  $\kappa(\mathbf{K})$  est grande. De plus, il est à noter que l'inégalité (2.7) peut devenir une égalité pour certains couples  $(\mathbf{y}, \delta\mathbf{y})$ . Par conséquent, les conditions de Hadamard ne sont pas suffisantes pour assurer la robustesse de la solution, un problème bien posé pouvant être mal conditionné.

En résumé, un problème inverse peut être :

- mal posé,
- bien posé (ce qui est le cas lors de l'utilisation de l'inverse généralisée en dimension finie) mais mal conditionné.

La régularisation permet d'apporter une solution satisfaisante à ces deux types de problèmes.

## 2.2.4 Régularisation par contrôle de la dimension

D'après le théorème 2.1, le mauvais conditionnement du problème est dû à la présence de petites valeurs singulières dans  $\mathbf{K}$  dont l'inversion a pour effet d'amplifier les perturbations. La régularisation par contrôle de la dimension (*Truncated Singular Value Decomposition* ou TSVD) consiste à tronquer le développement (2.4) à partir d'un certain ordre  $\sigma_{\text{seuil}}$  (Nashed, 1981). La solution TSVD s'exprime à partir de la décomposition de  $\mathbf{K}$  en éléments singuliers :

$$\mathbf{x}^{TSVD} = \sum_{i=1}^n \alpha_i \mathbf{v}_i \quad \text{avec} \quad \alpha_i = \begin{cases} \mathbf{u}_i^T \mathbf{y} / \sigma_i & \text{si } \sigma_i > \sigma_{\text{seuil}} \\ 0 & \text{sinon} \end{cases} \quad (2.8)$$

La troncature permet d'assurer la stabilité numérique de  $\mathbf{x}^{TSVD}$ . Cependant, elle présente deux inconvénients. Tout d'abord, cette méthode nécessite de choisir le seuil de troncature  $\sigma_{\text{seuil}}$ . Ce choix est difficile surtout lorsque le spectre des valeurs singulières est régulier. D'autre part, tronquer signifie que l'on renonce à restaurer les composantes hautes fréquences correspondant aux valeurs singulières  $\sigma_i \leq \sigma_{\text{seuil}}$ . Cela est problématique lorsque l'on sait *a priori* que l'objet à restaurer contient des zones irrégulières comme des contours dans une image. L'intégration de connaissances *a priori* constitue l'une des motivations principales de la méthode de régularisation par minimisation d'un critère pénalisé.

## 2.3 Régularisation par minimisation d'un critère pénalisé

### 2.3.1 Principe

Les problèmes inverses mal conditionnés sont intrinsèquement instables et leur inversion naïve à partir de la seule connaissance de l'observation  $\mathbf{y}$  conduit systématiquement à des difficultés. Le principe de *régularisation* consiste à renforcer certaines propriétés souhaitables de la solution, sur la base de connaissances *a priori* (Demoment, 1989). La solution du problème inverse est définie comme le minimiseur d'un critère composite

$$F(\mathbf{x}) = J(\mathbf{x}) + \lambda R(\mathbf{x}). \quad (2.9)$$

où  $J$ , appelé *critère d'adéquation aux données*, mesure l'erreur résiduelle entre  $\mathbf{y}$  et  $\mathbf{K}\mathbf{x}$ , et  $R$ , appelé *critère de régularisation*, introduit et contrôle l'*a priori* sur la solution. Le paramètre  $\lambda$  règle le compromis entre l'adéquation aux données et la fidélité aux connaissances *a priori*. Fixer  $\lambda = 0$  correspond à une solution non régularisée construite uniquement à partir des données. À l'inverse,  $\lambda = \infty$  revient à considérer uniquement la fidélité à l'*a priori*.

### 2.3.2 Formulation du critère d'adéquation aux données

#### 2.3.2.1 Moindres carrés pondérés

Repartons de l'équation d'observation (2.2). Puisque la solution exacte est inaccessible à partir de la seule connaissance des données  $\mathbf{y}$  et du modèle physique  $\mathbf{K}$ , une démarche naturelle consiste à chercher une solution qui permet une certaine adéquation entre les données  $\mathbf{y}$  et la sortie du modèle  $\mathbf{K}\mathbf{x}$ . De manière formelle, cela revient à choisir

$$J(\mathbf{x}) = D(\mathbf{K}\mathbf{x}, \mathbf{y}) \quad (2.10)$$

où  $D$  est une mesure vérifiant

$$D(\mathbf{u}, \mathbf{u}) = 0, \quad D(\mathbf{u}, \mathbf{v}) \geq 0 \quad (2.11)$$

qu'on choisira souvent de façon à ce que  $J$  soit strictement convexe et coercive (Titterton, 1985a). La fonction  $D$  va rappeler  $\mathbf{K}\mathbf{x}$  vers la cible  $\mathbf{y}$ . Le choix le plus courant pour mesurer l'adéquation aux données est la distance euclidienne (ou norme  $\ell_2$ )

$$D(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|_{\mathbf{P}}^2 = (\mathbf{u} - \mathbf{v})^T \mathbf{P} (\mathbf{u} - \mathbf{v}) \quad (2.12)$$

où  $\mathbf{P} = \mathbf{L}^T \mathbf{L}$  est une matrice symétrique définie positive (SPD), choisie pour renforcer certaines caractéristiques de la solution. Ce choix conduit au critère des *moindres carrés pondérés* défini par :

$$J(\mathbf{x}) = \|\mathbf{L}(\mathbf{y} - \mathbf{K}\mathbf{x})\|^2. \quad (2.13)$$

Lorsque  $\mathbf{P}$  est la matrice identité, la minimisation du critère non régularisé (2.13) conduit à la solution des moindres carrés  $\mathbf{x}^{\text{MC}}$  définie en section 2.2.3.1.

#### 2.3.2.2 Divergence de Kullback

La distance euclidienne n'est pas l'unique manière de mesurer la distance entre deux vecteurs. La norme  $\ell_p$  est également employée en particulier dans des applications relevant de l'estimation robuste (Yarlagadda *et al.*, 1985). D'autre part, lorsque les deux vecteurs sont positifs, on peut les identifier après normalisation comme des densités de probabilités et construire le critère à partir de mesures de distances entre lois de probabilité (Csiszàr, 1991). Une mesure de distance probabiliste classique est la divergence de Kullback, basée sur la notion d'entropie introduite dans (Jaynes, 1982) en théorie de l'information : La divergence de Kullback entre deux densités de probabilités  $\mathbf{p}$  et  $\mathbf{q}$  est définie par

$$D(\mathbf{p}, \mathbf{q}) = \sum_i p_i \log \frac{p_i}{q_i} + q_i - p_i, \quad (2.14)$$

sous réserve que  $q_i > 0$  pour tout  $i$  tel que  $p_i > 0$ . Elle est couramment utilisée pour construire le critère d'adéquation aux données en tomographie d'émission (De Pierro, 1995; Lange *et al.*, 1987; Vardi *et al.*, 1985) sous la forme inversée

$$J(\mathbf{x}) = D(\mathbf{y}, \mathbf{K}\mathbf{x}) = \sum_i y_i \log \frac{y_i}{[\mathbf{K}\mathbf{x}]_i} - y_i + [\mathbf{K}\mathbf{x}]_i. \quad (2.15)$$

La divergence de Kullback se généralise en deux principales classes de distances entre probabilités, les  $\phi$ -divergences et les divergences de Bregman. Ces fonctions sont également appelées *entropies généralisées*. Elles sont définies à partir d'une fonction  $\phi : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  continue, convexe et différentiable.

**Divergences de Bregman.** Les divergences de Bregman (Ben-Tal *et al.*, 1989; Bregman, 1967; Censor et Lent, 1981) sont définies par

$$D(\mathbf{p}, \mathbf{q}) = \sum_i \phi(p_i) - \phi(q_i) - \dot{\phi}(q_i)(p_i - q_i). \quad (2.16)$$

Pour  $\phi(u) = u \log u - u$ , on retrouve la divergence de Kullback et pour  $\phi(u) = -\log u$ , on obtient la divergence d'Itakura-Saito :

$$D(\mathbf{p}, \mathbf{q}) = \sum_i \frac{p_i}{q_i} - \log \frac{p_i}{q_i} - 1, \quad (2.17)$$

définie lorsque  $p_i, q_i > 0$  pour tout  $i$ , utilisée par exemple en imagerie astronomique (Cao *et al.*, 1999).

**$\phi$  divergences.** Les  $\phi$  divergences, ou distances d'Ali-Silvey-Csiszàr (Ali et Silvey, 1966; Csiszàr, 1963, 1967), sont définies par

$$D(\mathbf{p}, \mathbf{q}) = \sum_i q_i \phi(p_i/q_i) \quad (2.18)$$

avec

$$\lim_{u \rightarrow 0^+} \dot{\phi}(u) = -\infty, \quad (2.19)$$

$$\phi(1) = \dot{\phi}(1) = 0, \quad (2.20)$$

$$\ddot{\phi}(1) > 0. \quad (2.21)$$

Pour  $\phi(u) = u \log u - u + 1$ , on retrouve à nouveau la divergence de Kullback. De nombreuses autres distances bien connues sont éléments de cette classe. On pourra consulter à ce sujet les références (Ben-Tal *et al.*, 1991, 1989; Taneja et Krumar, 2004).

### 2.3.2.3 Lien avec le maximum de vraisemblance

L'introduction d'un cadre probabiliste permet d'interpréter le choix du terme d'adéquation aux données comme une modélisation statistique de la composante de bruit  $\epsilon$ . Cela permet en particulier de justifier l'utilisation d'une divergence de Kullback en tomographie d'émission.

Partant de l'hypothèse que le bruit  $\epsilon$  constitue une réalisation d'un vecteur aléatoire  $\mathbf{E}$  qui admet une loi à densité  $f_{\mathbf{E}}(\epsilon)$ , on en déduit que le vecteur des observées  $\mathbf{y}$  constitue une réalisation d'un vecteur aléatoire  $\mathbf{Y}$  dont la densité de probabilité se déduit de celle du bruit

$$f_{\mathbf{Y}}(\mathbf{y}; \mathbf{x}) = f_{\mathbf{E}}(\mathbf{y} - \mathbf{K}\mathbf{x}). \quad (2.22)$$

Cette densité de probabilité résume à elle seule l'information contenue dans les données sur l'objet déterministe  $\mathbf{x}$ . L'estimation au sens du maximum de vraisemblance consiste à définir la solution

comme le maximiseur  $\mathbf{x}^{MV}$  de la fonction  $f_{\mathbf{Y}}(\mathbf{y}; \mathbf{x})$ , appelée par la suite vraisemblance des données et notée  $V(\mathbf{y}; \mathbf{x})$ , ou de façon équivalente

$$\mathbf{x}^{MV} = \arg \min_{\mathbf{x}} -\log V(\mathbf{y}; \mathbf{x}). \quad (2.23)$$

Supposons que la loi du bruit suive une densité gaussienne centrée

$$f_{\mathbf{E}}(\epsilon) = \frac{1}{Z} \exp \left\{ -\frac{1}{2} \epsilon^T \mathbf{C}^{-1} \epsilon \right\} \quad (2.24)$$

où  $Z$  et  $\mathbf{C} \in \mathbb{R}^{n \times n}$  sont respectivement la constante de normalisation et la matrice de covariance symétrique définie positive. Alors le vecteur des observées  $\mathbf{y}$  constitue une réalisation d'un vecteur aléatoire  $\mathbf{Y}$  dont la densité de probabilité s'écrit

$$f_{\mathbf{Y}}(\mathbf{y}; \mathbf{x}) = \frac{1}{Z} \exp \left\{ -\frac{1}{2} (\mathbf{y} - \mathbf{K}\mathbf{x})^T \mathbf{C}^{-1} (\mathbf{y} - \mathbf{K}\mathbf{x}) \right\}. \quad (2.25)$$

Nous obtenons donc que trouver  $\mathbf{x}^{MV}$  revient à minimiser la fonction d'adéquation aux données des moindres carrés pondérés (2.13), avec  $\mathbf{P} = \mathbf{C}^{-1}$ .

De manière générale, on peut associer une vraisemblance à une fonction d'adéquation aux données pour peu que la loi de probabilité ci-dessous ait effectivement un sens

$$V(\mathbf{y}; \mathbf{x}) = \frac{1}{Z} \exp \{ -J(\mathbf{x})/T \} \quad (2.26)$$

avec  $Z$  la constante de normalisation et  $T$  un paramètre de température. Réciproquement, les solutions du maximum de vraisemblance correspondent aux solutions non régularisées minimisant le terme d'adéquation aux données tel que

$$-\log V(\mathbf{y}; \mathbf{x}) = J(\mathbf{x})/T. \quad (2.27)$$

Nous avons dit dans la section 2.3.2.2 que la divergence de Kullback était classiquement utilisée comme terme d'adéquations aux données en tomographie d'émission. Le cadre probabiliste permet de justifier un tel choix. Dans cette modalité d'imagerie, ainsi qu'en tomographie de transmission à faible dose ou dans certaines applications en astronomie, la faible quantité de mesures conduit à un modèle de bruit Poisson (Bertero *et al.*, 2009; Snyder *et al.*, 1993; Vardi *et al.*, 1985). En effet, le processus d'acquisition correspond à un comptage de particules. Lorsque le nombre de particules comptées est faible, les mesures  $y_j$  s'interprètent comme la réalisation d'un vecteur aléatoire  $[\mathbf{Y}]_j$  suivant la loi de Poisson de moyenne  $[\mathbf{K}\mathbf{x}]_j$ . La densité de probabilité du vecteur aléatoire  $\mathbf{Y}$  est

$$V(\mathbf{y}; \mathbf{x}) = \prod_{j=1}^m \frac{([\mathbf{K}\mathbf{x}]_j)^{y_j}}{y_j!} \exp(-[\mathbf{K}\mathbf{x}]_j). \quad (2.28)$$

La solution au sens du maximum de vraisemblance minimise le critère

$$J(\mathbf{x}) = -\log V(\mathbf{y}; \mathbf{x}) \quad (2.29)$$

$$= \sum_{j=1}^m (-y_j \log[\mathbf{K}\mathbf{x}]_j + [\mathbf{K}\mathbf{x}]_j + \log(y_j!)). \quad (2.30)$$

correspondant, à une constante additive près, à la divergence de Kullback entre  $\mathbf{y}$  et  $\mathbf{K}\mathbf{x}$ .

### 2.3.3 Modèle *a priori*

Une stabilisation correcte de l'inversion passe par l'ajout d'un terme de pénalisation  $R(\mathbf{x})$  modélisant un certain comportement attendu de l'objet à estimer (Idier, 2008). Nous présentons tout d'abord une classe de fonctions de pénalisations issue d'un *a priori* de régularité sur la solution. La fonction  $R$  somme les variations de  $\mathbf{x}$  à travers un opérateur de régularisation.

### 2.3.3.1 Méthode de Tikhonov - Mesure de régularité

La méthode de Tikhonov (Tikhonov, 1963) consiste à résoudre le problème continu (2.1) en minimisant le critère des moindres carrés pénalisé

$$\|\mathbf{y} - K(x(t))\|^2 + \lambda R(x(t)) \quad (2.31)$$

où  $R$  est appelée fonction stabilisante de Tikhonov et s'exprime par

$$R(x(t)) = \int \sum_{l=0}^d \alpha_l(t) \left( \frac{d^l x(t)}{dt^l} \right)^2 dt. \quad (2.32)$$

Dans cette expression, les fonctions  $\alpha_l(t)$  sont des fonctions réelles, non négatives et continues sur l'intervalle d'intégration. Le stabilisateur est dit d'ordre  $d$  car il contient la  $d$ ème dérivée de  $x(t)$ . Cette forme de pénalisation permet d'introduire un *a priori* sur la douceur de la solution. En effet, cela revient à rechercher des solutions dans l'espace des fonctions admettant des dérivées généralisées jusqu'à l'ordre  $d$ .

La formulation générale de Tikhonov est rarement utilisée dans le cadre de la résolution de problèmes inverses discrets et on emploie plutôt la forme simplifiée définie dans (Tikhonov et Arsénine, 1976)

$$R(\mathbf{x}) = \|\mathbf{D}\mathbf{x}\|^2 \quad (2.33)$$

avec  $\mathbf{D}$  la matrice de différentiation de taille  $n \times n$ , typiquement d'ordre  $d = 0, 1$  ou  $2$ . Lorsque  $d = 0$ , l'association de  $R$  au critère des moindres carrés s'interprète comme la relaxation de la contrainte d'être exactement la solution des moindres carrés  $\mathbf{x}^{MC}$  et assurer à la solution d'avoir une norme minimale. La solution est donc reliée à la solution généralisée  $\mathbf{x}^\dagger$ . En reconstruction d'image, le choix le plus répandu consiste à utiliser la dérivée première  $d = 1$ , ce qui a pour effet de favoriser les solutions constantes par morceaux. La figure 2.3 illustre les solutions obtenues pour  $d = 0$  et  $d = 1$ , sur le problème de déconvolution d'image. Sur cet exemple, le choix  $d = 1$  fournit un meilleur résultat du point de vue de la qualité visuelle.



(a) Régularisation d'ordre 0

(b) Régularisation d'ordre 1

FIGURE 2.3 – Déconvolution de l'image *cameraman* par minimisation du critère des moindres carrés pénalisé par la régularisation de Tikhonov (Soulez, 2008).

Une analyse fréquentielle montre que la pénalisation de Tikhonov entraîne un lissage systématique des hautes fréquences de l'objet, or dans certaines applications comme la détection de

contours en traitement d'images, il est désirable d'avoir des solutions non lisses. Pour faire face à cette perte de résolution, des pénalisations non quadratiques préservant les discontinuités ont été proposées

$$R(\mathbf{x}) = \sum_c \phi([\mathbf{D}\mathbf{x}]_c) \quad (2.34)$$

où  $\phi$  est la *fonction de pondération*. De nombreux candidats pour le choix de  $\phi$  ont été proposés. Il s'agit de remplacer la quadratique par une fonction mieux adaptée à la forme *a priori* de l'objet. Présentons quelques-unes de ces fonctions  $\phi$  non quadratiques en les classant en deux catégories distinctes suivant qu'elles soient convexes ou non. Cette séparation conditionnera en pratique la mise en œuvre de la stratégie d'optimisation du critère pénalisé.

**Pondérations convexes.** Les fonctions de pondération convexes non quadratiques sont classiquement des fonctions paires et coercives. Elles présentent un accroissement asymptotique moins important que la parabole tout en restant convexe. Par conséquent, la fonction de régularisation associée (2.34) est également convexe. Un sous ensemble notable rassemble les pondérations  $\ell_2 - \ell_1$ , c'est à dire se comportant de manière quadratique proche de l'origine et de manière linéaire à l'infini. Les plus connues sont la fonction de Huber (Huber, 1981) et la fonction hyperbolique (Charbonnier *et al.*, 1997). Ces fonctions sont au moins deux fois continûment différentiables et s'écrivent respectivement

$$\phi(u) = \begin{cases} u^2 & \text{si } |u| < \delta \\ \delta|u| - \frac{\delta^2}{2} & \text{sinon} \end{cases} \quad (2.35)$$

et

$$\phi(u) = \sqrt{u^2 + \delta^2}. \quad (2.36)$$

D'autres pénalisations convexes sont présentées dans (Lange, 1990; Zibulevsky et Elad, 2010).

Les fonctions  $\ell_2 - \ell_1$  dépendent d'un paramètre  $\delta$  positif qui conditionne le passage du régime quadratique au régime linéaire. Une autre catégorie importante de pondérations convexes sont les fonctions  $\ell_p$  introduites par (Bouman et Sauer, 1996). Elles sont définies par

$$\phi(u) = |u|^p, \quad 1 \leq p < 2. \quad (2.37)$$

Lorsque  $p = 1$ , on obtient la norme  $\ell_1$  non différentiable en 0. La non différentiabilité a pour effet de favoriser la parcimonie de  $\mathbf{D}\mathbf{x}$ . La pondération  $\ell_1$  est principalement utilisée en compressed sensing (Donoho, 2006; Figueiredo *et al.*, 2007) en association avec la différentiation d'ordre 0 (i.e  $\mathbf{D} = \mathbf{I}$ ). En restauration d'image, elle fournit de bons résultats en association avec un opérateur de gradient (Rudin *et al.*, 1992), ce choix conduisant à la régularisation par variation totale.

La figure 2.4 illustre différentes fonctions de pondération convexes. La solution du problème de déconvolution d'image est illustrée en figure 2.5 pour plusieurs fonctions de régularisation convexes associées à l'opérateur de différentiation d'ordre 1.

**Pondérations non convexes.** Les fonctions de régularisation non convexes les plus courantes sont les fonctions  $\ell_2 - \ell_0$ . Elles ont un comportement quadratique proche de l'origine et sont asymptotiquement constantes. La quadratique tronquée

$$\phi(u) = \min(u^2, \delta^2) \quad (2.38)$$

introduite par (Geman et Geman, 1984) appartient à cette famille. Une variante plus régulière est introduite dans (Geman et McClure, 1987) :

$$\phi(u) = \frac{u^2}{u^2 + \delta^2}. \quad (2.39)$$

Comme dans le cas convexe, ces fonctions de pondération dépendent d'un paramètre  $\delta$  correspondant au seuil entre le régime quadratique et le régime asymptotique borné.

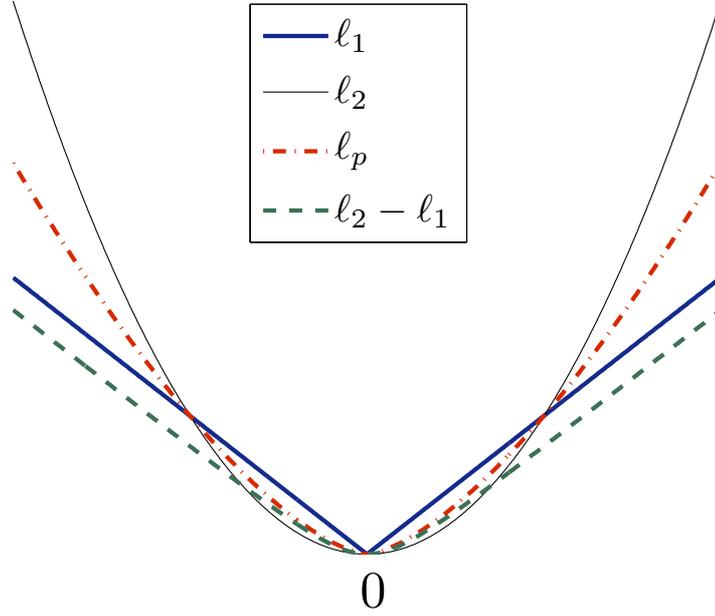


FIGURE 2.4 – Exemples de fonctions de pondération convexes. Proche de 0, la fonction  $\ell_2 - \ell_1$  se comporte comme une quadratique. A l'infini, elle tend vers une fonction  $\ell_1$ .

### 2.3.3.2 Critère pénalisé généralisé

La pénalisation (2.34) ne permet pas de prendre en compte certains modèles et dans certaines de situations, il est nécessaire de définir la forme pénalisée généralisée suivante (Aubert et Kornprobst, 2002; Nikolova, 2002)

$$R(\mathbf{x}) = \sum_{c=1}^C \phi(\|\mathbf{V}_c \mathbf{x} - \boldsymbol{\omega}_c\|) \quad (2.40)$$

où  $\mathbf{V}_c \in \mathbb{R}^{p \times n}$ ,  $\boldsymbol{\omega}_c \in \mathbb{R}^p$ , pour  $c = 1, \dots, C$ . Le vecteur  $\boldsymbol{\omega}_c \in \mathbb{R}^p$  joue un terme de rappel. En considérant  $\boldsymbol{\omega}_c = \mathbf{0}$  et  $p = 1$ , on retrouve bien (2.34).

En posant  $D(\mathbf{u}, \mathbf{v}) = \phi(\|\mathbf{u} - \mathbf{v}\|)$ , le critère (2.40) peut être vu sous la forme

$$R(\mathbf{x}) = \sum_{c=1}^C D(\mathbf{V}_c \mathbf{x}, \boldsymbol{\omega}_c) \quad (2.41)$$

Ainsi, la régularisation (2.40) a comme effet de minimiser les distances entre les vecteurs  $\mathbf{V}_c \mathbf{x}$  et des vecteurs fixes  $\boldsymbol{\omega}_c$  se comportant comme un terme de rappel. Cette formulation conduit à deux formes particulières de pénalisation.

**Distance à un objet cible.** Lorsque l'on connaît une estimation *a priori*  $\bar{\mathbf{x}}$  de l'objet à reconstruire, une façon simple de régulariser consiste à définir  $C = 1$ ,  $\mathbf{V}_1 = \mathbf{I}$  et  $\boldsymbol{\omega}_c = \bar{\mathbf{x}}$ . Le plus classique est de choisir la distance euclidienne associée ou non à une matrice de pondération. Un autre choix courant en reconstruction d'image est de prendre l'une des pénalisations suivantes, issues de la divergence de Kullback,

$$R(\mathbf{x}) = \sum_i x_i \log \frac{x_i}{\bar{x}_i} + \bar{x}_i - x_i \quad (2.42)$$

$$R(\mathbf{x}) = \sum_i \bar{x}_i \log \frac{\bar{x}_i}{x_i} + x_i - \bar{x}_i, \quad (2.43)$$

appelés d'entropies croisées directe et inverse (Byrne, 1993; Le Besnerais *et al.*, 1999). Ces critères sont utilisés afin de renforcer la positivité de la solution. En effet, la singularité du logarithme en zéro impliquera la positivité des minimiseurs du critère pénalisé. Notons que lorsque la cible  $\bar{\mathbf{x}}$  est trop éloignée de la solution du problème, pénaliser la distance entre  $\mathbf{x}$  et  $\bar{\mathbf{x}}$  peut conduire à de mauvais résultats.

**Distance à un objet variable.** Plaçons nous dans le cas où  $\mathbf{x}$  représente une image. Un *a priori* de douceur peut être apporté en minimisant

$$R(\mathbf{x}) = \sum_{c=1}^C D(\mathbf{V}_c \mathbf{x}, \mathbf{x}) \quad (2.44)$$

où  $\mathbf{V}_c \in \mathbb{R}^{n \times n}$ ,  $c = 1 \dots C$ , sont des opérateurs de décalage de l'image. Lorsque  $D$  est la distance euclidienne, on retrouve la pénalisation (2.34) avec  $\mathbf{D}$  opérateur de différentiation d'ordre 1. D'autres fonctions distances peuvent être considérées comme la distance de Kullback et la distance de Itakura Saito (O'Sullivan, 1995). Notons que le choix d'une distance appartenant à la famille des  $\phi$ -divergences présente l'avantage de contraindre la solution à être *positive*, à cause de la singularité de cette distance en 0. Par exemple, dans le cas d'une distance de Kullback, on obtient la régularisation

$$R(\mathbf{x}) = \sum_c \sum_i x_i \log \frac{x_i}{[\mathbf{V}_c \mathbf{x}]_i} + [\mathbf{V}_c \mathbf{x}]_i - x_i \quad (2.45)$$

qui, par son terme logarithmique, force le minimiseur du critère pénalisé à l'intérieur de l'orthant positif.

### 2.3.3.3 Maximum d'entropie

Une stratégie de régularisation également très utilisée en reconstruction du signal est la méthode du *maximum d'entropie* qui consiste à rechercher la solution comme minimiseur d'un critère des moindres carrés pénalisé par un terme d'entropie de Shannon

$$F(\mathbf{x}) = \|\mathbf{y} - \mathbf{K}\mathbf{x}\|^2 + \lambda \sum_i x_i \log x_i, \quad (2.46)$$

ou de Burg

$$F(\mathbf{x}) = \|\mathbf{y} - \mathbf{K}\mathbf{x}\|^2 - \lambda \sum_i \log x_i. \quad (2.47)$$

En traitement de l'information, l'entropie s'interprète comme une mesure de la quantité d'information dans les réalisations d'une variable aléatoire (Jaynes, 1982). Même si dans le cadre de la résolution d'un problème inverse la justification théorique de l'entropie n'est pas établie de façon formelle (Donoho *et al.*, 1990; Nityananda et Narayan, 1982), cette pénalisation fournit des solutions très satisfaisantes notamment en reconstruction d'images (Byrne *et al.*, 1983; Danielle et Gull, 1980; Gull et Skilling, 1984). Elle est donc utilisée dans de nombreuses applications telles que l'imagerie astronomique (Pichon et Thiébaud, 1998), la tomographie à rayons X (Dusaussoy et Abdou, 1995; Mohammad-Djafari et Demoment, 1988a) et l'imagerie par résonance magnétique (IRM) (Johnson et McGarry, 2003; Wang et Zhao, 2000). La raison principale de cet engouement est que la pénalisation entropique permet d'assurer la positivité de l'objet reconstruit à cause de sa singularité sur le bord de l'orthant positif.

### 2.3.3.4 Parcimonie dans un dictionnaire

Une autre façon de modéliser une image ou un signal consiste à considérer qu'il existe une base dans laquelle sa représentation est creuse. Ceci signifie que  $\mathbf{x}$  peut s'exprimer comme la somme pondérée d'un nombre fini (faible) d'atomes élémentaires d'un dictionnaire représenté par une matrice  $\mathbf{W}$  dont les colonnes sont constituées des fonctions de bases discrétisées :

$$\mathbf{x} = \mathbf{W}\mathbf{z}. \quad (2.48)$$

Le terme d'*a priori* est choisi afin de favoriser la parcimonie de la décomposition  $\mathbf{z}$ , c'est à dire de maximiser le nombre de termes  $z_i$  nuls. Deux approches peuvent être utilisées (Elad *et al.*, 2007).

L'approche par *analyse* consiste à appliquer l'*a priori* de parcimonie sur le produit  $\mathbf{W}^*\mathbf{x}$ , correspondant à la transformée de  $\mathbf{x}$  dans le dictionnaire (Aujol et Chambolle, 2005; Belge *et al.*, 2000; Chambolle *et al.*, 1998). La solution est obtenue en minimisant le critère pénalisé

$$F(\mathbf{x}) = D(\mathbf{K}\mathbf{x}, \mathbf{y}) + \lambda R(\mathbf{W}^*\mathbf{x}) \quad (2.49)$$

où  $R$  est typiquement une norme  $\ell_p$ ,  $p \geq 1$  afin de garantir la convexité du critère.

L'approche par *synthèse*, plus récente, consiste à travailler directement dans le domaine des coefficients, en minimisant le critère pénalisé

$$F(\mathbf{z}) = D(\mathbf{K}\mathbf{W}\mathbf{z}, \mathbf{y}) + \lambda R(\mathbf{z}) \quad (2.50)$$

où  $R$  favorise l'apparition de composantes nulles de  $\mathbf{z}$  (Chen *et al.*, 1998; Figueiredo *et al.*, 2007; Mallat et Zhang, 1993). Un choix simple consiste à définir  $R$  comme la norme  $\ell_0$  de  $\mathbf{z}$ . Cependant, minimiser une telle norme est un problème combinatoire connu pour être NP-complet. Il ne peut donc pas être résolu pour des problèmes de très grande dimension. On remplace donc souvent la norme  $\ell_0$  par son enveloppe convexe qui est la norme  $\ell_1$  (Chen *et al.*, 1998). Dans certains cas, on peut montrer que les deux normes fournissent des solutions identiques (Tropp, 2006).

La norme  $\ell_1$  conduit à un critère pénalisé non différentiable dont la minimisation peut être difficile. Dans le cadre de la reconstruction d'image par synthèse, Zibulevsky *et coll.* (Elad *et al.*, 2007; Zibulevsky et Elad, 2010) proposent de remplacer  $|u|$  par l'approximation différentiable paramétrée par un seuil  $\delta > 0$  :

$$\phi(u) = |u| - \delta \log(1 + |u|/\delta). \quad (2.51)$$

La question de savoir laquelle parmi l'approche par analyse et l'approche par synthèse fournit la meilleure reconstruction est un problème ouvert à ce jour (Chaâri *et al.*, 2009; Elad *et al.*, 2007; Weiss *et al.*, 2010). Du point de vue de l'optimisation, l'approche par synthèse est souvent préférée car elle mène à une fonction de pénalisation séparable, permettant l'application d'algorithmes simples de type seuillage itératif (Beck et Teboulle, 2009; Daubechies *et al.*, 2004).

### 2.3.3.5 Lien avec la loi *a priori*

Nous avons établi précédemment que les solutions minimisant le critère d'adéquation aux données correspondent aux solutions du maximum de vraisemblance. Le cadre probabiliste permet également d'établir un lien entre la fonction de régularisation et une loi *a priori*  $f_{\mathbf{X}}(\mathbf{x})$ , où  $\mathbf{x}$  constitue à présent une réalisation du vecteur aléatoire  $\mathbf{X}$ .

Sous réserve que l'on puisse établir des liens tels que

$$\begin{aligned} -\log V(\mathbf{y}; \mathbf{x}) &\leftrightarrow J(\mathbf{x}) \\ \exists \lambda \geq 0 \text{ tel que } -\log f_{\mathbf{X}}(\mathbf{x}) &\leftrightarrow R(\mathbf{x}) \end{aligned} \quad (2.52)$$

alors l'ensemble des solutions produites par la minimisation du critère pénalisé  $J + \lambda R$  correspond à l'ensemble des solutions fourni par l'estimateur du *maximum a posteriori* qui revient à considérer les minimiseurs de l'inverse de la log vraisemblance *a posteriori*

$$\min_{\mathbf{x}} (-\log V_p(\mathbf{y}; \mathbf{x}) = -\log V(\mathbf{y}; \mathbf{x}) - \log f_{\mathbf{X}}(\mathbf{x})). \quad (2.53)$$

Notons que le lien entre loi *a priori* et terme de régularisation déterministe utilisé en traitement d'image et de signal est souvent basé sur l'utilisation des champs aléatoires de Markov (Idier, 2008, Chap.7).

### 2.3.4 Réglage du paramètre de régularisation

La valeur du paramètre de régularisation  $\lambda$  influe de façon importante sur la reconstruction. Prenons l'exemple du problème de déconvolution de l'image `cameraman`. La solution, définie comme minimiseur d'un critère pénalisé  $J + \lambda R$ , dépend de  $\lambda$ . L'effet d'un mauvais choix du paramètre de régularisation est illustré sur la figure 2.6. Si  $\lambda$  est trop petit (sous-régularisation), les détails sont bien reconstruits mais l'image est noyée dans le bruit. A l'inverse, si  $\lambda$  est trop grand (sur-régularisation), l'image est trop lisse, les détails sont peu visibles et des oscillations de Gibbs apparaissent. Dans cet exemple simple, la solution exacte étant connue, il est possible d'estimer  $\lambda$  de manière supervisée, en modifiant manuellement  $\lambda$  pour minimiser l'erreur de reconstruction. Dans les problèmes pratiques, la solution exacte est bien sûr inconnue et il est nécessaire de trouver une façon de régler automatiquement le paramètre de régularisation. Les méthodes les plus courantes sont la validation croisée (Golub *et al.*, 1979) et la méthode de la courbe en L (Hansen et O'Leary, 1993; Hansen, 1992).

## 2.4 Conclusion

Un problème inverse mal posé se résout de façon efficace par la minimisation d'un critère pénalisé qui prend en compte conjointement les données observées et les informations préalables à la solution. Le choix du critère se base sur une nécessité de modélisation du processus de formation des données et des propriétés de la solution recherchée. Un autre point primordial à considérer est l'influence des propriétés mathématiques du critère sur la difficulté de la résolution. L'approche pénalisée nécessite la mise en œuvre d'un algorithme d'optimisation. Cet aspect fait l'objet du chapitre 3.

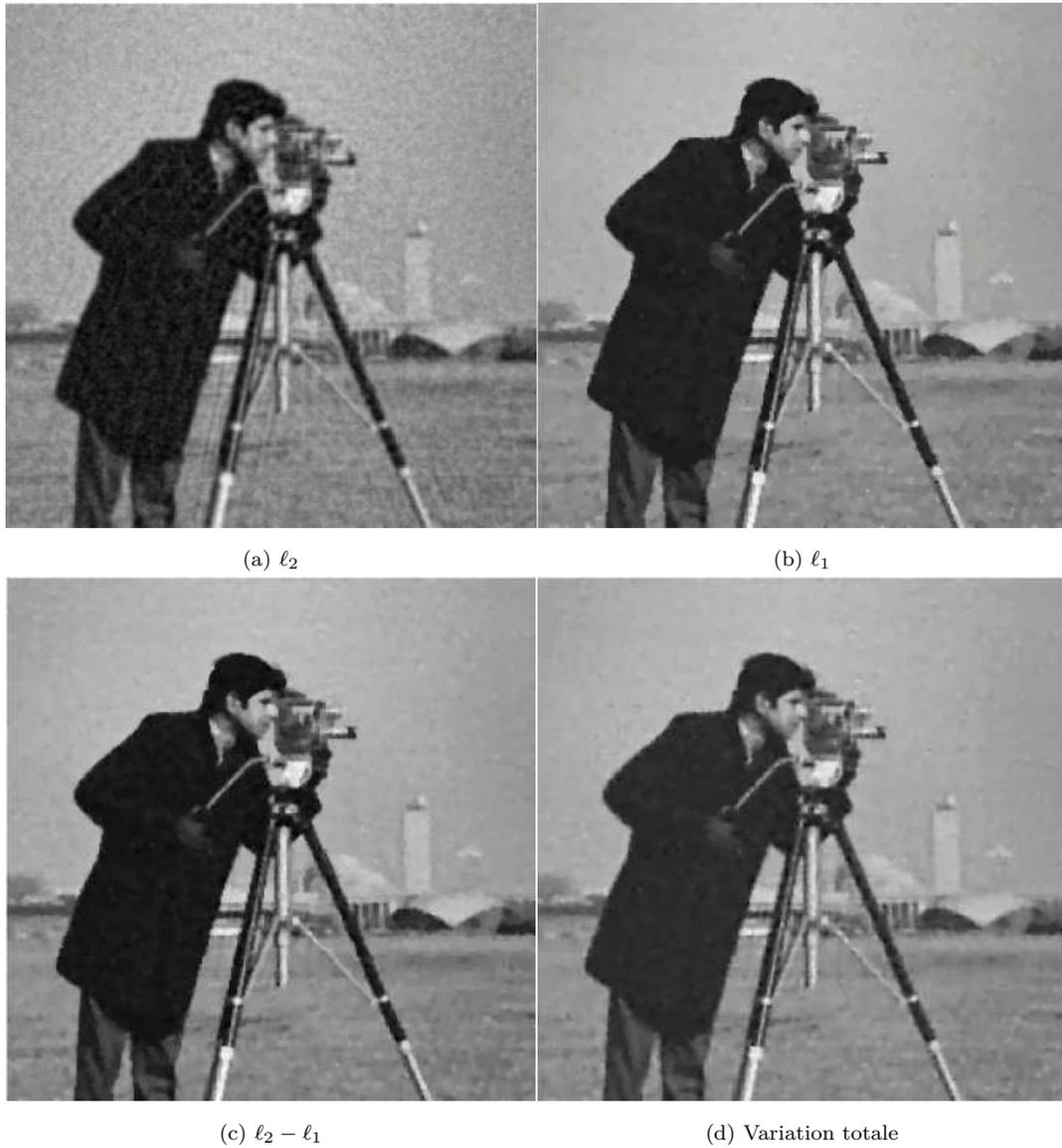


FIGURE 2.5 – Déconvolution de l'image *cameraman* par minimisation du critère des moindres carrés associé à différentes stratégies de régularisation opérant sur les différences entre pixels voisins de l'image (Soulez, 2008).



(a) Sous régularisé

(b) Sur régularisé

FIGURE 2.6 – Déconvolution de l'image `cameraman` par minimisation du critère des moindres carrés associé à la pénalisation de Tikhonov d'ordre 1 (Soulez, 2008).

# Chapitre 3

## Minimisation d'un critère pénalisé

### Sommaire

---

3.1	Introduction . . . . .	24
3.2	Formulation du problème d'optimisation . . . . .	24
3.2.1	Propriétés du critère . . . . .	24
3.2.2	Conditions d'optimalité . . . . .	27
3.3	Méthode de descente itérative . . . . .	28
3.3.1	Principe . . . . .	28
3.3.2	Méthodes de premier ordre pour la construction de la direction . . . . .	29
3.3.3	Méthodes de second ordre pour la construction de la direction . . . . .	31
3.3.4	Prise en compte de contraintes . . . . .	33
3.3.5	Discussion . . . . .	34
3.4	Méthode itérative de sous-espace . . . . .	34
3.4.1	Principe . . . . .	34
3.4.2	Construction du sous-espace . . . . .	34
3.4.3	Préconditionnement . . . . .	36
3.5	Conclusion . . . . .	37

---

### 3.1 Introduction

Une solution acceptable à un problème inverse mal posé s'obtient en résolvant le problème d'optimisation consistant à minimiser un critère pénalisé. Le cas le plus simple est celui d'un critère quadratique. En effet, son minimiseur s'obtient de façon analytique par la résolution d'un système linéaire. Néanmoins, lorsque le problème inverse est de grande taille, ce système ne peut être résolu que de façon approchée par une méthode itérative. De façon plus générale, la résolution du problème de la minimisation du critère pénalisé peut rarement s'effectuer de façon directe et une stratégie itérative doit être utilisée. Cette stratégie se base sur un algorithme d'optimisation qui construit une suite de points convergeant vers une solution du problème. Le recours à une méthode d'optimisation se rencontre très fréquemment en traitement du signal et de l'image. L'optimisation est un domaine des mathématiques appliquées très riche et très actif. De très nombreuses méthodes d'optimisation existent pour différents types de problèmes. Nous nous concentrons dans ce travail de thèse sur la minimisation de critères différentiables. Un algorithme fondamental pour résoudre ce type de problème est l'algorithme de descente itérative, dont sont issus par exemple les algorithmes standards de Newton et de gradient. Après avoir rappelées dans la section 3.2 les notions mathématiques indispensables à la construction d'un tel algorithme, nous décrivons dans la section 3.3 le schéma de descente itérative. Pour une présentation plus détaillée de cette famille de méthodes, le lecteur peut se référer aux ouvrages (Bertsekas, 1999; Nocedal et Wright, 1999; Sun et Yuan, 2006). Une accélération significative du schéma de descente itérative est obtenue en introduisant la notion de sous-espace (Conn *et al.*, 1994a; Narkiss et Zibulevsky, 2005) qui fera l'objet de la section 3.4. Les algorithmes de descente et de sous-espace nécessitent de définir une stratégie de calcul de pas. Cet aspect sera discuté de façon détaillée dans le chapitre 4.

### 3.2 Formulation du problème d'optimisation

Le problème d'optimisation qui nous intéresse consiste à minimiser une fonction à valeurs réelles sur son domaine de définition. Ce problème se formalise en

$$\inf_{\mathbf{x} \in \mathcal{D}_f} F(\mathbf{x}) \quad (3.1)$$

où  $\mathcal{D}_f \subset \mathbb{R}^n$  est le domaine de définition de la fonction  $F$ . En toute généralité,  $F$ , appelée *fonction objectif* ou *critère*, peut ne pas avoir de borne inférieure et si elle en admet une, cette borne peut ne pas être atteinte. Dans le cas contraire i.e., si  $F$  admet un minimum, le problème s'écrit

$$\min_{\mathbf{x} \in \mathcal{D}_f} F(\mathbf{x}). \quad (3.2)$$

Dans certains cas, il peut être nécessaire de considérer un problème de minimisation contraint. Au lieu de minimiser  $F$  sur  $\mathcal{D}_f$ , on la minimise seulement sur un ensemble  $\mathcal{E}$  délimité par des contraintes d'égalité et/ou d'inégalité. L'élaboration d'une méthode d'optimisation permettant la prise en compte de contraintes sera discutée à la fin de la section 3.3 ainsi que dans le chapitre 6.

#### 3.2.1 Propriétés du critère

Nous avons vu dans le chapitre 2 que le choix du critère est fortement lié aux particularités du problème à résoudre. Plus précisément, le terme d'adéquation aux données dépend des hypothèses sur le modèle de bruit (additif, multiplicatif, corrélé) et le terme de pénalisation dépend des propriétés de l'objet à reconstruire (parcimonie, rugosité). Il est également important de garder en mémoire lors de la construction du critère que ses propriétés mathématiques conditionnent le type de méthode employée pour la résolution du problème et influent sur la complexité du problème d'optimisation. Trois propriétés importantes à étudier sont la convexité, la différentiabilité et le caractère gradient Lipschitz (ou gradient borné) du critère.

### 3.2.1.1 Convexité

**Définition 3.1.** Soit  $F$  une fonction définie sur  $\mathcal{D}_f$ . Cette fonction est dite convexe sur  $\mathcal{V} \subset \mathcal{D}_f$  si  $\mathcal{V}$  est convexe et si pour tout  $\lambda \in [0, 1]$  et pour tout  $(\mathbf{x}, \mathbf{y}) \in \mathcal{V}$ ,

$$F(\lambda \mathbf{x} + (1 - \lambda) \mathbf{y}) \leq \lambda F(\mathbf{x}) + (1 - \lambda) F(\mathbf{y}) \quad (3.3)$$

La convexité du critère influe sur le nombre de solutions du problème d'optimisation. En effet, un critère non convexe est en général multimodal. Sa minimisation nécessite la mise en œuvre d'algorithmes d'optimisation mettant en œuvre une approche globale. On citera en particulier les approches par recuit simulé (Geman et Geman, 1984), les algorithmes génétiques (Holland, 1992) ou la non convexité graduelle (Nikolova *et al.*, 1998). A l'inverse, la minimisation d'un critère unimodal est un problème plus simple. En effet, le minimiseur global de  $F$  est caractérisé par une considération locale sur le critère. La mise en œuvre d'une méthode d'optimisation locale permettra de résoudre entièrement le problème.

Notons que les méthodes modernes d'optimisation globale font souvent appel à une ou plusieurs étapes d'optimisation locale. Ceci montre l'importance de l'optimisation locale, même pour un critère non convexe.

### 3.2.1.2 Différentiabilité

**Définition 3.2.** Soit  $F$  une fonction définie sur  $\mathcal{D}_f$ . Cette fonction est dite différentiable au sens de Gâteaux sur un ouvert  $\mathcal{V} \subset \mathcal{D}_f$  si pour tout  $\mathbf{x} \in \mathcal{V}$ , le gradient de  $F$  :

$$\nabla F(\mathbf{x}) = \begin{bmatrix} \frac{\partial F(\mathbf{x})}{\partial x_1} \\ \vdots \\ \frac{\partial F(\mathbf{x})}{\partial x_n} \end{bmatrix} \quad (3.4)$$

existe et vérifie pour tout  $\mathbf{y} \in \mathbb{R}^N$ ,

$$\nabla F(\mathbf{x})^T \mathbf{y} = \lim_{\alpha \searrow 0} \frac{F(\mathbf{x} + \alpha \mathbf{y}) - F(\mathbf{x})}{\alpha}. \quad (3.5)$$

La non différentiabilité du critère rend difficile la mise en œuvre de l'algorithme d'optimisation. Cependant, certaines applications nécessitent l'utilisation d'un terme de pénalisation non différentiable comme la variation totale pour la reconstruction d'images constantes par morceaux ou la norme  $\ell_1$  pour la reconstruction de représentations parcimonieuses. La minimisation du critère non différentiable qui en découle nécessite l'emploi d'une stratégie d'optimisation spécifique telle que celles proposées dans (Beck et Teboulle, 2009; Chaux *et al.*, 2007; Combettes et Pesquet, 2007; Daubechies *et al.*, 2004).

Une approche alternative consiste à approcher le terme non différentiable par une fonction différentiable paramétrée par un seuil  $\delta$ , quantifiant l'erreur d'approximation (Carlavan *et al.*, 2010; Nesterov, 2005; Weiss *et al.*, 2010; Zibulevsky et Elad, 2010). Le réglage du seuil résulte d'un compromis entre rapidité de l'optimisation et proximité avec la solution du critère non différentiable. En particulier, une très faible valeur de  $\delta$  conduit à un critère mal conditionné dont la minimisation par un algorithme itératif se caractérise par une lenteur importante (Nesterov, 2005).

En pratique, les solutions résultant du critère approché différentiable sont parfois meilleures que celles obtenues en minimisant le critère initial non différentiable (Carlavan *et al.*, 2010; Nikolova, 2004). Ceci est dû au fait que ce dernier ne modélise pas forcément de manière pertinente les images. Par exemple, la singularité de la norme  $\ell_1$  présente dans la variation totale introduit un effet de staircasing dans les images reconstruites (Nikolova, 2004). Dans le cadre des reconstructions basées sur un *a priori* de parcimonie, la question de savoir si la pénalisation  $\ell_1$  fournit de meilleurs résultats que sa version approchée reste un problème ouvert.

### 3.2.1.3 Gradient Lipschitz

**Définition 3.3.** Soit  $F$  une fonction définie sur  $\mathcal{D}_f$ . Cette fonction est dite à  $L$ -gradient Lipschitz sur le domaine  $\mathcal{V} \subset \mathcal{D}_f$  si pour tout  $(\mathbf{x}, \mathbf{y}) \in \mathcal{V}$ , les gradients  $\nabla F(\mathbf{x})$  et  $\nabla F(\mathbf{y})$  existent et vérifient

$$\|\nabla F(\mathbf{x}) - \nabla F(\mathbf{y})\| \leq L\|\mathbf{x} - \mathbf{y}\|. \quad (3.6)$$

Le caractère gradient Lipschitz est souvent nécessaire pour démontrer la convergence des algorithmes d'optimisation (Nocedal et Wright, 1999, Chap.3), ainsi que pour leur garantir une vitesse de convergence optimale (Nesterov, 2003, Chap.2). Voici quelques exemples de fonctions scalaires à gradient Lipschitz sur  $\mathcal{D}_f = \mathbb{R}$  :

$$\begin{aligned} F(x) &= (y - kx)^2 & \text{avec } L &= 2k^2 \\ F(x) &= \sqrt{x^2 + \delta^2} & \text{avec } L &= \frac{1}{\delta} \end{aligned} \quad (3.7)$$

Un résultat important concerne le caractère gradient Lipschitz des fonctions convexes :

**Lemme 3.1.** (Bertsekas, 2003)

Tout fonction convexe est à gradient Lipschitz sur  $\mathcal{V}$ , à condition que  $\mathcal{V}$  soit un fermé borné contenu dans le domaine de définition du gradient de  $F$ .

Les hypothèses sur  $\mathcal{V}$  sont essentielles comme on peut le voir sur les trois exemples suivants :

$$F_1(x) = x^3, \quad \mathcal{D}_f = \mathbb{R}, \quad \mathcal{V} = \mathbb{R}^+ \quad (3.8)$$

$F_1$  est convexe sur  $\mathcal{V}$  mais celui-ci n'est pas borné. Le gradient de  $F$  n'est pas Lipschitz sur  $\mathcal{V}$ .

$$F_2(x) = -\log x, \quad \mathcal{D}_f = (0, +\infty), \quad \mathcal{V} = (0, 1] \quad (3.9)$$

$F_2$  est convexe sur  $\mathcal{V}$  mais celui-ci est ouvert. Le gradient de  $F$  n'est pas Lipschitz sur  $\mathcal{V}$ .

$$F_3(x) = x \log x, \quad \mathcal{D}_f = [0, +\infty), \quad \mathcal{V} = [0, 1] \quad (3.10)$$

$F_3$  est convexe sur  $\mathcal{V}$  mais le gradient de  $F$  est infini en 0. Par conséquent,  $F$  n'est pas gradient Lipschitz sur  $\mathcal{V}$ .

Les fonctions  $F_2$  et  $F_3$  appartiennent à la classe particulière des *fonctions barrières* que nous détaillons dans le paragraphe suivant.

### 3.2.1.4 Fonction barrière

Le terme de *barrière* est employé pour décrire des fonctions possédant une asymptote verticale le long des frontières bornées de leur domaine de définition (Nesterov, 2003, Sec.4.2). Parmi les fonctions présentées dans la section 2.3, l'adéquation au données poissonienne (2.30) et les pénalisations par entropie (2.46), (2.47) sont des fonctions barrières. En effet, ces critères contiennent un terme de la forme

$$B(\mathbf{x}) = -\sum_i \log([\mathbf{C}\mathbf{x}]_i), \quad (3.11)$$

ou

$$B(\mathbf{x}) = \sum_i ([\mathbf{C}\mathbf{x}]_i) \log([\mathbf{C}\mathbf{x}]_i). \quad (3.12)$$

Comme l'illustre la figure 3.1(a), la fonction  $-\log u$  tend vers l'infini en zéro. Par conséquent, la fonction (3.11) est une fonction barrière sur le bord du domaine

$$\mathcal{C} = \{\mathbf{x} \mid [\mathbf{C}\mathbf{x}]_i > 0\}. \quad (3.13)$$

Il en est de même pour la fonction (3.12) car la dérivée de la fonction  $u \log u$  tend vers l'infini en zéro, comme l'illustre la figure 3.1(b).

Une telle propriété est intéressante car elle implique que tout critère contenant un terme de la forme (3.11) ou (3.12) aura ses minimiseurs à l'intérieur de  $\mathcal{C}$ . C'est particulièrement utile lorsque  $\mathcal{C}$  s'identifie avec l'orthant positif, la barrière contraignant la positivité sur la solution. Nous montrerons dans le chapitre 6 que la minimisation d'un critère contenant une fonction barrière nécessite une vigilance particulière en terme de choix d'algorithme et d'implémentation. L'objectif majeur de cette thèse sera alors de développer des stratégies d'optimisation efficaces pour traiter la minimisation de critère contenant une fonction barrière.

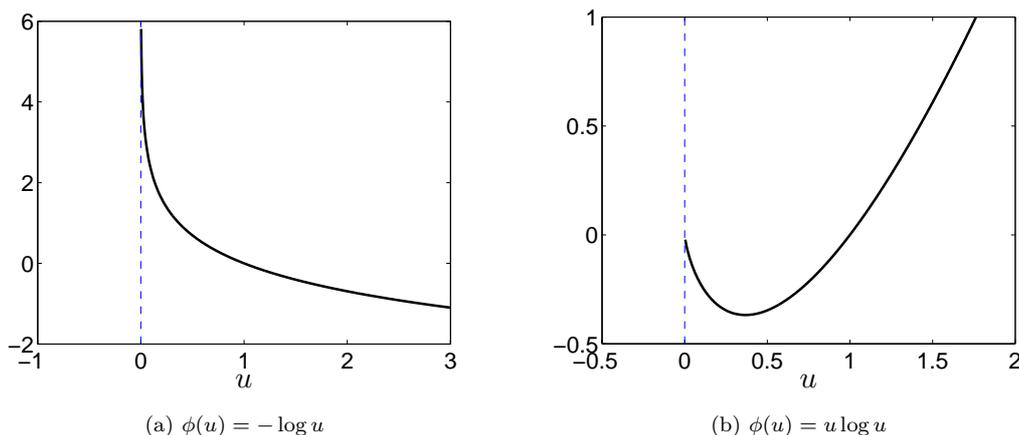


FIGURE 3.1 – Barrières logarithmique et entropique

### 3.2.2 Conditions d'optimalité

Généralement, on entend par solution de (3.2) un point associé à une valeur minimale de  $F$  sur  $\mathcal{D}_f$ . Ce point est appelé minimum global. Il vérifie

$$F(\mathbf{x}^*) \leq F(\mathbf{x}), \quad \forall \mathbf{x} \in \mathcal{D}_f. \quad (3.14)$$

Cependant, un tel point peut être difficile à trouver car notre connaissance de  $F$  est souvent limitée. En effet, les méthodes de résolution sont basées sur une évaluation de  $F$  en un nombre réduit de points. Il peut donc arriver que  $F$  présente un minimum dans une région inexplorée par la méthode. De ce fait, la plupart des stratégies de résolution déterminent des points où  $F$  atteint une valeur minimale dans un voisinage  $\mathcal{V}$  de  $\mathbf{x}^*$ , appelés solutions locales. On parle de minimum global ou local *strict* si on a l'égalité stricte dans (3.14) lorsque  $\mathbf{x} \neq \mathbf{x}^*$ . Les classifications que nous venons d'effectuer peuvent laisser croire que le seul moyen de vérifier que  $\mathbf{x}^*$  est solution de (3.2) est d'examiner tous les points de  $\mathcal{D}_f$  (ou du voisinage de  $\mathbf{x}^*$ ) afin de s'assurer qu'il n'existe pas un autre point où  $F$  prend une valeur plus petite. Lorsque la fonction  $F$  possède des propriétés de régularité (continuité, différentiabilité), il existe des conditions dites d'optimalité qui permettent de reconnaître une solution à partir de connaissances locales sur  $F$ . Les conditions d'optimalité sont de deux types : Condition nécessaire (CN) et condition suffisante (CS). La CN est une relation vérifiée par toute solution du problème (3.2). À l'inverse, la vérification de la CS assure à un point d'être solution de (3.2). On parle de conditions du premier ou second ordre suivant que celles-ci font intervenir uniquement les dérivées premières de  $F$  ou bien aussi les dérivées secondes.

#### **Théorème 3.1.** *CN du premier ordre*

*Si  $\mathbf{x}^*$  est un minimum local de (3.2) et  $F$  est différentiable en  $\mathbf{x}^*$ , alors  $\nabla F(\mathbf{x}^*) = \mathbf{0}$ .*

Un point vérifiant  $\nabla F(\mathbf{x}^*) = \mathbf{0}$  est appelé *point stationnaire*. Un point stationnaire peut être de trois types : minimum local, maximum local ou point selle.

**Théorème 3.2.** *CN du second ordre*

Si  $\mathbf{x}^*$  est un minimum local de (3.2) et  $F$  est deux fois différentiable en  $\mathbf{x}^*$ , alors  $\nabla F(\mathbf{x}^*) = \mathbf{0}$  et  $\nabla^2 F(\mathbf{x}^*)$  est semidéfinie positive, où  $\nabla^2 F(\mathbf{x}^*)$  est la matrice hessienne définie par les coefficients  $\frac{\partial^2 F}{\partial x_i \partial x_j}(\mathbf{x}^*)$ .

La contraposée de ce théorème permet de définir la CS d'optimalité du second ordre, dont la vérification assure à un point d'être solution de (3.2).

**Théorème 3.3.** *CS du second ordre*

Supposons que  $F$  soit deux fois différentiable en  $\mathbf{x}^*$ . Si  $\nabla F(\mathbf{x}^*) = \mathbf{0}$  et  $\nabla^2 F(\mathbf{x}^*)$  est définie positive, alors  $\mathbf{x}^*$  est un minimum local strict de (3.2).

Lorsque la fonction  $F$  est convexe, les minima locaux et globaux sont simples à caractériser :

**Théorème 3.4.** *Si  $F$  est convexe, tout minimum local  $\mathbf{x}^*$  de (3.2) est aussi un minimum global. Si, de plus,  $F$  est différentiable, alors tout point vérifiant  $\nabla F(\mathbf{x}^*) = \mathbf{0}$  est minimum global de (3.2).*

Les conditions d'optimalité servent de point de départ à l'élaboration de stratégies de résolution du problème (3.2). En général, il n'est pas possible d'obtenir une solution de (3.2) de façon directe ou même en un nombre fini d'opérations arithmétiques. De ce fait, la résolution passe par l'utilisation d'un *algorithme itératif* : A partir d'un point qui ne satisfait pas les conditions d'optimalité, on va produire un autre point, meilleur que le précédent. Puis on va recommencer, et engendrer une suite de points qui s'approche d'un optimum. L'élaboration de l'algorithme nécessite donc de répondre au problème formel suivant :

*A partir d'informations sur  $F$  au point  $\mathbf{x}_k$  (et éventuellement aux points précédents  $\mathbf{x}_0, \dots, \mathbf{x}_{k-1}$ ), comment trouver une nouvelle itérée  $\mathbf{x}_{k+1}$  plus proche de  $\mathbf{x}^*$  que  $\mathbf{x}_k$  ?*

L'objectif de l'algorithme étant de minimiser  $F$ , une approche naturelle consiste à générer  $\{\mathbf{x}_k\}_{k \geq 0}$  telle que pour tous ses membres, on ait la relation

$$F(\mathbf{x}_{k+1}) \leq F(\mathbf{x}_k), \quad (3.15)$$

on dit alors que l'algorithme est monotone. Une méthode très usuelle de construction d'algorithmes itératifs monotones est la *descente itérative* qui génère les itérés en se déplaçant le long de directions correspondant à une diminution locale du critère. Les développements effectués dans cette thèse vont contribuer à l'amélioration des performances, en terme de vitesse de convergence et de coût par itération, des algorithmes de descente itérative que nous présentons à présent.

### 3.3 Méthode de descente itérative

#### 3.3.1 Principe

Dans le cadre de l'optimisation différentiable, une famille fondamentale d'algorithmes est basée sur la stratégie de *descente itérative*. La méthode de descente itérative permet de générer une suite de points convergeant vers une solution du problème non contraint (3.2) tout en assurant la décroissance monotone du critère à chaque itération. La décroissance du critère est assurée en déplaçant l'itérée courante  $\mathbf{x}_k$  le long d'une *direction de descente*  $\mathbf{d}_k$ . On dit que  $\mathbf{d}_k$  est une direction de descente de  $F$  en  $\mathbf{x}_k$  si elle satisfait

$$\nabla F(\mathbf{x}_k)^T \mathbf{d}_k < 0 \quad (3.16)$$

Par définition du gradient, il revient au même de dire que  $\mathbf{d}_k$  fait avec  $-\nabla F(\mathbf{x}_k)$  un angle  $\theta_k$  strictement plus petit que 90 degrés, comme l'illustre la figure 3.2. L'ensemble des directions

vérifiant la relation de descente (3.16) forme un demi-espace ouvert de  $\mathbb{R}^n$ , dont les éléments  $\mathbf{d}_k$  sont tels que

$$F(\mathbf{x}_k + \alpha \mathbf{d}_k) < F(\mathbf{x}_k) \quad (3.17)$$

pour tout  $\alpha > 0$  suffisamment petit. De telles directions sont intéressantes car, pour faire décroître  $F$ , il suffit de se déplacer le long de  $\mathbf{d}_k$ .

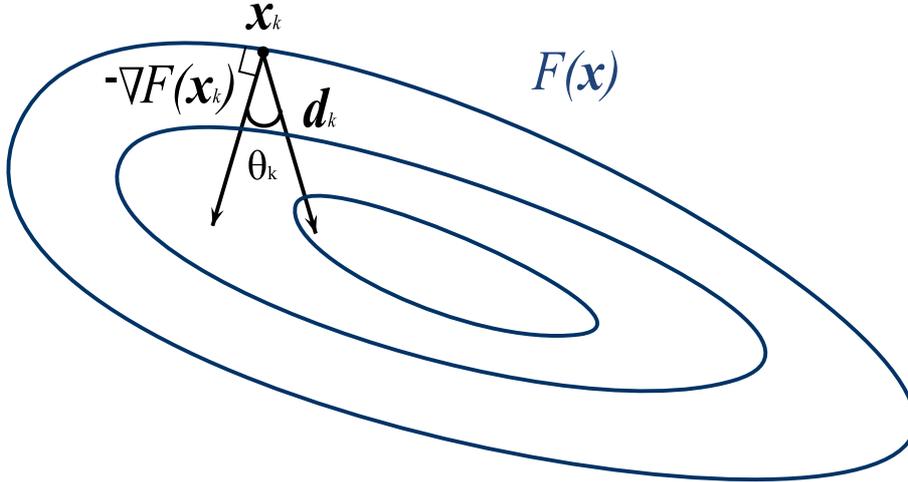


FIGURE 3.2 – La direction  $\mathbf{d}_k$  forme un angle  $\theta_k < 90^\circ$  avec la direction de plus grande pente  $-\nabla F(\mathbf{x}_k)$ . Il s'agit d'une direction de descente pour  $F$  en  $\mathbf{x}_k$ . En effet, elle est dirigée vers l'intérieur des lignes de niveau de  $F$ .

L'algorithme de descente utilise cette propriété pour minimiser la fonction  $F$ . La suite  $\{\mathbf{x}_k\}_{k \geq 0}$  est construite par la récurrence

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k \quad (3.18)$$

où  $\alpha_k > 0$  est le *pas* et  $\mathbf{d}_k$  est une direction de descente (Figure 3.3). Une telle méthode consiste à alterner la construction de  $\mathbf{d}_k$  et la détermination de  $\alpha_k$ . L'arrêt de la récurrence est contrôlé par un test portant généralement sur la petitesse du gradient  $\nabla F(\mathbf{x}_k)$ . C'est en effet ce que suggère la condition d'optimalité  $\nabla F(\mathbf{x}^*) = \mathbf{0}$ . Bien que l'annulation du gradient puisse intervenir en un maximum ou en un point selle de  $F$ , le choix de  $(\alpha_k, \mathbf{d}_k)$  va empêcher l'algorithme de converger vers un tel point. Pour mettre en œuvre une méthode de descente itérative, il faut donc spécifier deux choses :

- La construction de la direction  $\mathbf{d}_k$ . La manière de procéder donne le nom de l'algorithme. Elle est choisie à partir de considérations locales sur  $F$ . Elle doit vérifier (3.16) et encourager une convergence rapide de l'algorithme, en particulier lorsque l'on se rapproche de la solution.
- La détermination du pas  $\alpha_k$  par une stratégie de *recherche linéaire*. La valeur de  $\alpha_k$  résulte de la minimisation approchée de la fonction scalaire  $f(\alpha) = F(\mathbf{x}_k + \alpha \mathbf{d}_k)$ .

Les stratégies développées pour le calcul du pas feront l'objet du chapitre 4. Nous présentons à présent différentes stratégies employées pour construire la direction.

### 3.3.2 Méthodes de premier ordre pour la construction de la direction

#### 3.3.2.1 Algorithme de gradient

Considérons le gradient de  $F$  en  $\mathbf{x}_k$  :  $\nabla F(\mathbf{x}_k)$  que nous noterons par la suite  $\mathbf{g}_k$ . L'opposé de  $\mathbf{g}_k$  est un choix évident de direction de descente. Parmi toutes les directions provenant de  $\mathbf{x}_k$ , il s'agit de celle le long de laquelle  $F$  décroît le plus rapidement. Cette direction est orthogonale aux lignes de niveaux de la fonction. L'algorithme de descente utilisant

$$\mathbf{d}_k = -\mathbf{g}_k \quad (3.19)$$

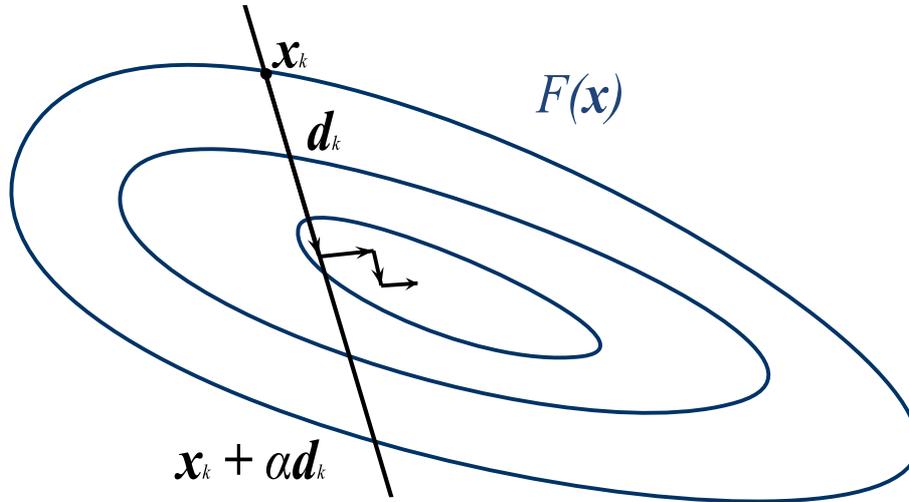


FIGURE 3.3 – La méthode de descente itérative alterne entre détermination d'une direction de descente  $\mathbf{d}_k$  et minimisation approchée de  $F$  le long de  $\mathbf{d}_k$  par une stratégie de recherche linéaire.

est appelé algorithme de la *plus grande pente* ou algorithme du gradient. Bien que séduisant par sa grande simplicité, ce choix de direction s'avère peu efficace en pratique et l'algorithme qui en résulte peut être très lent à converger. Par conséquent, on utilise plutôt la direction de *gradient conjugué* (GC).

### 3.3.2.2 Algorithme de gradient conjugué

L'algorithme de GC est originellement conçu pour résoudre l'équation linéaire  $\mathbf{K}\mathbf{x} = \mathbf{y}$ , où la matrice  $\mathbf{K}$  est symétrique et définie positive. L'extension de l'algorithme au cas non linéaire (GCNL) a été proposée par Fletcher et Reeves (Fletcher et Reeves, 1964) qui définissent

$$\mathbf{d}_k = \begin{cases} -\mathbf{g}_k & \text{si } k = 0 \\ -\mathbf{g}_k + \beta_k \mathbf{d}_{k-1} & \text{si } k \geq 1 \end{cases} \quad (3.20)$$

avec le coefficient de conjugaison

$$\beta_k^{\text{FR}} = \frac{\mathbf{g}_k^T \mathbf{g}_k}{\mathbf{g}_{k-1}^T \mathbf{g}_{k-1}}. \quad (3.21)$$

Lorsque la fonction à minimiser est quadratique, le coefficient  $\beta_k$  assure la conjugaison entre  $\mathbf{d}_k$  et  $\mathbf{d}_{k-1}$  et la méthode est assurée de converger en au plus  $n$  itérations, où  $n$  est la dimension du vecteur d'inconnues. Dans le cas non quadratique, il existe de nombreuses variantes à la méthode de Fletcher et Reeves, différant par le choix du paramètre  $\beta_k$  (Hager et Zhang, 2006b). De plus, lorsque  $\beta_k$  n'assure pas que  $\mathbf{d}_k$  soit une direction de descente, cette dernière est remplacée par son opposé  $-\mathbf{d}_k$ . Pour un problème non linéaire quelconque, il est conseillé dans (Nocedal et Wright, 1999) de choisir la formule de (Gilbert et Nocedal, 1992)

$$\beta_k^{\text{PRP}^+} = \max(\beta_k^{\text{PRP}}, 0), \quad (3.22)$$

où  $\beta_k^{\text{PRP}}$  est le coefficient de (Polak et Ribière, 1969; Polyak, 1969) :

$$\beta_k^{\text{PRP}} = \frac{\mathbf{g}_k^T (\mathbf{g}_k - \mathbf{g}_{k-1})}{\|\mathbf{g}_{k-1}\|^2}. \quad (3.23)$$

En pratique, la direction de gradient conjugué non linéaire est souvent bien plus efficace que la direction de plus grande pente en terme de vitesse de convergence et est toute aussi simple à calculer

(Nocedal et Wright, 1999). Une accélération de ces deux algorithmes s'obtient en introduisant des informations sur les dérivées secondes du critère. Une première stratégie est d'utiliser un préconditionneur.

### 3.3.2.3 Préconditionnement

Le conditionnement d'une matrice  $\mathbf{K}$ , noté  $\kappa(\mathbf{K})$ , est défini comme le rapport entre sa plus grande et sa plus petite valeur singulière. Nous avons vu dans le chapitre 2 que cette valeur quantifie l'effet de petites variations de  $\mathbf{y}$  sur  $\mathbf{x}^\dagger$ , inverse généralisé de  $\mathbf{y} = \mathbf{K}\mathbf{x}$ , un mauvais conditionnement (i.e.,  $\kappa(\mathbf{K}) \gg 1$ ) se traduisant par une amplification des erreurs numériques.

La vitesse de convergence d'un algorithme de descente itérative du premier ordre dépend du conditionnement du Hessien  $\nabla^2 F$ . La minimisation d'un critère ayant un Hessien mal conditionné est caractérisée par des valeurs de pas très petites, même si le pas minimise exactement le critère le long de la direction choisie. Par conséquent, la convergence des itérés vers la solution est très lente. Le préconditionnement consiste à opérer le changement de variable  $\tilde{F}(\mathbf{z}) = F(\mathbf{C}_k\mathbf{x})$  afin de modifier la répartition des valeurs propres de la matrice hessienne au point courant, dans le but d'accélérer la convergence de l'algorithme. La matrice  $\mathbf{C}_k$  est choisie afin que  $\mathbf{M}_k = \mathbf{C}_k\mathbf{C}_k^T$ , appelée *matrice de préconditionnement* ou *préconditionneur*, soit une approximation symétrique définie positive de l'inverse du Hessien  $\mathbf{H}_k^{-1}$ .

Le changement de variable se traduit par une modification du gradient du critère et par conséquent de l'expression de la direction de recherche. La direction de gradient devient

$$\mathbf{d}_k = -\mathbf{M}_k\mathbf{g}_k \quad (3.24)$$

et la direction GCNL devient

$$\mathbf{d}_k = -\mathbf{M}_k\mathbf{g}_k + \beta_k\mathbf{d}_{k-1} \quad (3.25)$$

où le coefficient de conjugaison  $\beta_k$  est modifié en remplaçant respectivement les termes  $\mathbf{d}_k$  et  $\mathbf{g}_k$  par  $\mathbf{C}_k^{-1}\mathbf{d}_k$  et  $\mathbf{C}_k^T\mathbf{g}_k$  ((Pytlak, 2009, Chap.4),(Hager et Zhang, 2006b)). Par exemple, le coefficient de Fletcher et Reeves devient

$$\beta_k = \frac{\mathbf{g}_k\mathbf{M}_k\mathbf{g}_k}{\mathbf{g}_{k-1}\mathbf{M}_{k-1}\mathbf{g}_{k-1}}. \quad (3.26)$$

Dans le cas quadratique, cette modification entraîne la conjugaison de deux directions consécutives dans l'espace de changement de variable  $\mathbf{z}$ .

Le préconditionneur  $\mathbf{M}_k$  est construit de façon à ce que le Hessien de  $\tilde{F}$  en  $\mathbf{x}_k$  soit bien conditionné afin d'entraîner que la direction de plus grande pente soit associée à une valeur de pas plus grande. Un bon préconditionneur exige un compromis entre une approximation satisfaisante de l'inverse du Hessien de  $F$  et un coût de calcul raisonnable. De nombreuses techniques de calcul de  $\mathbf{M}_k$  sont proposées dans la littérature, parmi lesquelles la méthode SSOR (Symmetric Successive Over-Relaxation) ainsi que la factorisation LU incomplète ou de Cholesky (Saad, 2003, Chap.10), (Chen, 2005). L'utilisation de l'expression exacte de l'inverse du Hessien conduit aux méthodes de descente itérative du second ordre.

## 3.3.3 Méthodes de second ordre pour la construction de la direction

### 3.3.3.1 Algorithme de Newton

La direction de Newton minimise une approximation quadratique correspondant au développement en série de Taylor au second ordre de  $F$  au point  $\mathbf{x}_k$  :

$$F(\mathbf{x}_k + \mathbf{d}) \approx F(\mathbf{x}_k) + \mathbf{d}^T \nabla F(\mathbf{x}_k) + \mathbf{d}^T \mathbf{H}_k \mathbf{d} \quad (3.27)$$

où  $\mathbf{H}_k \equiv \nabla^2 F(\mathbf{x}_k)$ . Elle s'écrit donc

$$\mathbf{d}_k = -\mathbf{H}_k^{-1}\mathbf{g}_k. \quad (3.28)$$

Lorsque la matrice  $\mathbf{H}_k$  est définie positive, la direction de Newton est une direction de descente et permet une convergence rapide de l'algorithme notamment lorsque le point courant  $\mathbf{x}_k$  est proche d'une solution  $\mathbf{x}^*$ . La direction de Newton possède cependant plusieurs défauts. Tout d'abord, elle n'est pas définie aux points où le Hessien est singulier. De plus, lorsque le Hessien n'est pas défini positif, la direction de Newton n'est pas nécessairement une direction de descente. Dans ce cas, la direction de *Newton modifiée* est utilisée

$$\mathbf{d}_k = -(\mathbf{H}_k + \max(-\sigma_{\min} + \epsilon, 0)\mathbf{I})^{-1} \mathbf{g}_k \quad (3.29)$$

où  $\sigma_{\min}$  est la plus petite valeur propre de  $\mathbf{H}_k$  et  $\epsilon$  une petite quantité positive. Pour un tel choix,  $\mathbf{H}_k$  sera égal au Hessien de  $F$  si celui-ci est défini positif et se verra ajouter un multiple de l'identité adéquat pour le rendre défini positif le cas échéant. La valeur exacte de  $\sigma_{\min}$  est souvent issue d'un calcul très coûteux du fait de la taille du problème et est classiquement approchée à l'aide d'une stratégie basée sur la factorisation de Choleski (Nocedal et Wright, 1999, Chap.6). Pour pallier le manque de robustesse de la méthode de Newton modifiée, celle-ci est souvent remplacée par la méthode de régions de confiance, où la modification du Hessien se base sur une interprétation géométrique (Conn et al., 1987).

### 3.3.3.2 Algorithme de Quasi-Newton

Un inconvénient de la direction de Newton est qu'elle nécessite la connaissance de Hessien de  $F$ . Le calcul de cette matrice peut être compliqué et coûteux. De plus, pour certains problèmes, le critère n'est pas deux fois différentiable. Ceci a motivé l'apparition des méthodes de *Quasi-Newton* qui définissent la direction par

$$\mathbf{d}_k = -\mathbf{B}_k^{-1} \mathbf{g}_k \quad (3.30)$$

où  $\mathbf{B}_k$  est une approximation du Hessien générée itérativement par une formule de *mise à jour* (Dennis et Moré, 1977). Les stratégies de mise à jour les plus populaires sont SR1 (symmetric-rank-one) (Khalfan et al., 1993) et BFGS du nom de ses auteurs Broyden, Fletcher, Goldfarb et Shanno. Elles se basent sur l'approximation suivante, appelée *équation sécante* :

$$\mathbf{B}_k(\mathbf{x}_k - \mathbf{x}_{k-1}) = \mathbf{g}_k - \mathbf{g}_{k-1}. \quad (3.31)$$

En effet, on peut montrer que cette relation est vérifiée par le Hessien d'un critère  $F$  quadratique. Les matrices SR1 et BFGS sont donc construites afin de vérifier (3.31). De plus, elles sont symétriques et telles que la différence  $\mathbf{B}_{k+1} - \mathbf{B}_k$  soit de rang faible (Nocedal et Wright, 1999, Chap.9). Les deux formules de mise à jour s'écrivent respectivement

$$\mathbf{B}_{k+1}^{SR1} = \mathbf{B}_k + \frac{(\boldsymbol{\delta}_k - \mathbf{B}_k \mathbf{z}_k)(\boldsymbol{\delta}_k - \mathbf{B}_k \mathbf{z}_k)^T}{(\boldsymbol{\delta}_k - \mathbf{B}_k \mathbf{z}_k)^T \mathbf{z}_k}, \quad (3.32)$$

et

$$\mathbf{B}_{k+1}^{BFGS} = \mathbf{B}_k - \frac{\mathbf{B}_k \mathbf{z}_k \mathbf{z}_k^T \mathbf{B}_k}{\mathbf{z}_k^T \mathbf{B}_k \mathbf{z}_k} + \frac{\boldsymbol{\delta}_k \boldsymbol{\delta}_k^T}{\boldsymbol{\delta}_k^T \mathbf{z}_k}, \quad (3.33)$$

avec

$$\mathbf{z}_k = \mathbf{x}_{k+1} - \mathbf{x}_k, \quad \boldsymbol{\delta}_k = \mathbf{g}_{k+1} - \mathbf{g}_k. \quad (3.34)$$

La direction de Quasi-Newton nécessite de calculer la matrice  $\mathbf{M}_k = \mathbf{B}_k^{-1}$ . Cette matrice s'obtient simplement par récurrence :

$$\mathbf{M}_{k+1}^{SR1} = \mathbf{M}_k + \frac{(\mathbf{z}_k - \mathbf{M}_k \boldsymbol{\delta}_k)(\mathbf{z}_k - \mathbf{M}_k \boldsymbol{\delta}_k)^T}{(\mathbf{z}_k - \mathbf{B}_k \boldsymbol{\delta}_k)^T \boldsymbol{\delta}_k}, \quad (3.35)$$

et

$$\mathbf{M}_{k+1}^{BFGS} = (\mathbf{I} - \rho_k \mathbf{z}_k \boldsymbol{\delta}_k^T) \mathbf{M}_k (\mathbf{I} - \rho_k \boldsymbol{\delta}_k \mathbf{z}_k^T) + \rho_k \mathbf{z}_k \mathbf{z}_k^T, \quad (3.36)$$

avec  $\rho_k = 1/(\boldsymbol{\delta}_k^T \mathbf{z}_k)$ . Le stockage de la matrice  $\mathbf{M}_k$  peut être impossible lorsque le problème est de très grande taille. Dans ce cas, des formules de mise à jour à mémoire limitée sont utilisées

donnant lieu aux algorithmes L-BFGS (Liu et Nocedal, 1989) et L-SR1 (Nocedal et Wright, 1999, Th.9.2).

Enfin, tout comme l'algorithme de gradient, l'algorithme de quasi-Newton peut être accéléré à l'aide d'un préconditionneur (Jiang *et al.*, 2004), mais ce sujet reste peu traité à ce jour.

### 3.3.3.3 Formes tronquées

Lorsque le problème est de grande taille, l'inversion de  $\mathbf{H}_k$  ou d'une approximation  $\mathbf{B}_k$  peut être très coûteuse ou même impraticable. Il faut alors se satisfaire d'un inverse approché obtenu par une méthode itérative (Dembo *et al.*, 1982). Cette approche est utilisée dans l'algorithme de Newton tronqué (Dembo et Steihaug, 1983; Nash, 2000) dans lequel la direction  $\mathbf{d}_k$  est issues d'itérations de gradient conjugué appliquées à la fonction quadratique (3.27) et stoppées avant convergence. L'étape de gradient conjugué peut être accélérée par l'utilisation d'un préconditionneur, donnant lieu à l'algorithme de Newton tronqué préconditionné.

### 3.3.4 Prise en compte de contraintes

L'algorithme de descente itérative permet également de résoudre le problème contraint

$$\min_{\mathbf{x} \in \mathcal{E}} F(\mathbf{x}), \quad (3.37)$$

avec  $\mathcal{E}$  un sous-ensemble convexe non vide de  $\mathbb{R}^n$ , à condition de modifier la direction  $\mathbf{d}_k$  afin de tenir compte des contraintes. Une condition nécessaire d'optimalité de (3.37) est la condition d'Euler

$$\nabla F(\mathbf{x}^*)^T (\mathbf{x} - \mathbf{x}^*) \geq 0, \quad \forall \mathbf{x} \in \mathcal{E}. \quad (3.38)$$

Afin de tenir compte des contraintes, l'algorithme de descente itérative doit assurer que la suite d'itérés converge vers un point vérifiant la condition (3.38). La notion de direction de descente *admissible* permet de construire une telle suite.

Etant donné  $\mathbf{x}_k \in \mathcal{E}$ , la direction  $\mathbf{d}_k$  est dite admissible si  $\mathbf{x}_k + \alpha \mathbf{d}_k$  est dans le domaine contraint pour tout  $\alpha > 0$  suffisamment petit. Toute direction admissible peut se mettre sous la forme

$$\mathbf{d}_k = \gamma(\bar{\mathbf{x}}_k - \mathbf{x}_k) \quad (3.39)$$

avec  $\gamma$  un réel strictement positif et  $\bar{\mathbf{x}}_k \in \mathcal{E}$ . Partant d'un point  $\mathbf{x}_0 \in \mathcal{E}$ , l'algorithme de directions admissibles génère une suite  $\{\mathbf{x}_k \in \mathcal{E}\}_{k \geq 0}$  par la récurrence

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k(\bar{\mathbf{x}}_k - \mathbf{x}_k) \quad (3.40)$$

avec  $\alpha_k$  un pas permettant d'assurer la convergence de l'algorithme, appartenant à  $(0, 1]$  de façon à maintenir les itérés dans  $\mathcal{E}$ . La construction de  $\mathbf{d}_k$  par (3.39) nécessite la connaissance de  $\bar{\mathbf{x}}_k \in \mathcal{E}$ . Le domaine contraint  $\mathcal{E}$  étant convexe, un tel point s'obtient simplement par l'application de l'opérateur de projection sur cet ensemble.

Une direction admissible très simple est la direction de *gradient projeté* définie par  $\gamma = 1$  et

$$\bar{\mathbf{x}}^k = \mathcal{P}_{\mathcal{E}}(\mathbf{x}_k - s_k \mathbf{g}_k) \quad (3.41)$$

avec  $s_k$  un paramètre de réglage et  $\mathcal{P}_{\mathcal{E}}$  la projection sur  $\mathcal{E}$ . Les propriétés de l'opérateur de projection entraînent que la direction de gradient projeté est une direction de descente pour  $F$ . La convergence de l'algorithme de gradient projeté dépend du choix des pas  $(s_k, \alpha_k)$  (Bertsekas, 1999, Chap.2). La méthode du gradient projeté est simple mais comme elle se base sur la direction de gradient, elle est très lente. Il existe donc des versions accélérées de cet algorithme basées sur la modification de la métrique de l'opérateur de projection (Bonettini *et al.*, 2009), sur le remplacement de la direction de plus grande pente  $-\mathbf{g}_k$  par une direction entraînant une convergence plus rapide (Ex : gradient préconditionné (Bertsekas, 1981), gradient conjugué (Pytlak, 1998), Quasi-Newton (Ni et Yuan, 1997; Xiao et Zengxin, 2007)) ou bien sur l'utilisation de régions de

confiance (Byrd *et al.*, 1987; Conn *et al.*, 1988). La complexité de ces méthodes est fortement liée au coût de l'opérateur  $\mathcal{P}_{\mathcal{E}}$ . Par conséquent, elles sont principalement utilisées quand  $\mathcal{E}$  est l'orthant positif car la projection sur  $\mathcal{E}$  se ramène simplement à une opération de seuillage. Comme souligné dans (Combettes et Pesquet, 2010, Sec.1.3), l'algorithme de gradient projeté correspond au cas particulier de l'application de l'algorithme implicite-explicite (*Forward-Backward* (Combettes et Wajs, 2005)) à la minimisation du critère composite

$$\mathcal{I}_{\mathcal{E}}(\mathbf{x}) + F(\mathbf{x}), \quad (3.42)$$

où  $\mathcal{I}_{\mathcal{E}}(\cdot)$  est la fonction indicatrice du convexe  $\mathcal{E}$ .

### 3.3.5 Discussion

L'algorithme de descente itérative est une méthode fondamentale d'optimisation différentiable. Dans le cadre de la résolution de problèmes inverses de grande taille, les directions de GCNL (Fessler et Booth, 1999; Khurd *et al.*, 2002; Mumcuoglu et Leahy, 1994), (Bertero et Boccacci, 1998, Chap.6) et de L-BFGS (Gilles *et al.*, 2002; Lu *et al.*, 2009) fournissent des algorithmes très efficaces.

Ces deux algorithmes possèdent une structure particulière. En effet, la direction de descente engendrée appartient à un sous-espace de faible dimension : la direction de GCNL s'écrit comme une combinaison de l'opposé du gradient et de la direction obtenue à l'itération précédente (Cragg et Levy, 1969; Miele et Cantrell, 1969; Yuan et Stoer, 1995). La direction de L-BFGS est une combinaison linéaire de  $2r + 1$  vecteurs,  $r$  étant la mémoire de l'algorithme (Wang *et al.*, 2004; Yuan, 2007). Cette constatation a mené à l'élaboration d'une stratégie d'accélération du schéma de descente itérative (Conn *et al.*, 1994a; Elad *et al.*, 2006; Zibulevsky et Elad, 2010), basée sur l'introduction d'un sous-espace. La présentation de cette technique fait l'objet de la section suivante.

## 3.4 Méthode itérative de sous-espace

### 3.4.1 Principe

La *méthode itérative de sous-espace* consiste à explorer le critère non plus le long d'une direction mais à l'intérieur d'un sous-espace engendré par un ensemble de directions. Le problème de minimisation de  $F$  est résolu par le schéma itératif

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{D}_k \mathbf{s}_k \quad (3.43)$$

où

$$\mathbf{D}_k = [\mathbf{d}_k^1, \dots, \mathbf{d}_k^r] \quad (3.44)$$

est un ensemble de directions de taille  $1 \leq r \ll n$  et  $\mathbf{s}_k$  est un pas multi-dimensionnel ayant pour objectif de minimiser  $f(\mathbf{s}) = F(\mathbf{x}_k + \mathbf{D}_k \mathbf{s})$ .

L'objectif de cette approche est d'accélérer la convergence de l'algorithme de descente itérative, tout en gardant un coût de calcul similaire par itération. Le schéma d'optimisation (3.43) a donné lieu à de nombreux algorithmes (Miele et Cantrell, 1969; Narkiss et Zibulevsky, 2005; Wolfe et Viazminsky, 1976; Zibulevsky, 2008) qui se distinguent par la procédure utilisée pour définir le sous-espace et pour déterminer le pas.

### 3.4.2 Construction du sous-espace

En augmentant le domaine de recherche à plusieurs directions, on espère entraîner une accélération de la convergence de l'algorithme de descente. Cependant, la minimisation du critère dans le sous-espace engendré risque d'être plus coûteuse qu'une recherche de pas scalaire. De ce fait,

Acronyme	Directions $\mathbf{D}_k$	Taille $r$
MG	$[-\mathbf{g}_k, \mathbf{d}_{k-1}]$	2
SMG	$[-\mathbf{g}_k, \mathbf{d}_{k-1}, \dots, \mathbf{d}_{k-m}]$	$m + 1$
SMD	$[\mathbf{p}_k, \mathbf{d}_{k-1}, \dots, \mathbf{d}_{k-m}]$	$m + 1$
GS	$[-\mathbf{g}_k, -\mathbf{g}_{k-1}, \dots, -\mathbf{g}_{k-m}]$	$m + 1$
ORTH	$[-\mathbf{g}_k, \mathbf{x}_k - \mathbf{x}_0, \sum_{i=0}^k w_i \mathbf{g}_i]$	3
SESOP	$[-\mathbf{g}_k, \mathbf{x}_k - \mathbf{x}_0, \sum_{i=0}^k w_i \mathbf{g}_i, \mathbf{d}_{k-1}, \dots, \mathbf{d}_{k-m}]$	$m + 3$
QNS	$[-\mathbf{g}_k, \boldsymbol{\delta}_{k-1}, \dots, \boldsymbol{\delta}_{k-m}, \mathbf{d}_{k-1}, \dots, \mathbf{d}_{k-m}]$	$2m + 1$
SESOP-TN	$[\mathbf{d}_k^\ell, \mathbf{Q}_k(\mathbf{d}_k^\ell), \mathbf{d}_k^\ell - \mathbf{d}_k^{\ell-1}, \mathbf{d}_{k-1}, \dots, \mathbf{d}_{k-m}]$	$m + 3$

TABLE 3.1 – Ensembles de directions dont sont issus la majorité des algorithmes de sous-espace existants. Les poids  $w_i$  et les vecteurs  $\boldsymbol{\delta}_i$  sont définis par (3.50) et (3.51), respectivement.  $\mathbf{Q}_k$  est définie par (3.52), et  $\mathbf{d}_k^\ell$  est la  $\ell$ ème itérée de l’algorithme de gradient conjugué résolvant  $\mathbf{Q}_k(\mathbf{d}) = \mathbf{0}$ .

la construction du sous-espace doit résulter d’un compromis entre un gain en terme de vitesse de convergence et un accroissement du coût de calcul par itération.

Nous présentons dans cette section les principaux algorithmes itératifs de sous-espaces recensés dans la littérature et leur ensemble de directions associé. Les différentes stratégies de constructions sont résumées dans le tableau 3.1. Le paramètre  $r$  correspond à la taille du sous-espace.

Comme pour l’algorithme de descente, nous distinguons les méthodes du premier ordre et celles du second ordre.

### 3.4.2.1 Sous-espace du premier ordre

Dans les six premiers algorithmes du tableau 3.1,  $\mathbf{D}_k$  est composé des gradients précédents et des directions précédentes définies par

$$\mathbf{d}_{k-i} = \mathbf{x}_{k-i+1} - \mathbf{x}_{k-i}, \quad i = 1, \dots, m. \quad (3.45)$$

A notre connaissance, le premier algorithme itératif de sous-espace est présenté dans (Miele et Cantrell, 1969), sous le nom de *mémoire de gradient* (MG). Dans le cas de la minimisation d’une fonction  $F$  quadratique, (Cantrell, 1969) démontre l’équivalence entre les algorithmes de gradient conjugué et de mémoire de gradient. (Ce résultat ne dépend pas du choix de la conjugaison dans le gradient conjugué, les méthodes étant toutes équivalentes dans le cas quadratique.)

Le sous-espace MG a été rapidement étendu dans (Cragg et Levy, 1969) en sous-espace de super mémoire de gradient (SMG), incluant plusieurs directions précédentes. De nombreux travaux récents utilisent ce sous-espace en association avec la recherche de pas par pondération analytique (Liu et al., 2009; Narushima et Yabe, 2006; Yu, 2008), que nous décrivons dans le chapitre suivant. L’algorithme SMG est généralisé sous la forme de super mémoire de descente (SMD) par (Wolfe et Viazminsky, 1976) en remplaçant la direction de plus grande pente par un vecteur quelconque  $\mathbf{p}_k$  non orthogonal au gradient.

Les algorithmes PCD-SESOP et SSF-SESOP développés par (Elad et al., 2006) pour la minimisation de critères de moindres carrés pénalisés sont des algorithmes SMD particuliers, où  $\mathbf{p}_k$  est égale respectivement à la direction de “parallel coordinate descent” (PCD)

$$\mathbf{p}_{i,k} = \arg \min_{\alpha} F(\mathbf{x}_k + \alpha \mathbf{e}_i), \quad i = 1, \dots, N, \quad (3.46)$$

et de “separable surrogate functional” (SSF)

$$\mathbf{p}_{i,k} = \arg \min_{\alpha} H(\mathbf{x}_k + \alpha \mathbf{e}_i, \mathbf{x}_k), \quad i = 1, \dots, N, \quad (3.47)$$

où  $H$  est une fonction séparable  $H(\mathbf{x}, \mathbf{y}) = \sum_i h_i(x_i, \mathbf{y})$  qui vérifie

$$H(\mathbf{x}, \mathbf{y}) \geq F(\mathbf{y}), \quad H(\mathbf{y}, \mathbf{y}) = F(\mathbf{y}), \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n. \quad (3.48)$$

Le sous-espace GS est présenté dans les travaux de Shi & Shen (Shi et Shen, 2005, 2006, 2007) sous le terme ‘super mémoire de gradient’. Cependant, il diffère de ce dernier car il est construit uniquement à partir des gradients et nous nous y référerons sous l’appellation *sous-espace de gradient* afin de le distinguer avec l’algorithme SMG.

Le sous-espace orthogonal (ORTH) algorithm est introduit dans (Nemirovski, 1982) dans l’objectif de construire un algorithme du premier ordre avec un taux de convergence théorique optimal. Le sous-espace ORTH est composé de la direction de plus grande pente, à laquelle deux directions dites de Nemirovski sont ajoutées,  $\mathbf{x}_k - \mathbf{x}_0$  et

$$\sum_{i=0}^k w_i \mathbf{g}_i, \quad (3.49)$$

où  $w_i$  sont des poids définis par la récurrence :

$$w_i = \begin{cases} 1 & \text{si } i = 0, \\ \frac{1}{2} + \sqrt{\frac{1}{4} + w_{i-1}^2} & \text{sinon.} \end{cases} \quad (3.50)$$

Dans (Narkiss et Zibulevsky, 2005), les directions précédentes sont rajoutées au sous-espace de Nemirovski. Cette modification conduit à l’algorithme SESOP dont la supériorité par rapport à ORTH est démontrée sur des exemples de restauration d’image. De plus, les tests menés par Zibulevsky *et coll.* démontrent que la présence des directions de Nemirovski dans le sous-espace SESOP n’améliore pas la vitesse de convergence pratique de l’algorithme. Par conséquent, dans (Zibulevsky et Elad, 2010), ces vecteurs supplémentaires ne sont plus utilisés, ce qui fait que l’algorithme SESOP actuel est identique à l’algorithme SMG de (Cragg et Levy, 1969).

### 3.4.2.2 Sous-espace du second ordre

Les deux derniers algorithmes du tableau 3.1 utilisent des directions du second ordre. Dans l’algorithme QNS proposé par (Wang *et al.*, 2004; Yuan, 2007),  $\mathbf{D}_k$  correspond au sous-espace SMG augmenté par

$$\boldsymbol{\delta}_{k-i} = \mathbf{g}_{k-i+1} - \mathbf{g}_{k-i}, \quad i = 1, \dots, m. \quad (3.51)$$

Cette construction est liée à l’algorithme L-BFGS (Liu et Nocedal, 1989), car ce dernier produit des directions dans l’espace généré par l’ensemble  $\mathbf{D}_k$  résultant de la construction QNS.

Le sous-espace SESOP-TN a été développé par (Zibulevsky, 2008) dans l’objectif de limiter la sensibilité de l’algorithme de Newton tronqué (TN) à un mauvais réglage de la troncature. La première composante de  $\mathbf{D}_k$  est la direction de Newton tronqué  $\mathbf{d}_k^\ell$  obtenues après  $\ell$  itérations de GC pour résoudre le système de Gauss-Newton  $\mathbf{Q}_k(\mathbf{d}) = \mathbf{0}$ , où

$$\mathbf{Q}_k(\mathbf{d}) = \nabla^2 F(\mathbf{x}_k) \mathbf{d} + \mathbf{g}_k. \quad (3.52)$$

Les deuxième et troisième composantes de SESOP-TN résultent aussi des itérations de gradient conjugué.

### 3.4.3 Préconditionnement

L’introduction d’un preconditionneur dans un algorithme de sous-espace intervient directement dans la construction du sous-espace lui-même. Une version preconditionnée de l’algorithme s’obtient en remplaçant dans  $\mathbf{D}_k$  tous les termes de gradients  $\mathbf{g}_i$  par  $\mathbf{M}_i \mathbf{g}_i$ , où  $\mathbf{M}_i$  est une version approchée de l’inverse du Hessien en  $\mathbf{x}_i$ . En ce qui concerne le sous-espace SESOP-TN, un preconditionneur peut être utilisé dans l’étape de gradient conjugué pour accélérer l’étape de calcul de la direction de Newton tronqué.

## 3.5 Conclusion

La résolution d'un problème inverse linéaire passe par la minimisation d'une fonction objectif constituée d'un terme d'adéquation aux données et d'un terme de pénalisation. Lorsque la fonction est différentiable, une stratégie d'optimisation efficace consiste à utiliser un schéma de descente itérative alliant construction d'une direction de descente et recherche de pas. La stratégie de construction de la direction est fixée en fonction des propriétés du problème (forme du critère, taille du problème) et des informations disponibles sur le critère (en particulier de son Hessien). L'algorithme de descente itérative peut être accéléré en remplaçant la direction par un ensemble de directions, usuellement issues des itérations précédentes, donnant lieu à l'algorithme itératif de sous-espace. Dans ce nouvel algorithme, la recherche de pas vise à minimiser le critère dans le sous-espace généré par les directions.

Le choix de la méthode de calcul du pas est important car la valeur de pas influe sur les propriétés de convergence et sur la complexité de l'algorithme. Le prochain chapitre est consacré à la présentation des stratégies de recherche de pas dans le cas scalaire et multi-dimensionnel. Nous discutons des stratégies de pas usuelles et de leurs propriétés de convergence. De plus, nous proposons une nouvelle procédure de pas multi-dimensionnelle pour l'algorithme de sous-espace.



# Chapitre 4

## Recherche de pas par Majoration-Minoration quadratique

### Sommaire

---

4.1	Introduction . . . . .	<b>40</b>
4.2	Recherche de pas pour l'algorithme de descente itérative . . . . .	<b>40</b>
4.2.1	Règles de pas exactes . . . . .	41
4.2.2	Règles de pas approchées . . . . .	41
4.2.3	Calcul du pas . . . . .	43
4.2.4	Condition de Zoutendijk . . . . .	46
4.3	Recherche de pas pour l'algorithme itératif de sous-espace . . . . .	<b>46</b>
4.3.1	Règle de pas multi-dimensionnelle . . . . .	46
4.3.2	Pondération analytique pour l'algorithme SMG . . . . .	47
4.3.3	Méthode de régions de confiance . . . . .	48
4.4	Algorithme de Majoration-Minoration - Application à la recherche de pas . . . . .	<b>50</b>
4.4.1	Approximation tangente majorante quadratique . . . . .	51
4.4.2	La récurrence de Majoration-Minoration . . . . .	52
4.4.3	Recherche de pas MMQ scalaire . . . . .	54
4.4.4	Recherche de pas MMQ multi-dimensionnelle . . . . .	55
4.5	Propriétés de convergence du pas MM quadratique . . . . .	<b>56</b>
4.5.1	Hypothèses . . . . .	56
4.5.2	Propriétés de la séquence de pas . . . . .	56
4.5.3	Convergence des algorithmes gradient reliés . . . . .	58
4.5.4	Convergence des algorithmes de gradient conjugué non linéaire . . . . .	62
4.6	Conclusion . . . . .	<b>63</b>

---

## 4.1 Introduction

Dans l'algorithme de descente itérative ainsi que dans sa forme accélérée par l'introduction d'un sous-espace, le schéma d'optimisation se formule comme la détermination à chaque itération d'un pas minimisant le critère dans un sous-espace généré par une ou plusieurs directions de recherche. Ce chapitre est consacré à la technique employée pour déterminer le pas, appelée *recherche de pas* ou *stratégie de pas*. Avoir une bonne stratégie de pas est fondamental car cela influe de façon importante sur les propriétés de convergence de l'algorithme.

Lorsque le sous-espace se limite à une unique direction, la recherche de pas consiste en la minimisation approchée d'une fonction scalaire. Nous présentons en section 4.2 les stratégies usuelles pour le calcul du pas scalaire, également appelées stratégies de recherche linéaire, en nous basant principalement sur les ouvrages (Bertsekas, 1999; Nocedal et Wright, 1999). La section 4.3 est consacrée aux recherches de pas utilisées dans l'implémentation de l'algorithme de sous-espace.

Les méthodes de recherche linéaire standards, basées sur des schémas itératifs, ont été récemment mises en concurrence avec une technique issue de l'optimisation par Majoration-Minoration quadratique (MMQ 1D), qui donne lieu à des formules de pas analytiques (Labat, 2006; Labat et Idier, 2008; Sun et Zhang, 2001). La stratégie de pas MMQ 1D fait l'objet des sections 4.4 et 4.5. Nous montrons dans la section 4.4 qu'elle s'étend de façon directe en une stratégie de pas multi-dimensionnelle MMQ  $rD$  pour les algorithmes de sous-espace. L'analyse des propriétés de convergence du pas MMQ 1D et  $rD$  est présentée en section 4.5. Nous étendons les résultats de convergence de la méthode MMQ 1D, établis dans (Labat et Idier, 2008) pour les algorithmes de gradient conjugué non linéaire, à la famille des algorithmes de descente itérative gradient reliés (Bertsekas, 1996). De plus, nous montrons que la stratégie de pas MMQ  $rD$  assure la convergence de nombreux algorithmes de sous-espace.

## 4.2 Recherche de pas pour l'algorithme de descente itérative

Dans l'algorithme de descente itérative, à chaque itération, après avoir construit une direction de descente  $\mathbf{d}_k$ , il faut déterminer le pas  $\alpha_k$  spécifiant le déplacement le long de la direction. Comme l'objectif de l'algorithme est de minimiser  $F$ , il semble naturel de définir  $\alpha_k$  comme solution du problème :

$$\min_{\alpha \geq 0} f(\alpha) \tag{4.1}$$

où  $f$  est la restriction de  $F$  à la droite  $\{\mathbf{x}_k + \alpha \mathbf{d}_k : \alpha \in \mathbb{R}\}$ , définie par

$$f(\alpha) = F(\mathbf{x}_k + \alpha \mathbf{d}_k). \tag{4.2}$$

Cependant, la résolution exacte peut demander un temps de calcul trop grand et ne peut de toute façon pas être faite avec une précision infinie. De plus, l'efficacité supplémentaire éventuellement apportée à l'algorithme par un pas exact permet rarement de compenser le temps consacré à déterminer un tel pas. Par conséquent, les stratégies usuelles pour le calcul du pas sont basées sur la vérification de conditions moins restrictives qui permettent toutefois de garantir la convergence des algorithmes.

Nous présentons dans cette section les règles de pas usuelles, numérotées de (a) à (i). Nous faisons la distinction entre les règles exactes visant à une minimisation exacte de la fonction  $f$  sur un intervalle et les règles approchées nécessitant une minimisation inexacte de  $f$ . Les règles approchées sont vérifiées par un ensemble de pas s'exprimant sous la forme d'un intervalle. Nous décrivons ensuite les méthodes itératives employées pour déterminer un pas à l'intérieur de cet intervalle. Remarquons que par définition de la fonction  $f$ , nous avons équivalence entre les notations

suivantes :

$$f(0) \equiv F(\mathbf{x}_k) \quad (4.3)$$

$$\dot{f}(\alpha) \equiv \nabla F(\mathbf{x}_k + \alpha \mathbf{d}_k)^T \mathbf{d}_k \quad (4.4)$$

$$\ddot{f}(\alpha) \equiv \mathbf{d}_k^T \nabla^2 F(\mathbf{x}_k + \alpha \mathbf{d}_k) \mathbf{d}_k \quad (4.5)$$

### 4.2.1 Règles de pas exactes

La détermination du minimiseur exact de  $f$  correspond à la règle de Cauchy. Lorsque la fonction  $f$  est multimodale, il peut être plus simple de rechercher le minimiseur dans un intervalle restreint (règle de Cauchy limitée) ou bien de déterminer le plus petit point stationnaire de  $f$  (règle de Curry) (Bertsekas, 1999). Ces règles, qualifiées d'exactes, ne sont utilisées que lorsque le minimiseur s'exprime de façon analytique.

(a) Règle de Cauchy

$$\alpha_k = \arg \min_{\alpha \in \mathbb{R}^+} F(\mathbf{x}_k + \alpha \mathbf{d}_k) \quad (4.6)$$

(b) Règle de Cauchy limitée

Le pas  $\alpha_k$  est défini par

$$\alpha_k = \arg \min_{\alpha \in [0, s]} F(\mathbf{x}_k + \alpha \mathbf{d}_k) \quad (4.7)$$

(c) Règle de Curry

$$\alpha_k = \min \{ \alpha > 0 \mid \nabla F(\mathbf{x}_k + \alpha \mathbf{d}_k)^T \mathbf{d}_k = 0 \} \quad (4.8)$$

### 4.2.2 Règles de pas approchées

#### 4.2.2.1 Condition de descente suffisante

La minimisation exacte n'implique pas forcément des performances optimales de l'algorithme. De plus, elle n'est pas nécessaire pour assurer la convergence, dès lors que des conditions sont assurées par le pas utilisé. Une condition naturelle est de demander que le pas entraîne une décroissance suffisante de la fonction  $f$ . Cela se traduit le plus souvent par une inégalité de la forme :

$$F(\mathbf{x}_k + \alpha_k \mathbf{d}_k) \leq F(\mathbf{x}_k) + (\text{terme} < 0) \quad (4.9)$$

appelée *condition de descente suffisante*. Une manière simple d'assurer la décroissance suffisante du critère est d'imposer que  $f$  décroisse autant qu'une proportion  $c_1$  de la décroissance de l'approximation linéaire de  $f$  en  $\alpha = 0$ . Cela conduit à l'inégalité suivante, appelée *première condition de Wolfe* ou *condition d'Armijo* :

$$F(\mathbf{x}_k + \alpha_k \mathbf{d}_k) - F(\mathbf{x}_k) \leq c_1 \alpha_k \mathbf{g}_k^T \mathbf{d}_k, \quad c_1 \in (0, 1). \quad (4.10)$$

Cette condition signifie qu'en  $\alpha_k$ , la fonction  $f$  doit prendre une valeur inférieure à celle prise par la fonction affine  $\alpha \mapsto f(0) + c_1 \alpha \mathbf{g}_k^T \mathbf{d}_k$ . En pratique, la constante  $c_1$  est prise très petite, de manière à favoriser l'acceptation de pas grands. Typiquement,  $c_1 = 10^{-4}$  ( Nocedal et Wright, 1999, Chap.1). Dans les algorithmes de second ordre, il est important de choisir  $c_1 < \frac{1}{2}$  pour que le pas unité soit accepté lorsque  $\mathbf{x}$  est proche d'une solution.

#### 4.2.2.2 Règles de pas monotones

Comme  $\mathbf{d}_k$  est une direction de descente, l'inégalité (4.10) sera satisfaite par tout pas arbitrairement petit. Afin d'éviter une 'fausse convergence' vers un point non stationnaire, une condition supplémentaire doit être vérifiée afin d'empêcher le pas  $\alpha_k > 0$  d'être trop proche de zéro. L'association de cette condition avec la première condition de Wolfe permet de définir une règle de pas dont la vérification assurera la convergence de l'algorithme de descente.

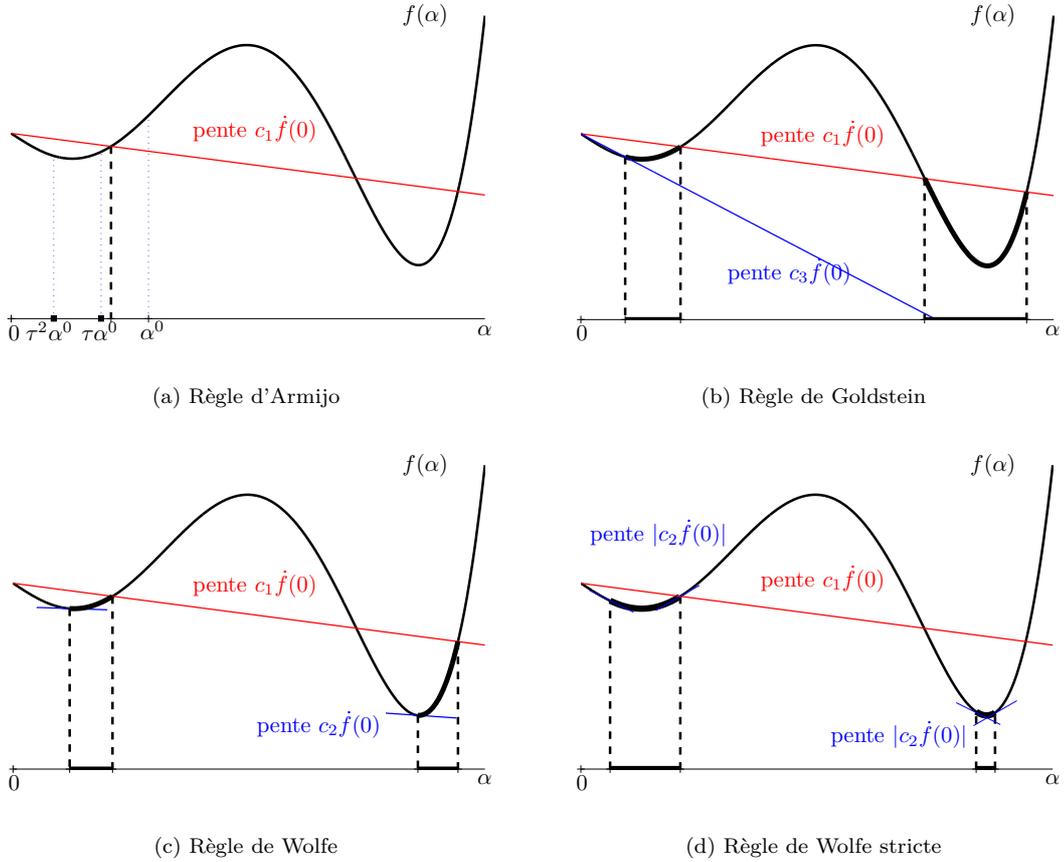


FIGURE 4.1 – Principe schématique des règles de pas usuelles. Les pas acceptables sont indiqués en trait épais.

(d) Règle d'Armijo (Figure 4.1(a)) ([Armijo, 1966](#))

Soient  $\tau \in (0, 1)$  et  $\alpha^0 > 0$ . La règle d'Armijo est vérifiée en  $\alpha_k = \alpha^0 \tau^m$  où  $m$  est le plus petit entier naturel tel que l'on ait (4.10). Dans les algorithmes de second ordre, il est opportun de prendre  $\alpha^0 = 1$  afin d'optimiser la vitesse de convergence de l'algorithme.

(e) Règle de Goldstein (Figure 4.1(b)) ([Goldstein et Price, 1967](#))

Elle est respectée si (4.10) est vérifiée et si il existe  $c_3 \in (c_1, 1)$  tel que

$$F(\mathbf{x}_k + \alpha_k \mathbf{d}_k) - F(\mathbf{x}_k) \geq c_3 \alpha_k \mathbf{g}_k^T \mathbf{d}_k \quad (4.11)$$

(f) Règle de Wolfe (Figure 4.1(c)) ([Wolfe, 1969](#))

Elle est respectée si (4.10) est vérifiée et si il existe  $c_2 \in (c_1, 1)$  tel que

$$\nabla F(\mathbf{x}_k + \alpha_k \mathbf{d}_k)^T \mathbf{d}_k \geq c_2 \mathbf{g}_k^T \mathbf{d}_k \quad (4.12)$$

(g) Règle de Wolfe stricte (Figure 4.1(d)) ([Wolfe, 1969](#))

Elle est respectée si (4.10) est vérifiée et si il existe  $c_2 \in (c_1, 1)$  tel que

$$|\nabla F(\mathbf{x}_k + \alpha_k \mathbf{d}_k)^T \mathbf{d}_k| \leq c_2 |\mathbf{g}_k^T \mathbf{d}_k| \quad (4.13)$$

La règle (4.13), plus difficile à vérifier que (4.12), est principalement utilisée en association avec la direction de gradient conjugué non linéaire ([Hager et Zhang, 2006b](#)). Pour certaines formules de conjugaison, elle assure que l'ensemble des directions générées sont des directions de descente ([Gilbert et Nocedal, 1992](#)).

### 4.2.2.3 Règles de pas non monotones

Dans certains algorithmes, il n'est pas nécessaire d'assurer la décroissance du critère pour obtenir la convergence. Il arrive même que la monotonie freine la progression des itérés en les bloquant dans des minimas locaux ou des points selles. Cette remarque a motivé l'introduction d'une nouvelle règle de pas (Grippo *et al.*, 1986) pouvant être vue comme une version non monotone de la règle d'Armijo et s'énonçant :

(h) Règle d'Armijo non monotone

Soient  $\tau \in (0, 1)$ ,  $\alpha^0 > 0$  et  $M \geq 0$ . La règle d'Armijo non monotone est vérifiée en  $\alpha_k = \alpha^0 \tau^m$  où  $m$  est le plus petit entier naturel tel que l'on ait :

$$F(\mathbf{x}_k + \alpha_k \mathbf{d}_k) \leq F_{\max}^M + c_1 \alpha_k \mathbf{g}_k^T \mathbf{d}_k \quad (4.14)$$

où  $F_{\max}^M$  est la valeur maximale prise par le critère sur les  $M$  dernières itérations de l'algorithme de descente.

Lorsque  $M$  est nul, on retrouve la règle d'Armijo classique. La recherche de pas non monotone est classiquement utilisée en association avec la direction de Barzilai-Borwein pour la résolution de problèmes non contraints (Fletcher, 2006; Raydan, 1997) ou contraints (Birgin *et al.*, 2000; Hager et Zhang, 2006a). Cependant, comme souligné dans (Dai et Yuan, 2002), cette méthode présente des inconvénients. Tout d'abord, son efficacité dépend de façon importante du choix de  $M$ . De plus, une bonne valeur du critère sera peu prise en compte à cause de l'opération max dans (4.14). Pour ces raisons, des versions hybrides, alternant règles monotone et non monotone ont été proposées dans (Dai et Yuan, 2002; Hager et Zhang, 2006a; Zhang et Hager, 2004).

## 4.2.3 Calcul du pas

La recherche de pas a pour objectif de déterminer un pas satisfaisant l'une des règles présentées précédemment. Les règles de pas exactes de (a) à (c) requièrent le calcul du minimiseur de  $f$  et nécessitent l'utilisation d'une stratégie d'optimisation. Les méthodes de recherche de pas exactes les plus courantes utilisent un algorithme de recherche de zéro, telles que la dichotomie, la section dorée, la méthode de la sécante et la méthode de Brent, appliqué à la fonction dérivée  $\dot{f}(\alpha)$  (Bonnans *et al.*, 2006, Chap.3). Chacune des règles inexactes de (d) à (h) définit un ensemble de pas acceptables. Ce dernier est non vide si  $f$  est bornée inférieurement (Nocedal et Wright, 1999) et la procédure de *recherche de pas* va consister à déterminer une valeur de  $\alpha_k$  à l'intérieur de cet ensemble. Nous présentons les méthodes usuelles permettant de déterminer un pas vérifiant l'une de ces cinq règles. Ces méthodes sont basées sur le schéma itératif suivant : Une valeur initiale de pas est choisie et tant que la règle n'est pas respectée, une technique d'interpolation/dichotomie est utilisée pour générer le prochain pas à tester.

### 4.2.3.1 Technique de rebroussement

La technique de rebroussement est une technique très simple qui permet d'obtenir un pas vérifiant la règle d'Armijo (d) et s'adapte de façon directe à la règle d'Armijo non monotone (h). On essaie d'abord un pas initial  $\alpha^0 > 0$ , et s'il ne vérifie pas la première condition de Wolfe, on rebrousse chemin en essayant des pas plus petits  $\alpha^0 \tau$ ,  $\alpha^0 \tau^2$ , etc avec  $\tau \in (0, 1)$ . Cette technique, appelée en anglais 'backtracking', est principalement utilisée dans les algorithmes de Newton et de gradient projeté (Iusem, 2003). Une version plus élaborée, présentée dans (Bonnans *et al.*, 2006), est décrite dans l'Algorithme 4.1.

### 4.2.3.2 Technique de dichotomie

La technique de dichotomie permet de déterminer un pas qui vérifie les règles (e), (f) ou (g) (respectivement Wolfe, Wolfe stricte ou Goldstein). La procédure est itérative et chaque itération se divise en deux phases. Tout d'abord, un intervalle  $[a, b]$  contenant au moins un pas acceptable est sélectionné. Puis ce pas est extrait de l'intervalle par une technique d'exploration qui résout

**Entrées :**  $\alpha^0 > 0$  et  $\tau \in (0, \frac{1}{2})$   
**Sorties :**  $\alpha$  vérifiant la règle d'Armijo  
**tantque** (4.10) n'est pas vérifiée **faire**  
     Choisir  $\alpha^j$  dans  $[\tau\alpha^j, (1 - \tau)\alpha^j]$  par interpolation  
      $j \leftarrow j + 1$   
**fin tantque**

Algorithme 4.1 – Algorithme de rebroussement pour la règle d'Armijo

la règle de pas par dichotomie dans l'intervalle  $[a, b]$ . Le découpage de l'intervalle est basé sur une interpolation de la fonction  $f$  entre  $a$  et  $b$ .

Les algorithmes 4.2 et 4.3 fournissent un pas vérifiant respectivement la règle de Wolfe et de Goldstein. La recherche d'un pas vérifiant la règle de Wolfe stricte est plus fastidieuse. En particulier, il est nécessaire que les bornes de l'intervalle  $[a, b]$  sélectionné vérifient les conditions suivantes (Nocedal et Wright, 1999) :

- Il existe au moins un pas  $\alpha \in [a, b]$  vérifiant la règle de Wolfe stricte
- L'inégalité  $\dot{f}(a)(b - a) < 0$  est vérifiée
- Parmi tous les pas générés par l'algorithme de recherche et qui vérifient la condition (4.10),  $a$  est celui qui entraîne la plus grande diminution de  $f$

Une implémentation efficace permettant d'obtenir un pas qui vérifie la règle de Wolfe stricte est décrite dans (Moré et Thuente, 1994).

**Entrées :**  $\alpha, a = 0$  et  $b = 0$   
**Sorties :**  $\alpha$  vérifiant la règle de Wolfe  
**répéter**  
     **si**  $f(\alpha) \leq f(0) + c_1\alpha\dot{f}(0)$  et  $\dot{f}(\alpha) \geq c_2\dot{f}(0)$  **alors**  
         Retourner  $\alpha$   
     **sinon**  
         **si**  $f(\alpha) > f(0) + c_1\alpha\dot{f}(0)$  **alors**  
              $b = \alpha$   
         **sinon si**  $f(\alpha) \leq f(0) + c_1\alpha\dot{f}(0)$  et  $\dot{f}(\alpha) < c_2\dot{f}(0)$  **alors**  
              $a = \alpha$   
     **finsi**  
**finsi**  
     **si**  $b = 0$  **alors**  
         Choisir  $\alpha > a$  par interpolation  
     **sinon si**  $b > 0$  **alors**  
         Choisir  $\alpha \in (a, b)$  par interpolation  
     **finsi**  
**jusqu'à Fin**

Algorithme 4.2 – Détermination d'un pas satisfaisant la règle de Wolfe

#### 4.2.3.3 Interpolation polynomiale

Les recherches de pas par rebroussement ou par dichotomie nécessitent le choix d'un pas à l'intérieur d'un intervalle  $[a, b]$ . La détermination de ce pas est basée sur une approximation du

```

Entrées :  $\alpha$ ,  $a = 0$  et  $b = 0$ 
Sorties :  $\alpha$  vérifiant la règle de Goldstein
répéter
  si  $f(\alpha) \leq f(0) + c_1\alpha\dot{f}(0)$  alors
    si  $f(\alpha) \geq f(0) + c_3\alpha\dot{f}(0)$  alors
      Retourner  $\alpha$ 
    sinon
       $a = \alpha$ 
    finsi
  sinon
     $b = \alpha$ 
  finsi
si  $b = 0$  alors
  Choisir  $\alpha > a$  par interpolation
sinon si  $b > 0$  alors
  Choisir  $\alpha \in (a, b)$  par interpolation
finsi
jusqu'à Fin

```

Algorithme 4.3 – Détermination d'un pas satisfaisant la règle de Goldstein

critère par une fonction d'interpolation (Press *et al.*, 1992, Chap.3). Les méthodes de recherche de pas standards utilisent des interpolations polynomiales cubiques (Dennis *et Schnabel*, 1996, Chap.6), (Fletcher, 1987, Chap.2)

$$f(\alpha) \approx h(\alpha) = f_0 + f_1\alpha + f_2\alpha^2 + f_3\alpha^3 \quad (4.15)$$

où la valeur des  $f_i$  dépend des points de coïncidence requis entre  $f$  et  $h$  :

- 1)  $f(a)$ ,  $f(b)$  et  $\dot{f}(a)$ .
- 2)  $f(a)$ ,  $\dot{f}(a)$  et  $\dot{f}(b)$ .
- 3)  $f(a)$ ,  $f(b)$  et  $f(c)$ .
- 4)  $f(a)$ ,  $f(b)$ ,  $\dot{f}(a)$  et  $\dot{f}(b)$ , avec  $\dot{f}(a)\dot{f}(b) < 0$ .

Dans les cas 1), 2) et 3), le coefficient  $f_3$  est fixé à zéro et la fonction d'interpolation est quadratique. Le minimiseur du polynôme interpolant s'obtient de façon analytique, comme le résume le tableau 4.1.

Cas	Interpolation	Minimiseur
1)	quadratique	$a - \frac{1}{2} \frac{(a-b)\dot{f}(a)}{\dot{f}(a) - ((\dot{f}(a) - \dot{f}(b))/(a-b))}$
2)	quadratique	$a - \frac{(a-b)\dot{f}(a)}{\dot{f}(a) - \dot{f}(b)}$
3)	quadratique	$\frac{1}{2}(a+b) + \frac{1}{2} \frac{(f(a)-f(b))(b-c)(c-a)}{(b-c)f(a) + (c-a)f(b) + (a-b)f(c)}$
4)	cubique	$a + \frac{-(b-a)\dot{f}(a)}{z+w-\dot{f}(a)}$ avec $z = 3\frac{\dot{f}(b)-\dot{f}(a)}{b-a} - \dot{f}(a) - \dot{f}(b)$ et $w^2 = z^2 - \dot{f}(a)\dot{f}(b)$

TABLE 4.1 – Fonctions interpolantes polynomiales (Sun *et Yuan*, 2006, Chap.2).

Le pas est ensuite choisi comme combinaison linéaire des quatre minimiseurs, afin de garantir son appartenance à l'intervalle  $[a, b]$  (Al-Baali et Fletcher, 1986; Moré et Thuente, 1994).

#### 4.2.4 Condition de Zoutendijk

La condition de Zoutendijk est un outil assez technique qui permet de démontrer l'apport de la stratégie de pas dans la convergence d'un algorithme de descente itérative. Nous donnons sa définition formelle ci-dessous.

**Définition 4.1.** *Condition de Zoutendijk*

On dit qu'une stratégie de pas satisfait la condition de Zoutendijk s'il existe une constante  $C > 0$  telle que pour tout indice  $k$ , on ait

$$F(\mathbf{x}_k + \alpha_k \mathbf{d}_k) \leq F(\mathbf{x}_k) - C \|\mathbf{g}_k\|^2 \cos^2 \theta_k \quad (4.16)$$

où  $\theta_k$  est l'angle que fait la direction  $\mathbf{d}_k$  avec  $-\mathbf{g}_k$ , défini par :

$$\cos \theta_k = \frac{-\mathbf{g}_k^T \mathbf{d}_k}{\|\mathbf{g}_k\| \|\mathbf{d}_k\|} \quad (4.17)$$

Lorsque la suite  $\{F(\mathbf{x}_k)\}_{k \geq 0}$  est bornée inférieurement, une condition équivalente très utilisée en pratique est (Bertsekas, 1999)

$$\sum_{k=0}^{\infty} \|\mathbf{g}_k\|^2 \cos^2 \theta_k < \infty \quad (4.18)$$

La condition de Zoutendijk est vérifiée lorsque le pas satisfait les règles de pas de (a) à (h), à condition que le critère soit à gradient Lipschitz et borné inférieurement (Bertsekas, 1999; Dai, 2002; Shi, 2004). Bien souvent, démontrer la convergence d'un algorithme de descente itérative consiste à associer la condition de Zoutendijk avec des propriétés de la direction de descente  $\mathbf{d}_k$  (Nocedal et Wright, 1999, Chap.3).

### 4.3 Recherche de pas pour l'algorithme itératif de sous-espace

Dans un algorithme de sous-espace, le calcul du pas consiste en une *recherche multi-dimensionnelle* visant à minimiser de façon approchée la fonction  $f(\mathbf{s}) = F(\mathbf{x}_k + \mathbf{D}_k \mathbf{s})$ , avec  $\mathbf{s} \in \mathbb{R}^r$ ,  $r$  étant la dimension du sous-espace. Comme dans le cas scalaire, la méthode de recherche de pas dans le sous-espace doit réaliser un compromis entre une résolution exacte et un coût d'implémentation faible.

#### 4.3.1 Règle de pas multi-dimensionnelle

Dans l'algorithme de descente itérative, le pas est obtenu par une procédure itérative qui génère des valeurs de pas jusqu'à vérifier des règles de convergence telles que Armijo, Wolfe ou Goldstein. Une formulation de ces conditions pour le cas multi-dimensionnel s'obtient facilement en remplaçant  $\mathbf{d}_k$  par  $\mathbf{D}_k \mathbf{s}_k$  et  $\alpha_k$  par 1 dans les inégalités. Par exemple, pour la règle de Goldstein, on obtient (Conn et al., 1994a) :

$$F(\mathbf{x}_k) + c_1 \nabla F(\mathbf{x}_k)^T \mathbf{D}_k \mathbf{s}_k \leq F(\mathbf{x}_{k+1}) \leq F(\mathbf{x}_k) + c_3 \nabla F(\mathbf{x}_k)^T \mathbf{D}_k \mathbf{s}_k \quad (4.19)$$

soit

$$f(\mathbf{0}) + c_1 \nabla f(\mathbf{0})^T \mathbf{s}_k \leq f(\mathbf{s}_k) \leq f(\mathbf{0}) + c_3 \nabla f(\mathbf{0})^T \mathbf{s}_k. \quad (4.20)$$

La convergence de l'algorithme peut alors être démontrée à condition que la règle de Zoutendijk reformulée en

$$\sum_{k=0}^{\infty} \left( \frac{-\nabla f(\mathbf{0})^T \mathbf{s}_k}{\|\mathbf{s}_k\|} \right)^2 < \infty \quad (4.21)$$

soit vérifiée. Cependant, l'élaboration d'un schéma itératif qui détermine un pas  $\mathbf{s}_k$  vérifiant la condition (4.20) est difficile (Conn *et al.*, 1994a). Par conséquent, des stratégies alternatives ont été proposées dans la littérature.

Une première méthode consiste à déterminer le pas à l'aide d'un algorithme itératif visant à minimiser  $f(\cdot)$  (Narkiss et Zibulevsky, 2005; Zibulevsky, 2008; Zibulevsky et Elad, 2010). La recherche de pas est effectuée par un algorithme de Newton associé à une stratégie de pas d'Armijo. Cependant, à moins d'avoir déterminé le minimiseur exact de  $f$  lors de la recherche de pas, la convergence de l'algorithme de sous-espace associée à cette stratégie de pas n'est pas garantie. Une façon simple d'assurer la convergence, proposée par (Conn *et al.*, 1994a), est d'initialiser la procédure par un pas  $\alpha_k$  résultant d'une recherche linéaire le long de la première direction  $\mathbf{d}_k^1$ . Si  $\mathbf{s}_k$  entraîne une plus grande décroissance du critère que  $\alpha_k$ , alors la convergence de l'algorithme peut être démontrée sous des hypothèses faibles sur  $\mathbf{d}_k^1$ . L'utilisation d'une stratégie de régions de confiance pour la minimisation de  $f$  permet également d'assurer la convergence de l'algorithme (Shi et Shen, 2007; Shi et Xu, 2009; Wang *et al.*, 2004). Enfin, dans le cas du sous-espace de super mémoire de gradient, une méthode très courante consiste à se ramener à un problème de recherche de pas scalaire en fixant la direction de recherche comme une combinaison linéaire des directions du sous-espace dont les poids sont obtenus par une formule analytique (Liu *et al.*, 2009; Yu, 2008) :

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \sum_{i=1}^r s_k^i \mathbf{d}_k^i. \quad (4.22)$$

Nous présentons ces deux dernières méthodes et les propriétés du pas résultant.

### 4.3.2 Pondération analytique pour l'algorithme SMG

La technique de pondération analytique est utilisée dans (Liu *et al.*, 2009; Narushima et Yabe, 2006; Yu, 2008) pour le calcul du pas dans l'algorithme de super mémoire de gradient (SMG). Elle consiste à transformer le problème du calcul de  $\mathbf{s}_k$  en une recherche de pas scalaire le long de la direction  $\mathbf{d}_k = \sum_{i=1}^r s_k^i \mathbf{d}_k^i$  où les  $s_k^i$  sont désormais obtenus par une formule analytique. Cette approche est utilisée exclusivement en association avec le sous-espace de mémoire de gradient. La valeur des poids  $s_k^i$  est fixée de telle façon que la séquence des directions entraîne la convergence de l'algorithme, lorsqu'elle est associée à un choix de pas classique.

**Théorème 4.1.** (Narushima et Yabe, 2006) *Soit un algorithme de descente itérative dont les directions de recherche sont définies par*

$$\mathbf{d}_k = -\gamma_k \mathbf{g}_k + \frac{1}{m} \sum_{i=1}^m \beta_k^i \mathbf{d}_{k-i}, \quad (4.23)$$

avec pour tout  $k, i$ ,

$$\beta_k^i = \|\mathbf{g}_k\|^2 \psi_k^{i*} \quad (4.24)$$

où  $a^*$  représente

$$a^* = \begin{cases} 0 & \text{if } a = 0 \\ \frac{1}{a} & \text{sinon.} \end{cases} \quad (4.25)$$

S'il existe une constante  $\gamma > 0$  telle que  $\gamma_k \geq \gamma$  et que les  $\psi_k^i$  vérifient :

$$\begin{cases} \psi_k^1 > \max(\mathbf{g}_k^T \mathbf{d}_{k-1} / \gamma_k, 0), & i = 1 \\ \psi_k^i \geq \max(\mathbf{g}_k^T \mathbf{d}_{k-i} / \gamma_k, 0), & i = 2, \dots, m \end{cases} \quad (4.26)$$

et

$$\begin{cases} \mathbf{g}_k^T \mathbf{d}_{k-1} + \|\mathbf{g}_k\| \|\mathbf{d}_{k-1}\| < \gamma_k \psi_k^1, & i = 1 \\ \mathbf{g}_k^T \mathbf{d}_{k-i} + \|\mathbf{g}_k\| \|\mathbf{d}_{k-i}\| \leq \gamma_k \psi_k^i, & i = 2, \dots, m \end{cases} \quad (4.27)$$

alors, pour tout  $k$ ,  $\mathbf{d}_k$  est une direction de descente. De plus, la séquence  $\{\mathbf{d}_k\}$  associée à une stratégie de pas vérifiant la condition de Zoutendijk (4.18), entraîne la convergence de l'algorithme au sens  $\lim_{k \rightarrow \infty} \|\mathbf{g}_k\| = 0$ .

Différentes formules analytiques pour les coefficients  $\psi_k^i$  vérifiant les hypothèses techniques du théorème 4.1 sont décrites et analysées dans (Liu *et al.*, 2009; Narushima et Yabe, 2006; Yu, 2008). La méthode n'est cependant pas entièrement analytique car l'étape de recherche de pas scalaire le long de  $\mathbf{d}_k$  est nécessaire pour assurer la convergence.

De notre point de vue, la méthodologie de pondération analytique s'éloigne de l'esprit des algorithmes de sous-espace puisque à aucun moment on ne cherche à minimiser la fonction dans le sous-espace. Nous l'interprétons plutôt comme une généralisation des formules de conjugaison de gradient conjugué non linéaire au cas où plusieurs directions précédentes sont prises en compte. Un autre point à souligner est que les formules obtenues sont valides uniquement pour le sous-espace SMG. En particulier, l'extension du théorème 4.1 au cas où le gradient est remplacé par sa version préconditionnée  $\mathbf{M}_k \mathbf{g}_k$  est un problème ouvert.

### 4.3.3 Méthode de régions de confiance

La méthode de régions de confiance est classiquement utilisée comme substitut à la recherche de pas dans les algorithmes de descente itérative (Conn *et al.*, 1987). Au lieu de rechercher la direction de recherche  $\mathbf{d}_k$  puis le pas  $\alpha_k$  comme le fait un algorithme de descente classique, la méthode de régions de confiance combine les deux étapes en posant  $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{d}_k$  avec

$$\mathbf{d}_k = \min_{\mathbf{d}} \left( Q_k(\mathbf{d}) = F(\mathbf{x}_k) + \mathbf{d}^T \mathbf{g}_k + \frac{1}{2} \mathbf{d}^T \mathbf{B}_k \mathbf{d} \right), \quad \text{s.c. } \|\mathbf{d}_k\| \leq \Delta_k, \quad (4.28)$$

où  $\mathbf{B}_k$  est une approximation du Hessien  $\mathbf{H}_k$  et  $\Delta_k$  le rayon de la région de confiance (figure 4.2). La solution  $\mathbf{d}_k$  de (4.28) est acceptée ou rejetée, suivant la valeur du résidu

$$r_k = \frac{F(\mathbf{x}_k) - F(\mathbf{x}_k + \mathbf{d}_k)}{Q_k(\mathbf{0}) - Q_k(\mathbf{d}_k)}, \quad (4.29)$$

comme le décrit l'algorithme 4.4 issu de (Conn *et al.*, 1987, Chap.6). La mise à jour de la taille de la région dépend également de  $r_k$ , comme l'illustre le schéma en figure 4.3.

L'application de la stratégie de régions de confiance pour la recherche de pas multi-dimensionnelle consiste à résoudre (4.28) sous la contrainte supplémentaire que  $\mathbf{d}_k$  appartienne au sous-espace engendré par les vecteurs de  $\mathbf{D}_k$  (Shi et Shen, 2007; Shi et Xu, 2009; Wang *et al.*, 2004), ce qui est équivalent à

$$\min_{\mathbf{s}} \left( q_k(\mathbf{s}) = f(\mathbf{0}) + \mathbf{s}^T \nabla f(\mathbf{0}) + \frac{1}{2} \mathbf{s}^T \mathbf{D}_k^T \mathbf{B}_k \mathbf{D}_k \mathbf{s} \right), \quad \text{s.c. } \|\mathbf{D}_k \mathbf{s}\| \leq \Delta_k. \quad (4.30)$$

La taille du sous-espace étant souvent très petite devant  $n$ , la minimisation (4.30) est moins coûteuse que la résolution de (4.28).

Les propriétés de convergence de l'algorithme de sous-espace associé à une stratégie de pas par région de confiance dépendent de la règle de choix de  $\Delta_k$ , des matrices  $\mathbf{B}_k$  et  $\mathbf{D}_k$  ainsi que de la précision de la résolution du problème (4.30). Par exemple, lorsque le pas est issu de l'algorithme 4.5 présenté dans (Shi et Xu, 2009), la convergence de l'algorithme de sous-espace est assurée (Shi et Xu, 2009, Theo.3.1) sous des hypothèses faibles sur l'approximation quadratique  $q_k$  et la direction  $\mathbf{d}_k$  choisie à l'intérieur du sous-espace.

A notre connaissance, la convergence de la stratégie de pas par régions de confiance n'est établie que si le problème contraint (4.30) est résolu de façon exacte. C'est le cas par exemple lorsque

**Entrées :**  $\mathbf{x}_0$ ,  $\Delta_0$  et des constantes  $\mu_1, \mu_2, \gamma_1, \gamma_2$  telles que

$$0 < \mu_1 \leq \mu_2 < 1, \quad \text{et} \quad 0 < \gamma_1 \leq \gamma_2 < 1.$$

**Sorties :** Minimiseur de  $F$

**tantque** Test d'arrêt non vérifié **faire**

Construire  $\mathbf{B}_k$

Obtenir  $\mathbf{d}_k$  en résolvant de façon approchée (4.28)

Calculer  $r_k$  par (4.29)

**si**  $r_k \geq \mu_1$  **alors**

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{d}_k$$

**sinon**

$$\mathbf{x}_{k+1} = \mathbf{x}_k$$

**finsi**

Choisir  $\Delta_{k+1}$  vérifiant :

$$\Delta_{k+1} \in \begin{cases} [\Delta_k, \infty) & \text{si } r_k \geq \mu_2 \\ [\gamma_2 \Delta_k, \Delta_k] & \text{si } r_k \in [\mu_1, \mu_2) \\ [\gamma_1 \Delta_k, \gamma_2 \Delta_k] & \text{si } r_k < \mu_1 \end{cases}$$

$k \leftarrow k + 1$

**fin tantque**

Algorithme 4.4 – Algorithme de régions de confiance (Conn *et al.*, 1987, Chap.6)

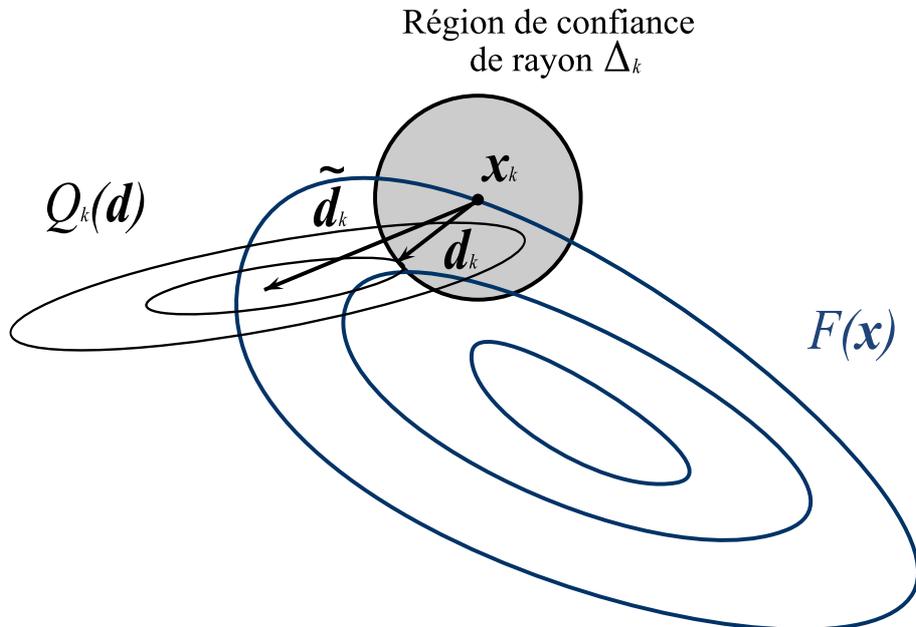


FIGURE 4.2 – Algorithme de régions de confiance. La minimisation non contrainte de l'approximation quadratique  $Q_k(\cdot)$  fournit la direction  $\tilde{\mathbf{d}}_k$ . La minimisation à l'intérieur de la région de confiance fournit la direction  $\mathbf{d}_k$ , impliquant une plus grande décroissance du critère.

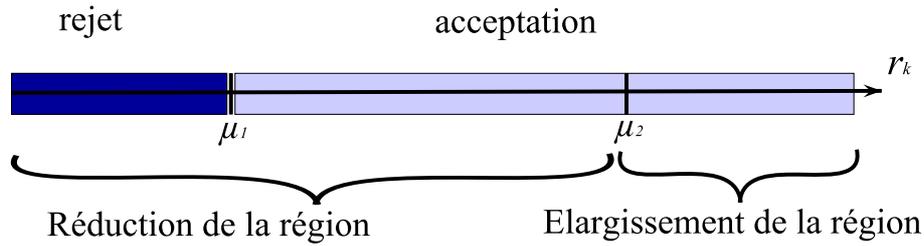


FIGURE 4.3 – Sélection de la région de confiance

l'approximation quadratique est séparable (Shi et Guo, 2008; Shi et Shen, 2007). Dans le cas de fonctions  $q_k$  plus générales, la résolution des sous-problèmes est réalisée de façon approchée par une stratégie itérative dédiée aux problèmes de région de confiance (Conn et al., 1987, Chap.7), (Golub et Matt, 1991; Moré et Sorensen, 1983), stoppée après un petit nombre d'itérations (typiquement  $2 \leq J \leq 5$ ) (Wang et al., 2004).

<p><b>Entrées :</b> <math>\mathbf{x}_k, \mathbf{D}_k</math> et les coefficients <math>\mu, \rho \in (0, 1)</math></p> <p><b>Sorties :</b> Pas <math>\mathbf{s}_k</math></p> <p>Construire <math>\mathbf{B}_k</math></p> <p>Initialiser <math>\Delta_k = -(\mathbf{g}_k^T \mathbf{d}_k) / (\mathbf{d}_k^T \mathbf{B}_k \mathbf{d}_k) \ \mathbf{d}_k\ </math> à partir de <math>\mathbf{d}_k \in \text{vect}(\mathbf{D}_k)</math>.</p> <p><b>répéter</b></p> <p>Calculer <math>\mathbf{s}_k</math> solution de (4.30)</p> <p>Calculer le résidu</p> $r_k = \frac{f(\mathbf{0}) - f(\mathbf{s}_k)}{f(\mathbf{0}) - q_k(\mathbf{s}_k)}. \quad (4.31)$ <p>Réduction de la région : <math>\Delta_k = \rho \Delta_k</math></p> <p><b>jusqu'à</b> <math>r_k \geq \mu</math></p>
---

Algorithme 4.5 – Recherche de pas multi-dimensionnelle par régions de confiance (Shi et Xu, 2009)

## 4.4 Algorithme de Majoration-Minoration - Application à la recherche de pas

Les stratégies de pas standards pour l'algorithme de descente sont itératives, des valeurs de pas sont générées jusqu'à vérifier la condition de convergence désirée. Le nombre d'itérations nécessaires à la vérification de ce critère d'arrêt influe de façon importante sur la complexité de l'algorithme, car chaque itération de recherche de pas nécessite une évaluation du critère et parfois du gradient. Cependant, ce nombre d'itérations n'est pas connu *a priori*.

Récemment, une nouvelle méthode de construction du pas pour l'algorithme de descente itérative a été proposée (Labat et Idier, 2008; Sun et Zhang, 2001). Le pas  $\mathbf{y}$  est obtenu par une formule analytique (Sun et Zhang, 2001), ou plus généralement par une récurrence simple dont le nombre d'itérations est fixé à l'avance (Labat et Idier, 2008). Ces méthodes utilisent le principe de Majoration-Minoration quadratique (MMQ) qui consiste à substituer une séquence de problèmes quadratiques au problème initial de minimisation de  $f(\alpha) = F(\mathbf{x}_k + \alpha \mathbf{d}_k)$ . Les itérés résultent de la minimisation successive d'approximations majorantes quadratiques de  $f$ , construites à partir d'informations sur sa structure analytique.

Après avoir introduit les notions d'approximation tangente majorante quadratique et de récurrence MMQ, nous décrivons la recherche de pas MMQ scalaire. Nous montrons que cette stratégie s'étend de façon directe en une méthode de recherche de pas multi-dimensionnelle pour l'algorithme de sous-espace. La section suivante sera consacrée à l'analyse de convergence des méthodes de pas MMQ.

#### 4.4.1 Approximation tangente majorante quadratique

Afin de décrire le principe de Majoration-Minoration, nous introduisons la notion d'approximation tangente majorante d'un critère en un point.

**Définition 4.2.** La fonction  $H(\mathbf{x}, \mathbf{y})$  est dite approximation tangente majorante de  $F(\mathbf{x})$  en  $\mathbf{y}$  sur le domaine  $\mathcal{D}_f$  si pour tout  $\mathbf{x} \in \mathcal{D}_f$ ,

$$\begin{cases} H(\mathbf{x}, \mathbf{y}) \geq F(\mathbf{x}) \\ H(\mathbf{y}, \mathbf{y}) = F(\mathbf{y}) \end{cases} \quad (4.32)$$

Lorsque  $H$  est différentiable sur  $\mathcal{D}_f$ , la définition (4.32) implique que la fonction  $F$  est tangente à  $H$  en  $\mathbf{y}$ , au sens

$$\nabla F(\mathbf{y}) = \nabla_1 H(\mathbf{y}, \mathbf{y}) \quad (4.33)$$

en notant  $\nabla_1$  le gradient de  $H$  par rapport à sa première variable. La figure 4.4 illustre le principe de la majoration sur un exemple bi-dimensionnel.

La Majoration-Minoration quadratique (MMQ) se base sur la construction d'approximations tangentes majorantes quadratiques du critère. D'après la définition 4.2, une approximation tangente majorante quadratique de  $F$  s'écrit

$$H(\mathbf{x}, \mathbf{y}) = F(\mathbf{y}) + \nabla F(\mathbf{y})^T (\mathbf{x} - \mathbf{y}) + \frac{1}{2} (\mathbf{x} - \mathbf{y})^T \mathbf{A}(\mathbf{y}) (\mathbf{x} - \mathbf{y}), \quad (4.34)$$

où  $\mathbf{A}(\mathbf{y}) \in \mathbb{R}^{n \times n}$  est une matrice définie positive assurant que (4.34) vérifie les conditions (4.32) pour tout  $\mathbf{y}$ . De façon générale, toute matrice  $\mathbf{A}(\mathbf{y})$  telle que  $\mathbf{A}(\mathbf{y}) - \nabla^2 F(\mathbf{y})$  soit définie positive entraîne que (4.34) est une approximation majorante de  $F$  (Hunter et L., 2004). Cependant, ce résultat ne permet pas de définir une méthode de construction pour les matrices  $\mathbf{A}(\mathbf{y})$ . En pratique, ces matrices sont construites à partir de propriétés sur la structure analytique du critère.

**Critère à gradient Lipschitz** Nous avons tout d'abord le résultat suivant, concernant les critères à gradient Lipschitz.

**Lemme 4.1.** Soit  $F$  une fonction à gradient Lipschitz sur  $\mathcal{D}_f$ , i.e.

$$\|\nabla F(\mathbf{x}) - \nabla F(\mathbf{y})\| \leq L \|\mathbf{x} - \mathbf{y}\|, \quad \forall \mathbf{x}, \mathbf{y} \in \mathcal{D}_f. \quad (4.35)$$

Si

$$\mathbf{A}(\mathbf{y}) = a\mathbf{I}, \quad a \geq L, \quad (4.36)$$

alors la fonction (4.34) est une approximation majorante de  $F$  en  $\mathbf{y} \in \mathcal{D}_f$ .

**Critère de Huber** La propriété suivante, issue de la régression robuste (Huber, 1981), permet de construire des approximations majorantes quadratiques sans connaître la constante de Lipschitz de  $F$ .

**Lemme 4.2.** Soit  $F$  une fonction potentiel de Huber, c'est à dire

–  $F$  est continue, dérivable et paire sur  $\mathbb{R}$

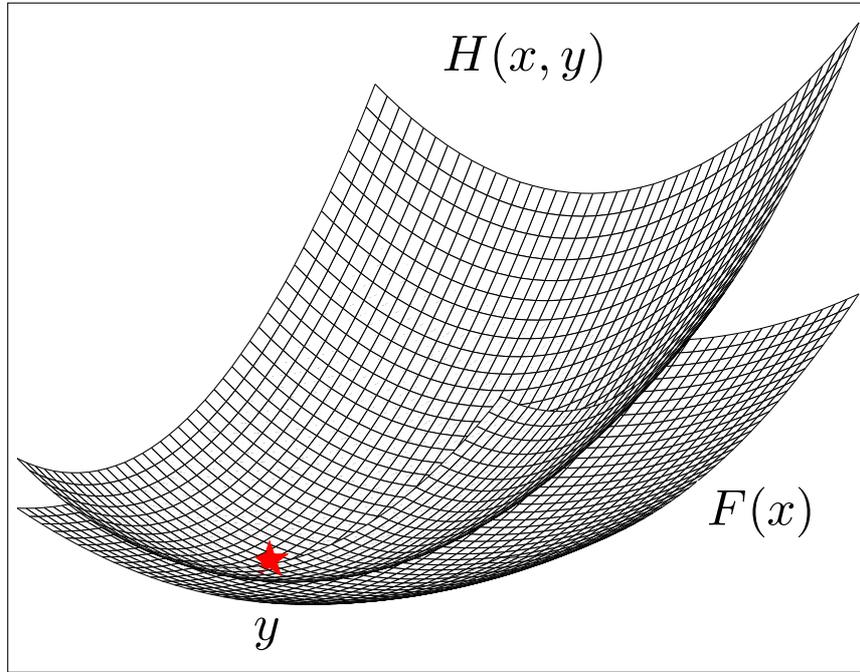


FIGURE 4.4 – Approximation tangente majorante  $H(x, y)$  de  $F$  en  $y$ . La fonction  $H$  est supérieure à  $F$  en tout point et est tangente à  $F$  en  $y$ .

- $F$  est croissante sur  $\mathbb{R}^+$
- $F(\sqrt{\cdot})$  est concave sur  $\mathbb{R}^+$
- $\omega(x) = \dot{F}(x)/x < \infty, \quad \forall x \in \mathbb{R}$

alors la fonction

$$F(y) + (x - y)\dot{F}(y) + \frac{1}{2}\omega(y)(x - y)^2 \quad (4.37)$$

est une approximation majorante de  $F$  en  $y \in \mathbb{R}$ .

Nous décrivons à présent comment utiliser ces approximations majorantes pour minimiser un critère.

#### 4.4.2 La récurrence de Majoration-Minoration

La minimisation du critère  $F$  par Majoration-Minoration consiste à résoudre le problème initial (3.2) par la règle de mise à jour MM

$$\mathbf{x}_{k+1} = \arg \min_{\mathbf{x} \in \mathcal{D}_f} H(\mathbf{x}, \mathbf{x}_k), \quad (4.38)$$

illustrée dans un cas scalaire en figure 4.5.

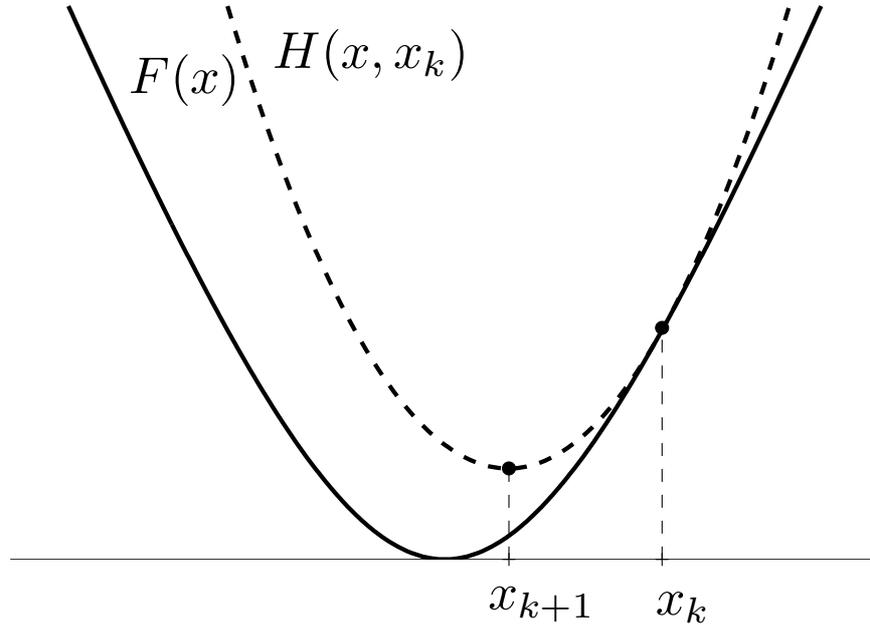


FIGURE 4.5 – Algorithme de Majoration-Minoration. Le nouvel itéré  $\mathbf{x}_{k+1}$  est le minimiseur de  $H(\cdot, \mathbf{x}_k)$ , approximation majorante de  $F$  en  $\mathbf{x}_k$ .

Le schéma de minimisation (4.38) apparaît pour la première fois dans l'ouvrage (Ortega et Rheinboldt, 1970). De nombreuses contributions algorithmiques dans le cadre de l'inversion pénalisée en traitement du signal et de l'image s'identifient avec un schéma de minimisation MM. C'est le cas notamment des algorithmes EM (Dempster *et al.*, 1977) et proximaux (Rockafellar, 1970), ces derniers étant basés sur la majoration décrite dans le lemme 4.1. La stratégie MM a été utilisée également comme principe constructif d'algorithmes, dans (Böhning et Lindsay, 1988) ainsi que dans les algorithmes développés par Fessler *et coll.* (Erdogan et Fessler, 1999; Fessler et Erdogan, 1998; Soththivirat et Fessler, 2002) en reconstruction tomographique, basés notamment sur le lemme de majoration 4.2. Les propriétés de l'algorithme (4.38) sont détaillées dans (Jacobson et Fessler, 2007). En particulier, la récurrence MM (4.38) entraîne la décroissance de la suite  $\{F(\mathbf{x}_k)\}$ . Ceci est une conséquence directe des relations (4.32). En effet,

$$F(\mathbf{x}_{k+1}) \underset{(1)}{\leq} H(\mathbf{x}_{k+1}, \mathbf{x}_k) \underset{(2)}{\leq} H(\mathbf{x}_k, \mathbf{x}_k) \underset{(3)}{=} F(\mathbf{x}_k), \quad (4.39)$$

où les relations (1) à (3) proviennent des propriétés de

- (1) Majoration
- (2) Minimisation
- (3) Tangence

Lorsque l'approximation majorante est quadratique, son minimiseur s'obtient de façon analytique. Il en résulte le schéma de minimisation très simple, que nous notons MMQ :

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{A}(\mathbf{x}_k)^{-1} \nabla F(\mathbf{x}_k). \quad (4.40)$$

Récemment, le schéma MMQ (4.40) a été utilisé pour définir une procédure de recherche de pas. La description de cette méthode fait l'objet du paragraphe suivant.

### 4.4.3 Recherche de pas MMQ scalaire

**Principe** La fonction majorante quadratique multi-dimensionnelle (4.34) permet de construire la fonction scalaire  $h(\alpha, \alpha_k^j)$ , approximation majorante quadratique de  $f$  en un pas  $\alpha_k^j$

$$h(\alpha, \alpha_k^j) = f(\alpha_k^j) + (\alpha - \alpha_k^j) \dot{f}(\alpha_k^j) + \frac{1}{2} m_k^j (\alpha - \alpha_k^j)^2 \quad (4.41)$$

avec

$$m_k^j = \mathbf{d}_k^T \mathbf{A}_k^j \mathbf{d}_k. \quad (4.42)$$

où

$$\mathbf{A}_k^j = \mathbf{A}(\mathbf{x}_k + \alpha_k^j \mathbf{d}_k) \quad (4.43)$$

est la courbure de la quadratique  $H(\mathbf{x}, \mathbf{x}_k + \alpha_k^j \mathbf{d}_k)$ . Ce constat amène naturellement à utiliser comme recherche de pas un algorithme MMQ scalaire. Cette approche, proposée dans (Fessler et Booth, 1999; Labat et Idier, 2008; Rivera et Marroquin, 2003), a pour intérêt de fournir de manière très simple une séquence  $\alpha_k^1, \dots, \alpha_k^J$  qui converge vers un minimum local de  $f$  (et donc vers le pas optimal lorsque  $F$  est convexe). La recherche de pas MMQ, notée MMQ 1D, est une récurrence MM relaxée appliquée aux fonctions majorantes  $h(\cdot, \alpha_k^j)$ . La récurrence est initialisée avec  $\alpha_k^0 = 0$  et s'écrit :

$$\begin{cases} \alpha_k^0 &= 0, \\ \alpha_k^{j+1} &= \alpha_k^j + \theta(\hat{\alpha}_k^{j+1} - \alpha_k^j), \quad j < J, \\ \alpha_k &= \alpha_k^j, \end{cases} \quad (4.44)$$

avec

$$\hat{\alpha}_k^{j+1} = \arg \min_{\alpha} h(\alpha, \alpha_k^j), \quad j < J, \quad (4.45)$$

et  $\theta \in (0, 2)$ . Pour une itération  $j$  donnée, choisir  $\theta > 1$  revient à choisir  $\alpha_k^{j+1}$  au delà du minimiseur de la majorante  $\hat{\alpha}_k^{j+1}$ . A l'inverse, quand  $\theta$  est inférieur à 1, la récurrence MM est freinée.  $\theta$  est maintenu dans  $(0, 2)$  afin de préserver la monotonie de l'algorithme, c'est à dire

$$f(\alpha_k) < f(0). \quad (4.46)$$

La figure 4.6 illustre les différentes valeurs de pas obtenues pour  $\theta \in (0, 2)$ .

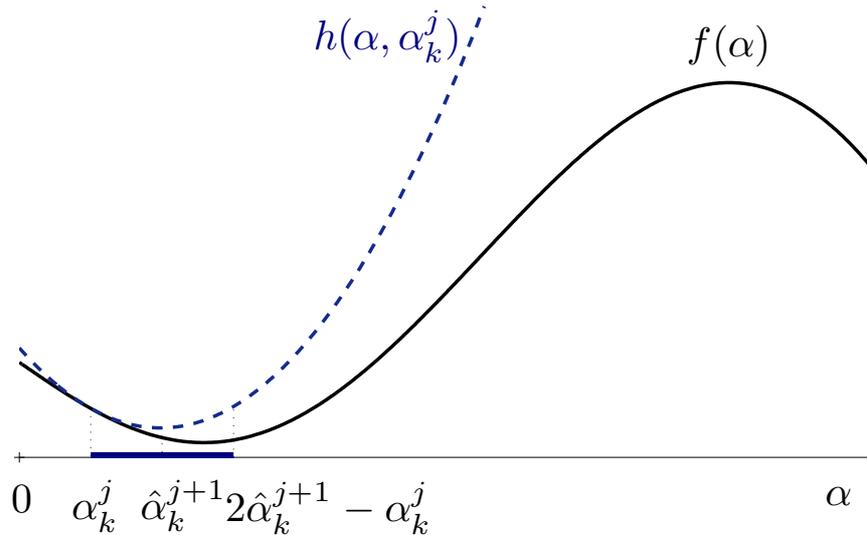


FIGURE 4.6 – Recherche de pas par Majoration-Minoration. Le pas  $\hat{\alpha}_k^{j+1}$  correspond au choix  $\theta = 1$ . Les pas obtenus en faisant varier  $\theta \in (0, 2)$  sont indiqués en trait épais.

**Lien avec les méthodes à pas constant** La forme (4.44) est une généralisation des stratégies de pas constant proposées dans (Bertsekas, 1999, Chap.1) et (Chen et Sun, 2002; Sun et Zhang, 2001) pour le but d'obtenir un algorithme de descente itérative sans recherche de pas pour les critères à gradient  $L$ -Lipschitz. Dans la méthode proposée par Bertsekas, lorsque  $\mathbf{d}_k$  est la direction de la plus grande pente, le pas est choisi tel que

$$\epsilon \leq \alpha_k \leq \frac{2 - \epsilon}{L}, \quad \epsilon > 0. \quad (4.47)$$

Cette formule se retrouve de façon simple en posant  $m_k^j = L\|\mathbf{d}_k\|^2$  (issue du lemme 4.1) dans (4.44). En effet, le minimiseur de la majorante de courbure  $L\|\mathbf{d}_k\|^2$  en 0 s'écrit

$$\hat{\alpha}_k^1 = -\frac{\mathbf{g}_k^T \mathbf{d}_k}{L\|\mathbf{d}_k\|^2} \quad (4.48)$$

$$= \frac{1}{L} \quad (4.49)$$

où la deuxième égalité provient du choix de la direction  $\mathbf{d}_k = -\mathbf{g}_k$ . Quelque soit  $\epsilon > 0$ , il existe  $\theta \in (0, 2)$  tel que le pas vérifiant (4.47) s'écrit

$$\alpha_k = \theta \hat{\alpha}_k^1, \quad (4.50)$$

ce qui correspond à  $J = 1$  itération du schéma (4.44). La stratégie de pas proposée dans (Chen et Sun, 2002; Sun et Zhang, 2001) consiste à choisir (4.50), avec  $\theta \in (0, \nu_{\min}^A)$ , où  $\nu_{\min}^A$  minore le spectre de valeurs propres des matrices  $\{\mathbf{A}_k^j\}$ . L'intervalle  $(0, \nu_{\min}^A)$  étant inclus dans  $(0, 2)$ , cette méthode s'identifie à nouveau avec  $J = 1$  itération de (4.44).

#### 4.4.4 Recherche de pas MMQ multi-dimensionnelle

Nous proposons dans cette section une nouvelle stratégie de pas multi-dimensionnelle pour l'algorithme itératif de sous-espace, basée sur une généralisation de la recherche de pas MMQ scalaire. Dans cette dernière méthode, la valeur de  $\alpha_k$  résulte de  $J$  minimisations successives de fonctions quadratiques tangentes majorantes à  $f(\alpha)$ . Afin de généraliser cette technique au cas multi-dimensionnel, on construit à partir de la fonction majorante quadratique du critère  $H(\cdot, \cdot)$ , une fonction quadratique tangente majorante de  $f(\mathbf{s})$  en  $\mathbf{s}_k^j$

$$h(\mathbf{s}, \mathbf{s}_k^j) = f(\mathbf{s}_k^j) + \nabla f(\mathbf{s}_k^j)^T (\mathbf{s} - \mathbf{s}_k^j) + \frac{1}{2} (\mathbf{s} - \mathbf{s}_k^j)^T \mathbf{B}_k^j (\mathbf{s} - \mathbf{s}_k^j), \quad (4.51)$$

avec

$$\mathbf{B}_k^j = \mathbf{D}_k^T \mathbf{A}_k^j \mathbf{D}_k \quad (4.52)$$

où

$$\mathbf{A}_k^j = \mathbf{A}(\mathbf{x}_k + \mathbf{D}_k \mathbf{s}_k^j) \quad (4.53)$$

est la courbure de la quadratique  $H(\mathbf{x}, \mathbf{x}_k + \mathbf{D}_k \mathbf{s}_k^j)$ . Nous définissons alors le pas MMQ multi-dimensionnel  $\mathbf{s}_k \in \mathbb{R}^r$ , notée MMQ rD, par la récurrence suivante correspondant à l'application d'une stratégie MM relaxée sur la fonction  $f(\mathbf{s})$

$$\begin{cases} \mathbf{s}_k^0 &= \mathbf{0}, \\ \mathbf{s}_k^{j+1} &= \mathbf{s}_k^j + \theta(\hat{\mathbf{s}}_k^{j+1} - \mathbf{s}_k^j) \\ \mathbf{s}_k &= \mathbf{s}_k^j, \end{cases} \quad (4.54)$$

avec

$$\hat{\mathbf{s}}_k^{j+1} = \arg \min_{\mathbf{s}} h(\mathbf{s}, \mathbf{s}_k^j), \quad j < J. \quad (4.55)$$

Afin de garantir la décroissance de la suite  $\{f(\mathbf{s}_k^j)\}$  et la convergence de la suite  $\{\mathbf{s}_k^j\}$  vers un point stationnaire de  $f(\mathbf{s})$ , le coefficient  $\theta$  est choisi à l'intérieur de  $(0, 2)$ . De plus, comme dans le cas scalaire, les termes de la récurrence se construisent de façon analytique par :

$$\mathbf{s}_k^{j+1} = \mathbf{s}_k^j - \theta(\mathbf{B}_k^j)^{-1} \nabla f(\mathbf{s}_k^j). \quad (4.56)$$

## 4.5 Propriétés de convergence du pas MM quadratique

Cette section est consacrée à l'analyse de convergence des algorithmes itératifs associés à la recherche de pas MM quadratique. La convergence d'une famille d'algorithmes de gradient conjugué non linéaire associés au pas MMQ 1D est établie dans (Labat et Idier, 2008). Notre contribution consiste à étendre ce résultat au cas des algorithmes à gradient relié (Bertsekas, 1996), regroupant de nombreux algorithmes de descente classiques tels que l'algorithme de gradient, de Newton et de Newton tronqué. Nous établissons également la convergence de l'algorithme de sous-espace associé au pas MMQ  $r$ D, sous une hypothèse faible sur l'ensemble des directions sélectionné.

### 4.5.1 Hypothèses

Afin d'établir les résultats de convergence, nous supposons que la recherche de pas est bien définie, c'est à dire qu'il existe des fonctions majorantes quadratiques du critère. De plus, nous supposons que la courbure des majorantes quadratiques est construite à partir de matrices à spectre borné.

**Hypothèse 4.1.** *Pour tout  $\mathbf{y} \in \mathcal{D}_f$ , il existe une matrice symétrique définie positive  $\mathbf{A}(\mathbf{y})$  telle que pour tout  $\mathbf{x} \in \mathcal{D}_f$ ,*

$$H(\mathbf{x}, \mathbf{y}) = F(\mathbf{y}) + (\mathbf{x} - \mathbf{y})^T \nabla F(\mathbf{y}) + \frac{1}{2} (\mathbf{x} - \mathbf{y})^T \mathbf{A}(\mathbf{y}) (\mathbf{x} - \mathbf{y}) \quad (4.57)$$

*est une approximation tangente majorante de  $F$  en  $\mathbf{y}$ . De plus, l'ensemble  $\{\mathbf{A}(\mathbf{x}) | \mathbf{x} \in \mathcal{D}_f\}$  est à spectre strictement positif borné par  $\nu_{\max}^A$  i.e., pour tout  $\mathbf{x} \in \mathcal{D}_f$ ,*

$$0 < \frac{\mathbf{v}^T \mathbf{A}(\mathbf{x}) \mathbf{v}}{\|\mathbf{v}\|^2} \leq \nu_{\max}^A, \forall \mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}\} \quad (4.58)$$

### 4.5.2 Propriétés de la séquence de pas

#### 4.5.2.1 Décroissance de l'algorithme

Nous établissons par le lemme suivant que la recherche de pas MMQ multi-dimensionnelle permet d'assurer la décroissance de la séquence  $\{F(\mathbf{x}_k)\}$  générée par l'algorithme de descente itérative.

**Lemme 4.3.** *Soit  $F$  une fonction vérifiant l'hypothèse 4.1. Alors pour tout  $j > 0$  et pour tout  $\theta \in (0, 2)$ , les pas générés par (4.54) vérifient*

$$f(\mathbf{s}_k^{j+1}) \leq f(\mathbf{s}_k^j). \quad (4.59)$$

*Par conséquent, pour tout  $k$ , la séquence produite par l'algorithme itératif de sous-espace (3.43) associé au pas  $\mathbf{s}_k^j$  vérifie*

$$F(\mathbf{x}_{k+1}) \leq F(\mathbf{x}). \quad (4.60)$$

*Démonstration.* La preuve s'inspire de la démonstration de (Allain, 2002, Prop.5). Tout d'abord,

$$h(\mathbf{s}_k^{j+1}, \mathbf{s}_k^j) = f(\mathbf{s}_k^j) + \theta (\hat{\mathbf{s}}_k^{j+1} - \mathbf{s}_k^j)^T \nabla f(\mathbf{s}_k^j) + \frac{\theta^2}{2} (\hat{\mathbf{s}}_k^{j+1} - \mathbf{s}_k^j)^T \mathbf{B}_k^j (\hat{\mathbf{s}}_k^{j+1} - \mathbf{s}_k^j) \quad (4.61)$$

De plus,

$$\hat{\mathbf{s}}_k^{j+1} = \mathbf{s}_k^j - (\mathbf{B}_k^j)^{-1} \nabla f(\mathbf{s}_k^j). \quad (4.62)$$

Donc

$$h(\mathbf{s}_k^{j+1}, \mathbf{s}_k^j) = f(\mathbf{s}_k^j) + \theta(\theta/2 - 1)(\hat{\mathbf{s}}_k^{j+1} - \mathbf{s}_k^j)^T \mathbf{B}_k^j(\hat{\mathbf{s}}_k^{j+1} - \mathbf{s}_k^j), \quad (4.63)$$

soit

$$h(\mathbf{s}_k^{j+1}, \mathbf{s}_k^j) - f(\mathbf{s}_k^j) = \theta(\theta/2 - 1)(\hat{\mathbf{s}}_k^{j+1} - \mathbf{s}_k^j)^T \mathbf{B}_k^j(\hat{\mathbf{s}}_k^{j+1} - \mathbf{s}_k^j). \quad (4.64)$$

Le paramètre  $\theta$  appartient à  $(0, 2)$ . De plus, la matrice  $\mathbf{B}_k^j$  est définie positive d'après l'hypothèse 4.1 donc le terme de droite de l'égalité est négatif ou nul. Par conséquent,

$$h(\mathbf{s}_k^{j+1}, \mathbf{s}_k^j) \leq f(\mathbf{s}_k^j). \quad (4.65)$$

De plus,  $h(\cdot, \mathbf{s}_k^j)$  est une fonction majorante de  $f$  en  $\mathbf{s}_k^j$  d'après l'hypothèse 4.1 donc  $f(\mathbf{s}_k^{j+1}) \leq h(\mathbf{s}_k^{j+1}, \mathbf{s}_k^j)$  d'où

$$f(\mathbf{s}_k^{j+1}) \leq f(\mathbf{s}_k^j). \quad (4.66)$$

Par récurrence sur  $j$ , on obtient

$$f(\mathbf{s}_k^J) \leq f(\mathbf{0}). \quad (4.67)$$

□

En considérant un sous-espace de dimension  $r = 1$ , on obtient que la recherche de pas MMQ scalaire assure la décroissance de l'algorithme itératif de sous-espace.

**Lemme 4.4.** *Soit  $F$  une fonction vérifiant l'hypothèse 4.1. Alors pour tout  $j > 0$  et pour tout  $\theta \in (0, 2)$ , les pas générés par (4.44) vérifient*

$$f(\alpha_k^{j+1}) \leq f(\alpha_k^j). \quad (4.68)$$

Par conséquent, pour tout  $k$ , la séquence produite par l'algorithme de descente itérative (3.18) associé au pas  $\alpha_k^J$  vérifie

$$F(\mathbf{x}_{k+1}) \leq F(\mathbf{x}). \quad (4.69)$$

#### 4.5.2.2 Condition de Zoutendijk

Le lemme suivant assure que la condition de Zoutendijk est vérifiée par la stratégie de pas MMQ scalaire. De plus, le résultat ne dépend pas du nombre de sous-itérations  $J$ .

**Lemme 4.5.** *Soit  $F$  une fonction vérifiant l'hypothèse 4.1. Alors pour tout  $J > 0$ , le pas  $\alpha_k^J$  obtenu par (4.44) où  $\theta \in (0, 2)$  vérifie la condition de Zoutendijk (4.16) avec  $C = (\theta/\nu_{\max}^A)(1 - \theta/2) \in (0, 1/2)$ .*

*Démonstration.* La démonstration est basée sur (Allain, 2002, Prop.6). Tout d'abord, pour tout  $j$ ,

$$h(\alpha_k^{j+1}, \alpha_k^j) - f(\alpha_k^j) = \theta\left(\frac{\theta}{2} - 1\right)m_k^j(\hat{\alpha}_k^{j+1} - \alpha_k^j)^2. \quad (4.70)$$

En prenant  $j = 1$  dans l'inégalité ci-dessus, on obtient que

$$h(\alpha_k^1, 0) - f(0) = -\delta m_k^1(\hat{\alpha}_k^1)^2, \quad (4.71)$$

en posant  $\delta = \theta(1 - \frac{\theta}{2}) > 0$ . De plus, d'après l'hypothèse 4.1,  $f(\alpha_k^1) \leq h(\alpha_k^1, 0)$  donc

$$f(\alpha_k^1) - f(0) \leq -\delta m_k^1(\hat{\alpha}_k^1)^2. \quad (4.72)$$

D'après l'expression de  $\hat{\alpha}_k^1$ ,

$$f(\alpha_k^1) - f(0) \leq -\delta \frac{(\mathbf{g}_k^T \mathbf{d}_k)^2}{m_k^1}. \quad (4.73)$$

De plus,  $m_k^1 = \mathbf{d}_k^T \mathbf{A}_k^0 \mathbf{d}_k$  donc, d'après l'hypothèse 4.1,

$$f(\alpha_k^1) - f(0) \leq -\delta \frac{(\mathbf{g}_k^T \mathbf{d}_k)^2}{\nu_{\max}^{\mathcal{A}} \|\mathbf{d}_k\|^2}. \quad (4.74)$$

D'après le lemme 4.4,  $f(\alpha_k^J) \leq f(\alpha_k^1)$  donc

$$f(\alpha_k^J) - f(0) \leq -\delta \frac{(\mathbf{g}_k^T \mathbf{d}_k)^2}{\nu_{\max}^{\mathcal{A}} \|\mathbf{d}_k\|^2}, \quad (4.75)$$

soit

$$F(\mathbf{x}_{k+1}) - F(\mathbf{x}_k) \leq -\delta \frac{(\mathbf{g}_k^T \mathbf{d}_k)^2}{\nu_{\max}^{\mathcal{A}} \|\mathbf{d}_k\|^2}. \quad (4.76)$$

□

Nous montrons dans la section suivante que le lemme 4.5 est suffisant pour démontrer la convergence d'une large famille d'algorithmes de descente itérative, en utilisant la notion de séquence de directions *gradient reliée*. Cette notion nous permet également d'établir la convergence des algorithmes de sous-espace associés au pas MMQ *rD*.

### 4.5.3 Convergence des algorithmes gradient reliés

#### 4.5.3.1 Séquence de directions gradient reliée

La notion de suite gradient reliée permet de démontrer la convergence de nombreux algorithmes de descente itérative, sous l'hypothèse que le pas vérifie la condition de Zoutendijk (4.16). Considérons la méthode itérative de descente de gradient. Dans celle-ci, la direction est égale à  $\mathbf{d}_k = -\mathbf{g}_k$  donc  $\cos \theta_k = 1$ . Par conséquent, (4.18) entraîne  $\|\mathbf{g}_k\| \rightarrow 0$ . On obtient donc directement la convergence de la méthode du gradient. Plus généralement, si l'angle entre  $\mathbf{d}_k$  et  $-\mathbf{g}_k$  est inférieur à  $\frac{\pi}{2}$ , l'algorithme est convergent. C'est précisément cette condition qui est vérifiée par une suite de directions gradient reliée : La notion de séquence gradient reliée introduite dans (Bertsekas, 1996) permet d'empêcher que  $\mathbf{d}_k$  devienne asymptotiquement orthogonale au gradient de  $F$  alors que  $\mathbf{x}_k$  tend vers un point qui n'est pas solution. Un tel cas de figure entraînerait un blocage de l'algorithme en ce point.

**Définition 4.3.** (Bertsekas, 1996)

Une séquence de directions  $\{\mathbf{d}_k\}_{k \geq 0}$  est dite gradient reliée à  $\{\mathbf{x}_k\}_{k \geq 0}$  si la propriété suivante est vérifiée : Pour toute suite extraite  $\{\mathbf{x}_k\}_{k \in \mathcal{K}}$  convergeant vers un point non stationnaire, la suite extraite correspondante  $\{\mathbf{d}_k\}_{k \in \mathcal{K}}$  est bornée et satisfait

$$\limsup_{k \rightarrow \infty, k \in \mathcal{K}} \mathbf{g}_k^T \mathbf{d}_k < 0.$$

Autrement dit, la direction  $\mathbf{d}_k$  n'est jamais 'trop petite' ou 'trop grande' par rapport à  $\mathbf{g}_k$  et l'angle entre les deux vecteurs n'est jamais trop proche de  $\frac{\pi}{2}$ . La propriété suivante établit deux conditions suffisantes pour qu'une séquence soit gradient reliée.

**Propriété 4.1.** (Bertsekas, 1999) S'il existe des réels  $\nu_1 > 0$ ,  $\nu_2 > 0$  tels que pour tout  $k$ ,  $\mathbf{d}_k$  vérifie

$$\nu_1 \|\mathbf{g}_k\|^2 \leq -\mathbf{g}_k^T \mathbf{d}_k, \quad \|\mathbf{d}_k\|^2 \leq \nu_2 \|\mathbf{g}_k\|^2, \quad (4.77)$$

alors la séquence  $\{\mathbf{d}_k\}_{k \geq 0}$  est gradient reliée à  $\{\mathbf{x}_k\}_{k \geq 0}$ .

De nombreux algorithmes de descente construisent des directions vérifiant les conditions (4.77), et par conséquent génèrent des séquences gradient reliées.

**Propriété 4.2.** (*Bertsekas, 1999*) Supposons que pour tout  $k$ ,  $\mathbf{d}_k = -\mathbf{M}_k \mathbf{g}_k$ . Si l'ensemble  $\{\mathbf{M}_k, k \geq 0\}$  est à spectre borné strictement positif, alors  $\mathbf{d}_k$  vérifie (4.77) et la suite  $\{\mathbf{d}_k\}_{k \geq 0}$  est gradient reliée à  $\{\mathbf{x}_k\}_{k \geq 0}$ .

La propriété 4.2 permet d'établir le caractère gradient relié des algorithmes de Newton dans le cas convexe, ainsi que de quasi-Newton. Deux familles d'algorithmes cependant manquent à l'appel : les méthodes de gradient conjugué non linéaire et les méthodes de Newton tronquées. Nous établissons le caractère gradient relié des directions issues de cette dernière famille. La démonstration de la propriété 4.3 se base sur des résultats développés dans (*Dembo et Steihaug, 1983*) au sujet de l'algorithme de Newton tronqué.

**Propriété 4.3.** Supposons que pour tout  $k$ ,  $\mathbf{d}_k$  résulte de  $I_k$  itérations de GC appliqué au système  $\mathbf{B}_k \mathbf{d} = -\mathbf{g}_k$ . Si l'ensemble de matrices  $\{\mathbf{B}_k, k \geq 0\}$  est à spectre borné strictement positif, alors  $\mathbf{d}_k$  vérifie (4.77) et la suite  $\{\mathbf{d}_k\}_{k \geq 0}$  est gradient reliée à  $\{\mathbf{x}_k\}_{k \geq 0}$ .

*Démonstration.* La direction  $\mathbf{d}_k$  résulte de  $I_k \geq 1$  itérations de GC appliqué au système  $\mathbf{B}_k \mathbf{d} = -\mathbf{g}_k$  où  $\{\mathbf{B}_k, k \geq 0\}$  est à spectre borné strictement positif. Alors, d'après (*Dembo et Steihaug, 1983, Th.A.1*) et (*Dembo et Steihaug, 1983, Lem.A.2*), il existe des constantes  $\nu_1 > 0$ ,  $\nu_2 > 0$  telles que  $\mathbf{d}_k$  vérifie (4.77) pour tout  $k$ . D'après la propriété 4.1, il s'agit d'une condition suffisante pour garantir que la suite  $\{\mathbf{d}_k\}_{k \geq 0}$  est gradient reliée à  $\{\mathbf{x}_k\}_{k \geq 0}$ .  $\square$

Le même résultat est établi pour la version préconditionnée de la direction tronquée, en interprétant le préconditionnement comme un changement de variable.

**Corollaire 4.1.** Supposons que pour tout  $k$ ,  $\mathbf{d}_k$  résulte de  $I_k$  itérations de GC appliqué au système  $\mathbf{B}_k \mathbf{d} = -\mathbf{g}_k$  preconditionné avec  $\mathbf{P}_k$ . Si  $\{\mathbf{B}_k, k \geq 0\}$  et  $\{\mathbf{P}_k, k \geq 0\}$  sont à spectres bornés strictement positifs, alors  $\mathbf{d}_k$  vérifie (4.77) et la suite  $\{\mathbf{d}_k\}_{k \geq 0}$  est gradient reliée à  $\{\mathbf{x}_k\}_{k \geq 0}$ .

*Démonstration.* Définissons  $\mathbf{C}_k$  telle que  $\mathbf{P}_k = \mathbf{C}_k^T \mathbf{C}_k$ . La direction  $\mathbf{d}_k$  est issue de  $I_k$  itérations de GC preconditionné avec  $\mathbf{P}_k$  appliqué à  $\mathbf{H}_k \mathbf{d} = -\mathbf{g}_k$ . D'après (*Nocedal et Wright, 1999*), elle vérifie

$$\hat{\mathbf{d}}_k = \mathbf{C}_k \mathbf{d}_k \quad (4.78)$$

où  $\hat{\mathbf{d}}_k$  est issue de  $I_k$  itérations de GC non preconditionné appliqué à

$$\mathbf{C}_k^{-T} \mathbf{H}_k \mathbf{C}_k \hat{\mathbf{d}} = -\mathbf{C}_k^{-T} \mathbf{g}_k \quad (4.79)$$

D'après (*Dembo et Steihaug, 1983, Th.A.1*) et (*Dembo et Steihaug, 1983, Lem.A.2*), il existe des constantes strictement positives  $\nu'_1 > 0$ ,  $\nu'_2 > 0$  telles que

$$(\mathbf{C}_k^{-T} \mathbf{g}_k)^T \hat{\mathbf{d}}_k \leq -\nu'_1 \|\mathbf{C}_k^{-T} \mathbf{g}_k\|^2 \quad (4.80)$$

et

$$\|\hat{\mathbf{d}}_k\|^2 \leq \nu'_2 \|\mathbf{C}_k^{-T} \mathbf{g}_k\|^2. \quad (4.81)$$

D'après (4.78),

$$(\mathbf{C}_k^{-T} \mathbf{g}_k)^T \hat{\mathbf{d}}_k = \mathbf{g}_k^T \mathbf{d}_k. \quad (4.82)$$

De plus, l'hypothèse sur le spectre de  $\{\mathbf{P}_k, k = 1, \dots, K\}$  implique que

$$-\|\mathbf{C}_k^{-T} \mathbf{g}_k\|^2 \leq -\frac{1}{\nu_{\max}^{\mathcal{P}}} \|\mathbf{g}_k\|^2, \quad (4.83)$$

$$\|\mathbf{C}_k^{-T} \mathbf{g}_k\| \leq \frac{1}{\sqrt{\nu_{\min}^{\mathcal{P}}}} \|\mathbf{g}_k\|, \quad (4.84)$$

$$\sqrt{\nu_{\min}^{\mathcal{P}}} \|\mathbf{d}_k\| \leq \|\mathbf{C}_k \mathbf{d}_k\| = \|\hat{\mathbf{d}}_k\|, \quad (4.85)$$

où  $(\nu_{\min}^{\mathcal{P}}, \nu_{\max}^{\mathcal{P}}) > 0$  sont les bornes du spectre de  $\{\mathbf{P}_k\}$ . Par conséquent, (4.77) est vérifiée en posant  $\nu_1 = \nu_1'/\nu_{\max}^{\mathcal{P}}$  et  $\nu_2 = \nu_2'/\nu_{\min}^{\mathcal{P}}$ , d'où le résultat en utilisant la propriété 4.1.  $\square$

#### 4.5.3.2 Résultat de convergence pour la recherche de pas MMQ 1D

Lorsque la séquence des directions est gradient reliée, ou plus précisément lorsque toute direction générée vérifie (4.77), nous établissons simplement la convergence de l'algorithme de descente lorsque le pas est obtenu par la stratégie MMQ 1D.

**Théorème 4.2.** *Soit une fonction  $F$  continûment différentiable et bornée inférieurement sur  $\mathcal{D}_f$ . On suppose que  $F$  vérifie l'hypothèse 4.1. Si pour tout  $k$ , la direction  $\mathbf{d}_k$  vérifie (4.77) et que  $\alpha_k$  est obtenu par (4.44) avec  $J > 0$  et  $\theta \in (0, 2)$  alors l'algorithme de descente itérative (3.18) converge au sens  $\lim_{k \rightarrow \infty} \|\mathbf{g}_k\| = 0$ .*

*Démonstration.* D'après le lemme 4.5, la condition de Zoutendijk (4.16) est vérifiée.  $F$  est bornée inférieurement donc on en déduit (4.18). Le théorème 4.2 est une application directe du résultat (Shi, 2004, Theo.5.1).  $\square$

#### 4.5.3.3 Résultat de convergence pour la recherche de pas MMQ $r$ D

La notion de direction gradient reliée nous permet également de démontrer la convergence de l'algorithme itératif de sous-espace, lorsqu'il est associé au pas MMQ multi-dimensionnel.

**Théorème 4.3.** *Soit une fonction  $F$  continûment différentiable et bornée inférieurement sur  $\mathcal{D}_f$ . On suppose que  $F$  vérifie l'hypothèse 4.1 sur  $\mathcal{D}_f$ . Si pour tout  $k \geq 0$ , la matrice  $\mathbf{D}_k$  se met sous la forme*

$$\mathbf{D}_k = [\mathbf{d}_k^1, \dots, \mathbf{d}_k^{r_k}], \quad 1 \leq r_k \leq n \quad (4.86)$$

où  $\mathbf{d}_k^1$  vérifie les conditions (4.77) alors pour tout  $J > 0$  et pour tout  $\theta \in (0, 2)$ , l'algorithme (3.43) associé à la recherche de pas MM (4.54) converge vers une solution de (3.2) au sens

$$\lim_{k \rightarrow \infty} \|\mathbf{g}_k\| = 0.$$

*Démonstration.* Définissons la fonction scalaire

$$q(\alpha) \triangleq h([\alpha, 0, \dots, 0]^T, \mathbf{0}), \quad \forall \alpha \in \mathbb{R}. \quad (4.87)$$

L'expression de  $h(\cdot, \mathbf{0})$  implique que

$$q(\alpha) = f(\mathbf{0}) + \alpha \mathbf{g}_k^T \mathbf{d}_k^1 + \frac{1}{2} \alpha^2 \mathbf{d}_k^{1T} \mathbf{A}_k^0 \mathbf{d}_k^1. \quad (4.88)$$

Le minimiseur  $\hat{\alpha}_k$  de  $q$  s'écrit

$$\hat{\alpha}_k = -\frac{\mathbf{g}_k^T \mathbf{d}_k^1}{\mathbf{d}_k^{1T} \mathbf{A}_k^0 \mathbf{d}_k^1}. \quad (4.89)$$

Donc,

$$q(\hat{\alpha}_k) = f(\mathbf{0}) + \frac{1}{2}\hat{\alpha}_k \mathbf{g}_k^T \mathbf{d}_k^1. \quad (4.90)$$

De plus, d'après l'expression de  $\hat{\mathbf{s}}_k^1$ ,

$$h(\hat{\mathbf{s}}_k^1, \mathbf{0}) = f(\mathbf{0}) + \frac{1}{2}\nabla f(\mathbf{0})^T \hat{\mathbf{s}}_k^1. \quad (4.91)$$

$\hat{\mathbf{s}}_k^1$  minimise  $h(\mathbf{s}, \mathbf{0})$  donc  $h(\hat{\mathbf{s}}_k^1, \mathbf{0}) \leq q(\hat{\alpha}_k)$ . Par conséquent, (4.90)-(4.91) conduisent à

$$\hat{\alpha}_k \mathbf{g}_k^T \mathbf{d}_k^1 \geq \nabla f(\mathbf{0})^T \hat{\mathbf{s}}_k^1. \quad (4.92)$$

D'après (4.88) et (4.89), le pas  $\alpha_k = \theta \hat{\alpha}_k$  vérifie

$$q(\alpha_k) = f(\mathbf{0}) + \delta \hat{\alpha}_k \mathbf{g}_k^T \mathbf{d}_k^1, \quad (4.93)$$

avec  $\delta = \theta(1 - \theta/2)$ . En outre,

$$h(\mathbf{s}_k^1, \mathbf{0}) = f(\mathbf{0}) + \delta \nabla f(\mathbf{0})^T \hat{\mathbf{s}}_k^1. \quad (4.94)$$

Donc, les égalités (4.92)-(4.93)-(4.94) impliquent que  $h(\mathbf{s}_k^1, \mathbf{0}) \leq q(\alpha_k)$  et

$$f(\mathbf{0}) - h(\mathbf{s}_k^1, \mathbf{0}) \geq -\delta \hat{\alpha}_k \mathbf{g}_k^T \mathbf{d}_k^1. \quad (4.95)$$

De plus,  $h(\mathbf{s}_k^1, \mathbf{0}) \geq f(\mathbf{s}_k^1) \geq f(\mathbf{s}_k)$  d'après l'hypothèse 4.1 et le lemme 4.3. Donc,

$$f(\mathbf{0}) - f(\mathbf{s}_k) \geq -\delta \hat{\alpha}_k \mathbf{g}_k^T \mathbf{d}_k^1 \quad (4.96)$$

D'après l'hypothèse 4.1,

$$\hat{\alpha}_k \geq -\frac{\mathbf{g}_k^T \mathbf{d}_k^1}{\nu_{\max}^{\mathcal{A}} \|\mathbf{d}_k^1\|^2}. \quad (4.97)$$

Par conséquent, en utilisant (4.96), (4.97) et (4.77), on déduit la relation

$$f(\mathbf{0}) - f(\mathbf{s}_k) \geq \frac{\delta \nu_1^2}{\nu_{\max}^{\mathcal{A}} \nu_2^2} \|\mathbf{g}_k\|^2, \quad (4.98)$$

qui s'écrit aussi

$$F(\mathbf{x}_k) - F(\mathbf{x}_{k+1}) \geq \frac{\delta \nu_1^2}{\nu_{\max}^{\mathcal{A}} \nu_2^2} \|\mathbf{g}_k\|^2. \quad (4.99)$$

De plus,  $F$  étant bornée inférieurement,  $\lim_{k \rightarrow \infty} F(\mathbf{x}_k)$  est finie. Soit

$$\infty > \left( \frac{\delta \nu_1^2}{\nu_{\max}^{\mathcal{A}} \nu_2^2} \right)^{-1} \left( F(\mathbf{x}_0) - \lim_{k \rightarrow \infty} F(\mathbf{x}_k) \right) \geq \sum_k \|\mathbf{g}_k\|^2,$$

et pour conclure

$$\lim_{k \rightarrow \infty} \|\mathbf{g}_k\| = 0.$$

□

#### 4.5.4 Convergence des algorithmes de gradient conjugué non linéaire

Le caractère gradient relié des directions de gradient conjugué non linéaire n'a pas été démontré à notre connaissance. La difficulté réside dans le fait que la direction GCNL tient compte du gradient courant mais aussi de la direction précédente  $\mathbf{d}_{k-1}$ . La convergence des algorithmes GCNL associés à la recherche de pas MMQ 1D ne s'obtient pas de façon directe. Elle est établie dans (Labat et Idier, 2008) sous l'hypothèse supplémentaire que le critère est à gradient Lipschitz sur la ligne de niveau initiale.

**Hypothèse 4.2.** Soit la ligne de niveau

$$\mathcal{L}_0 = \{\mathbf{x} \in \mathcal{D}_f | F(\mathbf{x}) \leq F(\mathbf{x}_0)\} \quad (4.100)$$

supposée bornée et  $\mathcal{V}$  un ouvert contenant  $\mathcal{L}_0$ . Les assertions suivantes sont vérifiées :

- $F$  est deux fois continument différentiable sur  $\mathcal{V}$ ,
- $F$  est à gradient  $L$ -Lipschitz sur  $\mathcal{V}$ ,
- $\mathcal{V}$  est borné.
- Pour tous  $\mathbf{x} \in \mathcal{V}$  et  $\mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ ,  $\frac{\mathbf{v}^T \mathbf{A}(\mathbf{x}) \mathbf{v}}{\|\mathbf{v}\|^2} \geq \nu_{\min}^{\mathbf{A}} > 0$

Le caractère borné de  $\mathcal{L}_0$ , et a fortiori du voisinage  $\mathcal{V}$  est une hypothèse classique en optimisation. Elle permet notamment d'assurer l'existence d'un minimiseur de  $F$  à l'intérieur strict de  $\mathcal{L}_0$ . Cette hypothèse est vérifiée par exemple lorsque la fonction  $F$  est coercive. L'analyse de convergence dépend du coefficient de conjugaison employé. Afin de vérifier la propriété de descente, la direction est définie par

$$\begin{aligned} \mathbf{c}_{k+1} &= -\mathbf{g}_{k+1} + \beta_{k+1} \mathbf{d}_k, \\ \mathbf{d}_{k+1} &= -\mathbf{c}_{k+1} \text{signe}(\mathbf{g}_{k+1}^T \mathbf{c}_{k+1}). \end{aligned} \quad (4.101)$$

De plus,  $\beta_k$  est supposé s'écrire sous la forme (Dai et Yuan, 2001) :

$$\beta_0 = 0, \quad \beta_k = \beta_k^{\mu_k, \omega_k} = \mathbf{g}_k^T \boldsymbol{\delta}_{k-1} / D_k, \quad \forall k > 0 \quad (4.102)$$

avec

$$\begin{aligned} D_k &= (1 - \mu_k - \omega_k) \|\mathbf{g}_{k-1}\|^2 + \mu_k \mathbf{d}_{k-1}^T \boldsymbol{\delta}_{k-1} - \omega_k \mathbf{d}_{k-1}^T \mathbf{g}_{k-1} \\ \boldsymbol{\delta}_{k-1} &= \mathbf{g}_k - \mathbf{g}_{k-1} \\ \mu_k &\in [0, 1], \quad \omega_k \in [0, 1 - \mu_k] \end{aligned}$$

La formule (4.102) couvre les différentes stratégies de conjugaisons :

$$\begin{array}{ll} \beta_k^{1,0} = \mathbf{g}_k^T \boldsymbol{\delta}_{k-1} / \mathbf{d}_{k-1}^T \boldsymbol{\delta}_{k-1} & \text{Hestenes-Stiefel (HS)} \\ \beta_k^{0,0} = \mathbf{g}_k^T \boldsymbol{\delta}_{k-1} / \|\mathbf{g}_{k-1}\|^2 & \text{Polak-Ribière-Polyak (PRP)} \\ \beta_k^{0,1} = -\mathbf{g}_k^T \boldsymbol{\delta}_{k-1} / \mathbf{d}_{k-1}^T \mathbf{g}_{k-1} & \text{Liu-Storey (LS)} \end{array}$$

**Théorème 4.4.** (Labat et Idier, 2008)

Supposons que le critère  $F$  vérifie les hypothèses 4.1 et 4.2. L'algorithme de descente itérative (3.18) associé à la direction (4.101) converge au sens  $\liminf_{k \rightarrow \infty} \mathbf{g}_k = 0$  si  $\alpha_k$  est obtenu par (4.44) et  $\beta_k$  provient de la méthode PRP ou LS, ou plus généralement de (4.102) avec  $\mu_k = 0$  and  $\omega_k \in [0, 1]$ . De plus, si  $F$  est fortement convexe sur  $\mathcal{V}$ , alors on a pour tout  $J > 0$  et  $\theta \in (0, 2)$ ,  $\liminf_{k \rightarrow \infty} \mathbf{g}_k = 0$  pour toute valeur de  $(\mu_k, \omega_k)$ .

*Démonstration.* La démonstration du théorème 4.4 est établie dans l'article (Labat et Idier, 2008). Elle utilise les propriétés suivantes du pas MM :

- Pour tout  $k$ , on a l'inégalité  $-\nu(\mathbf{g}_k^T \mathbf{d}_k)/\|\mathbf{d}_k\|^2 \leq \alpha_k^1 \leq -\nu'(\mathbf{g}_k^T \mathbf{d}_k)/\|\mathbf{d}_k\|^2$  avec  $(\nu, \nu') > 0$  paramètres dépendant de  $\nu_{\min}^A, \nu_{\max}^A$  et  $\theta$ . (Labat et Idier, 2008, Eq.13)
- La condition d'Armijo est vérifiée (Labat et Idier, 2008, Lem.3.5)
- Pour tout  $J, k$ , on a l'inégalité  $c_{\min} \alpha_k^1 \leq \alpha_k \leq c_{\max}^J \alpha_k^1$  avec  $(c_{\min}, c_{\max}^J) \in (0, 1)$  paramètres dépendant de  $L, \nu_{\min}^A, \nu_{\max}^A$  et  $\theta$ . (Labat et Idier, 2008, Theo.3.1)
- La condition de Zoutendijk est vérifiée (Labat et Idier, 2008, Lem.4.1)

□

## 4.6 Conclusion

La recherche de pas par Majoration-Minoration quadratique consiste à utiliser des informations structurelles sur le critère pour en construire des approximations quadratiques majorantes. Le pas résulte d'une récurrence très simple à implémenter puisque chaque itération correspond à la minimisation d'une fonction quadratique dans un espace de dimension réduite. Nous avons étendu dans ce chapitre la recherche de pas MMQ scalaire proposée par (Labat et Idier, 2008) en une recherche de pas MMQ multi-dimensionnelle pour l'algorithme de sous-espace. Des propriétés de convergence des schémas scalaires et multi-dimensionnels ont été établis. Le chapitre 5 est consacré à une analyse expérimentale des algorithmes associés à ces recherches de pas sur un problème de traitement d'image.



# Chapitre 5

## Application à la déconvolution d'image

### Sommaire

---

5.1	Introduction . . . . .	<b>66</b>
5.2	Le problème de déconvolution d'image . . . . .	<b>66</b>
5.3	Algorithmes comparés . . . . .	<b>66</b>
5.3.1	Algorithmes de gradient conjugué non linéaire . . . . .	66
5.3.2	Algorithme à super mémoire de gradient . . . . .	67
5.3.3	Préconditionnement . . . . .	68
5.4	Mise en œuvre des stratégies de pas MMQ . . . . .	<b>68</b>
5.4.1	Construction d'approximations majorantes quadratiques . . . . .	68
5.4.2	Réduction du coût de calcul . . . . .	70
5.5	Interprétation des résultats expérimentaux . . . . .	<b>72</b>
5.5.1	Algorithme GCNL . . . . .	72
5.5.2	Algorithme SMG . . . . .	73
5.5.3	Analyse comparée GCNL / SMG . . . . .	74
5.6	Résultats complémentaires . . . . .	<b>75</b>
5.6.1	Construction du sous-espace . . . . .	75
5.6.2	Reconstruction dans un dictionnaire . . . . .	76
5.7	Conclusion . . . . .	<b>76</b>

---

## 5.1 Introduction

La convergence de nombreux algorithmes de sous-espace associés à la recherche de pas MMQ multi-dimensionnelle a été établie dans le chapitre 4. Parmi ceux-ci, l'algorithme de super mémoire de gradient (SMG) consiste à utiliser un sous-espace construit à partir du gradient courant et de  $m$  directions précédentes. Lorsque  $m = 1$ , on retrouve un algorithme de structure très proche de celle des algorithmes de gradient conjugué non linéaire (GCNL). Il s'agit dans ce chapitre de comparer les performances des algorithmes SMG et GCNL, lorsque le calcul du pas est fait en utilisant l'approche MMQ. L'analyse est effectuée à travers la minimisation d'un critère pénalisé sur un problème de déconvolution d'image avec préservation des discontinuités.

## 5.2 Le problème de déconvolution d'image

On considère un problème de déconvolution d'image de grande taille. L'image à restaurer est une image convoluée et bruitée. Nos expérimentations ont été menées sur l'image `peppers` de taille  $n = 512 \times 512$ . Le noyau de convolution est gaussien avec un écart type de 2.24 et de taille  $17 \times 17$ . L'hypothèse de bord considérée est celle de Dirichlet (zéros à l'extérieur). Un bruit blanc gaussien est ajouté à l'image convoluée de telle sorte à avoir un rapport signal sur bruit de 40 dB.

L'image est déconvoluée en utilisant une approche pénalisée. Nous utilisons le critère pénalisé fréquemment utilisé en restauration d'images (Allain *et al.*, 2006; Charbonnier *et al.*, 1997; Nikolova *et Ng*, 2005)

$$F(\mathbf{x}) = \|\mathbf{y} - \mathbf{K}\mathbf{x}\|^2 + \lambda \sum_c \phi([\mathbf{V}\mathbf{x}]_c) \quad (5.1)$$

où  $\mathbf{V}$  est la matrice des différences du premier ordre avec quatre voisins et  $\phi$  est la fonction de régularisation hyperbolique  $\phi(u) = \sqrt{u^2 + \delta^2}$  (Bouman *et Sauer*, 1993). Cela correspond à un critère des moindres carrés associé à une pénalisation  $\ell_1 - \ell_2$  qui favorise les zones lisses dans l'image sans trop en dégrader les contours. Le critère est différentiable et strictement convexe, car le noyau de  $\mathbf{V}^T\mathbf{V}$  réduit aux images constantes n'est pas inclus dans le noyau de  $\mathbf{K}^T\mathbf{K}$  soit

$$\ker(\mathbf{H}^T\mathbf{H}) \cap \ker(\mathbf{V}^T\mathbf{V}) = \{\mathbf{0}\}. \quad (5.2)$$

La convexité du critère implique que l'annulation du gradient est une condition nécessaire et suffisante d'obtention d'un minimum local. De plus, la stricte convexité assure l'unicité du minimiseur de (5.1).

Le critère pénalisé dépend des paramètres  $\lambda$  et  $\delta$ . Ils sont réglés afin de maximiser la similarité, au sens de la norme quadratique, entre l'image déconvoluée et l'image originale, soit  $\delta = 8$  et  $\lambda = 0.2$ . Le résultat obtenu est illustré en figure 5.1.

## 5.3 Algorithmes comparés

L'objectif de ce chapitre est de comparer les performances de l'algorithme GCNL associé au pas MMQ scalaire, avec celles de l'algorithme SMG associé au pas MMQ multi-dimensionnel. Nous donnons à présent une description détaillée des algorithmes comparés.

### 5.3.1 Algorithmes de gradient conjugué non linéaire

L'algorithme GCNL appartient à la famille des méthodes de descente itérative. Il s'écrit sous la forme

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k, \quad (5.3)$$

avec

$$\mathbf{d}_k = -\mathbf{g}_k + \beta_k \mathbf{d}_{k-1}. \quad (5.4)$$



(a) Image convoluée et bruitée

(b) Image reconstruite

FIGURE 5.1 – Reconstruction de peppers avec  $(\lambda, \delta) = (0.2, 8)$ 

Différentes formules pour le coefficient de conjugaison  $\beta_k$  sont considérées (Hager et Zhang, 2006b) :

$\beta_k = \mathbf{g}_k^T \boldsymbol{\delta}_{k-1} / \ \mathbf{g}_{k-1}\ ^2$	Polak-Ribière-Polyak (PRP)	(Polak et Ribière, 1969; Polyak, 1969)
$\beta_k = \mathbf{g}_k^T \boldsymbol{\delta}_{k-1} / \mathbf{d}_{k-1}^T \boldsymbol{\delta}_{k-1}$	Hestenes-Stiefel (HS)	(Hestenes et Stiefel, 1952)
$\beta_k = -\mathbf{g}_k^T \mathbf{y}_{k-1} / \mathbf{d}_{k-1}^T \mathbf{g}_{k-1}$	Liu-Storey (LS)	(Liu et Storey, 1991)
$\beta_k = \ \mathbf{g}_k\ ^2 / \ \mathbf{g}_{k-1}\ ^2$	Fletcher-Reeves (FR)	(Fletcher et Reeves, 1964)
$\beta_k = \ \mathbf{g}_k\ ^2 / \mathbf{d}_{k-1}^T \boldsymbol{\delta}_{k-1}$	Dan-Yuan (DY)	(Dai et Yuan, 1999)

avec  $\boldsymbol{\delta}_{k-1} = \mathbf{g}_k - \mathbf{g}_{k-1}$ .

Concernant le pas  $\alpha_k$ , deux stratégies sont comparées :

- La recherche de pas MM quadratique de paramètre  $J = \{1, 2, 5, 10\}$  décrite par la récurrence

$$\begin{cases} \alpha_k^0 &= 0, \\ \alpha_k^{j+1} &= \alpha_k^j - \frac{f(\alpha_k^j)}{m_k^j}, j < J, \\ \alpha_k &= \alpha_k^J, \end{cases} \quad (5.5)$$

avec  $m_k^j = \mathbf{d}_k^T \mathbf{A}(\mathbf{x}_k + \alpha_k^j \mathbf{d}_k) \mathbf{d}_k$ .

- La recherche de pas itérative basée sur la règle de Wolfe stricte (4.10)-(4.13), paramétrée par les coefficients de Wolfe  $c_1$  et  $c_2$  qui doivent vérifier la condition  $0 < c_1 < c_2 < 1$ . Nous utilisons la procédure de recherche de pas correspondant aux algorithmes 3.2 et 3.3 de (Nocedal et Wright, 1999) initialisée avec un pas unitaire. Nous fixons  $c_1 = 0.01$  et nous testons  $c_2 = \{0.5, 0.9, 0.99\}$ .

### 5.3.2 Algorithme à super mémoire de gradient

L'algorithme SMG appartient à la classe des méthodes itératives de sous-espace dont une itération s'écrit sous la forme

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{D}_k \mathbf{s}_k. \quad (5.6)$$

Dans l'algorithme SMG, le sous-espace est engendré par le gradient courant et un petit nombre de directions précédentes :

$$\mathbf{D}_k = [-\mathbf{g}_k, \mathbf{d}_{k-1}, \dots, \mathbf{d}_{k-m}], \quad (5.7)$$

en utilisant la notation  $\mathbf{d}_{k-i} = \mathbf{x}_{k-i+1} - \mathbf{x}_{k-i}$ .

Nous testons les valeurs  $\{1, 2, 5, 10\}$  pour le paramètre  $m$  définissant la mémoire de l'algorithme. La taille du sous-espace associée est  $r = m+1$ . Nous comparons l'algorithme SMG avec l'algorithme ORTH de (Nemirovski, 1982) basé sur le sous-espace

$$\mathbf{D}_k = [-\mathbf{g}_k, \mathbf{x}_k - \mathbf{x}_0, \sum_{i=0}^k w_i \mathbf{g}_i] \quad (5.8)$$

où les poids  $w_i$  sont définis par (3.50), et avec la version accélérée de ORTH, présentée sous le nom de SESOP dans (Narkiss et Zibulevsky, 2005), où

$$\mathbf{D}_k = [-\mathbf{g}_k, \mathbf{x}_k - \mathbf{x}_0, \sum_{i=0}^k w_i \mathbf{g}_i, \mathbf{d}_{k-1}, \dots, \mathbf{d}_{k-m}]. \quad (5.9)$$

Concernant le calcul de  $\mathbf{s}_k$ , nous comparons la recherche de pas MM quadratique de paramètre  $J = \{1, 2, 5, 10\}$  décrite par

$$\begin{cases} \mathbf{s}_k^0 &= \mathbf{0}, \\ \mathbf{s}_k^{j+1} &= \mathbf{s}_k^j - (\mathbf{B}_k^j)^{-1} \nabla f(\mathbf{s}_k^j), \quad j < J, \\ \mathbf{s}_k &= \mathbf{s}_k^J, \end{cases} \quad (5.10)$$

où  $\mathbf{B}_k^j = \mathbf{D}_k^T \mathbf{A}(\mathbf{x}_k + \mathbf{D}_k \mathbf{s}_k^j) \mathbf{D}_k$ , avec la recherche de pas employée dans (Narkiss et Zibulevsky, 2005; Zibulevsky, 2008; Zibulevsky et Elad, 2010) correspondant à une récurrence de Newton stoppée lorsque

$$\frac{\|\nabla f(\mathbf{s}_k^j)\|}{\|\nabla f(\mathbf{0})\|} < 10^{-10}, \quad \text{ou } j > 7.$$

### 5.3.3 Préconditionnement

Afin d'accélérer la convergence des algorithmes, nous envisageons d'utiliser la version préconditionnée des algorithmes GCNL et SMG. Nous utilisons un préconditionneur basé sur la transformée en cosinus 2D, introduit dans (Ng *et al.*, 1999) pour le préconditionnement des matrices à structure Toeplitz. Bien que le Hessien du critère ne soit pas Toeplitz (à cause de sa dépendance en  $\mathbf{x}_k$ ), l'efficacité de ce préconditionneur pour la minimisation de (5.1) est soulignée dans (Labat et Idier, 2007, 2008).

## 5.4 Mise en œuvre des stratégies de pas MMQ

### 5.4.1 Construction d'approximations majorantes quadratiques

Afin de garantir les propriétés de convergence des stratégies de pas MMQ, il nous faut déterminer une famille de matrices  $\mathbf{A}(\cdot)$  qui vérifie l'hypothèse 4.1. Nous décrivons ici deux stratégies de construction de  $\mathbf{A}(\cdot)$ , issues des algorithmes semi-quadratiques de Geman & Yang (GY) (Geman et Yang, 1995) et de Geman & Reynolds (GR) (Geman et Reynolds, 1992). D'autres techniques de construction sont présentées dans (De Pierro, 1995; Erdogan et Fessler, 1999; Hunter et L., 2004; Lange et Fessler, 1995; Zheng *et al.*, 2004).

#### 5.4.1.1 Algorithmes semi-quadratiques

Les algorithmes semi-quadratiques (SQ) de GY et GR sont très utilisés dans la communauté du traitement de l'image pour la minimisation du critère pénalisé (5.1) (Allain *et al.*, 2006; Charbonnier *et al.*, 1997; Nikolova et Ng, 2005). Ils consistent à exploiter la forme analytique du critère pour remplacer sa minimisation par une séquences de problèmes plus simples.

L'algorithme SQ visant à minimiser (5.1) s'écrit

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \theta(\hat{\mathbf{x}}_{k+1} - \mathbf{x}_k), \quad (5.11)$$

$$\hat{\mathbf{x}}_{k+1} = \mathbf{x}_k - \mathbf{A}(\mathbf{x}_k)^{-1} \nabla F(\mathbf{x}_k), \quad (5.12)$$

avec  $\theta \in (0, 2)$  et  $\mathbf{A}(\cdot)$  égal à l'un des deux opérateurs suivants :

$$\mathbf{A}_{\text{GY}}^a = 2\mathbf{K}^T \mathbf{K} + \lambda/a \mathbf{V}^T \mathbf{V} \quad (5.13)$$

$$\mathbf{A}_{\text{GR}}(\mathbf{x}_k) = 2\mathbf{K}^T \mathbf{K} + \lambda \mathbf{V}^T \text{Diag} \left\{ \dot{\phi}([\mathbf{V} \mathbf{x}_k]) / ([\mathbf{V} \mathbf{x}_k]) \right\} \mathbf{V} \quad (5.14)$$

Les algorithmes SQ-GR et SQ-GY peuvent recevoir plusieurs interprétations possibles (Allain *et al.*, 2006). Nous nous intéressons à la propriété de majoration des constructions SQ qui va nous conduire à deux principes de constructions d'approximations majorantes quadratiques du critère (5.1).

#### 5.4.1.2 Fonctions majorantes semi-quadratiques

Dans (Chan et Mulet, 1999), il est démontré que l'algorithme GR peut être analysé comme un algorithme MMQ i.e., une minimisation itérative d'approximations majorantes quadratiques. Un tel résultat est aussi obtenu pour l'algorithme GY (Allain *et al.*, 2006). Ces propriétés sont établies sous plusieurs hypothèses sur la fonction de régularisation.

##### Hypothèse 5.1.

- (i)  $\phi$  est  $\mathcal{C}^1$
- (ii)  $\dot{\phi}$  est  $L$ -Lipschitz

##### Hypothèse 5.2.

- (i)  $\phi$  est  $\mathcal{C}^1$  et paire
- (ii)  $\phi(\sqrt{\cdot})$  est concave sur  $\mathbb{R}^+$
- (iii)  $0 < \dot{\phi}(t)/t < \infty, \quad \forall t \in \mathbb{R}$

Alors, le lemme suivant établit que les matrices normales  $\mathbf{A}_{\text{GY}}^a$  et  $\mathbf{A}_{\text{GR}}(\cdot)$  issues des constructions de GY et de GR induisent des approximations quadratiques majorantes pour le critère pénalisé (5.1).

##### Lemme 5.1. (Allain et al., 2006, Prop.1)

Soit  $F$  le critère pénalisé défini par (5.1). Supposons que l'hypothèse 5.1 (resp. hypothèse 5.2) est satisfait. Alors

$$H(\mathbf{x}, \mathbf{x}_k) = F(\mathbf{x}_k) + \nabla F(\mathbf{x}_k)^T (\mathbf{x} - \mathbf{x}_k) + \frac{1}{2} (\mathbf{x} - \mathbf{x}_k)^T \mathbf{A}(\mathbf{x}_k) (\mathbf{x} - \mathbf{x}_k), \quad (5.15)$$

est une approximation quadratique majorante du critère  $F$ , où  $\mathbf{A}(\cdot) = \mathbf{A}_{\text{GY}}^a$  avec  $0 < a < 1/L$  (resp.  $\mathbf{A}(\cdot) = \mathbf{A}_{\text{GR}}(\cdot)$ ).

Les deux variantes de l'algorithme SQ s'écrivent sous la forme

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \theta(\hat{\mathbf{x}}_{k+1} - \mathbf{x}_k), \quad (5.16)$$

$$\hat{\mathbf{x}}_{k+1} = \arg \min_{\mathbf{x}} H(\mathbf{x}, \mathbf{x}_k), \quad (5.17)$$

et peuvent donc être interprétés comme résultant de minimisations successives d'approximations quadratiques majorantes. Nous remarquons que les propriétés de majoration des opérateurs GY et GR se déduisent simplement des lemmes 4.1 et 4.2.

### 5.4.1.3 Comparaison entre majorations GY et GR

La fonction hyperbolique vérifie à la fois l'hypothèse 5.1 avec  $L = 1/\delta$  et l'hypothèse 5.2 donc le lemme 5.1 s'applique. De plus, la condition (5.2) assure que les matrices GY et GR sont à spectre borné sur  $\mathbb{R}^n$  (Allain *et al.*, 2006). Par conséquent, l'hypothèse 4.1 est vérifiée et les matrices GY et GR sont deux candidates acceptables pour la construction du pas MMQ. La question se pose de déterminer la meilleure stratégie de majoration, entre GY et GR.

La figure 5.2 illustre les fonctions majorantes GY et GR de la fonction hyperbolique centrée  $\phi(u) = \sqrt{u^2 + \delta^2} - \delta$  aux valeurs 1, 2, 4. Nous remarquons que la majorante issue de la méthode GR est plus proche visuellement du critère. Nous la retenons par la suite car elle conduit aux meilleurs résultats en terme de nombre d'itérations ainsi que de temps de calcul nécessaire à la convergence de l'algorithme.

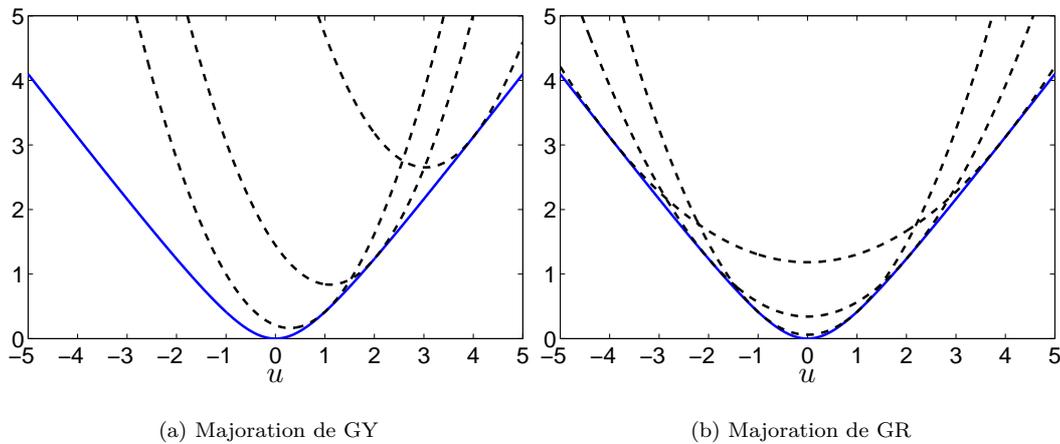


FIGURE 5.2 – Principe de construction des majorantes semi-quadratiques. La fonction hyperbolique  $\phi(u) = \sqrt{u^2 + 1} - 1$  (trait continu) est majorée en  $u = 1, 2$  et  $4$  par des fonctions quadratiques (trait pointillé) issues de la stratégie de majoration de Geman & Yang ou Geman & Reynolds.

### 5.4.1.4 Résultats de convergence

L'approximation majorante quadratique de Geman & Reynolds vérifie l'hypothèse 4.1. De plus,  $\phi$  étant coercive et à gradient Lipschitz sur  $\mathbb{R}$ , le critère vérifie l'hypothèse 4.2. Le théorème 4.4 garantit donc la convergence de l'algorithme GCNL associé à la conjugaison PRP, LS ou HS, et au pas MM quadratique scalaire. Le théorème 4.4 ne traite pas le cas des conjugaisons FR et DY. Cependant, pour tous nos tests, les algorithmes GCNL utilisant ces conjugaisons parviennent à vérifier le critère d'arrêt requis.

La direction  $\mathbf{d}_k^1 = -\mathbf{g}_k$  vérifie la condition (4.77). Par conséquent, le théorème 4.3 assure la convergence des algorithmes SMG, ORTH et SESOP associé au pas MM quadratique multi-dimensionnel.

Enfin, le choix d'un préconditionneur constant pour accélérer la convergence des algorithmes permet de conserver la validité des théorèmes 4.4 et 4.3.

## 5.4.2 Réduction du coût de calcul

Le problème étant de grande taille, il est nécessaire d'être vigilant dans l'implémentation des algorithmes afin de minimiser le coût de calcul par itération. Dans cet exemple, le coût de calcul est gouverné par l'application des opérateurs  $\mathbf{K}$ ,  $\mathbf{K}^T$ ,  $\mathbf{V}$  et  $\mathbf{V}^T$  sur un vecteur. Nous parlerons par la suite de "produits matrice-vecteur", bien que l'application de la matrice de convolution soit effectuée à l'aide de l'outil `conv2` de Matlab.

Dans les algorithmes GCNL et SMG, chaque itération se décompose en deux étapes consécutives, le calcul de la direction (resp. du sous-espace) et le calcul du pas. La construction de la direction GNCL et du sous-espace SMG nécessite de calculer le gradient du critère au point courant

$$\nabla F(\mathbf{x}_k) = 2\mathbf{K}^T(\mathbf{K}\mathbf{x}_k - \mathbf{y}) + \lambda\mathbf{V}^T\dot{\phi}(\mathbf{V}\mathbf{x}_k) \quad (5.18)$$

soit 4 produits matrice-vecteur. Nous proposons dans cette section de minimiser le nombre de produits matrice-vecteur lors de l'étape de recherche de pas MMQ en tirant partie de la structure du critère.

Etudions tout d'abord le coût de la recherche de pas MMQ scalaire. A chaque itération  $j$ , il faut calculer  $\dot{f}(\alpha_k^j)$  et  $m_k^j$  à partir des expressions

$$\nabla f(\alpha_k^j) = 2(\mathbf{K}\mathbf{d}_k)^T \left( \mathbf{K}(\mathbf{x}_k + \alpha_k^j \mathbf{d}_k) - \mathbf{y} \right) + \lambda(\mathbf{V}\mathbf{d}_k)^T \dot{\phi} \left( \mathbf{V}(\mathbf{x}_k + \alpha_k^j \mathbf{d}_k) \right) \quad (5.19)$$

et

$$m_k^j = 2(\mathbf{K}\mathbf{d}_k)^T(\mathbf{K}\mathbf{d}_k) + \lambda(\mathbf{V}\mathbf{d}_k)^T \Sigma_k^j(\mathbf{V}\mathbf{d}_k) \quad (5.20)$$

avec

$$\Sigma_k^j = \text{Diag} \left\{ \dot{\phi}([\mathbf{V}(\mathbf{x}_k + \alpha_k^j \mathbf{d}_k)]) / ([\mathbf{V}(\mathbf{x}_k + \alpha_k^j \mathbf{d}_k)]) \right\}, \quad (5.21)$$

soit apparemment  $4J$  produits matrice-vecteur. Le coût de calcul de  $\dot{f}(\alpha_k^j)$  et  $m_k^j$  peut être réduit simplement en remarquant que

$$\mathbf{K}(\mathbf{x}_k + \alpha_k^j \mathbf{d}_k) = (\mathbf{K}\mathbf{x}_k) + \alpha_k^j(\mathbf{K}\mathbf{d}_k). \quad (5.22)$$

En utilisant les produits  $\mathbf{K}\mathbf{x}_k$  et  $\mathbf{V}\mathbf{x}_k$  déjà calculés lors de la construction de la direction, on réduit le coût à seulement 2 produits matrice-vecteur  $\mathbf{K}\mathbf{d}_k$  et  $\mathbf{V}\mathbf{d}_k$  pour implémenter  $J$  itérations de recherche de pas. Notons que cette stratégie n'est pas nouvelle car elle est déjà proposée dans (Kaufman, 1987) pour l'implémentation d'une recherche de pas pour l'algorithme Richardson-Lucy (Lucy, 1974; Richardson, 1972) en tomographie d'émission.

Etudions maintenant le coût d'implémentation de la recherche de pas MMQ multi-dimensionnelle. A chaque itération  $j$ , il faut calculer  $\dot{f}(\mathbf{s}_k^j)$  et  $\mathbf{B}_k^j$  à partir des expressions

$$\nabla f(\mathbf{s}_k^j) = 2(\mathbf{K}\mathbf{D}_k)^T(\mathbf{K}\mathbf{x}_k + \mathbf{K}\mathbf{D}_k\mathbf{s}_k^j - \mathbf{y}) + \lambda(\mathbf{V}\mathbf{D}_k)^T \dot{\phi}(\mathbf{V}\mathbf{x}_k + \mathbf{V}\mathbf{D}_k\mathbf{s}_k^j) \quad (5.23)$$

et

$$\mathbf{B}_k^j = 2(\mathbf{K}\mathbf{D}_k)^T(\mathbf{K}\mathbf{D}_k) + \lambda(\mathbf{V}\mathbf{D}_k)^T \Sigma_k^j(\mathbf{V}\mathbf{D}_k) \quad (5.24)$$

avec

$$\Sigma_k^j = \text{Diag} \left\{ \dot{\phi}([\mathbf{V}\mathbf{x}_k + \mathbf{V}\mathbf{D}_k\mathbf{s}_k^j]) / ([\mathbf{V}\mathbf{x}_k + \mathbf{V}\mathbf{D}_k\mathbf{s}_k^j]) \right\}. \quad (5.25)$$

Une stratégie similaire au cas scalaire peut être appliquée, nécessitant de calculer au préalable les produits  $\mathbf{K}\mathbf{D}_k$  et  $\mathbf{V}\mathbf{D}_k$ . Nous ramenons ainsi le coût de calcul de la recherche de pas à  $2r$  produits matrice-vecteur. Cependant, le coût reste linéairement lié à la taille du sous-espace. Cela risque de contre-balancer le bon taux de convergence attendu de l'algorithme SMG en terme de nombre d'itérations. Face à ce constat, nous proposons de tirer parti de la structure récursive du sous-espace SMG pour réduire le coût d'implémentation de la recherche de pas multi-dimensionnelle. Notre approche reprend l'idée développée dans (Narkiss et Zibulevsky, 2005) pour l'implémentation de la méthode de Newton pour la recherche de pas dans l'algorithme SESOP.

D'après l'expression de  $\mathbf{D}_k$ , le produit  $\mathbf{P}_k \triangleq \mathbf{K}\mathbf{D}_k$  s'écrit

$$\mathbf{P}_k = [-\mathbf{K}\mathbf{g}_k, \mathbf{K}\mathbf{d}_{k-1}, \dots, \mathbf{K}\mathbf{d}_{k-m}]. \quad (5.26)$$

En utilisant l'expression de  $\mathbf{d}_{k-1} = \mathbf{D}_{k-1}\mathbf{s}_{k-1}$ , nous obtenons que

$$\mathbf{P}_k = [-\mathbf{K}\mathbf{g}_k, \mathbf{P}_{k-1}\mathbf{s}_{k-1}, \mathbf{P}_{k-1}(2:m)], \quad (5.27)$$

avec  $\mathbf{P}(i : j)$  la sous-matrice contenant les colonnes de  $i$  à  $j$  de  $\mathbf{P}$ . Par conséquent, la matrice  $\mathbf{P}_k$  se déduit de  $\mathbf{P}_{k-1}$  en effectuant un seul produit matrice-vecteur et des opération peu coûteuses sur les colonnes de  $\mathbf{P}_{k-1}$ . En utilisant la même procédure pour calculer le produit  $\mathbf{V}\mathbf{D}_k$ , nous réduisons le coût de la recherche de pas multi-dimensionnelle à 2 produits matrice-vecteur soit un coût de calcul similaire à celui de la recherche de pas scalaire. Cette stratégie se généralise facilement aux autres algorithmes de sous-espace présentés dans le tableau 3.1, en utilisant leur expression récursive donnée dans le tableau 5.1.

Acronyme	Expression récursive de $\mathbf{D}_k$	Coût
MG	$[-\mathbf{g}_k, \mathbf{D}_{k-1} \mathbf{s}_{k-1}]$	2
SMG	$[-\mathbf{g}_k, \mathbf{D}_{k-1} \mathbf{s}_{k-1}, \mathbf{D}_{k-1}(2 : m)]$	2
SMD	$[-\mathbf{p}_k, \mathbf{D}_{k-1} \mathbf{s}_{k-1}, \mathbf{D}_{k-1}(2 : m)]$	2
GS	$[-\mathbf{g}_k, \mathbf{D}_{k-1}(1 : m)]$	2
ORTH	$[-\mathbf{g}_k, \mathbf{x}_k - \mathbf{x}_0, \omega_k \mathbf{g}_k + \mathbf{D}_{k-1}(3)]$	4
SESOP	$[-\mathbf{g}_k, \mathbf{x}_k - \mathbf{x}_0, \omega_k \mathbf{g}_k + \mathbf{D}_{k-1}(3), \mathbf{D}_{k-1} \mathbf{s}_{k-1}, \mathbf{D}_{k-1}(4 : m + 2)]$	4
QNS	$[-\mathbf{g}_k, \mathbf{g}_k + \mathbf{D}_{k-1}(1), \mathbf{D}_{k-1}(2 : m), \mathbf{D}_{k-1} \mathbf{s}_{k-1}, \mathbf{D}_{k-1}(m + 2 : 2m)]$	2
SESOP-TN	$[\mathbf{d}_k^\ell, \mathbf{Q}_k(\mathbf{d}_k^\ell), \mathbf{d}_k^\ell - \mathbf{d}_k^{\ell-1}, \mathbf{D}_{k-1}(4 : m + 2)]$	6

TABLE 5.1 – Expression récursive des sous-espaces présentés dans le tableau 3.1 et coût de la recherche de pas MM multi-dimensionnelle en terme de produits matrice-vecteur nécessaires. Ici,  $\mathbf{D}(i : j)$  dénote la sous-matrice contenant les colonnes de  $i$  à  $j$  de  $\mathbf{D}$ .

## 5.5 Interprétation des résultats expérimentaux

Les expériences ont été menées avec Matlab 7 sur un PC P4 3.2GHz RAM 3Go. L'initialisation des algorithmes est systématiquement faite avec l'image nulle ( $\mathbf{x}_0 = \mathbf{0}$ ). On utilise comme critère d'arrêt pour tous les algorithmes comparés le seuil suivant sur la norme du gradient par pixel

$$\frac{\|\mathbf{g}_k\|}{\sqrt{n}} \leq 10^{-4}. \quad (5.28)$$

Pour chaque algorithme testé, nous indiquons le nombre d'itérations  $K$  et le temps  $T$  (en s.) nécessaires pour satisfaire le critère d'arrêt, sous la forme  $K/T$ .

### 5.5.1 Algorithme GCNL

**Recherche de pas de Wolfe.** Le tableau 5.2 présente les performances de l'algorithme GCNL associé à une recherche de pas basée sur la vérification des conditions de Wolfe stricte. Nous présentons uniquement les résultats obtenus pour  $c_1 = 0.01$  car nous avons observé que la valeur de  $c_1$  avait très peu d'influence sur les résultats. A l'opposé, les performances de l'algorithme dépendent de façon importante du paramètre  $c_2$ , correspondant à la condition de courbure (4.13). Cette condition est d'autant plus contraignante que  $c_2$  est petit. On s'attend donc à obtenir de meilleures performances pour le réglage  $c_2 = 0.5$ . C'est effectivement le cas pour les conjugaisons PRP, LS et FR où le nombre d'itérations croît avec  $c_2$ , de manière considérable. A l'inverse, pour les algorithmes GCNL-HS et GCNL-DY, il vaut mieux choisir  $c_2$  proche de un. Le meilleur choix des valeurs  $(c_1, c_2)$  correspond à un compromis qui n'est pas aisé à trouver et qui dépend de la règle de conjugaison de l'algorithme.

$(c_1, c_2)$	(0.01, 0.5)	(0.01, 0.9)	(0.01, 0.99)
GCNL-PRP	<b>79/268.3</b>	109/329.6	109/316
GCNL-HS	80/239.8	<b>71/227.2</b>	106/304.9
GCNL-LS	<b>80/243.1</b>	109/334.3	109/313.9
GCNL-FR	<b>164/521</b>	204/676.5	204/644
GCNL-DY	136/433.6	116/358.6	<b>116/333.3</b>

TABLE 5.2 – Comparaison entre différents algorithmes GCNL en termes de nombre d’itérations et de temps avant convergence  $K/T$ . Le pas résulte de la stratégie de (Moré et Thunente, 1994), satisfaisant la règle de Wolfe stricte de paramètres  $(c_1, c_2)$ .

**Recherche de pas MM.** Le tableau 5.3 présente les performances de l’algorithme GCNL associé à la recherche de pas MM quadratique non relaxée (i.e.,  $\theta = 1$ ). Les algorithmes GCNL obtiennent de meilleurs résultats en terme de temps d’optimisation lorsque  $J$  est très petit ( $J = 1$  ou  $J = 2$ ). Augmenter  $J$  n’apporte pas d’amélioration dans la mesure où le nombre d’itérations globales reste presque constant, voire même augmente légèrement. Cette dernière constatation est contre intuitive car augmenter  $J$  correspond à une minimisation plus précise du critère le long de la direction de GCNL. On s’attendrait donc à ce qu’une croissance de  $J$  s’accompagne d’une amélioration des performances de l’algorithme. Ce comportement, également souligné dans (Labat et Idier, 2008), n’est pas expliqué à ce jour.

Nous n’avons pas présenté les résultats obtenus pour un paramètre de relaxation  $\theta \neq 1$ . Nos conclusions sont identiques à celles de (Labat et Idier, 2008), à savoir que les meilleurs résultats sont obtenus pour  $\theta = 1$ . En particulier, une sur-relaxation de la recherche de pas n’entraîne pas d’amélioration en terme de temps d’optimisation ni de nombre d’itérations.

$J$	1	2	5	10
GCNL-PRP	<b>77/137</b>	69/139	75/202	77/273
GCNL-HS	<b>68/122</b>	67/134	75/191	77/289
GCNL-LS	82/149	<b>67/135</b>	74/190	76/266
GCNL-FR	<b>145/270</b>	137/279	143/379	143/515
GCNL-DY	234/447	<b>159/338</b>	144/387	143/516

TABLE 5.3 – Comparaison entre différents algorithmes GCNL en termes de nombre d’itérations et de temps avant convergence  $K/T$ . Le pas résulte de  $J$  itérations de récurrence MM quadratique.

**Influence de la formule de conjugaison.** Les tableaux 5.2 et 5.3 permettent également de comparer l’influence de la formule de conjugaison pour les algorithmes GCNL. Nous remarquons que les performances de l’algorithme GCNL varient de façon très importante en fonction du choix de la conjugaison. En particulier, un rapport de 4 en terme de nombre d’itérations est observé entre les conjugaisons *DY* et *HS*. Les meilleurs résultats sont obtenus pour la conjugaison de Hestenes et Siefel, pour les deux stratégies de pas. Un des avantages important de l’algorithme SMG est de s’affranchir de ce choix de formule de conjugaison.

### 5.5.2 Algorithme SMG

**Recherche de pas MM.** Le tableau 5.4 présente les performances de l’algorithme SMG associé à la recherche de pas MM quadratique multi-dimensionnelle non relaxée (i.e.,  $\theta = 1$ ). Les algorithmes SMG obtiennent de meilleurs résultats en terme de temps d’optimisation lorsque  $J$  est égal à un. Augmenter  $J$  entraîne un accroissement du coût de calcul par itération et n’améliore pas les performances des algorithmes, en terme de nombre d’itérations.

La recherche de pas de Newton s'avère très coûteuse en temps, malgré une troncature après 7 itérations. De plus, elle ne permet pas d'obtenir de meilleures performances de l'algorithme en terme de nombre d'itérations. Comme dans le cas scalaire, il est inutile et même contre productif de chercher à minimiser de façon précise la fonction  $f(\mathbf{s})$ .

**Taille de la mémoire.** Le tableau 5.4 permet également d'analyser l'influence de la mémoire  $m$  sur les performances de l'algorithme SMG. Les résultats pour  $m = 0$  ne sont pas présentés car l'algorithme s'identifie dans ce cas avec l'algorithme du gradient bien connu pour sa vitesse de convergence très lente (Pour cet exemple, plus de 300 itérations et plus de 500 s étaient nécessaires). Pour  $m \geq 1$ , le nombre d'itérations est presque constant. Les meilleurs résultats sont obtenus dès  $m = 1$ , c'est-à-dire pour le sous-espace de mémoire de gradient (MG).

SMG( $m$ )		1	2	5	10
Newton		76/578	75/630	76/701	74/886
MM ( $J$ )	1	<b>67/119</b>	68/125	67/140	67/163
	2	66/141	66/147	67/172	67/206
	5	74/211	72/225	71/255	72/323
	10	76/297	74/319	73/394	74/508

TABLE 5.4 – Comparaison entre recherche de pas MMQ et exactes, en termes de nombre d'itérations et de temps avant convergence  $K/T$  pour l'algorithme SMG.

### 5.5.3 Analyse comparée GCNL / SMG

Nous terminons cette étude par une comparaison des algorithmes GCNL et SMG, associés à leur meilleur réglage : Le pas résulte d'une itération MM quadratique et la taille de la mémoire est fixée à  $m = 1$ . Le tableau 5.5 présente les résultats obtenus pour l'image **peppers** ainsi que pour deux autres images **lena** et **boat**. A la fin du tableau, nous présentons les résultats obtenus pour l'algorithme MG sans utiliser l'expression récursive du sous-espace dans l'implémentation (MG+). Nous présentons également les performances des algorithmes ORTH, et SESOP associé à une mémoire  $m = 1$ .

	boat	lena	peppers
GCNL-PRP	40/74	55/99	77/137
GCNL-HS	39/71	50/93	68/122
GCNL-LS	42/81	57/103	82/149
GCNL-FR	77/141	98/179	145/270
GCNL-DY	86/161	127/240	234/447
MG	<b>37/67</b>	<b>47/85</b>	<b>67/119</b>
MG+	37/86	47/102	67/155
ORTH	149/330	234/524	391/847
SESOP( $m = 1$ )	37/88	47/109	67/156

TABLE 5.5 – Comparaison entre algorithmes MG et GCNL en termes de nombre d'itérations et de temps avant convergence  $K/T$ . Le pas est issu de  $J = 1$  itération MM quadratique.

Pour les trois problèmes de déconvolution, l'algorithme MG est le plus performant, aussi bien en terme de nombre d'itérations que de temps de calcul. Les résultats obtenus avec MG+ illustrent l'importance d'utiliser l'implémentation optimisée de la recherche de pas. Le cas échéant, l'algorithme de mémoire de gradient devient plus lent que le GCNL.

L'algorithme ORTH obtient des performances mauvaises, très similaires à celles de l'algorithme du gradient. Enfin, comme souligné dans (Narkiss et Zibulevsky, 2005), l'ajout des directions de Nemirovski dans l'algorithme MG, correspondant à SESOP, ne permet pas d'accélérer la convergence de l'algorithme. En effet, SESOP et MG nécessitent le même nombre d'itérations pour converger et, le sous espace MG étant de dimension plus faible, il conduit à un plus faible coût par itération.

Les algorithmes GCNL et MG possèdent une structure très proche lorsqu'ils sont associés à une itération de Majoration-Minoration. D'après l'expression du sous-espace de mémoire de gradient, une itération de l'algorithme MG s'écrit sous la forme

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k(-\mathbf{g}_k + \beta_k \mathbf{d}_{k-1}) \quad (5.29)$$

en posant

$$\alpha_k = \mathbf{s}_{1,k}, \quad \beta_k = \mathbf{s}_{2,k}/\mathbf{s}_{1,k}. \quad (5.30)$$

Lorsque le pas  $\mathbf{s}_k$  est issu d'une seule itération MM, alors on obtient les expressions analytiques :

$$\alpha_k = \frac{(\mathbf{d}_{k-1}^T \mathbf{A}_k \mathbf{d}_{k-1})(\mathbf{g}_k^T \mathbf{g}_k) - (\mathbf{g}_k^T \mathbf{A}_k \mathbf{d}_{k-1})(\mathbf{g}_k^T \mathbf{d}_{k-1})}{(\mathbf{d}_{k-1}^T \mathbf{A}_k \mathbf{d}_{k-1})(\mathbf{g}_k^T \mathbf{A}_k \mathbf{g}_k) - (\mathbf{g}_k^T \mathbf{A}_k \mathbf{d}_{k-1})^2},$$

et

$$\beta_k = \frac{(\mathbf{g}_k^T \mathbf{A}_k \mathbf{d}_{k-1})(\mathbf{g}_k^T \mathbf{g}_k) - (\mathbf{g}_k^T \mathbf{A}_k \mathbf{g}_k)(\mathbf{g}_k^T \mathbf{d}_{k-1})}{(\mathbf{d}_{k-1}^T \mathbf{A}_k \mathbf{d}_{k-1})(\mathbf{g}_k^T \mathbf{g}_k) - (\mathbf{g}_k^T \mathbf{A}_k \mathbf{d}_{k-1})(\mathbf{g}_k^T \mathbf{d}_{k-1})},$$

en notant  $\mathbf{A}_k = \mathbf{A}_k^0$ . Nous remarquons une proximité entre l'expression de  $\beta_k$  et celle de la conjugaison de (Daniel, 1967)

$$\beta_k^D = \frac{\mathbf{g}_k^T \mathbf{H}_k \mathbf{d}_{k-1}}{\mathbf{d}_{k-1}^T \mathbf{H}_k \mathbf{d}_{k-1}} \quad (5.31)$$

faisant intervenir le Hessien du critère. Les deux expressions sont identiques dès lors que  $\mathbf{A}_k = \mathbf{H}_k$  et que la condition de conjugaison  $\mathbf{g}_k^T \mathbf{d}_{k-1} = 0$  est vérifiée. D'autre part, l'analyse de (Cantrell, 1969) permet facilement de montrer que si le critère est quadratique et que  $\mathbf{A}_k = \mathbf{H}_k$ , les expressions de  $\alpha_k$  et  $\beta_k$  s'identifient avec celles utilisées dans l'algorithme de gradient conjugué.

## 5.6 Résultats complémentaires

Des tests additionnels ont été menés pour compléter notre analyse des performances de la recherche de pas MMQ  $rD$ . Afin d'alléger la lecture, nous présentons ici une synthèse des résultats obtenus et renvoyons le lecteur intéressé aux sections IV et V du rapport technique (Chouzenoux et al., 2010a).

### 5.6.1 Construction du sous-espace

Nous avons étudié le comportement de notre méthode de pas lorsque le sous-espace est construit à partir de la méthode GS, QNS ou SESOP-TN, toutes trois décrites dans le tableau 3.1. Les meilleures performances des algorithmes GS et QNS sont obtenues lorsqu'une seule itération de recherche de pas MMQ est effectuée. A l'inverse, pour le sous-espace SESOP-TN, un plus grand nombre d'itérations est nécessaire (typiquement,  $J = 10$ ). La recherche de pas de Newton proposée dans (Narkiss et Zibulevsky, 2005) conduit à de moins bons résultats que notre procédure, indépendamment du sous-espace utilisé.

Nous avons également analysé l'influence de la taille de la mémoire sur la vitesse de convergence de ces trois algorithmes. Pour QNS et SESOP-TN, les meilleurs résultats sont obtenus lorsqu'un faible nombre de directions est sauvegardé ( $m = 1$  ou  $2$ ). A l'opposé, l'algorithme GS obtient ses meilleures performances lorsque  $m \geq 15$ .

Enfin, parmi tous les algorithmes comparés, les plus faibles nombre d'itérations et temps de calcul sont obtenus avec l'algorithme de mémoire de gradient avec  $m = 1$  associé à  $J = 1$  itération de la recherche de pas MMQ.

### 5.6.2 Reconstruction dans un dictionnaire

Les performances de la méthode MMQ  $rD$  ont également été étudiées sur les trois problèmes de restauration d'image présentés dans (Zibulevsky et Elad, 2010). Ces problèmes sont résolus en utilisant l'approche par dictionnaire décrite dans la section 2.3.3.4. L'image est supposée s'écrire sous la forme  $\mathbf{x} = \mathbf{W}\mathbf{z}$  où  $\mathbf{z}$  est une décomposition parcimonieuse de  $\mathbf{x}$  dans le dictionnaire choisi  $\mathbf{W}$ . La reconstruction  $\mathbf{x}^*$  est définie comme  $\mathbf{W}\mathbf{z}^*$ , où  $\mathbf{z}^*$  minimise le critère

$$F(\mathbf{z}) = \|\mathbf{y} - \mathbf{K}\mathbf{W}\mathbf{z}\|^2 + \lambda \sum_i \phi(z_i) \quad (5.32)$$

avec  $\phi(u) = |u| - \delta \log(1 + |u|/\delta)$ , approximation différentiable de la valeur absolue introduisant un *a priori* de parcimonie sur les coefficients  $z_i$ .

Les algorithmes SESOP et PCD-SESOP sont comparés, en association avec la recherche de pas MMQ  $rD$  ou la recherche de pas de Newton. Les performances de SESOP et PCD-SESOP sont similaires en terme de vitesse de convergence et de temps de calcul. De plus, pour les deux sous-espaces, la méthode MMQ conduit à de meilleurs résultats que la méthode de Newton, grâce à un coût de calcul par itération plus faible.

## 5.7 Conclusion

Nous avons montré expérimentalement sur un problème de déconvolution d'image qu'il est largement préférable pour les algorithmes GCNL et SMG associés à la recherche de pas MMQ de ne pas effectuer un nombre trop grand de sous-itérations. Un constat similaire est obtenu pour les algorithmes GS, PCD-SESOP et QNS (Chouzenoux *et al.*, 2010a).

Lorsqu'elle est associée au sous-espace SMG, la stratégie MMQ  $rD$  s'avère plus performante que la stratégie de Newton utilisée dans (Narkiss et Zibulevsky, 2005; Zibulevsky, 2008; Zibulevsky et Elad, 2010). Les tests expérimentaux menés sur les sous-espaces GS, QNS, SESOP-TN et PCD-SESOP mènent à une conclusion identique (Chouzenoux *et al.*, 2010a). Alors que la recherche de pas de Newton ne s'accompagne d'aucun résultat de convergence, sauf dans le cas très particulier où le pas obtenu est le minimiseur exact du critère dans le sous-espace, la recherche de pas MMQ  $rD$  s'accompagne de résultats théoriques, et cela dès lors qu'une sous-itération est effectuée.

D'autre part, nous avons montré que l'algorithme de super mémoire de gradient montre de meilleures performances lorsque la taille de la mémoire est faible ( $m = 1$ ). En associant ce sous-espace à une recherche de pas correspondant à une seule itération MM, on obtient un algorithme entièrement analytique, dans le sens où il ne nécessite aucun réglage ni sous-itération, qui se démarque des autres algorithmes par son efficacité et sa simplicité. Cet algorithme nous semble être une référence à considérer systématiquement avant tout autre algorithme ou réglage.

# Chapitre 6

## Minimisation d'un critère présentant une barrière

### Sommaire

---

6.1	Introduction . . . . .	<b>78</b>
6.2	Fonction barrière . . . . .	<b>78</b>
6.2.1	Définition . . . . .	78
6.2.2	Fonctions barrières scalaires . . . . .	78
6.2.3	Critères pénalisés à barrière . . . . .	80
6.2.4	Fonction barrière en optimisation contrainte . . . . .	81
6.3	Minimisation d'un critère avec barrière . . . . .	<b>82</b>
6.3.1	Difficultés lors de la résolution du problème . . . . .	83
6.3.2	Restriction du pas . . . . .	84
6.3.3	Interpolation log-quadratique . . . . .	84
6.3.4	Formule de pas pour les fonctions auto-concordantes . . . . .	85
6.3.5	Pas MMQ pour les fonctions barrières . . . . .	87
6.4	Conclusion . . . . .	<b>88</b>

---

## 6.1 Introduction

Pour définir les recherches de pas MM scalaire ou multidimensionnelle quadratiques, nous avons supposé, par l'hypothèse 4.1, l'existence de fonctions quadratiques majorantes du critère à minimiser. Cependant, la résolution de certains problèmes nécessite la minimisation de critères à gradient non borné qui ne vérifient pas cette hypothèse.

Dans cette deuxième partie de la thèse, nous nous intéressons à la minimisation de critères contenant un terme de *barrière*, c'est-à-dire une fonction strictement convexe dont le gradient est de norme infinie aux bornes de son domaine de définition. Cette singularité implique que la restriction du critère à une droite peut présenter une ou plusieurs asymptotes verticales. Par conséquent, la majoration du critère par une parabole devient impossible. La contribution majeure de cette thèse consiste à développer une stratégie de recherche de pas MM pour les fonctions contenant un terme de barrière. Cette approche sera présentée dans le chapitre 7.

Ce chapitre 6 constitue une introduction à la présentation de la nouvelle stratégie de pas. La section 6.2 propose une définition formelle d'une fonction barrière. Nous montrons ensuite dans la section 6.3 qu'une attention particulière doit être portée lors de l'implémentation d'un algorithme de descente itérative pour minimiser un critère avec barrière, en particulier lors du calcul du pas.

## 6.2 Fonction barrière

### 6.2.1 Définition

Soit  $B$  une fonction définie sur  $\mathcal{D}_b$ . Cette fonction est dite *fonction barrière associée à  $\mathcal{D}_b$*  si les deux conditions suivantes sont vérifiées (Nesterov, 2003, Sec.4.2) :

- $B$  est convexe sur  $\mathcal{D}_b$
- Toute suite  $\{\mathbf{x}_k\}$  convergeant vers un point  $\bar{\mathbf{x}}$  appartenant à la frontière  $\partial\mathcal{D}_b$  de  $\mathcal{D}_b$  et de norme finie vérifie

$$\|\nabla B(\mathbf{x}_k)\| \rightarrow +\infty. \quad (6.1)$$

Si  $\mathcal{D}_b$  est ouvert,  $B$  n'est pas définie en  $\bar{\mathbf{x}}$  et la convexité de  $B$  implique que

$$B(\mathbf{x}_k) \rightarrow +\infty. \quad (6.2)$$

On dit alors que  $B$  est une *barrière stricte* associée à  $\mathcal{D}_b$ .

Le caractère non borné du gradient de  $B$  implique que ses minimiseurs appartiennent à l'ouvert  $\mathcal{C}$  défini par

$$\mathcal{C} \triangleq \mathcal{D}_b / \partial\mathcal{D}_b. \quad (6.3)$$

Cette propriété se transmet à tout critère contenant une fonction barrière associée à  $\mathcal{D}_b$ . Par exemple, les minimiseurs de  $F(\mathbf{x}) = F_0(\mathbf{x}) + \sum_{i=1}^n x_i \log x_i$  appartiennent à

$$\mathcal{C} = \{\mathbf{x} | x_i > 0, i = 1, \dots, n\}, \quad (6.4)$$

car  $B(\mathbf{x}) = \sum_{i=1}^n x_i \log x_i$  est une fonction barrière associée à l'orthant positif

$$\mathcal{D}_b = \{\mathbf{x} | x_i \geq 0, i = 1, \dots, n\}. \quad (6.5)$$

### 6.2.2 Fonctions barrières scalaires

Le tableau 6.1 et la figure 6.1 présentent quelques exemples de fonctions scalaires  $B(u)$  qui sont des barrières associées au domaine  $\mathcal{D}_b = [0, +\infty)$ . Les barrières logarithmique et inverse sont des barrières strictes. En effet, elles tendent vers l'infini en  $u = 0$ . À l'inverse, les barrières entropique et hyperbolique sont définies en 0 et valent 0.

À partir des fonctions du tableau 6.1, il est possible de construire une infinité de fonctions barrières en rajoutant un terme convexe à  $\psi$ . Une famille particulière de fonctions barrières s'écrit

$$B(u) = a\phi(u/a) \quad (6.6)$$

Nom	Barrière $B(u)$
Barrière logarithmique	$-\log u$
Barrière inverse	$1/u$
Barrière entropique	$u \log u$
Barrière hyperbolique	$-\sqrt{u}$

TABLE 6.1 – Exemples de fonctions barrières scalaires associées à l'orthant positif.

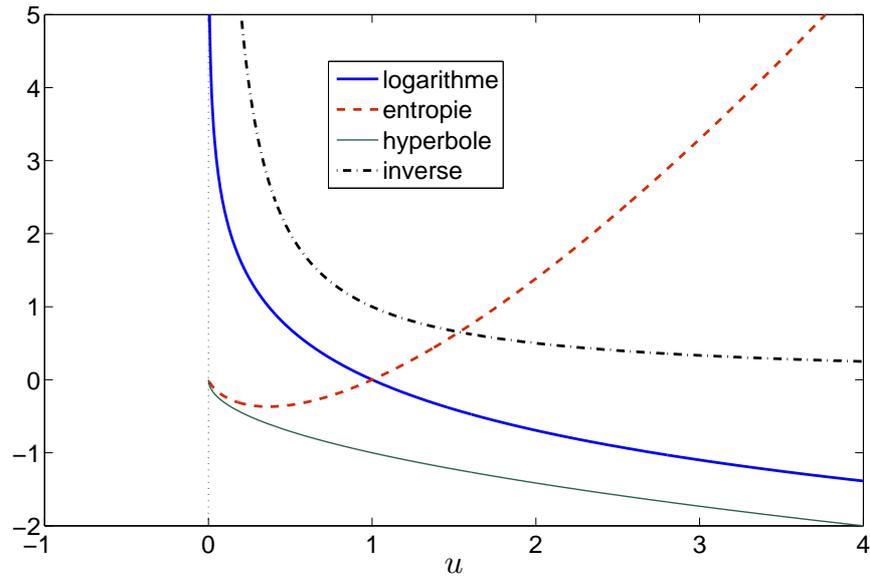


FIGURE 6.1 – Exemples de fonctions barrières scalaires

avec  $\phi : \mathbb{R}^+ \rightarrow \mathbb{R}$  convexe telle que

$$\lim_{u \rightarrow 0^+} \dot{\phi}(u) = -\infty \quad (6.7)$$

$$\phi(1) = \dot{\phi}(1) = 0 \quad (6.8)$$

$$\ddot{\phi}(1) > 0 \quad (6.9)$$

On dit que  $B$  est une  $\phi$ -divergence entre les variables  $u$  et  $a$  (Csiszàr, 1974). Le caractère non borné de la dérivée de  $\phi$  en  $0^+$  se transmet à  $B$ . En effet,

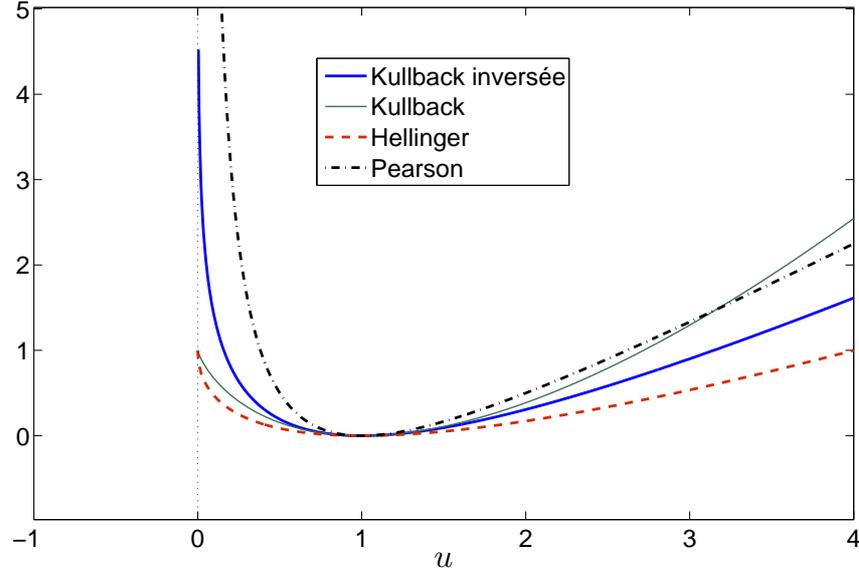
$$\dot{B}(u) = \dot{\phi}(u/a) \quad (6.10)$$

donc

$$\lim_{u \rightarrow 0^+} \dot{B}(u) = -\infty. \quad (6.11)$$

Par conséquent,  $B$  est une fonction barrière pour  $\mathcal{D}_b = [0, +\infty)$ . Nous présentons dans le tableau 6.2 et la figure 6.2 plusieurs choix possible pour  $\phi$  et la fonction  $B$  associée. Nous distinguons à nouveau des barrières strictes (Kullback inversée et Pearson) et des barrières non strictes (Kullback et Hellinger).

Nom	Fonction $\phi(t)$	$\phi$ -divergence $B(u)$
Kullback	$t \log t - t + 1$	$u \log u/a - u + a$
Kullback 'inversée'	$-\log t + t - 1$	$a \log a/u + u - a$
Hellinger	$(\sqrt{t} - 1)^2$	$(\sqrt{u} - \sqrt{a})^2$
Pearson	$t + t^{-1} - 2$	$(u - a)^2/u$

 TABLE 6.2 – Exemples de  $\phi$ -divergences entre  $u$  et  $a > 0$ .

 FIGURE 6.2 – Exemples de  $\phi$ -divergences entre  $u$  et  $a = 1$ .

### 6.2.3 Critères pénalisés à barrière

Dans le cadre de l'inversion pénalisée, de nombreux choix du terme d'adéquation aux données  $J(\mathbf{x})$  ou du terme de régularisation  $R(\mathbf{x})$  présentés dans le chapitre 2 s'expriment sous la forme

$$\sum_{i=1}^m \kappa_i \psi_i(C_i(\mathbf{x})), \quad \kappa_i > 0 \quad (6.12)$$

avec pour tout  $i$ ,  $\psi_i$  fonction barrière issue du tableau 6.1 ou  $\phi$ -divergence issue du tableau 6.2 associée au paramètre  $a_i > 0$ , et

$$C_i(\mathbf{x}) = \mathbf{c}_i^T \mathbf{x} + \rho_i, \quad i = 1, \dots, m, \quad (6.13)$$

avec  $\mathbf{c}_i \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ ,  $\rho_i \in \mathbb{R}$ . Toute fonction de la forme (6.12) étant une barrière associée au domaine

$$\mathcal{D}_b = \{\mathbf{x} \in \mathbb{R}^n | C_i(\mathbf{x}) \geq 0, \quad \forall i \in \{1, \dots, m\}\}, \quad (6.14)$$

il en résulte que tout minimiseur d'un critère pénalisé contenant un terme de cette forme appartient au domaine strict

$$\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^n | C_i(\mathbf{x}) > 0, \quad \forall i \in \{1, \dots, m\}\}. \quad (6.15)$$

Citons par exemple la fonction d'adéquation aux données associée à un modèle de bruit Poisson qui s'écrit

$$J(\mathbf{x}) = \sum_i (-y_i \log[\mathbf{K}\mathbf{x}]_i + [\mathbf{K}\mathbf{x}]_i). \quad (6.16)$$

correspondant à (6.12) avec  $\kappa_i = 1$ ,  $\psi_i(u) = a_i \log a_i/u + u - a_i$ ,  $a_i = y_i$  et  $C_i(\mathbf{x}) = [\mathbf{K}\mathbf{x}]_i$ . Si  $R$  est définie sur  $\mathbb{R}^n$  alors tout minimiseur de  $J + \lambda R$  vérifie  $\mathbf{K}\mathbf{x}^* > 0$ .

De même, la pénalisation entropique de Shannon

$$R(\mathbf{x}) = \sum_i x_i \log x_i \quad (6.17)$$

s'identifie avec (6.12) en posant  $\kappa_i = 1$ ,  $\psi_i(u) = u \log u$  et  $C_i(\mathbf{x}) = x_i$ . Par conséquent, les minimiseurs d'un critère pénalisé par (6.17) sont strictement positifs. Cette pénalisation sera utilisée pour la régularisation du problème de reconstruction traité dans le chapitre 9.

### 6.2.4 Fonction barrière en optimisation contrainte

Les propriétés des fonctions barrières sont utilisées dans la construction des algorithmes de points intérieurs pour l'optimisation contrainte. Supposons que l'on sache *a priori* que la solution d'un problème inverse vérifie les contraintes d'inégalités

$$C_i(\mathbf{x}) \geq 0, \quad i = 1, \dots, m. \quad (6.18)$$

Si le critère pénalisé contient un terme de la forme (6.12), ses minimiseurs *non contraints* vérifieront par construction les inégalités (6.18). Par contre, s'il n'y a pas de barrière dans le critère, il faut résoudre le problème *contraint*

$$\min_{\mathbf{x}} F(\mathbf{x}) \quad \text{s.c.} \quad C_i(\mathbf{x}) \geq 0, \quad i = 1, \dots, m \quad (6.19)$$

pour obtenir la solution souhaitée. Une stratégie algorithmique efficace, appelée méthode de *points intérieurs* (Fiacco et McCormick, 1967; Nesterov et Nemirovski, 1994; Roos *et al.*, 2006; Wright, 1991), consiste à rajouter artificiellement au critère le terme de barrière manquant afin de forcer la vérification de la contrainte. Le problème (6.19) est remplacé par une *séquence* de problèmes auxiliaires non contraints

$$\min_{\mathbf{x} \in \mathbb{R}^n} (F_{\mu_t}(\mathbf{x}) = F(\mathbf{x}) + \mu_t B(\mathbf{x})) \quad (6.20)$$

où  $B(\mathbf{x})$  est une fonction barrière associée au domaine réalisable

$$\mathcal{D}_b = \{\mathbf{x} \in \mathbb{R}^n \mid C_i(\mathbf{x}) \geq 0, \quad \forall i \in \{1, \dots, m\}\}. \quad (6.21)$$

Chaque sous-problème est paramétré par un coefficient  $\mu_t > 0$ . La résolution des problèmes augmentés (6.20) pour une suite donnée  $\{\mu_t\}$  définit la récurrence

$$\mathbf{x}_{t+1} = \arg \min F_{\mu_t}(\mathbf{x}). \quad (6.22)$$

La présence de la barrière implique que tous les itérés générés par (6.22) appartiennent au domaine contraint strict

$$\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^n \mid C_i(\mathbf{x}) > 0, \quad \forall i \in \{1, \dots, m\}\}. \quad (6.23)$$

La convergence de la séquence  $\{\mathbf{x}_t\}$  vers une solution de (6.19) dépend du choix de la barrière  $B$  et de la technique de mise à jour du paramètre  $\mu_t$ .

#### 6.2.4.1 Algorithme de barrière logarithmique

Le premier algorithme de points intérieurs est l'*algorithme de barrière logarithmique*, développé par (Fiacco et McCormick, 1967). La barrière utilisée est la fonction logarithmique

$$B(\mathbf{x}) = - \sum_{i=1}^m \log(C_i(\mathbf{x})), \quad (6.24)$$

et  $\{\mu_t\}$  est une suite de paramètres positifs tendant vers 0. La méthode de barrière logarithmique est principalement utilisée en programmation linéaire ( $F$  et  $C_i$  linéaires) (Roos *et al.*, 2006). La convergence de l'algorithme est établie dans (Fiacco et McCormick, 1967) sous l'hypothèse forte que la minimisation (6.20) est faite de façon exacte. En pratique, cette résolution est effectuée de façon approchée par une méthode de descente itérative. L'analyse de convergence de cette version inexacte est complexe et nécessite de se placer dans le cadre de fonctions  $F$  particulières dites *auto-concordantes* (Hertog, 1994; Nesterov et Nemirovski, 1994). L'algorithme de barrière logarithmique sera appliqué dans le chapitre 8 à un ensemble de problèmes de programmation quadratique contraints.

#### 6.2.4.2 Algorithme de barrière proximale

L'algorithme de barrière proximale (ou barrière adaptative) (Auslender et Haddou, 1995; Lange, 1994; Zibulevsky, 1996) est basé sur une récurrence MM. La fonction  $B$  est une fonction barrière pour le domaine  $\mathcal{D}_b$  délimité par les contraintes  $C_i$ . De plus, elle dépend du point courant  $\mathbf{x}_t$  et est construite de telle façon que la fonction augmentée

$$F_{\mu_t}(\mathbf{x}) = F(\mathbf{x}) + \mu_t B(\mathbf{x}, \mathbf{x}_t) \quad (6.25)$$

soit une approximation majorante de  $F$  au point courant  $\mathbf{x}_t$ . Les propriétés de convergence de l'algorithme dépendent du choix de  $\mu_t$ , de  $B(\mathbf{x}, \mathbf{y})$  et des contraintes  $C_i$ . D'après la définition d'une fonction majorante, la fonction  $B$  doit vérifier pour tout  $(\mathbf{x}, \mathbf{x}_t) \in \mathcal{C}$ ,

$$B(\mathbf{x}, \mathbf{x}_t) \geq 0, \quad (6.26)$$

$$B(\mathbf{x}_t, \mathbf{x}_t) = 0. \quad (6.27)$$

Il est donc naturel de choisir  $B$  s'écrivant comme une distance entre  $C_i(\mathbf{x})$  et  $C_i(\mathbf{x}_t)$ . Les propriétés de convergence de l'algorithme sont étudiées pour  $B$  distance de Kullback entre  $C_i(\mathbf{x})$  et  $C_i(\mathbf{x}_t)$  (Auslender et Haddou, 1995; Eggermont, 1990) et plus généralement pour  $B$   $\phi$ -divergence entre  $C_i(\mathbf{x})$  et  $C_i(\mathbf{x}_t)$  (Iusem *et al.*, 1994; Teboulle, 1997). Dans tous les cas, l'analyse de convergence s'applique uniquement au cas de contraintes linéaires. Par contre, contrairement à l'algorithme de barrière logarithmique, l'hypothèse que  $\mu_t$  tende vers 0 n'est pas requise. En particulier, les résultats de convergence sont valables si  $\mu_t$  est constant par rapport à  $t$ . Ce paramètre agit comme un coefficient de réglage influant sur la vitesse de convergence (Auslender et Haddou, 1995).

Nous soulignons que cet algorithme est qualifié de *proximal* car il s'identifie à l'algorithme proximal quadratique de (Rockafellar, 1970, 1976) à condition de remplacer la distance euclidienne par une distance faisant intervenir une barrière (Tseng, 2004). De plus, tout comme l'algorithme proximal quadratique est le dual de l'algorithme du Lagrangien augmenté, au sens de la dualité de Fenchel-Rockafellar (Rockafellar, 1973), l'algorithme de barrière proximale se dualise en la méthode des multiplicateurs exponentiels (Tseng et Bertsekas, 1993; Zibulevsky, 1996).

### 6.3 Minimisation d'un critère avec barrière

La résolution de nombreux problèmes inverses de traitement du signal nécessitent de minimiser un critère contenant un terme de barrière, soit parce qu'il est présent intrinséquement dans le critère pénalisé, soit parce qu'il est utilisé comme outil pour une résolution sous contraintes. Nous nous intéressons donc dans cette section au problème d'optimisation générique

$$\min_{\mathbf{x} \in \mathcal{D}_f} F(\mathbf{x}) = P(\mathbf{x}) + \mu B(\mathbf{x}), \quad \mu > 0 \quad (6.28)$$

où  $B$  s'écrit sous la forme (6.12) avec  $\psi_i$  une fonction barrière pour l'orthant positif. En outre, nous supposons que les contraintes  $C_i$  sont linéaires. Bien que la présence de la barrière contraigne les solutions de (6.28) à l'intérieur strict de  $\mathcal{D}_b$ , le problème (6.28) est un problème d'optimisation non contraint résolu classiquement par un algorithme de descente itérative. Cet algorithme présenté

dans le chapitre 2 alterne une étape de construction de la direction et une étape de recherche de pas le long de cette direction. A cause de la présence de la barrière dans le critère, des difficultés sont rencontrées lors de l'implémentation de l'algorithme de descente, en particulier lors du calcul du pas.

### 6.3.1 Difficultés lors de la résolution du problème

Deux problèmes majeurs ont été diagnostiqués lors de l'implémentation de l'algorithme de descente itérative pour résoudre (6.28) (Forsgren *et al.*, 2002, Sec.4), (Conn *et al.*, 1994b; Wright, 1995) :

1. Les itérés peuvent sortir du domaine de définition ou de différentiabilité du critère
2. La convergence est très lente

Dans les deux cas, l'algorithme doit être arrêté de façon prématurée, avant convergence.

La raison du premier problème est simple. Du fait de la présence de la fonction barrière, le critère à minimiser n'est différentiable que sur l'ensemble restreint  $\mathcal{C}$ . Partant d'un point  $\mathbf{x}$  et d'une direction  $\mathbf{d}$ , la stratégie de pas employée doit donc générer un pas qui assure que

$$C_i(\mathbf{x} + \alpha \mathbf{d}) > 0, \quad i = 1, \dots, m. \quad (6.29)$$

Lorsque le domaine associé à la barrière s'exprime sous la forme de contraintes linéaires

$$C_i(\mathbf{x}) = \mathbf{c}_i^T \mathbf{x} + \rho_i, \quad i = 1, \dots, m, \quad (6.30)$$

alors la contrainte (6.29) s'écrit  $\alpha \in (\alpha_-, \alpha_+)$  avec

$$\begin{cases} \alpha_- &= \max_{i|\delta_i > 0} -\frac{\theta_i}{\delta_i}, \\ \alpha_+ &= \min_{i|\delta_i < 0} -\frac{\theta_i}{\delta_i}, \end{cases} \quad (6.31)$$

en notant

$$\theta_i = \mathbf{c}_i^T \mathbf{x} + \rho_i, \quad (6.32)$$

$$\delta_i = \mathbf{c}_i^T \mathbf{d}. \quad (6.33)$$

Le premier problème est résolu dès lors que cette contrainte est prise en compte dans l'implémentation de la stratégie de pas.

Intéressons nous maintenant au second problème. La lenteur de la convergence peut être due à une stratégie de pas renvoyant une valeur trop éloignée du minimiseur de la fonction scalaire  $f(\alpha) = F(\mathbf{x} + \alpha \mathbf{d})$ . Une solution simple consiste à être plus exigeant dans la règle de pas choisie. Cependant, du fait de la présence de la barrière, la dérivée de la fonction  $f$  tend vers  $-\infty$  lorsque  $\alpha$  est égal à l'une des deux valeurs limites  $\alpha_-$  et  $\alpha_+$ . Par conséquent, une stratégie basée sur l'interpolation cubique ou quadratique de  $f$  sera peu efficace et un grand nombre d'itérations seront nécessaires à la vérification de la règle de pas.

Un autre motif courant de ralentissement de la convergence est le mauvais conditionnement du problème. Cela se traduit par des valeurs de pas extrêmement petites entraînant un blocage de l'algorithme avant la convergence. Le conditionnement du problème (6.28) dépend de la répartition des valeurs propres de  $\nabla^2 F$  et en particulier de  $\nabla^2 B$ . Le comportement des valeurs propres d'une fonction barrière est analysé de façon détaillée dans (Murray, 1971). En particulier, il est démontré que le Hessien d'une fonction barrière est très mal conditionné lorsque  $C_i(\mathbf{x})$  est proche de zéro pour un nombre d'indices  $i$  inférieur à  $m$ . Dans le cadre de l'optimisation par points intérieurs, plusieurs stratégies de préconditionnement ont été proposées pour pallier les difficultés liées à la présence de la barrière. Les techniques présentées dans (Nash et Sofer, 1993; Wright, 1994, 1998) se basent sur une décomposition du Hessien en blocs permettant de distinguer les contraintes proches de l'activation. Une technique plus simple et efficace en pratique consiste à approcher le Hessien par sa diagonale (Johnson et McGarry, 2003; Johnson *et al.*, 2000).

Nous présentons à présent les stratégies recensées dans la littérature pour calculer le pas lors de la minimisation d'un critère avec barrière. Partant d'un point  $\mathbf{x}$  et d'une direction  $\mathbf{d}$ , l'objectif de la recherche de pas est de déterminer  $\alpha$  qui minimise de façon approchée la fonction  $f(\alpha) = F(\mathbf{x} + \alpha\mathbf{d})$ . Dans le cadre de la résolution du problème (6.28), la recherche de pas doit fournir un pas appartenant à l'intervalle  $(\alpha_-, \alpha_+)$ . Nous faisons l'hypothèse que  $\mathbf{d}$  est une direction de descente. Par conséquent, la recherche de pas se limite à l'intervalle  $(0, \alpha_+)$ . Dans le cas contraire, les stratégies décrites se transposent simplement en posant

$$\begin{cases} \alpha & \leftarrow -\alpha \\ \mathbf{d} & \leftarrow -\mathbf{d} \end{cases} \quad (6.34)$$

### 6.3.2 Restriction du pas

Le vecteur  $\mathbf{d}$  est une direction de descente. Par conséquent, la singularité de la barrière implique l'existence d'un minimiseur local de  $f$  sur l'intervalle  $(0, \alpha_+)$ . La stratégie de restriction du pas, très courante dans les algorithmes de points intérieurs (Wright, 1991), consiste à définir le pas comme

$$\alpha = \theta\alpha_+, \quad (6.35)$$

avec  $\theta \in (0, 1)$  proche de 1. Ainsi, le critère est bien défini au nouvel itéré  $\mathbf{x} + \alpha\mathbf{d}$ . Cependant, cette approche simple n'assure pas la monotonie de l'algorithme. De plus, elle peut entraîner un 'blocage' de la suite d'itérés au bord du domaine  $\mathcal{C}$  (Murray et Wright, 1994). Une alternative permettant d'assurer la convergence est d'utiliser (6.35) comme pas initial de la stratégie de rebroussement décrite en table 4.1 (Segalat, 2002) afin d'obtenir un pas qui vérifie la règle d'Armijo .

### 6.3.3 Interpolation log-quadratique

Dans la section 4.2, nous avons présenté les stratégies de pas itératives classiques qui se basent sur une interpolation polynomiale de  $f$ . Cependant, ces méthodes ne sont pas efficaces de par le comportement singulier de  $f$  en  $\alpha_-$  et  $\alpha_+$ . Par conséquent, (Murray et Wright, 1994) proposent de remplacer l'interpolation polynomiale par

$$f(\alpha) \approx h(\alpha) = f_0 + f_1\alpha + f_2\alpha^2 - \mu \log(f_3 - \alpha), \quad (6.36)$$

afin de l'intégrer dans un schéma itératif de recherche de pas standard. Différentes méthodes de calcul des paramètres  $f_i$  de la fonction log-quadratique  $h$  sont proposées suivant les points de coïncidence requis entre  $f$  et  $h$

- 1)  $f(0)$ ,  $f(a)$  et  $\dot{f}(0)$ .
- 2)  $f(a)$ ,  $f(b)$  et  $f(c)$  avec  $c \in [a, b]$ .
- 3)  $f(a)$ ,  $f(b)$ ,  $\dot{f}(a)$  et  $\dot{f}(b)$ .
- 4)  $f(a)$ ,  $f(b)$ ,  $\dot{f}(a)$  et  $\dot{f}(a)$ .

Dans les cas 1) et 2), seulement trois équations lient  $f$  et la fonction interpolante. Par conséquent, le coefficient  $f_2$  est fixé à zéro. Dans les quatre cas, la détermination de  $f_3$  nécessite de résoudre une équation scalaire  $\phi(y) = 0$ . L'existence d'une solution est conditionnée par le respect d'une contrainte de la forme  $\psi(a, b) > 0$  ou  $\psi(a, b, c) > 0$  pour le cas 2). Lorsque cette contrainte est respectée, la racine de  $\phi$  est calculée de façon approchée par la méthode de Newton. Les autres coefficients sont ensuite obtenus par des formules analytiques. Dans le cas où  $\psi(a, b) \leq 0$ , l'interpolation log-quadratique n'est pas définie et une fonction interpolante classique de type polynomiale ou cubique doit être utilisée. Le tableau 6.3 indique pour chaque cas la forme de la fonction interpolante, la fonction  $\phi$  à annuler et la contrainte  $\psi$  à respecter.

L'article de (Murray et Wright, 1994) ne fournit pas d'algorithme de recherche de pas dans lequel intégrer les fonctions interpolantes. Dans (Nash et Sofer, 1993), la fonction interpolante issue du cas 3) est utilisée afin de déterminer un pas vérifiant la règle de Wolfe stricte. Si un tel pas n'est pas obtenu après 3 itérations, un schéma de rebroussement classique est utilisé, renvoyant un pas qui vérifie la règle d'Armijo. L'utilisation de ce schéma conduit à une accélération

Cas	Interpolation	Equation $\phi(y) = 0, y \in (0, 1)$	Contrainte $\psi(a, b, c) > 0$
1)	lin + log	$y + e^{-y}e^{-\kappa} - 1$	$f(a) - f(0) - a\dot{f}(0)$
2)	lin + log	$1 - y - \beta(1 - y/\nu)^\nu$	$(f(c) - f(a))/c - (f(b) - f(a))/b$
3)	quad + log	$\log y + (1/2)(1/y - y) - \kappa$	$(b/2)(\dot{f}(a) + \dot{f}(b)) - (f(b) - f(a))$
4)	quad + log	$\log y + 1 - y + (1/2)(1 - y)^2 + \kappa$	$f(b) - f(a) - b\dot{f}(a) - (1/2)b^2\ddot{f}(a)$

TABLE 6.3 – Récapitulatif des différentes fonctions interpolantes présentées dans (Murray et Wright, 1994). Lorsque l'interpolation est de la forme 'lin + log', le coefficient  $f_2$  est pris égal à 0.

significative en terme d'évaluation de critère lors de la recherche de pas dans l'algorithme de barrière logarithmique.

Les interpolations proposées par Murray *et coll.* ne nécessitent pas la connaissance de  $\alpha_+$ . Ceci est particulièrement intéressant dans le cas d'une fonction barrière associée à des contraintes non linéaires dont les valeurs limites ( $\alpha_-, \alpha_+$ ) ne s'obtiennent pas de façon analytique. Cependant, un inconvénient des fonctions interpolantes obtenues est que  $f_3$ , et par conséquent le minimiseur de  $h$ , peut être supérieur à  $\alpha_+$ . Lorsque ce problème se présente, le pas est diminué par rebroussement, jusqu'à vérifier la contrainte  $\mathbf{x} + \alpha\mathbf{d} \in \mathcal{C}$ .

Lorsque les contraintes sont linéaires, la valeur de  $\alpha_+$  est connue de façon exacte. Dans ce cas, une stratégie simple consiste à remplacer  $f_3$  par  $\alpha_+$  (Doyle, 2003; Lin, 2002). Dans les thèses de Lin et Doyle, un schéma complet basé sur la technique de dichotomie fournissant un pas satisfaisant la règle de Wolfe stricte est proposé. Les fonctions interpolantes utilisées sont de la forme

$$f(\alpha) \approx h(\alpha) = h_p(\alpha) + \mu h_b(\alpha) \quad (6.37)$$

avec  $h_p$  une fonction polynomiale d'ordre 1 ou 2 correspondant à l'interpolation de  $p(\alpha)$  et

$$h_b(\alpha) = f_0 + f_1\alpha + f_2\alpha^2 - f_3 \log(\alpha_+ - \alpha). \quad (6.38)$$

De même que dans (Murray et Wright, 1994), le calcul des  $f_i$  dépend des points de coïncidence requis entre  $f$  et  $h$ . Le calcul est cependant plus simple car les  $f_i$  sont obtenus de façon analytique, et aucune contrainte n'est faite sur les valeurs des pas.

### 6.3.4 Formule de pas pour les fonctions auto-concordantes

La notion de fonction auto-concordante, issue de l'ouvrage (Nesterov et Nemirovski, 1994), permet d'établir la convergence des algorithmes de points intérieurs dans leur version inexacte. Les propriétés sont établies sous l'hypothèse que la fonction augmentée est minimisée à l'aide d'un algorithme de Newton associé à une recherche de pas particulière que nous décrivons dans cette section.

#### 6.3.4.1 Fonction auto-concordante

La stratégie de pas développée par Nesterov *et coll.* suppose que  $F$  est une fonction *auto-concordante*, c'est-à-dire vérifie

- $F$  est une fonction strictement convexe définie sur un ouvert  $\mathcal{D}_f$  de  $\mathbb{R}^n$

- Pour tout couple  $(\mathbf{x}, \mathbf{d})$  de  $\mathcal{D}_f \times \mathbb{R}^n$ , la fonction  $f(\alpha) = F(\mathbf{x} + \alpha \mathbf{d})$  vérifie

$$|\ddot{f}(0)| \leq 2(\dot{f}(0))^{3/2} \quad (6.39)$$

L'auto-concordance est simple à vérifier pour les fonctions scalaires. La construction de fonctions auto-concordantes en dimension quelconque s'effectue en utilisant les propriétés de stabilité de l'auto-concordance par linéarité et composition, résumées par exemple dans (Roos *et al.*, 2006, Chap.3). Donnons quelques exemples de problèmes (6.28) dans lesquels  $F$  est auto-concordante. Dans tous les cas, nous supposons que  $P$  est quadratique convexe et que  $\mu = 1$ . Alors, les fonctions  $\psi_i$  suivantes assurent que  $F$  est auto-concordante :

- Barrière logarithmique  $-\log u$
- Somme d'une barrière logarithmique et d'une barrière entropique  $-\log u + u \log u$
- Divergence de Kullback inversée  $a \log a/u + u - a$

Notons que lorsque  $\mu$  est différent de 1, la propriété d'auto-concordance est vérifiée par  $(1/\mu)F$ . Les fonctions barrières entropiques, hyperboliques, et inverses, de même que les distances associées, ne vérifient pas la propriété d'auto-concordance.

### 6.3.4.2 Propriétés d'une fonction auto-concordante

Une fonction auto-concordante vérifie un grand nombre de propriétés. Nous en présentons uniquement un résumé, indispensable à la compréhension de la technique de pas. Nous renvoyons le lecteur à l'ouvrage de Nesterov et Nemirovskii ainsi qu'aux documents (Glineur, 2000; Hertog, 1994; Jarre, 1995; Roos *et al.*, 2006).

Soit  $F$  une fonction auto-concordante sur  $\mathcal{D}_f$ . Soit  $\mathbf{x} \in \mathcal{D}_f$ ,  $\mathbf{d} \in \mathbb{R}^n$  et un pas  $\alpha \in \mathbb{R}$ . Alors, nous avons les propriétés suivantes :

- (P1) Si  $\alpha$  est tel que  $\alpha^2 \ddot{f}(0) < 1$  alors

$$\mathbf{x} + \alpha \mathbf{d} \in \mathcal{D}_f \quad (6.40)$$

- (P2) Si  $\alpha$  est tel que  $\alpha^2 \ddot{f}(0) < r^2$  avec  $r \in [0, 1)$  alors

$$(1-r)^2 \ddot{f}(0) \leq \ddot{f}(\alpha) \leq \frac{1}{(1-r)^2} \ddot{f}(0). \quad (6.41)$$

- (P3) Si  $\alpha$  est tel que  $\mathbf{x} + \alpha \mathbf{d} \in \mathcal{D}_f$ , alors

$$f(\alpha) \geq f(0) + \alpha \left( \dot{f}(0) + [\ddot{f}(0)]^{1/2} \right) - \log \left( 1 + \alpha [\ddot{f}(0)]^{1/2} \right) \quad (6.42)$$

- (P4) Si  $\alpha$  est tel que  $\alpha^2 \ddot{f}(0) < 1$ , alors

$$f(\alpha) \leq f(0) + \alpha \left( \dot{f}(0) - [\ddot{f}(0)]^{1/2} \right) - \log \left( 1 - \alpha [\ddot{f}(0)]^{1/2} \right) \quad (6.43)$$

- (P5) Si  $\alpha$  est tel que  $0 < \alpha^2 \ddot{f}(0) < r^2$  avec  $r \in [0, 1)$  alors

$$\left( 1 - r + \frac{r^2}{3} \right) \ddot{f}(0) \leq \frac{\dot{f}(\alpha) - \dot{f}(0)}{\alpha} \leq \frac{1}{1-r} \ddot{f}(0). \quad (6.44)$$

### 6.3.4.3 Calcul du pas

Lorsque la fonction  $F$  du problème (6.28) est auto-concordante, la théorie de Nesterov *et coll.* permet de construire une stratégie de pas. Plus précisément, la propriété P1 assure que  $\alpha_+$  est supérieur à

$$\bar{\alpha} = \frac{1}{(\ddot{f}(0))^{1/2}} = [\ddot{f}(0)]^{-2}. \quad (6.45)$$

De plus, la propriété P3 implique que la fonction  $f$  est majorée sur  $[0, \bar{\alpha}]$  par

$$h(\alpha) = f(0) + \alpha (\dot{f}(0) - \bar{\alpha}^{-1}) - \log(1 - \alpha \bar{\alpha}^{-1}). \quad (6.46)$$

Le minimiseur de la fonction majorante  $h$  fournit la valeur de pas

$$\alpha = \frac{-\dot{f}(0)}{\bar{\alpha}^{-1}(\bar{\alpha}^{-1} - \dot{f}(0))}. \quad (6.47)$$

Lorsque la direction  $\mathbf{d}$  est la direction de Newton, on a la relation  $-\dot{f}(0) = \bar{\alpha}^{-2}$ . Par conséquence, la formule de pas se simplifie en

$$\alpha = \frac{1}{1 + \bar{\alpha}^{-1}}. \quad (6.48)$$

Les travaux de (Nesterov et Nemirovski, 1994) démontrent que cette valeur de pas assure la convergence de l'algorithme de Newton vers un minimiseur de  $F$ . La preuve de convergence est étendue dans (Ji *et al.*, 2007) au cas du gradient conjugué non linéaire avec la formule de conjugaison CD ('Conjugate Descent'). Le pas (6.47) correspond au résultat d'une itération MM appliquée à la fonction (6.46), tangente majorante de  $f$  en  $\alpha = 0$ . La forme de la majorante s'apparente, de plus, à la fonction interpolante (6.36).

### 6.3.5 Pas MMQ pour les fonctions barrières

La recherche de pas MMQ, présentée dans le chapitre 4, est caractérisée par sa simplicité d'implémentation et ses propriétés de convergences. Intéressons-nous à son application pour le problème (6.28). La stratégie MMQ 1D repose sur la minimisation de fonctions tangentes majorantes quadratiques  $h(\alpha, \alpha^j)$  de la fonction  $f(\alpha) = F(\mathbf{x} + \alpha \mathbf{d})$ . D'après (6.28),  $f$  se décompose sous la forme

$$f(\alpha) = P(\mathbf{x} + \alpha \mathbf{d}) + \mu B(\mathbf{x} + \alpha \mathbf{d}). \quad (6.49)$$

De plus, d'après l'expression (6.12) de  $B$ , s'il existe un  $i \in \{1, \dots, m\}$  tel que  $\psi_i$  soit une barrière stricte en  $u > 0$ , alors la fonction  $B$  est une barrière stricte. Pour  $(\mathbf{x}, \mathbf{d})$  donné, la fonction  $f(\alpha)$  est non bornée en  $\alpha$  tel que l'un des  $C_i(\mathbf{x} + \alpha \mathbf{d})$  s'annule. Par conséquent, quelle que soit la courbure de la fonction quadratique  $h(\cdot, \alpha^j)$  tangente à  $f$  en  $\alpha^j$ , il existe toujours des valeurs de  $\alpha$  telles que

$$f(\alpha) > h(\alpha, \alpha^j), \quad (6.50)$$

ce qui signifie que  $h$  n'est pas une fonction majorante de  $f$ . Ce problème est illustré dans l'exemple 6.1 traitant de la majoration de la fonction scalaire  $-\log u$ . Notons que la majoration de  $f$  sur la ligne de niveau

$$\ell_0 = \{\alpha \mid f(\alpha) \leq f(\alpha^j)\}, \quad (6.51)$$

est possible car la courbure de  $f$  est bornée sur  $\ell_0$ . Cependant, la construction de la majorante est très difficile car l'intervalle  $\ell_0$  n'est pas connu de façon explicite.

**Exemple 6.1.** *Considérons la fonction scalaire*

$$f(u) = -\log u. \quad (6.52)$$

*La dérivée seconde de  $f$  est  $1/u^2$ , fonction non bornée au voisinage de 0. Nous avons tracé en figure 6.3(a) plusieurs fonctions quadratiques tangentes à  $f$  en  $u = 0.5$ . Quelle que soit la valeur de la courbure choisie, la courbe de l'approximation quadratique croise la courbe de la fonction  $f$  en un  $\bar{u} > 0$ . Plus on augmente la courbure de la quadratique, plus  $\bar{u}$  se rapproche de 0. Lorsque la courbure est égale à l'infini, les deux courbes ne se croisent plus.*

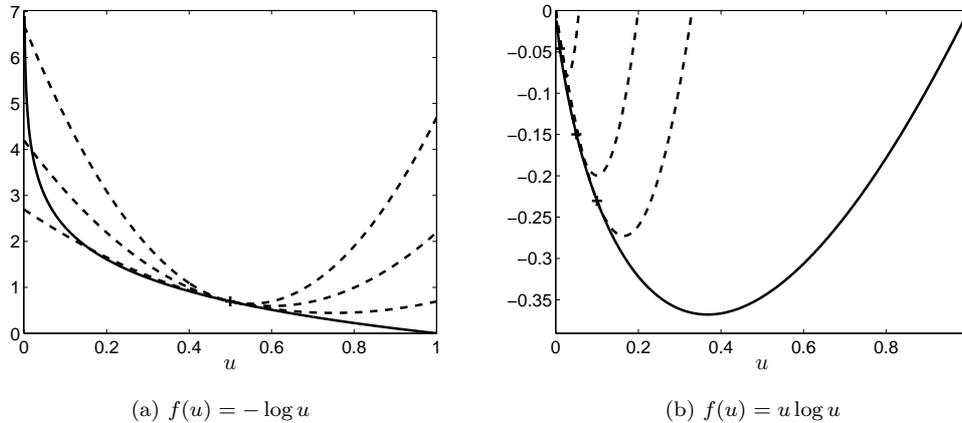


FIGURE 6.3 – Fonction barrière (trait continu) et approximations tangentes quadratiques (trait discontinu).

Si aucune des barrières  $\psi_i$  n'est stricte, alors  $B$  est également une fonction barrière non stricte. Par conséquent, la fonction objectif est bornée sur son domaine de définition. Pour  $(\mathbf{x}, \mathbf{d})$  donné, la fonction  $f(\alpha)$  est donc bornée. Il est possible de la majorer par une fonction quadratique en tout point  $\alpha^j$  tel que  $C_i(\mathbf{x} + \alpha^j \mathbf{d}) > 0$ . Cependant, si  $C_i(\mathbf{x} + \alpha^j \mathbf{d})$  est très proche de 0, la valeur de la courbure de la majorante peut être très importante et la fonction quadratique majorante sera une mauvaise approximation de  $f$ . La séquence  $\{\alpha^j\}$ , générée par la méthode MMQ 1D convergera alors très lentement vers le minimiseur de  $f$ . L'exemple 6.2 illustre ce phénomène lors de la majoration de la fonction scalaire  $u \log u$ .

**Exemple 6.2.** *Considérons la fonction scalaire*

$$f(u) = u \log u. \quad (6.53)$$

La dérivée seconde de  $f$  est  $1/u$ , fonction non bornée en 0. Cependant,  $f$  est finie en 0 et vaut 0. Il est donc possible de construire une tangente majorante quadratique de  $f$  en tout point  $u > 0$ . Nous avons tracé en figure 6.3(b) des fonctions quadratiques tangentes majorante de  $f$  en  $u = 0.01, 0.1, 0.5$ . A cause de la singularité de  $f$  en 0, la courbure minimale assurant la majoration est plus importante lorsque  $u$  se rapproche de 0. Le minimiseur d'une quadratique de courbure importante est très proche du point de tangence. La récurrence MM va avancer très lentement.

En conclusion, la recherche de pas MMQ 1D est difficile à appliquer à cause de la présence de la barrière. Nous proposons dans le chapitre 7 d'utiliser une nouvelle forme de fonction majorante, inspirée des approximations log-quadratiques (6.36), (6.37) et (6.46), afin de mieux approcher le critère avec barrière.

## 6.4 Conclusion

Dans de nombreux exemples, la résolution d'un problème inverse nécessite la minimisation d'un critère contenant une fonction barrière. Lorsqu'un algorithme de descente itérative est utilisé pour l'optimisation, la présence de la barrière nécessite l'adaptation des stratégies de pas classiques. Ce problème est souvent mis en avant dans les travaux traitant des algorithmes de points intérieurs (Wright, 1997). Cependant, la technique la plus utilisée dans ce contexte, forçant la restriction des itérés à l'intérieur du domaine réalisable, ne permet pas d'assurer la décroissance du critère et par conséquent la convergence. Les méthodes plus élaborées basées sur l'interpolation log-quadratique

ou la théorie des fonctions auto-concordantes sont rarement employées. La complexité de mise en œuvre de la première méthode et la théorie assez peu connue sur laquelle se base la deuxième sont deux sources probables de cet oubli. Nous avons montré précédemment que la stratégie MMQ permettait de construire des stratégies de pas simples et associées à des résultats de convergence forts. La présence de la barrière nécessite cependant de modifier cette technique afin de tenir compte du caractère non gradient borné du critère. Notre objectif est de définir un schéma de pas s'accompagnant des bonnes propriétés qui caractérisent la technique de pas MM quadratique.



# Chapitre 7

## Recherche de pas MM pour les fonctions avec barrières

### Sommaire

---

7.1	Introduction . . . . .	<b>92</b>
7.2	Description de la stratégie de pas . . . . .	<b>92</b>
7.2.1	Cadre de l'étude . . . . .	92
7.2.2	Fonction majorante log-quadratique . . . . .	92
7.2.3	Construction de la fonction majorante . . . . .	93
7.2.4	Expression du minimiseur . . . . .	98
7.3	Propriétés de convergence de la recherche de pas MMLQ 1D . . . . .	<b>99</b>
7.3.1	Rappel des hypothèses . . . . .	99
7.3.2	Propriétés de la séquence de pas . . . . .	100
7.3.3	Résultats de convergence . . . . .	102
7.4	Conclusion . . . . .	<b>104</b>

---

## 7.1 Introduction

Nous nous plaçons dans le cadre défini en section 6.3 correspondant à la minimisation d'un critère  $F$  de la forme  $P + \mu B$  où  $B$  est une barrière associée à un ensemble  $\mathcal{D}_b$  décrit par des contraintes d'inégalités linéaires. Ce problème non contraint se résout classiquement par une méthode de descente itérative alternant construction de la direction et recherche de pas le long de cette direction. Différentes stratégies de pas adaptées à la minimisation de critère contenant une barrière ont été proposées dans la littérature. Cependant, ces méthodes sont rarement employées notamment à cause de leur complexité de mise en œuvre.

La stratégie MM quadratique présentée dans le chapitre 4 conduit à une stratégie de pas simple et bénéficie de résultats de convergence forts. Cependant, la construction du pas MM quadratique nécessite de savoir construire des approximations majorantes quadratiques du critère scalaire  $f(\alpha) = F(\mathbf{x} + \alpha \mathbf{d})$ . Lorsque le critère contient une barrière, cette construction est difficile voire impossible. Nous proposons de modifier la technique de pas MM quadratique afin de tenir compte de la présence de la barrière du critère. La modification repose sur la définition d'une nouvelle forme de fonction majorante contenant elle-même un terme de barrière. Nous décrivons en section 7.2 la nouvelle fonction majorante log-quadratique et la stratégie de pas MMLQ 1D associée. La section 7.3 présente les propriétés de convergence de la recherche de pas MMLQ 1D. Pour des raisons de lisibilité, les démonstrations de convergence sont fournies en annexe A.

## 7.2 Description de la stratégie de pas

### 7.2.1 Cadre de l'étude

Nous nous intéressons à la minimisation du critère

$$F(\mathbf{x}) = P(\mathbf{x}) + \mu B(\mathbf{x}) \quad (7.1)$$

où  $B$  est la fonction barrière

$$B(\mathbf{x}) = \sum_{i=1}^m \psi_i(C_i(\mathbf{x})), \quad C_i(\mathbf{x}) = \mathbf{c}_i^T \mathbf{x} + \rho_i. \quad (7.2)$$

Lors de la minimisation de  $F$  par un algorithme de descente itérative, la stratégie de pas doit assurer que  $\mathbf{x}$  soit maintenu à l'intérieur du domaine

$$\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^n \mid C_i(\mathbf{x}) > 0, \quad \forall i \in \{1, \dots, m\}\}. \quad (7.3)$$

Nous proposons une nouvelle technique de recherche de pas MM adaptée à la présence de la barrière en définissant une nouvelle forme de fonction majorante.

### 7.2.2 Fonction majorante log-quadratique

Posons  $\mathbf{x} \in \mathcal{C}$ ,  $\mathbf{d}$  une direction de recherche non identiquement nulle. L'expression de  $B(\mathbf{x})$  implique que  $f(\alpha) = F(\mathbf{x} + \alpha \mathbf{d})$  présente une barrière  $\alpha_-$  et en  $\alpha_+$  définis par

$$\begin{cases} \alpha_- &= \max_{i \mid \delta_i > 0} -\frac{\theta_i}{\delta_i}, \\ \alpha_+ &= \min_{i \mid \delta_i < 0} -\frac{\theta_i}{\delta_i}, \end{cases} \quad \theta_i = \mathbf{c}_i^T \mathbf{x} + \rho_i, \quad \delta_i = \mathbf{c}_i^T \mathbf{d}. \quad (7.4)$$

Afin de prendre en compte cette singularité, nous proposons de modifier la méthode de pas MM en remplaçant la quadratique par une fonction majorante de la forme

$$h(\alpha) = h_0 + h_1 \alpha + h_2 \alpha^2 - h_3 \log(h_4 - \alpha) \quad (7.5)$$

définie sur  $(\alpha_-, \alpha_+)$ . Le pas sera généré par une récurrence MM, de la même façon que dans la stratégie de pas MM quadratique. Notons que, comme dans le cas quadratique, le minimiseur de  $h$  peut se calculer de façon explicite car il résulte de la résolution d'un polynôme de second ordre.

L'approximation (7.5) s'apparente aux approximations (6.36), (6.37) et (6.46) décrites au chapitre 6. L'originalité de notre approche est de choisir les paramètres  $h_i$  afin d'assurer la propriété de majoration en  $\alpha^j$ . Les égalités

$$\begin{cases} h(\alpha^j, \alpha^j) &= f(\alpha^j) \\ \dot{h}(\alpha^j, \alpha^j) &= \dot{f}(\alpha^j), \end{cases} \quad (7.6)$$

nous conduisent à choisir

$$\begin{aligned} h(\alpha, \alpha^j) &= f(\alpha^j) + (\alpha - \alpha^j)\dot{f}(\alpha^j) + \frac{1}{2}m^j(\alpha - \alpha^j)^2 \\ &+ \gamma^j \left[ (\bar{\alpha}^j - \alpha^j) \log\left(\frac{\bar{\alpha}^j - \alpha^j}{\bar{\alpha}^j - \alpha}\right) - \alpha + \alpha^j \right]. \end{aligned} \quad (7.7)$$

Cette formulation est inspirée de la barrière adaptative décrite dans (Hunter et L., 2004; Lange, 1994). Les expressions (7.7) et (7.5) sont équivalentes si l'on pose

$$h_0 = f(\alpha^j) - \alpha^j \dot{f}(\alpha^j) + \frac{1}{2}m^j(\alpha^j)^2 + \gamma^j(\bar{\alpha}^j - \alpha^j) \log(\bar{\alpha}^j - \alpha^j) + \gamma^j \alpha^j \quad (7.8)$$

$$h_1 = \dot{f}(\alpha^j) - m^j \alpha^j - \gamma^j \quad (7.9)$$

$$h_2 = \frac{1}{2}m^j \quad (7.10)$$

$$h_3 = \gamma^j(\bar{\alpha}^j - \alpha^j) \quad (7.11)$$

$$h_4 = \bar{\alpha}^j \quad (7.12)$$

Il reste désormais à choisir des paramètres  $\bar{\alpha}^j, m^j$  et  $\gamma^j$  permettant d'assurer la propriété de majoration

$$h(\alpha, \alpha^j) \geq f(\alpha), \quad \forall \alpha \in (\alpha_-, \alpha_+). \quad (7.13)$$

### 7.2.3 Construction de la fonction majorante

Nous décrivons dans cette section une technique de construction de  $h(\cdot, \alpha^j)$ , tangente majorante de  $f$  en  $\alpha^j \in (\alpha_-, \alpha_+)$ . Notre méthode nécessite de supposer que les fonctions  $\psi_i$  composant la barrière  $B$  vérifient

$$-\frac{2}{u}\ddot{\psi}_i(u) \leq \ddot{\psi}_i(u) \leq 0, \quad \forall u > 0. \quad (7.14)$$

Cette hypothèse est vérifiée par les barrières logarithmiques, entropiques et hyperboliques. Par contre, elle exclut la barrière inverse dont la majoration reste un problème ouvert. Afin de nous concentrer sur le terme de barrière, nous supposons en outre que nous savons construire des fonctions majorantes quadratiques du premier terme  $P(\mathbf{x} + \alpha \mathbf{d})$ . En d'autres termes, nous supposons que l'hypothèse suivante est vérifiée.

**Hypothèse 7.1.** Soit  $\mathcal{V}$  un ensemble ouvert, borné, inclus dans  $\mathcal{D}_p$ . Pour tout  $\mathbf{x}' \in \mathcal{V}$ , il existe une matrice symétrique définie positive  $\mathbf{A}(\mathbf{x}')$  telle que pour tout  $\mathbf{x} \in \mathcal{V}$ ,

$$H(\mathbf{x}, \mathbf{x}') = P(\mathbf{x}') + (\mathbf{x} - \mathbf{x}')^T \nabla P(\mathbf{x}') + \frac{1}{2}(\mathbf{x} - \mathbf{x}')^T \mathbf{A}(\mathbf{x}')(\mathbf{x} - \mathbf{x}') \quad (7.15)$$

est une approximation tangente majorante de  $P$  en  $\mathbf{x}'$ . De plus, l'ensemble  $\{\mathbf{A}(\mathbf{x}) | \mathbf{x} \in \mathcal{V}\}$  est à spectre borné strictement positif avec les bornes  $(\nu_{\min}^{\mathbf{A}}, \nu_{\max}^{\mathbf{A}})$ , i.e. pour tout  $\mathbf{x} \in \mathcal{V}$ ,

$$0 < \nu_{\min}^{\mathbf{A}} \leq \frac{\mathbf{v}^T \mathbf{A}(\mathbf{x}) \mathbf{v}}{\|\mathbf{v}\|^2} \leq \nu_{\max}^{\mathbf{A}}, \quad \forall \mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}\} \quad (7.16)$$

La fonction  $f$  s'écrivant  $p + \mu b$  avec

$$p(\alpha) = P(\mathbf{x} + \alpha \mathbf{d}), \quad b(\alpha) = B(\mathbf{x} + \alpha \mathbf{d}), \quad (7.17)$$

nous proposons de construire la majorante sous la forme

$$h(\alpha, \alpha^j) = h_p(\alpha, \alpha^j) + \mu h_b(\alpha, \alpha^j) \quad (7.18)$$

où  $h_p$  et  $h_b$  majorent respectivement  $p$  et  $b$ .

D'après l'hypothèse 4.1, une fonction majorante de  $p$  en  $\alpha^j$  est

$$h_p(\alpha, \alpha^j) = p(\alpha^j) + (\alpha - \alpha^j)\dot{p}(\alpha^j) + \frac{1}{2}m_p^j(\alpha - \alpha^j)^2 \quad (7.19)$$

avec la courbure  $m_p^j = \mathbf{d}^T \mathbf{A}(\mathbf{x} + \alpha^j \mathbf{d}) \mathbf{d}$ .

Nous devons désormais construire la fonction  $h_b$ . Nous proposons de définir cette fonction par morceaux, sur les intervalles  $(\alpha_-, \alpha^j]$  et  $[\alpha^j, \alpha_+)$ . Sur chacun des intervalles, la majorante sera de la forme

$$\begin{aligned} h_b(\alpha, \alpha^j) &= b(\alpha^j) + (\alpha - \alpha^j)\dot{b}(\alpha^j) + \frac{1}{2}m_b^j(\alpha - \alpha^j)^2 \\ &+ \gamma_b^j \left[ (\bar{\alpha}^j - \alpha^j) \log \left( \frac{\bar{\alpha}^j - \alpha^j}{\bar{\alpha}^j - \alpha} \right) - \alpha + \alpha^j \right], \end{aligned} \quad (7.20)$$

où la valeur de  $\bar{\alpha}^j$  correspond à la position de la barrière sur l'intervalle courant :

$$\bar{\alpha}^j = \alpha_- \quad \forall \alpha \in (\alpha_-, \alpha^j] \quad (7.21)$$

$$\bar{\alpha}^j = \alpha_+ \quad \forall \alpha \in [\alpha^j, \alpha_+) \quad (7.22)$$

La majorante globale  $h(\cdot, \alpha^j)$ , schématisée en figure 7.1 s'écrit alors sous la forme (7.7) avec

$$\begin{aligned} m^j &= m_p^j + \mu m_b^j \\ \gamma^j &= \mu m_b^j \end{aligned}$$

et  $\bar{\alpha}^j$  égal à  $\alpha_-$  ou  $\alpha_+$ , suivant la valeur de  $\alpha$ .

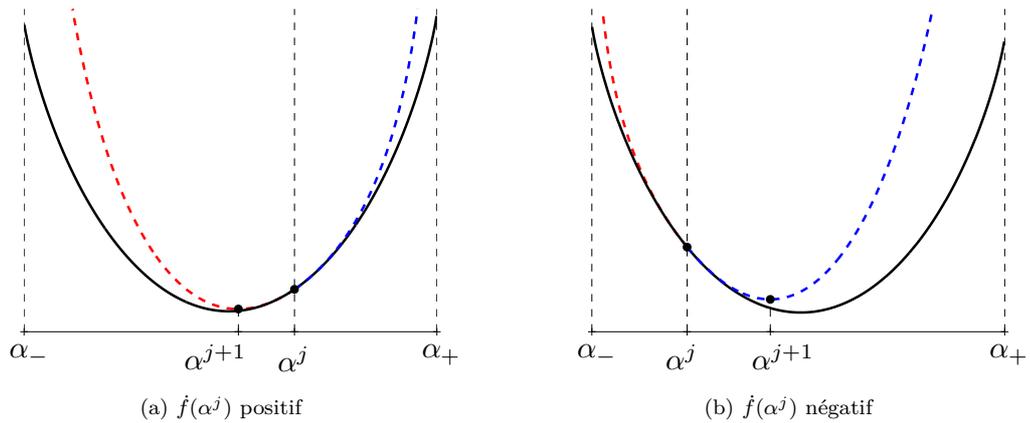


FIGURE 7.1 – Principe de la recherche de pas MMLQ pour un critère à barrière non stricte. La fonction tangente majorante  $h(\cdot, \alpha^j)$  (pointillés) pour  $f$  (ligne continue) en  $\alpha_j$  est définie par morceaux sur  $(\alpha_-, \alpha_j]$  et  $[\alpha_j, \alpha_+)$ . Son minimiseur fournit la prochaine itérée  $\alpha_{j+1}$ . Il s'obtient en minimisant la portion gauche (Cas (a)) ou droite (Cas (b)) suivant le signe de  $\dot{f}(\alpha^j)$

**Remarque 7.1.** Lorsque l'un des pas  $\alpha_-$  ou  $\alpha_+$  est infini, le coefficient  $\gamma^j$  n'est pas défini et la portion de majorante correspondant est quadratique de courbure  $m^j$ . Il peut arriver qu'à la fois  $\alpha_-$  et  $\alpha_+$  soient infinis. Alors, la fonction majorante  $h$  est quadratique sur  $\mathbb{R}$ .

La position de la barrière  $\bar{\alpha}^j$  étant fixée, il nous reste à déterminer les coefficients  $(m_b^j, \gamma_b^j)$ . La stratégie de détermination de ces coefficients que nous proposons est basée sur l'hypothèse que

$$-\frac{2}{u}\ddot{\psi}_i(u) \leq \ddot{\psi}_i(u) \leq 0, \quad \forall i = 1, \dots, m. \quad (7.23)$$

Nous décomposons tout d'abord la fonction  $b$  comme la somme de trois termes

$$b(\alpha) = b_1(\alpha) + b_2(\alpha) + \text{cste} \quad (7.24)$$

avec

$$b_1(\alpha) = \sum_{i|\delta_i > 0} \psi_i(\theta_i + \alpha\delta_i)$$

$$b_2(\alpha) = \sum_{i|\delta_i < 0} \psi_i(\theta_i + \alpha\delta_i)$$

Les stratégies employées pour majorer chacun de ces termes seront différentes. Nous présentons de façon détaillée la construction de la majorante sur la portion  $\alpha \in [\alpha^j, \alpha_+)$ . Le cas de l'intervalle  $[\alpha_-, \alpha_j]$  se déduit par symétrie axiale autour de  $\alpha^j$ .

### 7.2.3.1 Majoration de $b_1$ sur $[\alpha^j, \alpha_+)$

La fonction  $b_1$  résulte de la somme de

$$\begin{aligned} \psi_i(\theta_i + \alpha\delta_i) &= \psi_i\left(\delta_i\left(\alpha + \frac{\theta_i}{\delta_i}\right)\right) \\ &= \psi_i(\delta_i(\alpha - t_i)) \end{aligned}$$

en posant  $t_i = -\frac{\theta_i}{\delta_i} \leq \alpha_-$ , où  $\delta_i > 0$ . La forme de ces fonctions est illustrée sur la figure 7.2. Elles possèdent une barrière en  $t_i$  situé pour tout  $i$  hors de l'intervalle  $[\alpha^j, \alpha_+)$ . De plus, leur dérivée seconde est strictement décroissante. Par conséquent, la courbure de la fonction  $b_1$  résultante est bornée sur  $[\alpha^j, \alpha_+)$ , égale à  $\ddot{b}_1(\alpha^j)$ . Nous proposons donc de majorer la fonction  $b_1$  en  $\alpha^j$  par la quadratique

$$h_{b,1}(\alpha, \alpha^j) = b_1(\alpha^j) + (\alpha - \alpha^j)\dot{b}_1(\alpha^j) + \frac{1}{2}m_b^j(\alpha - \alpha^j)^2, \quad (7.25)$$

avec

$$m_b^j = \ddot{b}_1(\alpha^j) \quad (7.26)$$

$$= \sum_{i|\delta_i > 0} \delta_i^2 \ddot{\psi}_i(\theta_i + \alpha^j \delta_i). \quad (7.27)$$

### 7.2.3.2 Majoration de $b_2$ sur $[\alpha^j, \alpha_+)$

La fonction  $b_2$  résulte de la somme de

$$\begin{aligned} \psi_i(\theta_i + \alpha\delta_i) &= \psi_i\left(\delta_i\left(\alpha - \frac{\theta_i}{\delta_i}\right)\right) \\ &= \psi_i(\delta_i(\alpha - t_i)) \\ &= \phi_i(\alpha) \end{aligned}$$

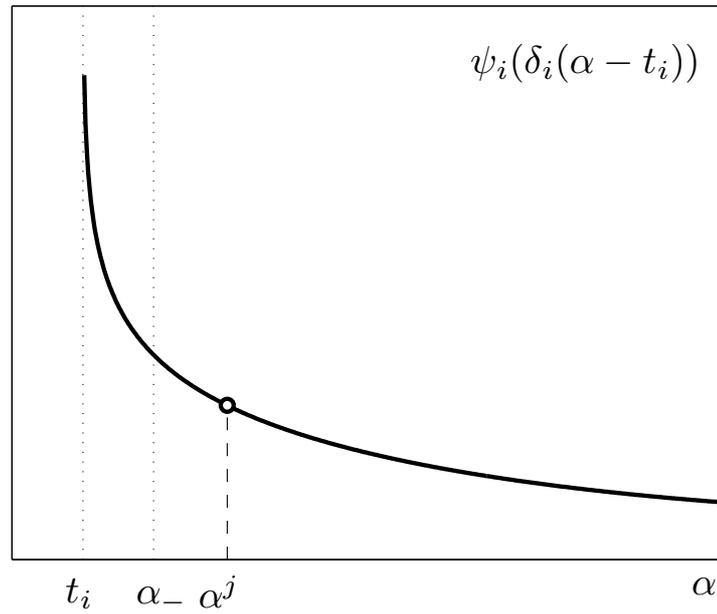


FIGURE 7.2 – Forme des termes composants la fonction  $b_1$ . Nous devons les majorer en  $\alpha^j$  sur  $[\alpha^j, \alpha_+)$ . Les fonctions ont une courbure décroissante sur cet intervalle.

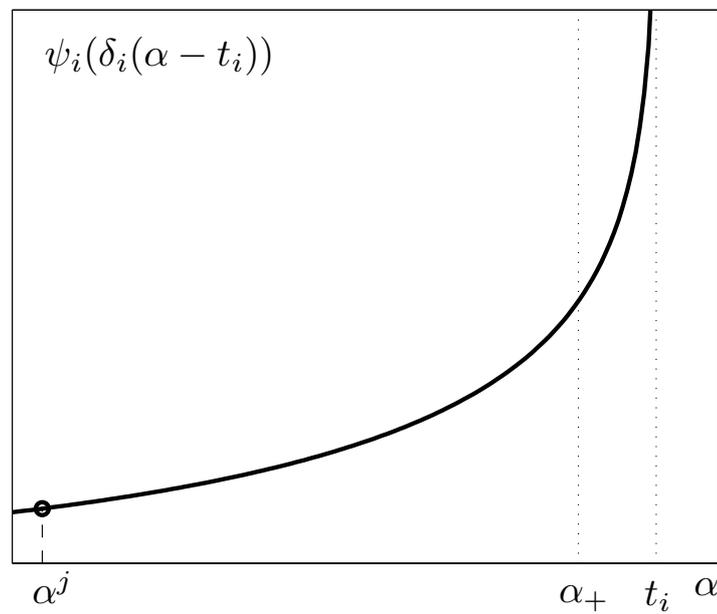


FIGURE 7.3 – Forme des termes composants la fonction  $b_2$ . Nous devons les majorer en  $\alpha^j$  sur  $[\alpha^j, \alpha_+)$ . Les fonctions ont une courbure croissante sur cet intervalle. De plus, lorsque  $i$  est tel que  $t_i = \alpha_+$ , cette courbure est non bornée.

en posant  $t_i = -\frac{\delta_i}{\theta_i} \leq \alpha_+$ , où  $\delta_i < 0$ . La forme de ces fonctions est illustrée sur la figure 7.3. A cause du signe négatif de  $\delta_i$ , ces fonctions possèdent une dérivée seconde strictement croissante et la stratégie utilisée pour majorer  $b_1$  ne peut pas être appliquée ici.

Afin de tenir compte de la barrière en  $t_i$ , nous proposons de majorer  $b_2$  par la fonction log-linéaire

$$h_{b,2}(\alpha, \alpha^j) = b_2(\alpha^j) + (\alpha - \alpha^j)\dot{b}_2(\alpha^j) + \gamma_b^j \left[ (\alpha_+ - \alpha^j) \log \frac{\alpha_+ - \alpha^j}{\alpha_+ - \alpha} + \alpha^j - \alpha \right], \quad (7.28)$$

le coefficient  $\gamma_b^j$  étant choisi afin d'assurer que  $h_{b,2}$  majore  $b_2$  sur l'intervalle  $[\alpha^j, \alpha_+)$ . La stratégie de détermination du coefficient  $\gamma_b^j$  que nous proposons est basée sur la relation (7.23). Elle nous permet en effet de montrer que

$$\gamma_b^j = (\alpha_+ - \alpha^j)\ddot{b}_2(\alpha^j) \quad (7.29)$$

$$= (\alpha_+ - \alpha^j) \sum_{i|\delta_i < 0} \delta_i^2 \ddot{\psi}_i(\theta_i + \alpha^j \delta_i) \quad (7.30)$$

assure que  $h_{b,2}(\cdot, \alpha^j)$  est une fonction tangente majorante de  $b_2$  en  $\alpha^j$ . Les outils utilisés dans notre démonstration sont inspirés de l'article (Erdogan et Fessler, 1999). Définissons les fonctions

$$\chi(\alpha) = \dot{b}_2(\alpha)(\alpha_+ - \alpha), \quad (7.31)$$

$$l(\alpha) = \dot{b}_2(\alpha^j)(\alpha_+ - \alpha) + \gamma_b^j(\alpha - \alpha^j). \quad (7.32)$$

Etant donnée l'expression de la majorante, la fonction linéaire  $l$  s'écrit aussi

$$l(\alpha) = \dot{h}_{b,2}(\alpha, \alpha^j)(\alpha_+ - \alpha). \quad (7.33)$$

Par conséquent, on a

$$l(\alpha^j) = \chi(\alpha^j). \quad (7.34)$$

De plus, d'après l'expression de  $\gamma_b^j$ ,

$$\dot{l}(\alpha^j) = \dot{\chi}(\alpha^j). \quad (7.35)$$

D'autre part,

$$\ddot{\chi}(\alpha) = \ddot{b}_2(\alpha)(\alpha_+ - \alpha) - 2\ddot{b}_2(\alpha) \quad (7.36)$$

$$= \sum_{i|\delta_i < 0} \delta_i^3 \ddot{\psi}_i(\theta_i + \alpha \delta_i)(\alpha_+ - \alpha) - 2\delta_i^2 \ddot{\psi}_i(\theta_i + \alpha \delta_i) \quad (7.37)$$

D'après la définition de  $\alpha_+$ ,

$$(\alpha_+ - \alpha) < -(\theta_i + \alpha \delta_i)/\delta_i, \quad \forall i \text{ tel que } \delta_i < 0 \quad (7.38)$$

Les inégalités (7.23) et (7.38) entraînent

$$\ddot{\chi}(\alpha) < \sum_{i|\delta_i < 0} \delta_i^2 [-\ddot{\psi}_i(\theta_i + \alpha \delta_i)(\theta_i + \alpha \delta_i) - 2\ddot{\psi}_i(\theta_i + \alpha \delta_i)] < 0$$

Donc  $\chi$  est concave. La fonction linéaire  $l$  étant tangente à  $\chi$ , on a

$$l(\alpha) \geq t(\alpha), \quad \forall \alpha \in [\alpha_j, \alpha^+]. \quad (7.39)$$

Puisque  $\alpha_+ > \alpha$ , (7.39) s'écrit aussi

$$\dot{h}_{b,2}(\alpha, \alpha^j) \geq \dot{b}_2(\alpha), \quad \forall \alpha \in [\alpha_j, \alpha^+]. \quad (7.40)$$

Par conséquent,  $h_{b,2}(\cdot, \alpha^j)$  majore  $b_2$  sur  $[\alpha_j; \alpha^+)$ .

### 7.2.3.3 Conclusion

La fonction  $h_b(\cdot, \alpha^j) = h_{b,1}(\cdot, \alpha^j) + h_{b,2}(\cdot, \alpha^j)$  majore  $b$  sur  $\alpha \geq \alpha_j$ . En utilisant les mêmes arguments, nous établissons l'expression de coefficients  $(m_b^j, \gamma_b^j)$  assurant la majoration de  $b$  par  $h_b$  sur l'intervalle  $(\alpha_-, \alpha^j]$  :

$$m_b^j = \ddot{b}_2(\alpha^j), \quad (7.41)$$

$$\gamma_b^j = (\alpha_- - \alpha^j) \ddot{b}_1(\alpha^j). \quad (7.42)$$

La forme globale de la majorante de  $f$  résultante est résumée dans le théorème 7.1.

**Théorème 7.1.** Soit  $F = P + \mu B$  où

$$B(\mathbf{x}) = \sum_{i=1}^m \psi_i(C_i(\mathbf{x})), \quad C_i(\mathbf{x}) = \mathbf{c}_i^T \mathbf{x} + \rho_i. \quad (7.43)$$

avec

$$-\frac{2}{u} \ddot{\psi}_i(u) \leq \ddot{\psi}_i(u) \leq 0, \quad \forall i = 1, \dots, m, \quad \forall u > 0 \quad (7.44)$$

et  $P$  est une fonction vérifiant l'hypothèse 4.1. Soit  $(\mathbf{x}, \mathbf{d})$  deux vecteurs. Notons

$$\begin{cases} \alpha_- = \max_{i|\delta_i > 0} -\frac{\theta_i}{\delta_i}, \\ \alpha_+ = \min_{i|\delta_i < 0} -\frac{\theta_i}{\delta_i}, \end{cases}, \quad \theta_i = \mathbf{c}_i^T \mathbf{x} + \rho_i, \quad \delta_i = \mathbf{c}_i^T \mathbf{d}. \quad (7.45)$$

la position des barrières de  $f$ . Alors, la fonction (7.7) de paramètres

$$\begin{cases} \bar{\alpha}^j = \alpha_- \\ m^j = m_p^j + \mu \sum_{i|\delta_i < 0} \phi_i(\alpha^j) \\ \gamma^j = \mu(\alpha_- - \alpha^j) \sum_{i|\delta_i > 0} \phi_i(\alpha^j) \end{cases} \quad \forall \alpha \in (\alpha_-; \alpha^j] \quad (7.46)$$

et

$$\begin{cases} \bar{\alpha}^j = \alpha_+ \\ m^j = m_p^j + \mu \sum_{i|\delta_i > 0} \phi_i(\alpha^j) \\ \gamma^j = \mu(\alpha_+ - \alpha^j) \sum_{i|\delta_i < 0} \phi_i(\alpha^j) \end{cases} \quad \forall \alpha \in [\alpha^j; \alpha_+) \quad (7.47)$$

où

$$\phi_i(\alpha) = \delta_i^2 \ddot{\psi}_i(\theta_i + \alpha \delta_i), \quad m_p^j = \mathbf{d}^T \mathbf{A}(\mathbf{x} + \alpha^j \mathbf{d}) \mathbf{d} \quad (7.48)$$

est une fonction tangente majorante de  $f(\alpha) = F(\mathbf{x} + \alpha \mathbf{d})$  en  $\alpha^j$ .

## 7.2.4 Expression du minimiseur

Le pas  $\alpha_k$  est obtenu par la récurrence MMLQ

$$\begin{cases} \alpha_k^0 = 0, \\ \alpha_k^{j+1} = \arg \min_{\alpha} h(\alpha, \alpha_k^j), \quad j < J, \\ \alpha_k = \alpha_k^J. \end{cases} \quad (7.49)$$

Cette récurrence nécessite de calculer le minimiseur de  $h(\alpha, \alpha^j)$  pour  $j \in \{1, \dots, J\}$ . Le lemme 7.1 assure la stricte convexité de la fonction majorante.

**Lemme 7.1.** *La fonction  $h(\cdot, \alpha^j)$  est  $C^2$  et strictement convexe.*

*Démonstration.* Tout d'abord,  $h(\cdot, \alpha^j)$  est  $C^\infty$  sur  $(\alpha_-; \alpha^j)$  et  $(\alpha^j; \alpha_+)$ . De plus, les expressions (7.46) et (7.47) assurent la continuité de  $h$  et des dérivées premières et secondes en  $\alpha^j$ . Donc,  $h(\cdot, \alpha^j)$  est  $C^2$  sur  $(\alpha_-; \alpha_+)$ . D'après (7.46) et (7.47), la dérivée seconde de  $h(\cdot, \alpha^j)$  est

$$\ddot{h}^j(\alpha, \alpha^j) = \begin{cases} m_p^j + \mu \sum_{i|\delta_i < 0} \phi_i(\alpha^j) + \mu \sum_{i|\delta_i > 0} \phi_i(\alpha^j) \frac{(\alpha_- - \alpha^j)^2}{(\alpha_- - \alpha)^2} & \forall \alpha \in (\alpha_-, \alpha^j] \\ m_p^j + \mu \sum_{i|\delta_i > 0} \phi_i(\alpha^j) + \mu \sum_{i|\delta_i < 0} \phi_i(\alpha^j) \frac{(\alpha_+ - \alpha^j)^2}{(\alpha_+ - \alpha)^2} & \forall \alpha \in [\alpha^j, \alpha_+) \end{cases}$$

Les  $\psi_i$  sont des fonctions barrières. En particulier, elles sont strictement convexes. Donc  $\phi(\cdot)$  est une fonction à valeurs strictement positives. De plus,  $P$  vérifie (7.15) donc  $\mathbf{A}^j$  est définie positive. Donc,  $h(\cdot, \alpha^j)$  est strictement convexe.  $\square$

La fonction majorante  $h(\cdot, \alpha^j)$  est strictement convexe donc elle possède un unique minimiseur. L'expression du minimiseur est analytique et dépend de la valeur de  $\dot{f}(\alpha^j)$  et de  $(m^j, \gamma^j, \bar{\alpha}^j)$ . Ces coefficients étant définis par morceaux, il nous faut préciser leur intervalle de définition. Comme schématisé sur la figure 7.1, ce choix est fonction du signe de  $\dot{f}(\alpha^j)$  :

- La dérivée  $\dot{f}(\alpha^j)$  est positive (Fig. 7.1 (a)) donc  $\alpha^{j+1}$  appartient à  $(\alpha_-; \alpha^j]$  et la fonction à minimiser est la portion de la majorante définie sur  $(\alpha_-; \alpha^j]$ .
- La dérivée  $\dot{f}(\alpha^j)$  est négative (Fig. 7.1 (b)) donc  $\alpha^{j+1}$  appartient à  $[\alpha^j; \alpha_+)$  et la fonction à minimiser est la portion de la majorante définie sur  $[\alpha^j; \alpha_+)$ .

Dans les deux cas, l'expression du minimiseur de la majorante log-quadratique est :

$$\alpha^{j+1} = \begin{cases} \alpha^j - \frac{2q_3}{q_2 + \sqrt{q_2^2 - 4q_1q_3}} & \text{si } \dot{f}(\alpha^j) \leq 0 \\ \alpha^j - \frac{2q_3}{q_2 - \sqrt{q_2^2 - 4q_1q_3}} & \text{si } \dot{f}(\alpha^j) > 0 \end{cases} \quad (7.50)$$

avec

$$\begin{cases} q_1 = -m^j \\ q_2 = \gamma^j - \dot{f}(\alpha^j) + m^j(\bar{\alpha}^j - \alpha^j) \\ q_3 = (\bar{\alpha}^j - \alpha^j)\dot{f}(\alpha^j) \end{cases} \quad (7.51)$$

où  $(m^j, \gamma^j, \bar{\alpha}^j)$  sont donnés par le théorème 7.1.

## 7.3 Propriétés de convergence de la recherche de pas MMLQ 1D

Nous présentons dans cette section l'analyse de convergence du schéma de pas MMLQ 1D décrit dans la section précédente. Afin de faciliter la lecture de cette section, nous avons reporté les démonstrations non immédiates dans l'annexe A.

### 7.3.1 Rappel des hypothèses

La construction de la majorante est établie sous l'hypothèse que la fonction à minimiser s'écrit  $F = P + \mu B$  avec  $B$  une fonction barrière sur le domaine  $\mathcal{D}_b$  délimité par des contraintes linéaires et  $P$  une fonction dont on sait construire des majorantes quadratiques (hypothèse 7.1). Notre analyse de convergence requiert en outre que la fonction globale  $F$  est à gradient  $L$ -Lipschitz sur un voisinage ouvert et borné  $\mathcal{V}$  de la ligne de niveau  $\mathcal{L}_0$ .

**Hypothèse 7.2.** *Soit la ligne de niveau*

$$\mathcal{L}_0 = \{\mathbf{x} \in \mathcal{D}_f | F(\mathbf{x}) \leq F(\mathbf{x}_0)\} \quad (7.52)$$

*supposée bornée et  $\mathcal{V}$  un ouvert contenant  $\mathcal{L}_0$ . Les assertions suivantes sont vérifiées :*

- $F$  est deux fois continument différentiable sur  $\mathcal{V}$
- $F$  est à gradient  $L$ -Lipschitz sur  $\mathcal{V}$
- $\mathcal{V}$  est borné

L'hypothèse 7.2 sur le critère global peut sembler contradictoire du fait de la présence de la fonction barrière. Sa restriction à un voisinage  $\mathcal{V}$  de la ligne de niveau permet cependant d'assurer sa vérification comme l'illustre l'exemple scalaire présenté en figure 7.4. D'autre part, l'hypothèse gradient Lipschitz assure l'existence de  $\eta > 0$  tel que

$$\|\nabla F(\mathbf{x})\| \leq \eta, \forall \mathbf{x} \in \mathcal{V}. \quad (7.53)$$

De plus, le gradient de  $B$  étant infini à la frontière de  $\mathcal{C}$ , (7.53) entraîne l'existence de  $\epsilon_0 > 0$  tel que

$$C_i(\mathbf{x}) \geq \epsilon_0, \forall \mathbf{x} \in \mathcal{V}, \forall i = 1, \dots, m. \quad (7.54)$$

Enfin, le caractère borné de  $\mathcal{V}$  implique qu'il existe  $E_0 > 0$  tel que

$$C_i(\mathbf{x}) \leq E_0, \forall \mathbf{x} \in \mathcal{V}, \forall i = 1, \dots, m. \quad (7.55)$$

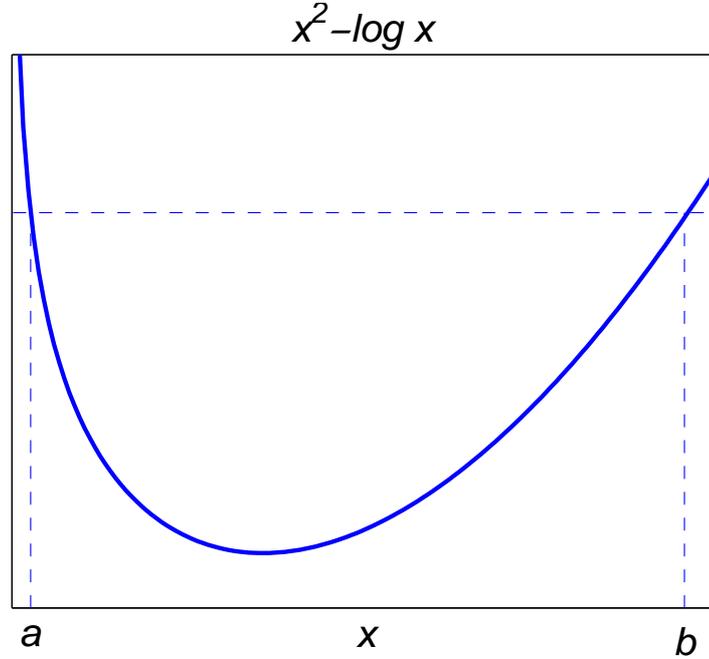


FIGURE 7.4 – La fonction  $F(x) = x^2 - \log x$  contient une barrière stricte pour  $x > 0$ . Cependant, elle est à gradient Lipschitz sur l'ouvert  $\mathcal{V} = (a - \epsilon, b + \epsilon)$  contenant la ligne de niveau  $[a, b]$ .

### 7.3.2 Propriétés de la séquence de pas

Cette section présente les propriétés vérifiées par la séquence (7.49) qui vont nous permettre d'établir les conditions de convergence de l'algorithme de descente. On considère  $\mathbf{x} \in \mathcal{V}$  et  $\mathbf{d}$  une direction de descente. Nous établissons tout d'abord quelques lemmes préliminaires.

**Lemme 7.2.** *La récurrence MM (7.49) entraîne*

$$\text{signe}(\alpha^{j+1} - \alpha^j) = -\text{signe}(f'(\alpha^j)), \quad \forall j \geq 0. \quad (7.56)$$

*Démonstration.* Par construction, on a la relation

$$\dot{f}(\alpha^j) = \dot{h}(\alpha^j, \alpha^j). \quad (7.57)$$

La relation (7.56) est démontrée en remarquant que  $\alpha^{j+1}$  est le minimiseur de  $h(\alpha, \alpha^j)$ .  $\square$

**Lemme 7.3.** Soit  $\ell(\tau)$  la ligne de niveau

$$\ell(\tau) = \{\alpha \in (\alpha_-, \alpha_+) | f(\alpha) \leq \tau\}, \quad (7.58)$$

avec  $\tau \in \mathbb{R}$  différent du minimiseur global de  $f$  sur  $(\alpha_-, \alpha_+)$ . S'il existe un indice  $j$  tel que  $\alpha^j$  appartienne à un ensemble  $\mathcal{I}(\tau)$  inclus dans  $\ell(\tau)$  alors

$$\alpha^i \in \mathcal{I}(\tau), \quad \forall i > j. \quad (7.59)$$

Si l'ensemble  $\mathcal{I}(\tau)$  est borné et qu'il contient un unique point stationnaire  $\alpha^*$  de  $f$  en dehors de ses bornes, alors la séquence  $\{\alpha^j\}$  générée par (7.49) converge vers ce point.

*Démonstration.* Le lemme 7.3 est une application directe des résultats (Jacobson et Fessler, 2007, Th.6.4) et (Jacobson et Fessler, 2007, Th.6.5).  $\square$

Une conséquence directe du lemme 7.3 est que la récurrence MM produit des valeurs décroissantes de  $\{f(\alpha^j)\}$ . Il permet également de montrer la positivité des pas générés.

**Lemme 7.4.** La récurrence MM (7.49) entraîne

$$\alpha^j > 0, \quad \forall j \geq 1. \quad (7.60)$$

*Démonstration.* La dérivée  $\dot{f}(0)$  est strictement négative donc  $\alpha^1 > 0$  d'après le lemme 7.2. De plus, d'après le lemme 7.3,  $f(\alpha^1) < f(0)$ . Posons  $\tau = f(\alpha^1)$  et

$$\mathcal{I}(\tau) = \ell(\tau) \cap (0, \alpha_+). \quad (7.61)$$

Le pas  $\alpha^1$  appartient à  $\mathcal{I}(\tau)$ . D'après le lemme 7.3, tout  $\alpha^j$  appartient donc à  $\mathcal{I}(\tau)$ . En particulier,  $\alpha^j$  est strictement positif pour tout  $j$  supérieur ou égal à 1.  $\square$

**Lemme 7.5.** Courbure de la majorante

Soit  $j \in \{0, \dots, J-1\}$ . Sous les hypothèses 7.1 et 7.2, il existe  $\nu_{\min}, \nu_{\max}, 0 < \nu_{\min} \leq \nu_{\max}$ , tels que pour tout  $\mathbf{x} \in \mathcal{V}$  et pour toute direction de descente  $\mathbf{d}$  en  $\mathbf{x}$  :

$$\nu_{\min} \|\mathbf{d}\|^2 \leq \ddot{h}(\alpha^j, \alpha^j) \leq \nu_{\max} \|\mathbf{d}\|^2, \quad \forall j \geq 0$$

*Démonstration.* La démonstration est reportée en annexe A.1.2.  $\square$

**Lemme 7.6.** Encadrement de la première itérée

Sous les hypothèses 7.1 et 7.2, il existe  $\nu, \nu' > 0$  tels que

$$\frac{-\mathbf{g}^T \mathbf{d}}{\nu \|\mathbf{d}\|^2} \leq \alpha^1 \leq \frac{-\mathbf{g}^T \mathbf{d}}{\nu' \|\mathbf{d}\|^2}. \quad (7.62)$$

*Démonstration.* La démonstration est reportée en annexe A.1.3.  $\square$

Les lemmes 7.5 et 7.6 conduisent aux propriétés 7.1, 7.2 et enfin 7.3.

**Propriété 7.1.** *Majoration du pas*

Sous les hypothèses 7.1 et 7.2, pour tout  $j \in \{1, \dots, J\}$ ,

$$\alpha^j \leq c_{\max}^j \alpha^1, \quad (7.63)$$

avec

$$c_{\max}^j = \left(1 + \frac{2\nu_{\max}L}{\nu_{\min}^2}\right)^{j-1} \left(1 + \frac{\nu}{L}\right) - \frac{\nu}{L} \geq 1. \quad (7.64)$$

*Démonstration.* La démonstration est reportée en annexe A.1.4. □

**Propriété 7.2.** *Minoration du pas*

Sous les hypothèses 7.1 et 7.2, pour tout  $j \in \{1, \dots, J\}$ ,

$$\alpha^j \geq c_{\min} \alpha^1, \quad (7.65)$$

avec

$$c_{\min} = \frac{\sqrt{1 + 2L/\nu_{\min}} - 1}{2L/\nu_{\min}} \in (0, 1/2) \quad (7.66)$$

*Démonstration.* La démonstration est reportée en annexe A.1.5. □

**Propriété 7.3.** *Condition d'Armijo*

Sous les hypothèses 7.1 et 7.2, pour tout  $j \in \{1, \dots, J\}$ ,

$$f(\alpha^j) - f(0) \leq c_1^j \alpha^j f'(0) \quad (7.67)$$

avec

$$c_1^j = (2c_{\max}^j)^{-1} \in (0, 1). \quad (7.68)$$

*Démonstration.* La démonstration est reportée en annexe A.1.6. □

### 7.3.3 Résultats de convergence

Cette section présente les résultats de convergence de l'algorithme de descente itérative (3.18), lorsque  $\mathbf{d}_k$  est une direction de descente et  $\alpha_k$  est obtenu par la stratégie MMLQ 1D (7.49). Nous établissons tout d'abord que le pas MM assure la vérification de la condition de Zoutendijk.

**Théorème 7.2.** *Soit  $\alpha_k$  défini par (7.49). Sous les hypothèses 7.1 et 7.2, la condition de Zoutendijk (4.16) est vérifiée avec*

$$C = \frac{c_{\min} c_1^j}{\nu} > 0. \quad (7.69)$$

*Démonstration.* Tout d'abord, pour tout  $k$ ,  $\mathbf{d}_k$  est une direction de descente soit  $\mathbf{g}_k^T \mathbf{d}_k < 0$ . Donc

$$\mathbf{d}_k \neq 0, \quad \forall k. \quad (7.70)$$

D'après la propriété 7.3, la première condition de Wolfe est vérifiée pour  $c_1 = c_1^J$  :

$$F(\mathbf{x}_k) - F(\mathbf{x}_{k+1}) \geq -c_1^J \alpha_k \mathbf{g}_k^T \mathbf{d}_k \quad (7.71)$$

D'après la propriété 7.2 et le lemme 7.6 :

$$\alpha_k \leq -c_{\min} \frac{\mathbf{g}_k^T \mathbf{d}_k}{\nu \|\mathbf{d}_k\|^2} \quad (7.72)$$

Soit :

$$F(\mathbf{x}_{k+1}) \leq F(\mathbf{x}_k) - C \frac{(\mathbf{g}_k^T \mathbf{d}_k)^2}{\|\mathbf{d}_k\|^2} \quad (7.73)$$

avec  $C = (c_{\min} c_1^J) / \nu > 0$ . Donc (4.16) est vérifiée, en utilisant la définition (4.17).  $\square$

### 7.3.3.1 Algorithmes gradients reliés

Le théorème 7.2 permet d'établir la convergence des algorithmes de type gradients reliés lorsqu'ils sont associés à la recherche de pas MMLQ 1D. Il s'agit d'un résultat très général. En particulier, il englobe les algorithmes de gradients, de Newton, de quasi-Newton, de Newton tronqué ainsi que leurs versions préconditionnées.

**Lemme 7.7.** *Supposons que le critère  $F$  vérifie les hypothèses 7.1 et 7.2. Si pour tout  $k$ , le pas est obtenu par (7.49) et que la suite  $\{\mathbf{d}_k\}_{k \geq 0}$  est gradient reliée à  $\{\mathbf{x}_k\}_{k \geq 0}$ , alors l'algorithme de descente itérative (3.18) converge au sens où  $\lim_{k \rightarrow \infty} \|\mathbf{g}_k\| = 0$ .*

*Démonstration.* D'après le théorème 7.2, l'hypothèse 7.2 entraîne que la condition de Zoutendijk (4.18) est vérifiée. Le lemme s'obtient alors en appliquant le théorème 4.2.  $\square$

### 7.3.3.2 Gradient conjugué non linéaire

La séquence des directions de gradient conjugué non linéaire n'est pas garantie d'être gradient reliée. L'analyse de convergence de cette famille de méthode est plus complexe et dépend du coefficient de conjugaison employé. Afin de vérifier la propriété de descente, nous supposons qu'à chaque itération, la direction s'écrit

$$\begin{aligned} \mathbf{c}_{k+1} &= -\mathbf{g}_{k+1} + \beta_{k+1} \mathbf{d}_k, \\ \mathbf{d}_{k+1} &= -\mathbf{c}_{k+1} \text{signe}(\mathbf{g}_{k+1}^T \mathbf{c}_{k+1}). \end{aligned} \quad (7.74)$$

Pour les coefficients de conjugaison de la forme (4.102), les résultats suivants

- L'encadrement du pas  $\alpha_k^1$  (Lemme 7.6)
- La majoration de  $\alpha_k$  (Propriété 7.1)
- La minoration de  $\alpha_k$  (Propriété 7.2)
- La vérification de la condition de Zoutendijk (Theorème 7.2)

nous permettent d'appliquer une démonstration de convergence identique à celle développée dans (Labat et Idier, 2008, Sec.4) et d'établir le théorème 7.3.

**Théorème 7.3.** *Supposons que les hypothèses 7.1 et 7.2 soient vérifiées. L'algorithme de descente itérative (3.18) associé à la direction (7.74) converge au sens  $\liminf_{k \rightarrow \infty} \mathbf{g}_k = 0$  si  $\alpha_k$  est obtenu par (7.49) et  $\beta_k$  provient de la méthode PRP ou LS, ou plus généralement de (4.102) avec  $\mu_k = 0$  and  $\omega_k \in [0, 1]$ . De plus, si  $F$  est strictement convexe sur  $\mathcal{V}$ , alors on a  $\liminf_{k \rightarrow \infty} \mathbf{g}_k = 0$  pour toute valeur de  $(\mu_k, \omega_k)$ .*

Le théorème 7.4 étend ce résultat pour les formules de conjugaison suivantes

$$\begin{aligned} \beta_k &= \max(\mathbf{g}_{k+1}^T(\mathbf{g}_{k+1} - \mathbf{g}_k)/\|\mathbf{g}_k\|, 0) && \text{Polak-Ribière-Polyak modifié (PRP+)} \\ \beta_k &= \|\mathbf{g}_{k+1}\|^2/\|\mathbf{g}_k\|^2 && \text{Fletcher-Reeves (FR)} \\ \beta_k &= \|\mathbf{g}_{k+1}\|^2/\mathbf{d}_k^T(\mathbf{g}_{k+1} - \mathbf{g}_k) && \text{Dai-Yuan (DY)} \end{aligned}$$

à l'aide d'une hypothèse supplémentaire sur la courbure de la fonction quadratique majorante de  $P$ .

**Théorème 7.4.** *Supposons que le critère  $F$  vérifie les hypothèses 7.1 et 7.2. Si pour tout  $j \in \{0, \dots, J-1\}$  et pour tout  $\mathbf{d} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ ,*

$$\mathbf{d}^T (\nabla P(\mathbf{x} + \alpha \mathbf{d}) - \nabla P(\mathbf{x} + \alpha^j \mathbf{d})) \leq (\alpha - \alpha^j) \mathbf{d}^T \mathbf{A}(\mathbf{x} + \alpha^j \mathbf{d}) \mathbf{d}, \quad \forall \alpha \geq \alpha^j, \quad (7.75)$$

alors l'algorithme de descente itérative (3.18) associé à la direction (7.74), au pas (7.49) et à la conjugaison PRP+ ou FR converge au sens  $\liminf_{k \rightarrow \infty} \|\mathbf{g}_k\| = 0$ . De plus, si le critère  $F$  est convexe sur  $\mathcal{V}$ , alors la convergence est également vérifiée pour la méthode DY.

*Démonstration.* La démonstration est reportée en annexe A.2. □

Soulignons que si  $\dot{p}$  est concave alors l'inégalité (7.75) est vérifiée dès que

$$m_p^j \geq \ddot{p}(\alpha^j). \quad (7.76)$$

## 7.4 Conclusion

La stratégie de pas par Majoration-Minoration quadratique présentée dans le chapitre 4 possède des propriétés nombreuses de convergence et est très simple à implémenter. Cette stratégie est basée sur des approximations majorantes quadratiques du critère le long de la direction de descente. Par conséquent, la présence d'une barrière dans le critère rend cette procédure inefficace voire parfois impossible à mettre en œuvre.

La stratégie développée dans ce chapitre tient compte de cette singularité en se basant sur des approximations majorantes contenant un terme logarithmique. Le pas est toujours obtenu par une récurrence MM très simple, et bénéficie de propriétés de convergence fortes. Nous illustrons à présent son application à la résolution de problèmes quadratiques par la méthode des points intérieurs.

# Chapitre 8

## Résultats expérimentaux sur les algorithmes de points intérieurs

### Sommaire

---

8.1	Introduction . . . . .	<b>106</b>
8.2	La programmation quadratique . . . . .	<b>106</b>
8.3	Algorithmes comparés . . . . .	<b>106</b>
8.3.1	Algorithme primal . . . . .	106
8.3.2	Algorithme primal-dual . . . . .	108
8.3.3	Stratégies de recherche de pas . . . . .	109
8.4	Contexte expérimental . . . . .	<b>110</b>
8.4.1	Génération des problèmes quadratiques . . . . .	110
8.4.2	Paramètres des algorithmes . . . . .	111
8.5	Analyse des résultats . . . . .	<b>111</b>
8.5.1	Recherche de pas MM log-quadratique . . . . .	111
8.5.2	Comparaison avec les méthodes de recherche de pas existantes . . . . .	112
8.6	Conclusion . . . . .	<b>113</b>

---

## 8.1 Introduction

Nous présentons dans ce chapitre les résultats expérimentaux obtenus lorsque la méthode de recherche de pas MM log-quadratique, décrite dans le chapitre 7, est intégrée dans un algorithme de points intérieurs. Ce dernier est un algorithme d'optimisation sous contraintes, basé sur l'introduction dans le critère d'une fonction barrière contraignant la solution à appartenir au domaine contraint. Nous analysons les performances de deux algorithmes de points intérieurs, l'un appartenant à la classe des méthodes primales, l'autre à la classe des méthodes primales-duales, appliqués à la résolution de problèmes de programmation quadratique.

## 8.2 La programmation quadratique

Le problème de programmation quadratique est un problème d'optimisation sous contraintes, où la fonction objectif est quadratique et les contraintes sont linéaires. Dans notre étude, nous considérons les problèmes de programmation quadratique de la forme

$$\begin{aligned} \min_{\mathbf{x}} \left\{ F(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{C}_0 \mathbf{x} + \mathbf{c}_0^T \mathbf{x} + \rho_0 \right\} \\ \text{s.c. } \mathbf{C}(\mathbf{x}) = \mathbf{C} \mathbf{x} + \boldsymbol{\rho} \geq \mathbf{0}, \end{aligned} \quad (8.1)$$

où  $\mathbf{C}_0$  est une matrice symétrique définie positive de  $\mathbb{R}^{n \times n}$ ,  $\mathbf{C}$  est une matrice de  $\mathbb{R}^{m \times n}$  et  $\boldsymbol{\rho}$  est un vecteur de  $\mathbb{R}^m$ . Dans le cadre de l'inversion pénalisée, cette forme de problème se retrouve par exemple lorsque la solution est définie comme le minimiseur d'un critère de moindres carrés ou plus généralement de Tikhonov sous contraintes de positivité (Butler *et al.*, 1981; Daube-Witherspoon et Muehllehner, 1986).

Nous proposons de résoudre ce problème par deux algorithmes de points intérieurs :

- l'algorithme primal de barrière logarithmique (Boyd et Vandenberghe, 2004, Alg.11.1)
- l'algorithme primal-dual de barrière logarithmique (Armand *et al.*, 2000)

L'approche primale consiste à remplacer le problème d'optimisation contraint par une séquence de problèmes d'optimisation non contraints s'identifiant avec la minimisation du critère augmenté d'une fonction barrière. L'approche primale-duale remplace le problème initial par une séquence de systèmes linéaires à résoudre s'identifiant avec des systèmes de Karush Kuhn Tucker relaxés. On qualifie cette méthode de primale-duale car elle permet d'estimer conjointement la solution primale et la solution duale, au sens de la dualité de Lagrange. A l'opposé, l'approche primale estime uniquement la solution dans le domaine primal. Ces deux méthodes font intervenir la fonction barrière logarithmique

$$F_\mu(\mathbf{x}) = F(\mathbf{x}) - \mu \sum_{i=1}^m \log([\mathbf{C} \mathbf{x} + \boldsymbol{\rho}]_i)$$

où  $\mu$ , appelé paramètre de barrière, est strictement positif. La singularité de la fonction barrière sur le bord du domaine contraint permet d'assurer les contraintes d'inégalités de façon simple tout au long de l'algorithme.

Notons que le problème (8.1) peut également être résolu par un algorithme de contraintes actives ou de gradient projeté lorsque les contraintes se limitent à des contraintes de bornes (Nocedal et Wright, 1999, Chap.16), (Gould *et al.*, 2005, Sec.4), (Moré et Toraldo, 1989).

Nous décrivons à présent les deux algorithmes de barrières sur lesquels portent notre analyse.

## 8.3 Algorithmes comparés

### 8.3.1 Algorithme primal

Les méthodes de points intérieurs primales se déroulent comme suit. A partir d'une valeur initiale de  $\mu$ , on minimise de manière approchée la fonction barrière  $F_\mu$ . Du fait de la présence de

la barrière, on obtient un itéré vérifiant strictement les contraintes. On calcule ensuite un nouveau paramètre de perturbation  $\mu > 0$  inférieur au précédent. On obtient un nouveau problème à résoudre. On le résout et on fait à nouveau décroître  $\mu$  de telle manière que la suite de ces paramètres tende vers 0. On va donc résoudre une suite de problèmes barrières à  $\mu$  fixé de manière approchée jusqu'à l'obtention d'une solution du problème initial (8.1).

La construction d'un algorithme primal nécessite de spécifier la méthode employée pour minimiser  $F_\mu$  et le critère d'arrêt associé, ainsi que la règle de mise à jour de  $\mu$  et le critère d'arrêt global de l'algorithme. Nous proposons d'utiliser l'algorithme 8.1 issu de (Boyd et Vandenberghe, 2004) à cause de sa simplicité d'implémentation. Pour chaque  $\mu$ , le critère augmenté  $F_\mu(\mathbf{x})$  est minimisé à l'aide de la méthode itérative de Newton

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k, \quad \text{avec} \quad \mathbf{d}_k = -\nabla^2 F_\mu^{-1}(\mathbf{x}_k) \nabla F_\mu(\mathbf{x}_k)$$

stoppée lorsque

$$\sqrt{-\nabla F_\mu(\mathbf{x}_k)^T \mathbf{d}_k} \leq \tau. \quad (8.2)$$

La mise à jour de  $\mu$  se fait par la règle géométrique

$$\mu \leftarrow (1 - \theta)\mu \quad (8.3)$$

et l'algorithme global arrêté lorsque  $\mu \leq \mu_{\min}$ .

**Entrées :**  $\mu_0 > 0$ ,  $\theta \in (0, 1)$ ,  $\tau > 0$ ,  $\mu_{\min}$  et  $\mathbf{x}_0$  tel que  $C(\mathbf{x}_0) > 0$ .

**Sorties :** Minimiseur contraint de (3.2)

Initialisation :  $t = 0$

**répéter**

$k \leftarrow 0$

$\mathbf{x}_{t+1}^0 \leftarrow \mathbf{x}_t$

**tantque**  $(\nabla F_{\mu_t}(\mathbf{x}_{t+1}^k)^T \mathbf{d}_t^k)^2 > \tau^4$  **faire**

*Boucle interne : Minimisation de  $F_{\mu_t}(\mathbf{x})$*

Calculer  $\mathbf{d}_t^k$  direction de Newton de  $F_{\mu_t}(\mathbf{x}_{t+1}^k)$

Déterminer  $\alpha_t^k$

$\mathbf{x}_{t+1}^{k+1} \leftarrow \mathbf{x}_{t+1}^k + \alpha_t^k \mathbf{d}_t^k$

$k \leftarrow k + 1$

**fin tantque**

$\mathbf{x}_{t+1} \leftarrow \mathbf{x}_{t+1}^k$

$\mu_{t+1} \leftarrow (1 - \theta)\mu_t$

Accroître  $t$

**jusqu'à**  $\mu_t < \mu_{\min}$

Algorithme 8.1 – Algorithme primal de points intérieurs

La convergence de l'algorithme 8.1 est établie dans le théorème 8.1 à condition que le pas  $\alpha_k$  assure la convergence des itérés  $\mathbf{x}_k$ .

**Théorème 8.1.** (Nesterov et Nemirovski, 1994) Si les conditions suivantes sont vérifiées

$$\left\{ \begin{array}{l} \tau < 1 \\ \alpha_k^t \text{ tel que} \\ \lim_{k \rightarrow \infty} \mathbf{x}_k^t = \mathbf{x}_{\mu_t}^* \end{array} \right., \quad (8.4)$$

alors la séquence  $\{\mathbf{x}_t\}$  générée par l'algorithme 8.1 converge vers une solution de (8.1).

### 8.3.2 Algorithme primal-dual

La méthode primale-duale consiste à estimer de façon jointe  $\mathbf{x}$  et les multiplicateurs de Lagrange associés  $\boldsymbol{\lambda}$ . Le problème (8.1) étant convexe, une condition nécessaire et suffisante d'optimalité est obtenue par la vérification des conditions de Karush-Kuhn-Tucker (KKT). La méthode primale-duale résout une version perturbée des conditions KKT

$$C(\mathbf{x}) \geq \mathbf{0}, \quad (8.5)$$

$$\boldsymbol{\lambda} \geq \mathbf{0}, \quad (8.6)$$

$$\boldsymbol{\Lambda}C(\mathbf{x}) = \mu\mathbf{e}, \quad (8.7)$$

$$C_0\mathbf{x} + \mathbf{c}_0 - C^T\boldsymbol{\lambda} = \mathbf{0} \quad (8.8)$$

où  $\boldsymbol{\Lambda} \triangleq \text{Diag}(\boldsymbol{\lambda})$ , pour une suite de paramètres positifs  $\mu$  convergeant vers 0. Ceci équivaut à annuler le résidu

$$r_\mu(\mathbf{x}, \boldsymbol{\lambda}) = \begin{pmatrix} C_0\mathbf{x} + \mathbf{c}_0 - C^T\boldsymbol{\lambda} \\ \boldsymbol{\Lambda}C(\mathbf{x}) - \mu\mathbf{e} \end{pmatrix}. \quad (8.9)$$

Lorsque la résolution de  $r_\mu(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{0}$  est effectuée de manière exacte pour une séquence  $\{\mu_t\}$  tendant vers 0, la suite  $\{\mathbf{x}_t\}$  générée est rigoureusement identique à celle obtenue par l'algorithme primal exact. Nous utilisons ici une version plus 'pratique' de l'algorithme primal-dual dans laquelle la résolution des KKT perturbées est effectuée par un algorithme itératif visant à annuler de façon approchée le résidu  $r_\mu(\mathbf{x}, \boldsymbol{\lambda})$ .

Comme pour l'algorithme primal, il nous faut spécifier la méthode utilisée pour résoudre le système  $r_\mu(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{0}$  et le critère d'arrêt associé, ainsi que la règle de mise à jour de  $\mu$  et le critère d'arrêt global de l'algorithme.

Nous proposons d'utiliser l'algorithme primal-dual 8.2 mis à disposition par Peter Carbonetto (Department of Computer Science, University of British Columbia) sur le site <http://www.cs.ubc.ca/~pcarbo/convexprog.html>. Il s'agit de l'algorithme NOPTIQ (Armand *et al.*, 2000) dans lequel la méthode de Quasi-Newton utilisée pour résoudre le système primal-dual a été remplacée par la méthode de Newton. Cette modification est justifiée dans notre cas car le critère à minimiser est convexe.

**Entrées :**  $\lambda_0 > 0$ ,  $\mu_{\min}$ ,  $\sigma_{\max}$ ,  $\epsilon$  et  $\mathbf{x}_0$  tel que  $C(\mathbf{x}_0) > 0$ .

**Sorties :** Minimiseur contraint de (3.2)

Initialisation :  $k = 0$

**répéter**

Calcul de  $\mu_k$

$$\eta_k = \|r_0(\mathbf{x}_k, \boldsymbol{\lambda}_k)\| / (m + n)$$

$$\sigma_k = \min(0.5, \sqrt{\eta_k})$$

$$\mu_k = \max(\mu_{\min}, \sigma_k C(\mathbf{x}_k)^T \boldsymbol{\lambda}_k / m)$$

Résolution de  $r_{\mu_k}(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{0}$

Calculer  $(\mathbf{d}_{\mathbf{x}_k}, \mathbf{d}_{\boldsymbol{\lambda}_k})$ , solution du système (8.10)

Déterminer  $\alpha_k$ , minimiseur approché de (8.13) le long de  $(\mathbf{d}_{\mathbf{x}}, \mathbf{d}_{\boldsymbol{\lambda}})$

$$\mathbf{x}_{k+1} \leftarrow \mathbf{x}_k + \alpha_k \mathbf{d}_{\mathbf{x}_k}$$

$$\boldsymbol{\lambda}_{k+1} \leftarrow \boldsymbol{\lambda}_k + \alpha_k \mathbf{d}_{\boldsymbol{\lambda}_k}$$

$$k \leftarrow k + 1$$

**jusqu'à**  $\eta_k < \epsilon$

Algorithme 8.2 – Algorithme primal-dual de points intérieurs

La direction de Newton  $(\mathbf{d}_{\mathbf{x}}, \mathbf{d}_{\boldsymbol{\lambda}})$  appliquée à l'équation  $r_\mu(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{0}$  s'obtient en résolvant le

système

$$\begin{pmatrix} \mathbf{C}_0 & -\mathbf{C}^T \\ \mathbf{\Lambda}\mathbf{C} & \text{Diag}(C(\mathbf{x})) \end{pmatrix} \begin{pmatrix} \mathbf{d}_x \\ \mathbf{d}_\lambda \end{pmatrix} = -r_\mu(\mathbf{x}, \boldsymbol{\lambda}). \quad (8.10)$$

En développant l'égalité (8.10), on obtient le système

$$\begin{cases} \mathbf{C}_0\mathbf{d}_x - \mathbf{C}^T\mathbf{d}_\lambda = -(\mathbf{C}_0\mathbf{x} + \mathbf{c}_0) + \mathbf{C}^T\boldsymbol{\lambda} \\ \mathbf{\Lambda}\mathbf{C}\mathbf{d}_x + \text{Diag}(C(\mathbf{x}))\mathbf{d}_\lambda = -\mathbf{\Lambda}\mathbf{C}(\mathbf{x}) + \mu\mathbf{e} \end{cases} \quad (8.11)$$

Le calcul des directions se fait en deux temps : On calcule d'abord la direction primale  $\mathbf{d}_x$  puis la direction duale  $\mathbf{d}_\lambda$ , à partir du système équivalent

$$\begin{cases} (\mathbf{C}_0 + \mathbf{C}^T\text{Diag}(C(\mathbf{x}))^{-1}\mathbf{\Lambda}\mathbf{C})\mathbf{d}_x = -(\mathbf{C}_0\mathbf{x} + \mathbf{c}_0) + \mu\mathbf{C}^T\text{Diag}(C(\mathbf{x}))^{-1}\mathbf{e} \\ \mathbf{d}_\lambda = \text{Diag}(C(\mathbf{x}))^{-1}(-\mathbf{\Lambda}\mathbf{C}(\mathbf{x}) + \mu\mathbf{e} - \mathbf{\Lambda}\mathbf{C}\mathbf{d}_x) \end{cases} \quad (8.12)$$

La direction de Newton est associée à une recherche linéaire faisant intervenir la fonction de mérite

$$\psi_\mu(\mathbf{x}, \boldsymbol{\lambda}) = F_\mu(\mathbf{x}) + \boldsymbol{\lambda}^T\mathbf{x} - \mu \sum_{i=1}^m \log(\lambda_i[\mathbf{C}\mathbf{x} + \boldsymbol{\rho}]_i). \quad (8.13)$$

Le pas vise à minimiser de façon approchée la fonction scalaire  $\psi_\mu(\mathbf{x} + \alpha\mathbf{d}_x, \boldsymbol{\lambda} + \alpha\mathbf{d}_\lambda)$ . Enfin, afin d'assurer la convergence de l'algorithme 8.2, la mise à jour de  $\mu$  est basée sur les conditions de  $\mu$ -criticité faisant intervenir la norme du résidu  $\|r_0(\mathbf{x}, \boldsymbol{\lambda})\|$  (Armand *et al.*, 2000).

### 8.3.3 Stratégies de recherche de pas

**Majoration-Minoration log-quadratique.** Les algorithmes primal et primal-dual considérés font tous les deux intervenir une étape de recherche de pas visant à minimiser de façon approchée un critère de la forme (7.1) contenant une barrière logarithmique le long d'une direction de descente. Dans le cas de l'algorithme primal, la fonction à courbure bornée s'identifie avec le critère quadratique et la barrière s'identifie avec le terme logarithmique. Le lien entre la fonction de mérite (8.13) et la formulation (7.1) s'établit en définissant

$$P(\mathbf{x}, \boldsymbol{\lambda}) = F(\mathbf{x}) + \boldsymbol{\lambda}^T\mathbf{x}, \quad (8.14)$$

$$B(\mathbf{x}, \boldsymbol{\lambda}) = -2 \sum_{i=1}^m \log([\mathbf{C}\mathbf{x} + \boldsymbol{\rho}]_i) - \sum_{i=1}^m \log(\lambda_i). \quad (8.15)$$

Nous pouvons donc appliquer la stratégie MM log-quadratique pour la recherche de pas dans l'algorithme primal et dans l'algorithme primal-dual.

Nous comparons les performances de l'algorithme associé au pas MMLQ 1D avec plusieurs méthodes de recherche de pas usuelles dans le contexte des points intérieurs. Nous décrivons à présent le principe des différentes méthodes testées, et renvoyons le lecteur à leur présentation détaillée au chapitre 6.

**Rebroussement avec restriction du pas.** Cette procédure très simple consiste à partir d'un pas initial correspondant à la frontière du domaine de la barrière, par exemple  $\alpha = 0.99\alpha_+$ , et à le réduire jusqu'à vérifier la condition d'Armijo (4.10) pour un  $c_1$  fixé a priori.

**Interpolation log-quadratique.** Nous allons également tester deux schémas de recherche de pas basés sur une interpolation log-quadratique du critère le long de la direction de recherche. Tout d'abord, nous proposons d'intégrer les stratégies d'interpolation de (Murray et Wright, 1994), décrites dans le tableau 6.3, dans la procédure de pas de (Moré et Thunente, 1994). Le pas obtenu vérifie les conditions de Wolfe strictes (4.10) et (4.12) pour un couple  $(c_1, c_2)$  fixé a priori.

Nous analysons également les performances des algorithmes lorsque la recherche de pas décrite dans les thèses (Doyle, 2003; Lin, 2002) est utilisée. Il s'agit d'une stratégie itérative basée sur une interpolation log-quadratique de la forme (6.37), visant aussi à vérifier les conditions de Wolfe strictes pour un couple  $(c_1, c_2)$ .

## 8.4 Contexte expérimental

### 8.4.1 Génération des problèmes quadratiques

#### 8.4.1.1 Série de tests n°1 : Problèmes quadratiques à contraintes de bornes (Moré et Toraldo, 1989)

Nous analysons tout d'abord les performances de l'algorithme sur un ensemble de problèmes quadratiques sous contraintes de bornes, générés selon la stratégie décrite dans (Moré et Toraldo, 1989, Sec.5). Les problèmes générés s'écrivent sous la forme

$$\begin{aligned} \min_{\mathbf{x}} \left\{ F(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{C}_0 \mathbf{x} + \mathbf{c}_0^T \mathbf{x} \right\} \\ \text{s.c } l_i \leq x_i \leq u_i, \quad i = 1, \dots, n \end{aligned} \quad (8.16)$$

La matrice  $\mathbf{C}_0$  est une matrice symétrique définie positive définie par

$$\mathbf{C}_0 = \mathbf{Y} \mathbf{D} \mathbf{Y}, \quad \mathbf{Y} = \mathbf{I} - \frac{2}{\|\mathbf{y}\|} \mathbf{y} \mathbf{y}^T \quad (8.17)$$

où  $y_i$  est généré aléatoirement dans l'intervalle  $(-1, 1)$ . La matrice  $\mathbf{D}$  est une matrice diagonale dont le  $i$ ème terme diagonal est

$$\log d_i = \left( \frac{i-1}{n-1} \right) n_c. \quad (8.18)$$

Cette expression permet de spécifier le conditionnement de  $\mathbf{C}_0$ , à partir du paramètre

$$n_c = \log(\kappa(\mathbf{C}_0)). \quad (8.19)$$

Les composantes de  $\mathbf{x}^*$  sont générées aléatoirement dans l'intervalle  $(-1, 1)$ . Le vecteur  $\mathbf{c}_0$  ainsi que les bornes  $\mathbf{l}$  et  $\mathbf{u}$  sont choisies afin que  $n_a$  contraintes soient actives en  $\mathbf{x}^*$ . On pose  $\mathbf{c}_0 = \mathbf{r} - \mathbf{C}_0 \mathbf{x}^*$ , où  $\mathbf{r}$  est un vecteur à définir. Le signe des composantes de  $\mathbf{r}$  et la valeur des bornes sont fixées par

$$l_i = -1, \quad u_i = 1, \quad r_i = 0 \quad (8.20)$$

pour  $i \notin \mathcal{A}(\mathbf{x}^*)$ , et

$$l_i = x_i^*, \quad u_i = 1, \quad r_i > 0 \quad (8.21)$$

ou

$$l_i = -1, \quad u_i = x_i^*, \quad r_i < 0 \quad (8.22)$$

pour  $i \in \mathcal{A}(\mathbf{x}^*)$ . En pratique, pour environ la moitié des indices de  $\mathcal{A}(\mathbf{x}^*)$ , on pose  $l_i = x_i^*$ , et pour les indices restants,  $u_i = x_i^*$ . De plus, on fixe

$$|r_i| = 10^{-\nu_i n_d} \quad (8.23)$$

avec  $\nu$  un vecteur généré aléatoirement dans  $(0, 1)$  et  $n_d$  un paramètre qui spécifie la dégénérescence du problème, c'est-à-dire la valeur du gradient du critère à la solution, à travers  $\nabla F(\mathbf{x}^*) = \mathbf{r}$ .

Nous poserons par la suite  $n = 100$ ,  $n_d = 5$  et  $n_a = 50$ . Afin d'étudier la sensibilité des algorithmes au conditionnement de  $\mathbf{C}_0$ , nous testerons les valeurs  $n_c = (0, 3, 6, 9, 12)$ . Enfin, les algorithmes sont initialisés avec un point initial  $\mathbf{x}_0$  correspondant au centre du domaine contraint, c'est-à-dire

$$x_{i,0} = (u_i + l_i)/2. \quad (8.24)$$

#### 8.4.1.2 Série de tests n°2 : Problèmes CUTER (Maros et Mészáros, 1999)

Nous avons également analysé les performances des algorithmes de barrière primal et primal-dual sur un ensemble de problèmes issus de la bibliothèque CUTER mise à disposition sur le site <http://cuter.rl.ac.uk/cuter-www/index.html> par Nicholas I. M. Gould, Dominique Orban et Philippe L. Toint. Les problèmes de programmation quadratique proposés dans CUTER proviennent de l'article (Maros et Mészáros, 1999). Nous avons sélectionné un sous-ensemble de ces problèmes, correspondant aux problèmes quadratiques sans contrainte d'égalité.

Les problèmes issus de CUTER sont de la forme générale (8.1). Par conséquent, l'obtention d'un point initial appartenant au domaine strict des contraintes n'est pas directe. Nous avons donc utilisé la procédure d'initialisation décrite dans (Boyd et Vandenberghe, 2004, Sec.11.4). Cette procédure consiste à définir  $\mathbf{x}_0$  comme la solution du problème d'optimisation suivant, de variables  $[s, \mathbf{x}]$ .

$$\min s, \quad \text{s.c.} \quad C_i(\mathbf{x}) \geq -s, \quad s > 0 \quad (8.25)$$

Les contraintes étant linéaire, l'algorithme de programmation linéaire `linprog` de Matlab a été utilisé pour résoudre le problème d'initialisation (8.25). Notons que, malgré leur forme compatible avec (8.1), les problèmes `hs118` et `powell120` de CUTER n'ont pas été testés car la procédure d'initialisation n'a pas permis d'obtenir un point vérifiant les contraintes strictes du domaine.

### 8.4.2 Paramètres des algorithmes

Dans l'algorithme primal, le paramètre de barrière initial  $\mu_0$  est pris égal à 1 et sa décroissance suit une série géométrique de raison 0.2 (soit  $\theta = 0.8$ ). L'algorithme est arrêté lorsque  $\mu \leq \mu_{\min}$ . Nous fixons les tolérances  $\mu_{\min} = 10^{-8}$  et  $\tau = 10^{-2}$ . En ce qui concerne l'algorithme primal-dual, nous fixons également  $\mu_{\min} = 10^{-8}$  et nous choisissons la tolérance  $\epsilon = 10^{-6}$ .

## 8.5 Analyse des résultats

$K$  dénote le nombre total de mise à jours de  $\mathbf{x}$  et  $T$  le temps avant convergence. Les résultats sont présentés sous forme de moyenne et écart-type. Dans les tableaux, les meilleurs résultats en terme de temps de calcul avant convergence sont indiqués en gras.

### 8.5.1 Recherche de pas MM log-quadratique

A partir de 100 problèmes aléatoires de la série de tests n°1, nous analysons l'influence du paramètre  $J$  de la recherche de pas MMLQ sur les performances des deux algorithmes. Les résultats en termes de moyennes et d'écart types du nombre d'itérations et du temps de calcul sont présentés sur la figure 8.1, pour différentes valeurs du paramètre de conditionnement  $n_c$ . L'influence de  $J$  sur le nombre d'itérations diffère suivant que l'on utilise la méthode primale ou la méthode primale-duale. Dans la méthode primale, le plus petit nombre d'itérations est obtenu pour  $J = 1$ . Ce même comportement a été observé lors de tests avec l'algorithme du gradient conjugué pour la minimisation d'un critère Poissonien (Chouzenoux *et al.*, 2009a) et la minimisation d'un critère entropique (Chouzenoux *et al.*, 2009b). Notons qu'il avait été également observé dans le cas de la recherche de pas MM quadratique dans le chapitre 5. A l'inverse, pour minimiser le nombre d'itérations de l'algorithme primal-dual, il faut choisir une valeur de  $J$  plus importante, par exemple  $J = 5$  ou 10. Cependant, pour les deux algorithmes, les meilleurs résultats en terme de temps de calcul sont toujours obtenus avec  $J = 1$ . Il n'est donc pas avantageux d'être très exigeant lors de la recherche de pas, même si cela pourrait avoir comme effet une diminution du nombre d'itérations. Augmenter  $n_c$  dégrade le conditionnement du problème. Cela entraîne un accroissement du nombre d'itérations globales et du temps de calcul. Cependant, l'influence du paramètre  $J$  n'en est pas modifiée. Nous conseillons donc de choisir  $J = 1$ , indépendamment de la valeur de  $n_c$ , pour les deux algorithmes.

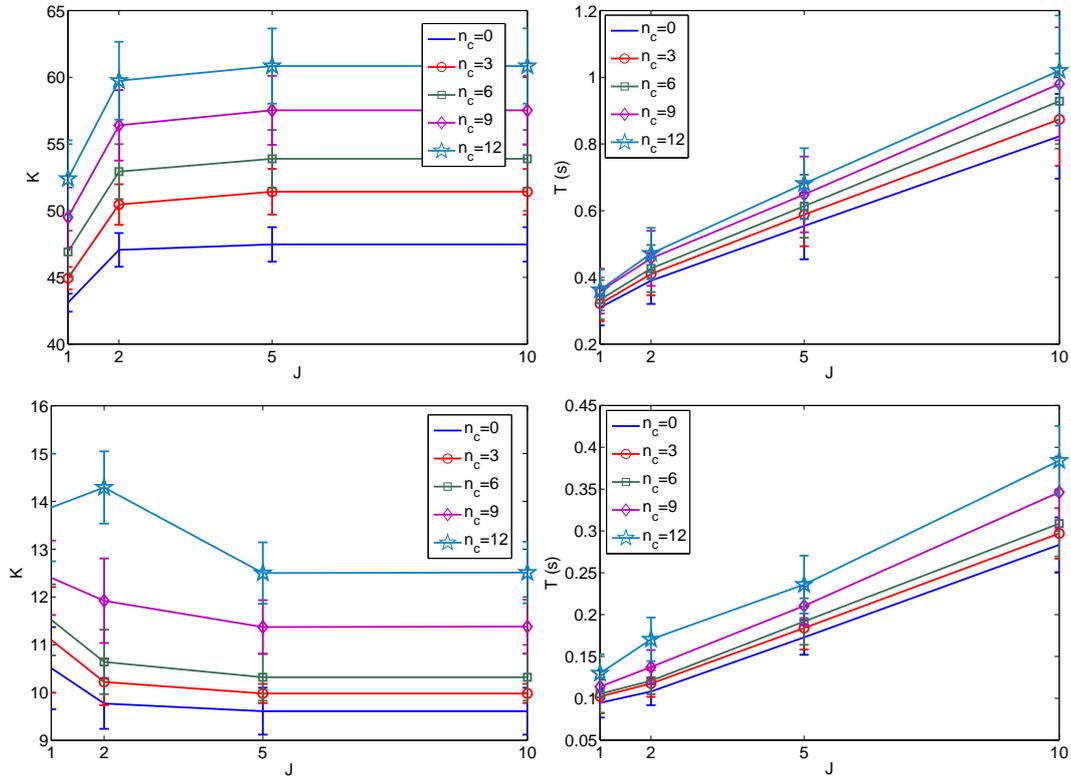


FIGURE 8.1 – Test n°1 : Performances de l’algorithme de barrière primal (haut) et primal-dual (bas) en fonction de  $J$ , nombre de sous itérations MM, et  $n_c$ , conditionnement du problème, sur 100 problèmes quadratiques avec contraintes de bornes générés aléatoirement.

### 8.5.2 Comparaison avec les méthodes de recherche de pas existantes

Nous comparons à présent la recherche de pas MM log-quadratique, associée au réglage optimal  $J = 1$  (MMLQ(1)), avec la méthode de rebroussement basée sur la règle d’Armijo de coefficient  $c_1$  (Armijo( $c_1$ )), et avec les méthodes d’interpolations log-quadratiques basées sur la règle de Wolfe stricte de paramètres  $(c_1, c_2)$  (MW( $c_1, c_2$ ) et LD( $c_1, c_2$ )). Nous analysons tout d’abord l’influence des paramètres de Wolfe sur les performances des algorithmes, à travers la série de test n°1, où nous avons posé  $n_c = 0$ . Les résultats des approches primales et primales-duales sont présentés dans le tableau 8.1.

Les paramètres  $c_1$  et  $c_2$  influent de façon très importante sur les performances de l’algorithme primal. Les résultats obtenus par la méthode de rebroussement sont stables à partir d’une valeur  $c_1$  maximale aux environs de 0.2. Pour les deux autres méthodes, le paramètre  $c_2$  est le plus influent, mais les conclusions diffèrent suivant la stratégie de pas utilisée. Dans la méthode de Murray et Wright, il est préférable d’être exigeant dans la condition d’arrêt, en choisissant  $c_2$  petit. A l’inverse, les meilleures performances de la méthode de Lin et Doyle sont obtenues pour  $c_2$  proche de 1. Pour l’algorithme primal, les meilleurs résultats en terme de nombre d’itérations et de temps de calcul sont obtenus par la méthode MMLQ. La méthode MW fournit également de bons résultats. Par contre, les méthodes d’Armijo et surtout LD requièrent un très grand nombre d’itérations avant la satisfaction du critère d’arrêt.

Les performances de l’algorithme primal-dual n’obéissent pas aux mêmes règles. Les paramètres  $c_1$  et  $c_2$  ont très peu d’influence. De plus, la méthode d’Armijo devient la meilleure en terme de nombre d’itérations et surtout de temps de calcul. La méthode de Lin et Doyle obtient de meilleurs résultats que dans le cas primal, dépassant la méthode MW. La recherche de pas MMLQ nécessite

	Primal		Primal-dual	
	$K$	$T(s)$	$K$	$T(s)$
MMLQ(1)	<b>43.1 ± 0.7</b>	<b>0.31 ± 0.05</b>	10.5 ± 0.9	0.09 ± 0.02
Armijo(0.5)	523 ± 29	2.6 ± 0.2	<b>9.7 ± 0.5</b>	<b>0.06 ± 0.02</b>
Armijo(0.2)	<b>110 ± 5</b>	<b>0.6 ± 0.1</b>	<b>9.7 ± 0.5</b>	<b>0.06 ± 0.02</b>
Armijo(0.1)	<b>100 ± 3</b>	<b>0.6 ± 0.1</b>	<b>9.7 ± 0.5</b>	<b>0.06 ± 0.02</b>
Armijo(0.01)	<b>100 ± 3</b>	<b>0.6 ± 0.1</b>	<b>9.6 ± 0.5</b>	<b>0.06 ± 0.02</b>
MW(0.1, 0.5)	<b>47 ± 1</b>	<b>0.5 ± 0.05</b>	9.7 ± 0.4	0.13 ± 0.04
MW(0.1, 0.9)	<b>50 ± 1</b>	<b>0.48 ± 0.07</b>	10.4 ± 0.5	0.14 ± 0.05
MW(0.1, 0.99)	104 ± 8	0.91 ± 0.16	12.2 ± 0.8	0.15 ± 0.05
MW(0.01, 0.5)	<b>48 ± 1</b>	<b>0.52 ± 0.06</b>	9.7 ± 0.4	0.13 ± 0.03
MW(0.01, 0.9)	<b>50 ± 1</b>	<b>0.47 ± 0.05</b>	10.4 ± 0.5	0.14 ± 0.04
MW(0.01, 0.99)	112 ± 9	0.97 ± 0.17	12.2 ± 0.8	0.15 ± 0.05
LD(0.1, 0.5)	604 ± 146	3.6 ± 0.9	9.7 ± 0.5	0.12 ± 0.03
LD(0.1, 0.9)	587 ± 146	3.5 ± 0.9	9.7 ± 0.4	0.12 ± 0.02
LD(0.1, 0.99)	593 ± 148	3.5 ± 0.9	9.7 ± 0.4	0.12 ± 0.03
LD(0.01, 0.5)	604 ± 146	3.7 ± 0.9	9.7 ± 0.5	0.12 ± 0.03
LD(0.01, 0.9)	542 ± 137	3.2 ± 0.8	9.7 ± 0.4	0.12 ± 0.02
LD(0.01, 0.99)	527 ± 136	3.1 ± 0.7	9.7 ± 0.4	0.12 ± 0.02

TABLE 8.1 – Test n°1 : Comparaison entre différentes stratégies de pas pour l’algorithme de barrière primal et primal-dual sur 100 problèmes quadratiques sous contraintes de bornes générés aléatoirement ( $n_c = 0$ ).

un temps de calcul légèrement supérieur à la méthode de rebroussement. Comme nous l’avons vu précédemment, augmenter  $J$  n’est pas efficace car cela a pour effet de diminuer le nombre d’itérations, mais aussi d’augmenter le temps de calcul.

Nous terminons par les résultats obtenus sur la série de tests n°2, dans les tableaux 8.2 et 8.3. La recherche de pas MMLQ permet d’obtenir les meilleurs résultats pour l’algorithme primal sur la majorité des problèmes. Lorsque l’approche primale-duale est utilisée, bien que la technique de rebroussement entraîne une convergence plus rapide, la recherche de pas MMLQ reste compétitive. Nous concluons en soulignant que sur ces deux séries d’exemples, l’algorithme primal-dual nécessite moins d’itérations et moins de temps de calcul que l’algorithme primal.

## 8.6 Conclusion

Nous avons analysé les performances de deux algorithmes de points intérieurs associés à la stratégie de pas MMLQ, sur un ensemble de problèmes de programmation quadratique. Dans le cas de l’algorithme primal, la stratégie MMLQ s’avère bien plus performante que les stratégies d’Armijo et d’interpolation log-quadratique conseillées dans la littérature. A l’opposé, dans le cas de l’algorithme primal-dual, les meilleurs résultats aussi bien en terme de temps qu’en terme de nombre d’itérations sont obtenus par une méthode basique de maintien du pas à l’intérieur du domaine contraint. Un constat similaire est effectué dans (Roos *et al.*, 2006, Sec.20.6.2) dans le contexte de la programmation linéaire.

D’autre part, comme dans l’application traitée au chapitre 5, les meilleures performances des algorithmes sont obtenues lorsqu’une seule sous-itération est effectuée lors de la recherche de pas MMLQ.

Problème	n	m	MMLQ(1)	Armijo(0.01)	MW (0.01, 0.99)	LD (0.01, 0.99)
hs21	2	5	<b>18/0.1</b>	176/0.3	349/4.9	1114/2.3
hs35	3	4	<b>26/0.1</b>	646/1.2	131/1.2	721/1.6
hs76	4	7	<b>27/0.1</b>	<b>91/0.1</b>	37/0.3	139/0.3
hs268	5	10	<b>28/0.1</b>	923/2.2	252/2.2	125/0.4
ksip	20	1001	<b>49/0.6</b>	81/0.8	131/1.9	645/7.8
liswet1	10002	10000	<b>138/6.3</b>	650/37.1	384/17.9	$\infty$
liswet2	10002	10000	<b>65/3.2</b>	647/35.1	187/9.3	1109/43.2
liswet3	10002	10000	<b>67/3.1</b>	459/25	95/4	744/28.9
liswet4	10002	10000	<b>70/3.4</b>	553/37.9	174/9.5	840/33.7
liswet5	10002	10000	<b>64/3.1</b>	372/19	93/4.2	744/28.6
liswet6	10002	10000	71/3.9	462/26.4	<b>80/3.2</b>	782/34.2
liswet7	10002	10000	<b>318/16.6</b>	751/39.7	458/27.2	1117/74.2
liswet8	10002	10000	<b>223/9</b>	753/47.2	572/31	$\infty$
liswet9	10002	10000	<b>256/11.7</b>	733/27.1	377/15.5	1019/46.2
liswet10	10002	10000	<b>189/8.7</b>	743/43	302/13.1	1109/50.9
liswet11	10002	10000	<b>100/5.1</b>	746/36.3	194/9.2	1019/49
liswet12	10002	10000	288/13.2	774/42	<b>287/12.7</b>	$\infty$
mosarqp1	2500	3200	<b>62/1.4</b>	111/2.2	89/2	630/12.5
mosarqp2	900	1500	<b>65/1</b>	112/1.4	110/1.6	661/8.9
primal1	325	86	<b>51/1.9</b>	90/2.9	102/3.7	565/17.4
primal2	649	97	<b>52/6.4</b>	88/9.5	80/8.5	469/42.5
primal3	745	112	<b>53/18</b>	96/28	93/27.3	739/181.2
primal4	1489	76	<b>50/66</b>	96/114	91/113	$\infty$
primalc1	230	224	278/5.3	254/3.9	<b>210/3.4</b>	316/5.2
primalc2	231	236	<b>52/1</b>	132/2.1	62/1.2	313/11.7
primalc5	287	286	<b>42/2</b>	111/3.1	63/2.2	130/4.6
primalc8	520	511	407/53	175/18	133/16.1	<b>116/12.3</b>
s268	5	5	<b>19/0.1</b>	1007/2.3	25/0.2	29/0.2
zecevic2	2	6	<b>21/0.1</b>	<b>91/0.1</b>	<b>31/0.1</b>	<b>24/0.1</b>

 TABLE 8.2 – Test n°2 : Comparaison des recherches de pas sur la liste de problèmes CUTeR en terme de nombre d'itérations et de temps de calcul  $K/T$  pour l'algorithme de barrière primal.

Problème	n	m	MMLQ(1)	Armijo(0.01)	MW(0.01, 0.99)	LD(0.01, 0.99)
hs21	2	5	16/0.2	<b>14/0.1</b>	27/0.2	18/0.1
hs35	3	4	12/0.2	<b>12/0.1</b>	33/0.3	27/0.2
hs76	4	7	9/0.2	<b>8/0.1</b>	19/0.2	24/0.2
hs268	5	10	<b>14/0.2</b>	144/1.1	$\infty$	260/1.7
ksip	20	1001	45/0.9	<b>25/0.4</b>	33/0.7	64/1.7
liswet1	10002	10000	49/2.9	<b>43/2</b>	43/2.4	45/3.2
liswet2	10002	10000	27/1.7	<b>20/0.9</b>	20/1.2	20/1.4
liswet3	10002	10000	19/1.2	<b>15/0.7</b>	15/1	15/1
liswet4	10002	10000	19/1.2	<b>15/0.7</b>	14/1.2	15/1.1
liswet5	10002	10000	26/1.6	<b>19/0.9</b>	19/1	19/1.3
liswet6	10002	10000	27/1.6	<b>20/0.9</b>	20/1.1	20/1.4
liswet7	10002	10000	113/6.4	<b>96/5</b>	97/7.7	92/6.6
liswet8	10002	10000	190/10.6	<b>149/7.6</b>	$\infty$	147/10
liswet9	10002	10000	$\infty$	$\infty$	$\infty$	$\infty$
liswet10	10002	10000	116/6.8	<b>91/4.6</b>	98/8.4	85/6.1
liswet11	10002	10000	65/3.7	<b>56/2.6</b>	56/3	59/4.4
liswet12	10002	10000	611/33.8	<b>433/20</b>	391/22.7	400/29.1
mosarqp1	2500	3200	17/0.5	<b>16/0.3</b>	16/0.6	16/0.4
mosarqp2	900	1500	16/0.4	<b>15/0.2</b>	15/0.4	<b>14/0.2</b>
primal1	325	86	29/0.8	<b>25/0.5</b>	26/0.8	44/1.3
primal2	649	97	23/1.6	<b>23/1.4</b>	23/1.6	29/2.2
primal3	745	112	<b>23/5.3</b>	24/5.4	24/5.7	51/13
primal4	1489	76	24/14.1	<b>22/12.9</b>	22/13.4	54/34.7
primalc1	230	224	74/1.7	<b>47/0.7</b>	47/1.1	65/1.4
primalc2	231	236	47/0.9	<b>22/0.4</b>	22/0.5	155/4.2
primalc5	287	286	29/0.8	<b>20/0.5</b>	20/0.8	<b>21/0.5</b>
primalc8	520	511	60/5.2	41/3.3	<b>36/3</b>	44/3.9
s268	5	5	19/0.2	<b>24/0.1</b>	$\infty$	<b>16/0.1</b>
zecevic2	2	6	10/0.2	<b>9/0.1</b>	10/0.3	96/0.7

TABLE 8.3 – Test n°2 : Comparaison des recherches de pas sur la liste de problèmes CUTeR en terme de nombre d'itérations et de temps de calcul  $K/T$  pour l'algorithme de barrière primal-dual.



# Chapitre 9

## Application à la reconstruction de spectres RMN 2D

### Sommaire

---

9.1	Introduction . . . . .	<b>118</b>
9.2	La résonance magnétique nucléaire . . . . .	<b>118</b>
9.2.1	Les phénomènes de relaxation . . . . .	118
9.2.2	Mesure des temps de relaxation . . . . .	121
9.2.3	Cas de systèmes hétérogènes . . . . .	121
9.3	Méthodes d'inversion pénalisée existantes . . . . .	<b>125</b>
9.3.1	Méthode de Tikhonov sous contraintes de positivité . . . . .	125
9.3.2	Reconstruction par maximum d'entropie . . . . .	126
9.4	Algorithme de minimisation ME proposé . . . . .	<b>128</b>
9.4.1	Description de l'algorithme . . . . .	128
9.4.2	Réduction du coût de calcul . . . . .	131
9.5	Application à la reconstruction de données réelles . . . . .	<b>132</b>
9.5.1	Réglage du coefficient de régularisation . . . . .	132
9.5.2	Réglage du paramètre d'angle de bascule . . . . .	134
9.5.3	Application à l'analyse d'échantillons de pomme par RMN . . . . .	135
9.6	Conclusion . . . . .	<b>136</b>

---

## 9.1 Introduction

La résonance magnétique nucléaire (RMN) est une méthode permettant d'accéder à des informations structurales aux échelles microscopiques ou moléculaires de produits biologiques (Canet *et al.*, 2002). Par exemple, dans le domaine de l'agro-alimentaire, cette technique est employée pour déterminer la teneur en eau ou en matière grasse d'un échantillon de matière (Mariette, 2004). Nous nous intéressons à l'estimation d'un spectre de corrélation  $T_1 - T_2$  à partir de mesures RMN. Les difficultés de l'estimation sont liées au caractère mal-posé de ce problème inverse et à la taille importante des données à traiter. La pénalisation par maximum d'entropie (ME) s'avère être une méthode de régularisation adéquate pour ce problème (Laue *et al.*, 1985; Mariette *et al.*, 1996). Nous présentons dans ce chapitre une stratégie efficace de minimisation du critère ME reposant sur l'emploi d'un algorithme de Newton tronqué. La singularité du terme entropique motive l'utilisation de la recherche de pas MMLQ 1D développée dans le chapitre 7.

Nous présentons tout d'abord dans la section 9.2 le principe de la mesure de données, pour l'estimation de spectres de corrélation  $T_1 - T_2$  en RMN, ainsi que le problème inverse associé. Le caractère mal-posé du problème conduit à l'utilisation d'une méthode d'inversion pénalisée. Nous présentons en section 9.3 les différentes stratégies de pénalisation employées dans la littérature pour résoudre ce problème. Nous retenons par la suite la pénalisation ME et nous proposons dans la section 9.4 un algorithme de minimisation efficace du critère pénalisé résultant de cette approche. Les performances de la méthode proposée sont analysées à travers des tests sur des données réelles, dont les résultats sont présentés en section 9.5.

## 9.2 La résonance magnétique nucléaire

Le phénomène physique sur lequel repose la RMN a été étudié en 1945 par des groupes de physiciens dirigés par Edward M. Purcell (Université de Harvard) et Felix Bloch (Université de Stanford) (Purcell, 1952). Depuis cette date, la technique a connu un développement spectaculaire dans différents domaines. La RMN apporte des renseignements intéressants sur la structure des molécules en se basant sur des propriétés quantiques des noyaux. Plus précisément, la RMN quantifie les phénomènes de relaxation qui interviennent au sein du spin des noyaux d'un échantillon lorsque celui-ci est soumis à un champ magnétique.

### 9.2.1 Les phénomènes de relaxation

La résonance magnétique nucléaire exploite une propriété particulière des particules quantiques qu'on appelle le spin. L'effet du spin est similaire à celui d'un moment magnétique avec lequel un champ magnétique extérieur peut interagir. La RMN consiste à modifier le moment magnétique nucléaire en appliquant des champs magnétiques à l'échantillon étudié (Canet *et al.*, 2002; Ernst *et al.*, 1997; Van de Ven, 1995).

En absence de champ magnétique externe, les moments magnétiques des spins d'un échantillon tissulaire sont orientés de façon aléatoire dans l'espace, comme le schématise la figure 9.1(1). La somme des vecteurs d'aimantation est nulle ainsi que le vecteur d'aimantation macroscopique résultant ( $M = 0$ ). Lorsqu'un champ magnétique externe  $B_0$  d'intensité significative est appliqué, les moments des spins s'orientent dans le sens parallèle à la direction du champ appliqué ou dans le sens anti-parallèle à cette direction (figure 9.1(2)). Le nombre de spins s'orientant parallèlement au champ  $B_0$  étant supérieur, il en résulte une aimantation induite  $M_0$  non nulle.

Les moments des spins sont animés d'un mouvement de précession autour de l'axe  $Oz$ , qui s'apparente à la rotation d'une toupie autour de la verticale. Ce mouvement est caractérisé par sa fréquence de rotation, appelée *fréquence de résonance* ou *fréquence de Larmor*. En appliquant un deuxième champ de radio-fréquence  $B_1$ , de fréquence égale à la fréquence de résonance des spins, on va provoquer des transitions entre les spins orientés parallèlement et ceux orientés anti-parallèlement. Cela se traduit par un basculement des spins, générant une nouvelle aimantation  $M_1$ . L'angle de bascule  $\Phi$  entre  $M_0$  et  $M_1$  dépend de l'intensité et de la durée d'application du

champ  $B_1$ . La figure 9.1(3) illustre le cas d'un champ  $B_1$  entraînant un basculement de  $\Phi = 90^\circ$ , c'est-à-dire la projection de  $M_0$  dans le plan normal  $Oxy$ . A l'arrêt de l'excitation, les spins tendent à revenir à leur état initial (figure 9.1(4)).

Ce retour à l'équilibre n'est pas instantané mais s'étale dans le temps : c'est le phénomène de *relaxation*. On peut décomposer le mouvement de  $M_1$  vers  $M_0$  en relaxation longitudinale pour ce qui concerne le retour vers  $M_0$  de la composante longitudinale  $M_z$  de cette aimantation et relaxation transversale pour ce qui concerne le retour à 0 de la composante transversale  $M_{xy}$ . Les phénomènes de relaxation sont dus à deux phénomènes quantiques différents. La relaxation longitudinale est provoquée par les transitions énergétiques des protons tandis que la relaxation transverse est due au déphasage des spins. Les figures 9.2(a) et 9.2(b) illustrent l'évolution de l'aimantation  $M(t)$  après une excitation  $B_1$  engendrant un basculement d'angle  $\Phi = 90^\circ$ .

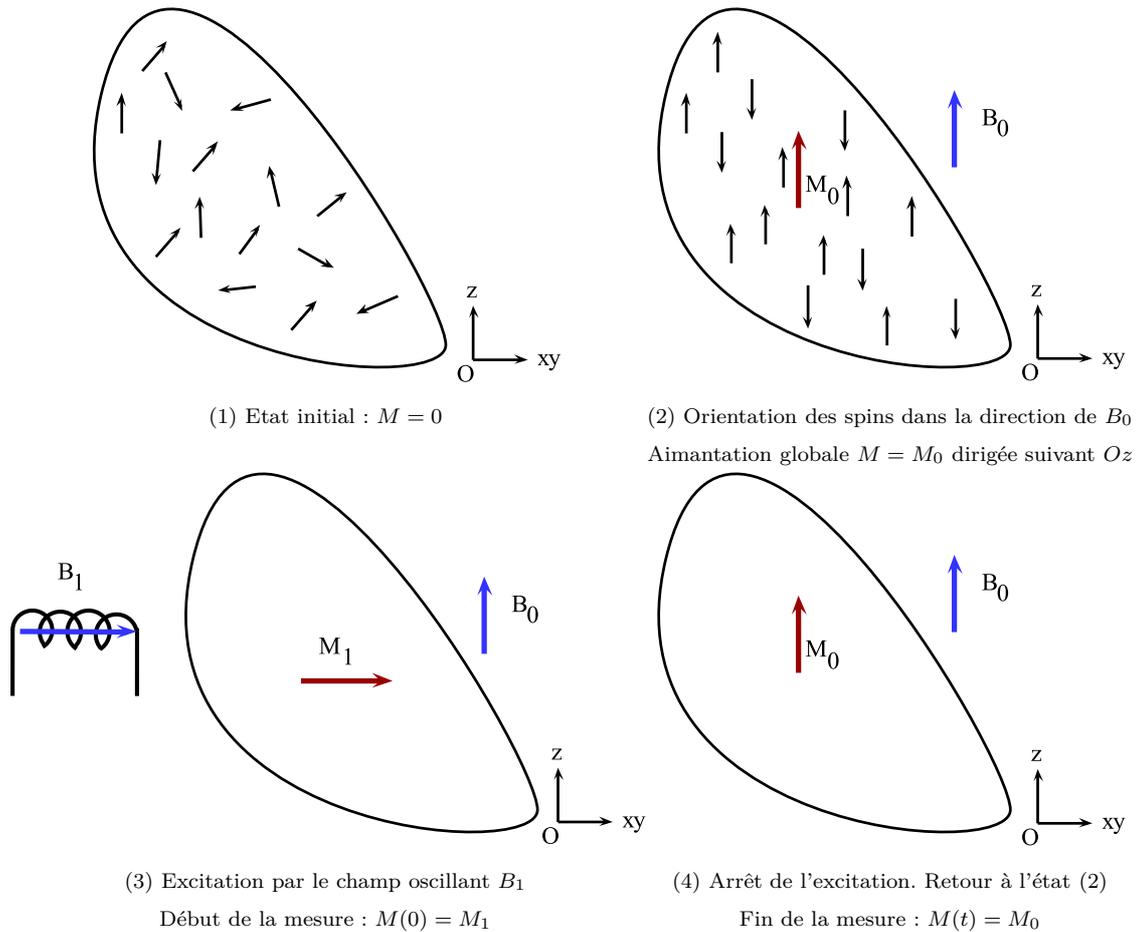
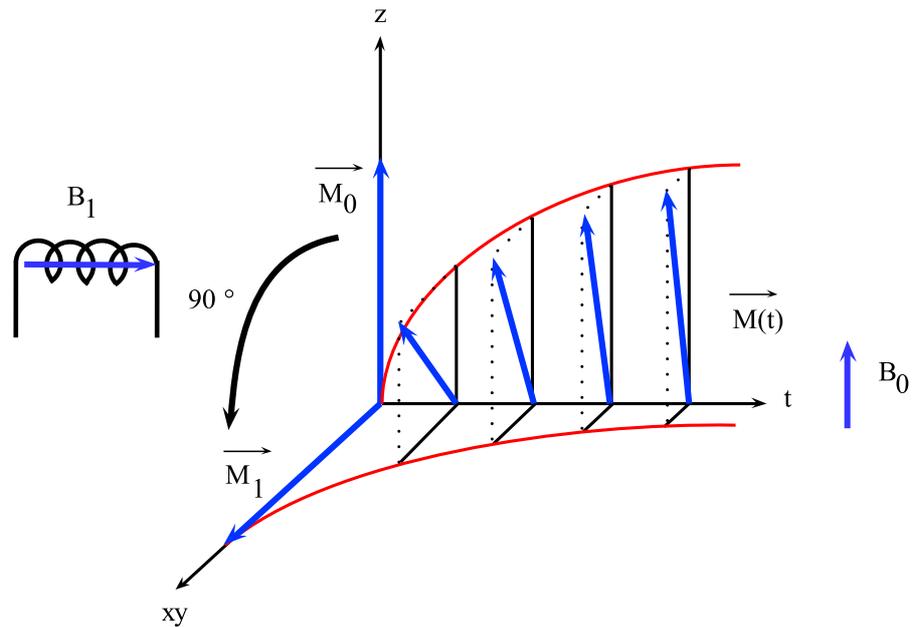


FIGURE 9.1 – Phénomène de résonance magnétique nucléaire

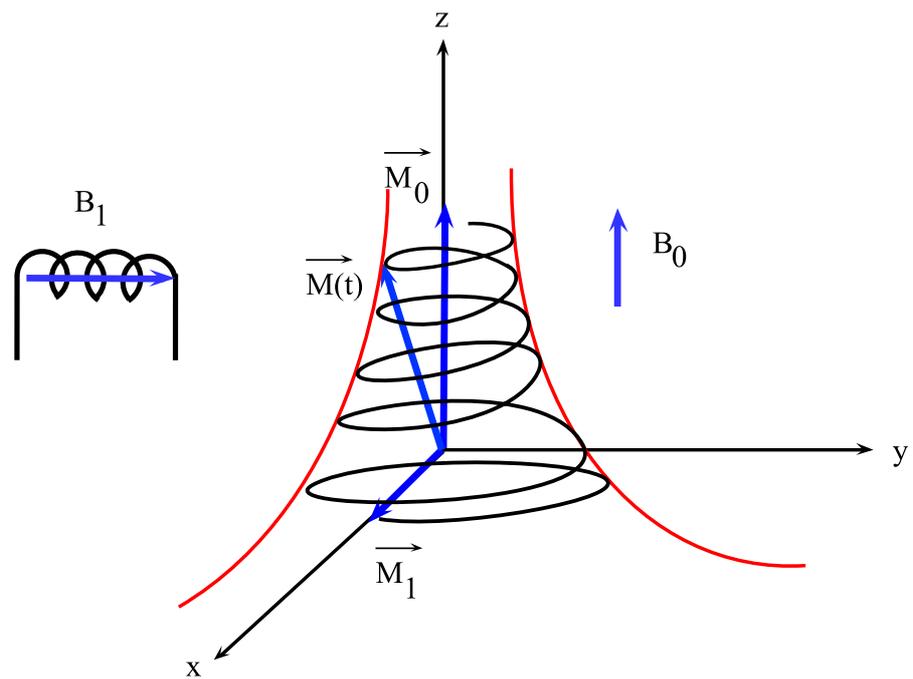
Le retour de  $M_z(t)$  à sa valeur initiale avant excitation  $M_0$  est exponentiel :

$$M_z(t) = M_0(1 - \gamma e^{-\frac{t}{T_1}}), \quad \gamma = 1 - \cos \Phi, \quad (9.1)$$

et dépend de l'angle de bascule  $\Phi$  à travers le paramètre  $\gamma$ . Par exemple, si  $B_1$  provoque un basculement de  $90^\circ$ ,  $\gamma$  est égal à 1. Cette relaxation longitudinale dite également relaxation  $T_1$  ou encore relaxation 'spin-réseau' correspond au retour à l'équilibre énergétique du système après l'excitation. La constante de temps  $T_1$  dépend de la mobilité des atomes ou de celle des molécules où ils sont engagés. Plus précisément,  $T_1$  sera d'autant plus court que la fluctuation des champs électromagnétiques locaux sera proche de la fréquence de résonance, la fluctuation des champs



(a) Evolution des composantes longitudinale ( $Oz$ ) et transversale ( $Oxy$ ) du moment magnétique



(b) Evolution tri-dimensionnelle du moment magnétique

FIGURE 9.2 – Relaxation du moment magnétique  $M(t)$  après un basculement de  $\Phi = 90^\circ$ .

electromagnétique locaux étant dépendante de la mobilité ou du temps de corrélation des molécules. Par exemple, les molécules liquides ayant des temps de corrélation longs par rapport à la fréquence de résonance sont caractérisées par un  $T_1$  grand (supérieur à 1.5 secondes). A l'inverse, le temps de relaxation  $T_1$  des molécules de graisse est de l'ordre de 0.3 secondes.

Le retour de  $M_{xy}(t)$  vers 0 est également exponentiel :

$$M_{xy}(t) = M_0 e^{-\frac{t}{T_2}}. \quad (9.2)$$

La relaxation transversale est due au déphasage des aimantations dans leur mouvement autour du champ  $B_0$ . Ce déphasage est lié aux interactions entre les aimantations des noyaux voisins qui créent des modifications locales du champ magnétique. Le temps de relaxation  $T_2$  est encore appelé temps de relaxation ‘spin-spin’. Il est toujours inférieur au temps de relaxation  $T_1$  et dépend lui aussi de la mobilité des atomes ou des molécules sur lesquelles ces atomes sont engagés. Il est d’autant plus court que la taille de la molécule est importante. Le temps de relaxation  $T_2$  des liquides est supérieur à 1 seconde et celui des graisses est de l’ordre de 40 ms.

L’analyse RMN s’intéresse à la caractérisation des constantes de temps  $T_1$  et  $T_2$ . Ces paramètres physiques sont fondamentaux. En effet, leur détermination permet de préciser l’environnement électronique et donc chimique du noyau étudié. Leur connaissance permet des applications nombreuses : En médecine, les temps de relaxation sont à l’origine du contraste des images IRM. En agro-alimentaire, on peut utiliser le fait que, dans un système complexe, le temps de relaxation des protons de l’eau soit différent du temps de relaxation des protons de la matière grasse, afin d’en déterminer la teneur en eau et en matière grasse. Nous expliquons à présent le protocole de mesure des temps de relaxation d’un échantillon.

## 9.2.2 Mesure des temps de relaxation

### 9.2.2.1 Temps de relaxation longitudinal

La relaxation du champ dépend de l’angle de bascule  $\Phi$  de l’aimantation résultante par rapport à la position initiale. Deux expériences typiques de mesure de  $T_1$  sont décrites sur la figure 9.3. Elles consistent à appliquer un champ  $B_1$  entraînant un basculement de  $\Phi = 90^\circ$  ou de  $\Phi = 180^\circ$ , ce qui correspond respectivement à la mesure par *saturation-récupération* ou par *inversion-récupération*. Afin de distinguer les variations de la composante longitudinale de l’aimantation, on produit à la fin de l’expérience un basculement de  $90^\circ$  de celle-ci pour pouvoir la lire dans le plan transversal. La durée d’acquisition, notée  $\tau_1$ , est appelée *temps de répétition*. La constante de temps  $T_1$  se déduit de l’intensité du signal obtenu au cours de l’expérience, à partir du modèle (9.1).

### 9.2.2.2 Temps de relaxation transversal

Si le champ  $B_0$  était parfaitement homogène et si l’échantillon placé dans ce champ n’induisait aucune inhomogénéité, la courbe de décroissance de la composante transversale après un simple basculement de  $90^\circ$  permettrait d’avoir accès directement au temps de relaxation  $T_2$ . En fait, le champ  $B_0$  n’est jamais parfaitement homogène et la désynchronisation des aimantations transversales s’effectue plus vite que ne le voudrait le  $T_2$  : le  $T_2^*$  qui correspond à la décroissance directement observée du signal RMN est toujours inférieur au  $T_2$ . Ce  $T_2^*$  est d’autant plus court que les inhomogénéités de construction sont importantes.

La technique d’*écho de spin* permet de s’affranchir de ces inhomogénéités du champ  $B_0$  (figure 9.4). Elle consiste à faire un premier basculement de  $90^\circ$  pour amener l’aimantation longitudinale dans le plan transversal, selon l’axe  $Oy$  par exemple. Puis à appliquer une séquence d’échos, espacés de  $\tau_e$ , qui basculent les aimantations de  $180^\circ$  par rapport au plan transversal. Sur une séquence de ce type, on montre que la décroissance des échos au cours du temps d’écho  $\tau_2 = k\tau_e$  se fait selon une exponentielle dont la constante de temps correspond cette fois au vrai  $T_2$ , ce qui correspond au modèle (9.2).

## 9.2.3 Cas de systèmes hétérogènes

Les composants hétérogènes sont composés d’un mélange de groupes d’atomes individuels possédant des couples  $(T_1, T_2)$  différents. Ce mélange est caractérisé globalement par une distribution

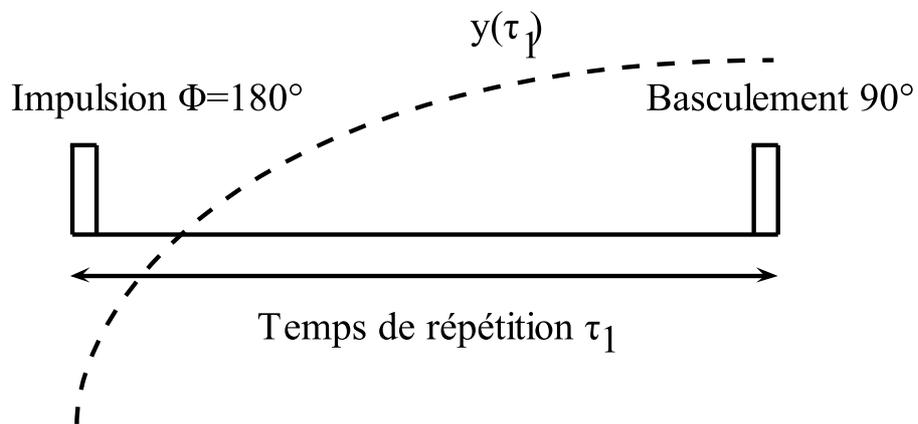
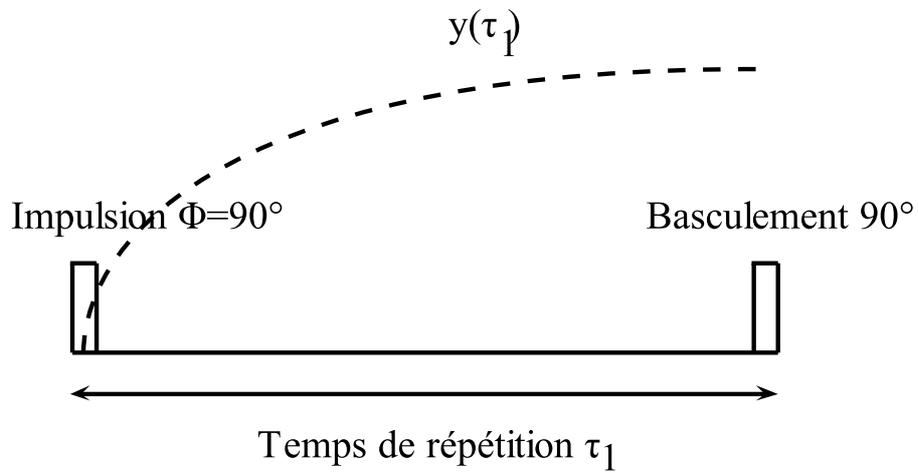


FIGURE 9.3 – Acquisition du temps de relaxation longitudinal

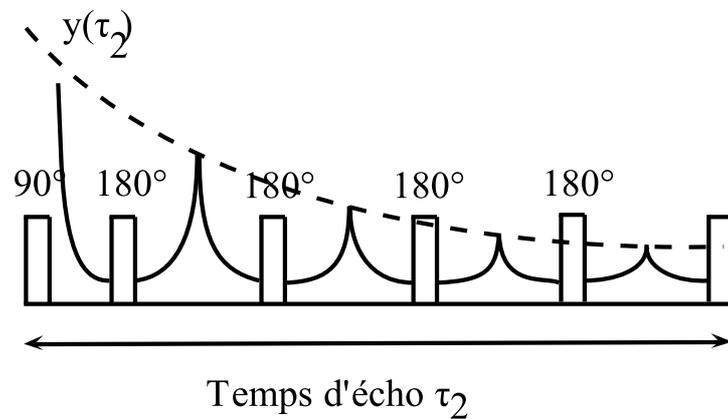


FIGURE 9.4 – Acquisition du temps de relaxation transversal : Echo de spin

$X(T_1, T_2)$  permettant d'identifier et de quantifier les différents modes de relaxation présents dans le composant. Cette cartographie bi-dimensionnelle est privilégiée par rapport aux spectres indépendants  $X(T_1)$  et  $X(T_2)$  car elle permet une mise en évidence de couplages entre les paramètres de relaxation  $T_1$  et  $T_2$  aidant notablement à leur interprétation (English *et al.*, 1991; Song *et al.*, 2002). La technique permettant de déterminer la distribution couplée  $X(T_1, T_2)$ , nommée *RMN bi-dimensionnelle* ou RMN 2D, fait l'objet de notre étude.

### 9.2.3.1 Principe de la mesure

Comme le décrit le schéma 9.5, l'expérience RMN 2D consiste à réaliser une séquence de saturation-récupération (ou inversion-récupération), suivie d'une séquence d'écho de spin. L'expérience 2D est donc caractérisée par trois paramètres : le temps de répétition  $\tau_1$  et l'angle de bascule  $\Phi$  agissant sur la relaxation longitudinale, et le temps d'écho  $\tau_2$  agissant sur la relaxation transversale.

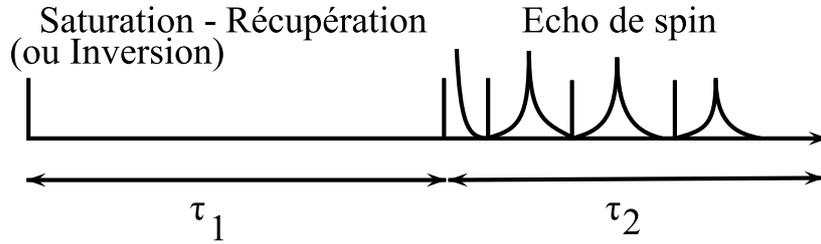


FIGURE 9.5 – Acquisition des mesures en RMN 2D

La mesure enregistrée  $Y(\tau_1, \tau_2, \Phi)$  est liée au spectre  $X(T_1, T_2)$  par une intégrale de Fredholm du premier type (Hürlimann et Venkataramanan, 2002; Song *et al.*, 2002)

$$Y(\tau_1, \tau_2, \Phi) = \int \int k_1(\tau_1, T_1, \Phi) X(T_1, T_2) k_2(\tau_2, T_2) dT_1 dT_2 \quad (9.3)$$

avec les noyaux

$$\begin{cases} k_1(\tau_1, T_1, \Phi) &= 1 - \gamma e^{-\tau_1/T_1} \\ k_2(\tau_2, T_2) &= e^{-\tau_2/T_2} \end{cases}, \quad \gamma = 1 - \cos \Phi.$$

Le spectre  $X(T_1, T_2)$  est déterminé en effectuant plusieurs expériences RMN avec des paramètres d'acquisition différents. Plus précisément, l'angle  $\Phi$  est maintenu constant (Typiquement  $\Phi = 90^\circ$  ou  $180^\circ$ ) et les mesures expérimentales sont collectées sur une grille de  $m_1 \times m_2$  valeurs discrètes du couple  $(\tau_1, \tau_2)$ , ce qui fournit une matrice de données  $\mathbf{Y} \in \mathbb{R}^{m_1 \times m_2}$ .

Les distributions mono-dimensionnelles  $X(T_1)$  et  $X(T_2)$  peuvent s'obtenir de deux façons différentes. Elles peuvent être déduites de  $X(T_1, T_2)$  par marginalisation :

$$X(T_1) = \int X(T_1, T_2) dT_2, \quad X(T_2) = \int X(T_1, T_2) dT_1. \quad (9.4)$$

Elles peuvent également être obtenues par inversion 1D des mesures  $Y(\tau_1, \tau_2, \Phi)$  obtenues pour  $\tau_2$  proche de 0 et pour  $\tau_1$  très grand. Cette stratégie se base sur les propriétés suivantes

$$Y(\tau_1, 0, \Phi) = \int \int k_1(\tau_1, T_1, \Phi) X(T_1, T_2) \underbrace{k_2(0, T_2)}_1 dT_1 dT_2, \quad (9.5)$$

$$= \int k_1(\tau_1, T_1, \Phi) \underbrace{\left( \int X(T_1, T_2) dT_2 \right)}_{X(T_1)} dT_1, \quad (9.6)$$

et

$$Y(\infty, \tau_2, \Phi) = \int \int \underbrace{k_1(\infty, T_1, \Phi)}_1 X(T_1, T_2) k_2(\tau_2, T_2) dT_1 dT_2, \quad (9.7)$$

$$= \int k_2(\tau_2, T_2) \underbrace{\left( \int X(T_1, T_2) dT_1 \right)}_{X(T_2)} dT_2. \quad (9.8)$$

### 9.2.3.2 Le problème de reconstruction

L'objectif du traitement est d'estimer le spectre  $X(T_1, T_2)$  à partir des mesures  $\mathbf{Y}$  en utilisant le modèle (9.3) pour une valeur du paramètre  $\Phi$  donnée. La quantité à estimer représentant une distribution, elle doit respecter la contrainte  $X \geq 0$ . Ce problème de reconstruction, équivalent à une inversion numérique d'une transformée de Laplace 2D, est un problème inverse mal posé (Sternin, 2008). Les méthodes développées pour la résolution de (9.3) peuvent se classer en deux familles.

**Approche paramétrique.** L'approche paramétrique consiste à rechercher le spectre sous la forme d'une somme d'un nombre connu d'impulsions caractérisés par leur position et leur amplitude (Borgia *et al.*, 1998; Bretthorst *et al.*, 2005; Gallegos et Smith, 1988; Naranayan *et al.*, 1995). Les mesures sont décrites comme une somme pondérée d'exponentielles décroissantes. La valeur des poids, correspondant aux amplitudes des différents pics, est estimée par un algorithme itératif visant à minimiser un critère des moindres carrés non linéaires (Gallegos et Smith, 1988; Naranayan *et al.*, 1995). A notre connaissance, l'approche paramétrique n'est utilisée que pour la reconstruction RMN mono-dimensionnelle.

**Approche non paramétrique.** L'approche non paramétrique estime la distribution  $X(T_1, T_2)$  sur une grille régulière en  $T_1$  et  $T_2$  de taille  $n_1 \times n_2$ . Pour cela, les noyaux  $k_1$  et  $k_2$  sont mis sous forme discrète  $\mathbf{K}_1 \in \mathbb{R}^{m_1 \times n_1}$  et  $\mathbf{K}_2 \in \mathbb{R}^{m_2 \times n_2}$ . Cette modélisation se traduit par l'équation linéaire

$$\mathbf{Y} = \mathbf{K}_1 \mathbf{X} \mathbf{K}_2^t + \mathbf{E} \quad (9.9)$$

où  $\mathbf{E}$  concentre l'erreur de modélisation et également le bruit lié à la mesure. Le problème inverse résultant de ce modèle linéaire est l'estimation de  $\mathbf{X}$  à partir de  $\mathbf{Y}$  sous la contrainte  $\mathbf{X} \geq 0$  (au sens où tous les éléments de  $\mathbf{X}$  sont positifs).

L'approche paramétrique est basée sur l'hypothèse peu réaliste que le spectre peut être représenté par une somme d'impulsions. L'utilisation d'une fonction de base plus élaborée, par exemple une gaussienne, est rendue difficile par l'absence d'expression analytique de son intégrale de Fredholm (Bose *et al.*, 2003). Nous optons donc dans cette thèse pour une résolution par l'approche non paramétrique. Une difficulté majeure concerne la dimension du problème. En effet, les méthodes standard d'inversion requièrent la réécriture du modèle direct sous la forme

$$\mathbf{y} = \mathbf{K} \mathbf{x} + \boldsymbol{\epsilon}$$

avec  $\mathbf{y} = \text{vect}(\mathbf{Y})$ ,  $\mathbf{x} = \text{vect}(\mathbf{X})$ ,  $\boldsymbol{\epsilon} = \text{vect}(\mathbf{E})$  et

$$\mathbf{K} = \mathbf{K}_1 \otimes \mathbf{K}_2 \quad (9.10)$$

où l'opérateur  $\text{vect}$  correspond à la transformation matrice-vecteur dans l'ordre lexicographique (c'est-à-dire par concaténation des éléments de la matrice parcourue ligne par ligne) et  $\otimes$  le produit de Kronecker (Van Loan, 2000). Néanmoins, ces méthodes ne sont plus applicables dès lors que la taille du problème augmente. En effet, une configuration typique est  $m_1 = 50$ ,  $m_2 = 10^4$  et  $n_1 \times n_2 = 200 \times 200$ , ce qui conduit à une matrice  $\mathbf{K}$  de  $2 \cdot 10^{10}$  éléments dont la construction et le stockage sont presque impossibles. Une solution consiste à réduire la taille des données par une étape de compression (Venkataramanan *et al.*, 2002), avec le risque de perdre des informations.

Une de nos contributions est d'utiliser la forme factorisée (9.9) pour résoudre le problème afin d'éviter de prétraiter les données.

Une autre difficulté est liée au mauvais conditionnement du modèle à inverser. En effet, à cause de la forme exponentielle des noyaux, les matrices  $\mathbf{K}_1$  et  $\mathbf{K}_2$  sont de rang déficient et très mal conditionnées (Butler *et al.*, 1981). Par conséquent, une approche régularisée doit être envisagée.

## 9.3 Méthodes d'inversion pénalisée existantes

Dans cette section, nous présentons deux stratégies qui ont été employées pour résoudre le problème inverse de la détermination de  $\mathbf{X}$  connaissant  $\mathbf{Y}$ ,  $\mathbf{K}_1$  et  $\mathbf{K}_2$ . Chacunes basées sur leur propre principe de régularisation, elles ont pour objectif de compenser le caractère mal posé du modèle d'observation.

### 9.3.1 Méthode de Tikhonov sous contraintes de positivité

La méthode employée dans (Butler *et al.*, 1981; Lamanna, 2005; Venkataramanan *et al.*, 2002) repose sur la minimisation contrainte de la fonction de Tikhonov

$$\min_{\mathbf{X} \geq 0} F(\mathbf{X}) = J(\mathbf{X}) + \lambda R(\mathbf{X}) \quad (9.11)$$

où  $J$  est le critère des moindres carrés

$$J(\mathbf{X}) = \frac{1}{2} \|\mathbf{y} - \mathbf{K}\mathbf{x}\|^2 = \frac{1}{2} \|\mathbf{Y} - \mathbf{K}_1 \mathbf{X} \mathbf{K}_2^t\|_F^2,$$

et  $R$  est le carré de la norme  $\ell_2$  du spectre

$$R(\mathbf{X}) = \frac{1}{2} \|\mathbf{x}\|^2 = \frac{1}{2} \|\mathbf{X}\|_F^2. \quad (9.12)$$

La vérification de la contrainte de positivité est désirable du point de vue de l'interprétation physique, mais a aussi un effet favorable sur la résolution du spectre reconstruit. En particulier, la contrainte de positivité permet de limiter la présence d'oscillations, en particulier autour des pics reconstruits (Chiang *et al.*, 2005). Cependant, à cause de la présence de la contrainte, il n'existe pas d'expression directe du minimiseur et la solution doit être calculée de façon itérative.

L'algorithme de Butler-Reeds-Dawson (BRD) (Butler *et al.*, 1981) permet de résoudre simplement (9.11) en se basant sur les conditions d'optimalité de ce problème. Il s'agit d'un algorithme de Newton appliqué à la fonction

$$L(\mathbf{z}) = \frac{1}{2} \mathbf{z}^T (G(\mathbf{z}) + \lambda \mathbf{I}) \mathbf{z} - \mathbf{z}^T \mathbf{y}, \quad (9.13)$$

où la variable  $\mathbf{z}$  est liée à  $\mathbf{x}$  à travers la relation

$$\mathbf{x} = \max(\mathbf{0}, \mathbf{K}^T \mathbf{z}), \quad (9.14)$$

et où  $G(\mathbf{z}) = \mathbf{K}^T \text{Diag}(\mathbb{H}(\mathbf{K}^T \mathbf{z})) \mathbf{K}$  avec

$$\mathbb{H}(u) = \begin{cases} u & \text{si } u > 0, \\ 0 & \text{sinon.} \end{cases} \quad (9.15)$$

L'algorithme BRD s'interprète également comme la minimisation de la fonction duale de  $F$ , au sens de la dualité Legendre-Fenchel (Chouzenoux *et al.*, 2010b, App.A).

Un inconvénient de cet algorithme est que son implémentation nécessite l'inversion d'un système de taille  $m \times m$  à chaque itération, où  $m$  est la taille des mesures. Dans le cas de la RMN 2D,  $m = m_1 m_2$ , et les valeurs usuelles de  $m_1$  et de  $m_2$  entraînent un coût prohibitif de la méthode. Afin de pallier ce problème, une étape de compression des données est proposée dans (Venkataramanan

*et al.*, 2002), effectuée préalablement à l'application de l'algorithme BRD. A partir des TSVD de  $\mathbf{K}_1$  et  $\mathbf{K}_2$  :

$$\mathbf{K}_i \approx \mathbf{U}_i \boldsymbol{\Sigma}_i \mathbf{V}_i^t, \quad i = 1, 2, \quad \text{avec} \quad \tilde{m}_i = \text{rang}(\mathbf{K}_i) \ll m_i, \quad (9.16)$$

le terme de fidélité aux données est approché par

$$\tilde{J}(\mathbf{X}) = \frac{1}{2} \|\tilde{\mathbf{Y}} - \tilde{\mathbf{K}}_1 \mathbf{X} \tilde{\mathbf{K}}_2^t\|_F^2, \quad (9.17)$$

où  $\tilde{\mathbf{K}}_1 = \boldsymbol{\Sigma}_1 \mathbf{V}_1^t$ ,  $\tilde{\mathbf{K}}_2 = \boldsymbol{\Sigma}_2 \mathbf{V}_2^t$  et  $\tilde{\mathbf{Y}} = \mathbf{U}_1^t \mathbf{Y} \mathbf{U}_2$  sont de taille  $\tilde{m}_1 \times n_1$ ,  $\tilde{m}_2 \times n_2$  et  $\tilde{m}_1 \times \tilde{m}_2$ , respectivement. Le réglage des niveaux de compression  $\tilde{m}_1$  et  $\tilde{m}_2$  dépend des propriétés spectrales des matrices  $\mathbf{K}_1$  et  $\mathbf{K}_2$ . Il doit résulter d'un compromis entre un gain de temps de calcul et une perte de résolution de la reconstruction.

### 9.3.2 Reconstruction par maximum d'entropie

L'utilisation d'une méthode itérative de résolution permet une plus grande souplesse pour le choix de la fonction de régularisation dans le critère (9.11). La pénalisation par maximum d'entropie (ME) (Eggermont, 1993; Skilling et Bryan, 1984), reposant sur l'entropie de Shannon, est reconnue comme une stratégie efficace pour la reconstruction RMN 1D (Laue *et al.*, 1985; Mariette *et al.*, 1996). Dans le cas 2D, elle s'écrit

$$R(\mathbf{X}) = \sum_{ij} X_{ij} \log X_{ij}. \quad (9.18)$$

La norme du gradient de  $R$  étant infinie sur le bord de l'orthant positif, la pénalisation entropique va assurer la positivité de la solution de façon implicite. La minimisation du critère pénalisé peut donc être effectuée sans tenir compte de la contrainte. Cependant, à cause de la propriété de barrière de la pénalisation entropique, les algorithmes classiques d'optimisation non linéaire s'avèrent extrêmement lents et ont été rapidement remplacés par des méthodes dédiées à la forme spécifique du critère.

#### 9.3.2.1 Résolution sous contraintes d'égalité

Les algorithmes initialement développés pour le problème du maximum d'entropie résolvaient le problème contraint

$$\min_{\mathbf{x} \in \mathbb{R}^n} \left( R(\mathbf{x}) = \sum_{i=1}^n x_i \log x_i \right) \quad \text{s.c.} \quad \mathbf{y} = \mathbf{K} \mathbf{x} \quad (9.19)$$

Deux classes d'algorithmes itératifs ont été principalement proposées pour résoudre (9.19) : La méthode affine primale ('affine scaling') (Gonzaga et Carlos, 1990; Gorodnitsky, 1999; Shi, 1995) et la reconstruction algébrique multiplicative (MART) (Byrne, 1996; Censor, 1983; Willis *et al.*, 2000).

**Méthode affine primale.** La stratégie affine primale consiste à déplacer l'itéré  $\mathbf{x}_k \geq \mathbf{0}$  dans la direction de gradient projeté calculée dans un espace transformé. La transformation est construite afin de maximiser la valeur de pas réalisable le long de la direction, sans sortir de l'orthant positif. L'algorithme comprend les trois étapes suivantes : le calcul de la transformation affine, le calcul de la direction et le calcul du pas. En combinant ces trois étapes, la récurrence s'écrit (Gonzaga et Carlos, 1990)

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k \quad (9.20)$$

avec  $\mathbf{d}_k$  la direction

$$\mathbf{d}_k = -\mathbf{D}_k \mathbf{P}_{\mathbf{K} \mathbf{D}_k} (\mathbf{D}_k \nabla R(\mathbf{x}_k)), \quad \mathbf{D}_k = \text{Diag}(\mathbf{x}_k) \quad (9.21)$$

où  $\mathbf{P}_M(\mathbf{v})$  est l'opérateur de projection

$$\mathbf{P}_M(\mathbf{v}) = (\mathbf{I} - \mathbf{M}^T(\mathbf{M}\mathbf{M}^T)^{-1}\mathbf{M})\mathbf{v}. \quad (9.22)$$

Le pas est choisi afin de maintenir le nouveau point  $\mathbf{x}_{k+1}$  à l'intérieur de l'orthant positif.

**Reconstruction algébrique multiplicative.** L'algorithme MART (Censor, 1983; Elfving, 1980) et ses variantes SMART et BI-SMART (Byrne, 1993, 1996; Trummer, 1983) appartiennent à la famille des méthodes par 'row action', méthodes itératives très simples dans laquelle la mise à jour des itérés correspond à un parcours des lignes de la matrice  $\mathbf{K}$ . Ces algorithmes sont construits sous l'hypothèse forte

$$y_j > 0, K_{ij} \geq 0, s_i = \sum_j K_{ji} > 0 \quad \forall i, j. \quad (9.23)$$

La forme la plus générale BI-SMART correspond à une mise à jour par blocs de lignes d'indices  $\mathcal{I}_k$ , tels que l'union des  $\mathcal{I}_k$  forme l'ensemble complet  $\{1, \dots, m\}$ . Partant d'un itéré  $\mathbf{x}_k$  et d'un bloc d'indices  $\mathcal{I}_k$ , le nouvel itéré  $\mathbf{x}_{k+1}$  s'écrit

$$x_{i,k+1} = x_{i,k} \exp \left( \gamma_i \delta_k \sum_{j \in \mathcal{I}_k} \alpha_{kj} K_{ji} \log \frac{y_j}{[\mathbf{K}\mathbf{x}_k]_j} \right), \quad (9.24)$$

où  $\gamma_i$ ,  $\delta_k$  et  $\alpha_{kj}$  sont des coefficients positifs. La convergence de ces algorithmes est établie dans (Byrne, 1996) sous l'hypothèse qu'il existe un  $\mathbf{x} \geq \mathbf{0}$  tel que  $\mathbf{y} = \mathbf{K}\mathbf{x}$ . La solution obtenue par BI-SMART dépend du choix de  $\mathbf{x}_0$ . La solution du problème (9.19) s'obtient en choisissant  $\mathbf{x}_0 = \exp(-1)\mathbf{e}$ , en notant  $\mathbf{e} = (1, \dots, 1)^T$  (Censor, 1983).

L'algorithme général BI-SMART est plus couramment utilisé sous les formulations particulières : MART et SMART. Ces algorithmes très simples sont définis par :

MART : Mise à jour séquentielle

$\mathcal{I}_k = \{j \bmod m\}$ ,  $\gamma_i = \delta_k = 1$  et  $\alpha_{kj} = m_j^{-1}$  où le coefficient de relaxation  $m_j$  est tel que

$$m_j^{-1} K_{ji} \leq 1 \quad (9.25)$$

SMART : Mise à jour simultanée

$\mathcal{I}_k = \{1, \dots, m\}$ ,  $\alpha_{kj} = \delta_k = 1$  et  $\gamma_i = s_i^{-1}$ .

L'algorithme MART s'interprète également comme une méthode de point fixe appliquée à l'optimisation de la fonction duale du problème (9.19) (au sens de la dualité de Lagrange) (Lent, 1991; Willis, 2000).

Les méthodes de 'row-action' nécessitant un parcours des lignes de la matrice  $\mathbf{K}$ , elles sont difficiles à mettre en œuvre pour la résolution de problèmes de grande taille utilisant des modèles implicites dont on sait uniquement calculer la projection  $\mathbf{K}\mathbf{u}$  et la rétroprojection  $\mathbf{K}^T\mathbf{v}$ . C'est le cas par exemple lorsque  $\mathbf{K}$  est opérateur de convolution. Une stratégie d'implémentation efficace de MART est décrite dans (Willis, 2000) lorsque  $\mathbf{K}$  est une matrice de convolution 2D ayant une structure bloc Toeplitz.

### 9.3.2.2 Résolution pénalisée

La forme très simple des algorithmes existants pour l'optimisation ME contrainte est attrayante mais, comme le soulignent (Johnson et McGarry, 2003), la présence de bruit dans les mesures rend préférable une résolution basée sur la minimisation du critère pénalisé

$$F(\mathbf{x}) = \|\mathbf{y} - \mathbf{K}\mathbf{x}\|^2 + \lambda \sum_i x_i \log x_i. \quad (9.26)$$

La minimisation de (9.26) peut être réalisée par une stratégie usuelle d'optimisation itérative. Dans le cadre spécifique du maximum d'entropie, Skilling et Bryan (Skilling et Bryan, 1984) ont

proposé un algorithme de sous-espace de la forme (3.43). Le sous-espace  $\mathbf{D}_k$  contient  $r \geq 3$  vecteurs  $\{\mathbf{d}_k^1 \dots \mathbf{d}_k^r\}$  définis par

$$\begin{cases} \mathbf{d}_k^1 = \mathbf{M}_k \nabla J(\mathbf{x}_k) \\ \mathbf{d}_k^2 = \mathbf{M}_k \nabla R(\mathbf{x}_k) \\ \mathbf{d}_k^j = \nabla^2 J(\mathbf{x}_k) \mathbf{M}_k \left( \frac{\mathbf{d}_k^{(j-2)}}{\|\mathbf{d}_k^{(j-2)}\|_{\mathbf{x}_k}} - \frac{\mathbf{d}_k^{j-1}}{\|\mathbf{d}_k^{j-1}\|_{\mathbf{x}_k}} \right), \quad j = 3, \dots, r \end{cases} \quad (9.27)$$

en introduisant les coefficients de normalisation

$$\|\mathbf{d}^j\|_{\mathbf{x}} = \left( \sum_{i=1}^n x_i (\mathbf{d}_i^j)^2 \right)^{1/2} \quad (9.28)$$

et le préconditionneur  $\mathbf{M}_k = \text{Diag}(\mathbf{x}_k)$ . Ce choix de préconditionneur permet de favoriser la positivité des itérés (Lanteri *et al.*, 2001; Pichon et Thiébaud, 1998). Le pas  $\mathbf{s}_k$  résulte d'une itération de Newton appliquée au critère dans le sous-espace. Il est donc solution de l'équation

$$(\mathbf{D}_k^t \mathbf{H}_k \mathbf{D}_k) \mathbf{s}_k = -\mathbf{D}_k^t \nabla F(\mathbf{x}_k), \quad (9.29)$$

avec  $\mathbf{H}_k \triangleq \nabla^2 F(\mathbf{x}_k)$ . Enfin, afin de préserver la positivité des itérés, une opération de projection sur l'orthant positif est appliquée :

$$\mathbf{x}_{k+1} = \max(\mathbf{x}_k + \mathbf{D}_k \mathbf{s}_k, \epsilon) \quad (9.30)$$

où  $\epsilon$  est une petite valeur strictement positive.

D'après (Press *et al.*, 1992, p. 1022), la convergence de cet algorithme n'est pas établie. En particulier, la recherche de pas par Newton et l'opération de projection peuvent entraîner une croissance du critère entre deux itérés consécutifs. Nous avons proposé dans (Chouzenoux *et al.*, 2009b) une modification de l'algorithme permettant de garantir la décroissance du critère. Notre stratégie est basée sur l'ajout d'une recherche de pas scalaire le long de la direction  $\mathbf{x}_{k+1} - \mathbf{x}_k$ . Cependant, cette modification ne permet toujours pas d'assurer la convergence des itérés en pratique, au sens où le critère d'arrêt portant sur la norme du gradient n'est jamais atteint. Nous proposons donc dans ce chapitre une nouvelle stratégie de minimisation du critère pénalisé ME (9.26).

## 9.4 Algorithme de minimisation ME proposé

### 9.4.1 Description de l'algorithme

La méthode proposée est un algorithme de Newton tronqué (TN) (Dembo *et al.*, 1982; Nash, 2000). L'algorithme TN est basé sur le schéma de descente itérative

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k, \quad (9.31)$$

où la direction de recherche  $\mathbf{d}_k$  est obtenue par résolution approchée des équations de Newton :

$$\mathbf{H}_k \mathbf{d}_k = -\mathbf{g}_k. \quad (9.32)$$

Cet algorithme est fréquemment utilisé dans le contexte des algorithmes de points intérieurs utilisant des barrières logarithmiques (Bellavia, 1998; Nash et Sofer, 1998) et entropiques (Johnson et McGarry, 2003). En pratique, la direction  $\mathbf{d}_k$  résulte d'itérations de gradient conjugué sur le système (9.32) arrêtées avant convergence. Le test d'arrêt est (Nash, 2000) :

$$\|\mathbf{g}_k + \mathbf{H}_k \mathbf{d}_k\| \leq \eta \|F(\mathbf{x}_k)\|. \quad (9.33)$$

D'autre part, le pas  $\alpha_k$  est obtenu par une méthode de recherche linéaire appliquée à la fonction scalaire  $f(\alpha) = F(\mathbf{x}_k + \alpha \mathbf{d}_k)$ .

Nous proposons d'adapter cet algorithme à la minimisation de (9.26) pour la RMN 2D à travers les modifications suivantes : La présence de la barrière entropique est prise en compte dans le choix de la stratégie de pas. L'effet du mauvais conditionnement de la matrice  $\mathbf{K}$  et du Hessien du terme entropique est réduit par l'ajout d'un préconditionneur. Enfin, le coût de calcul de l'algorithme lié à la grande taille du problème est réduit par une stratégie d'implémentation originale.

#### 9.4.1.1 Recherche de pas MM log-quadratique

La fonction entropie appartenant à la classe des fonctions barrières, nous proposons d'utiliser la recherche de pas MMLQ 1D décrite au chapitre 7 pour la détermination de  $\alpha_k$  :

$$\begin{cases} \alpha_k^0 = 0, \\ \alpha_k^{j+1} = \arg \min_{\alpha} h(\alpha, \alpha_k^j), \quad j < J, \\ \alpha_k = \alpha_k^J. \end{cases} \quad (9.34)$$

La fonction majorante  $h(\cdot, \alpha_k^j)$  est construite à partir de la procédure décrite dans le théorème 7.1 en identifiant  $P$  avec le terme des moindres carrés,  $\mu$  avec  $\lambda$  et  $B$  avec la pénalisation entropique. La fonction  $P$  est quadratique donc

$$m_p^j = \mathbf{d}_k^t \mathbf{K}^t \mathbf{K} \mathbf{d}_k. \quad (9.35)$$

On en déduit les valeurs des paramètres  $\bar{\alpha}^j$ ,  $m^j$ , et  $\gamma^j$  de la fonction majorante. On notera  $\phi_i(\alpha) = d_{k,i}^2 / (s_i + \alpha d_{k,i})$  dans les deux cas.

**Cas**  $\alpha \in (\alpha_-; \alpha^j]$

$$\begin{cases} \bar{\alpha}^j = \alpha_- \\ m^j = \mathbf{d}_k^t \mathbf{K}^t \mathbf{K} \mathbf{d}_k + \lambda \sum_{i|d_{k,i} < 0} \phi_i(\alpha^j) \\ \gamma^j = \lambda (\alpha_- - \alpha^j) \sum_{i|d_{k,i} > 0} \phi_i(\alpha^j) \end{cases} \quad (9.36)$$

**Cas**  $\alpha \in [\alpha^j; \alpha_+)$

$$\begin{cases} \bar{\alpha}^j = \alpha_+ \\ m^j = \mathbf{d}_k^t \mathbf{K}^t \mathbf{K} \mathbf{d}_k + \lambda \sum_{i|d_{k,i} > 0} \phi_i(\alpha^j) \\ \gamma^j = \lambda (\alpha_+ - \alpha^j) \sum_{i|d_{k,i} < 0} \phi_i(\alpha^j) \end{cases} \quad (9.37)$$

La mise en œuvre de la recherche de pas MM nécessite le réglage du paramètre  $J$ . L'influence de ce paramètre sur les performances de l'algorithme est discutée dans (Chouzenoux *et al.*, 2010b,c) à partir de reconstructions sur deux jeux de données synthétiques. Une conclusion identique à celle effectuée dans le chapitre 8 dans le cadre des algorithmes de points intérieurs est obtenue, à savoir que les meilleures performances de l'algorithme en terme de nombre d'itérations et de temps de calcul sont réalisées lorsque l'on fait une seule itération de recherche de pas MMLQ ( $J = 1$ ).

#### 9.4.1.2 Préconditionnement

Comme le souligne (Hansen, 1998), le Hessien du terme entropique est très mal conditionné pour des valeurs proches de la frontière de l'orthant positif, car certaines de ces valeurs propres tendent vers l'infini. De plus, les termes exponentiels dans les noyaux  $k_1$  et  $k_2$  impliquent que  $\mathbf{K}_1$  et  $\mathbf{K}_2$  sont aussi très mal conditionnés. L'utilisation d'un préconditionneur  $\mathbf{M}_k$  approchant l'inverse de la matrice  $\mathbf{H}_k$  dans l'étape de gradient conjugué est nécessaire pour avoir une convergence rapide

de l'algorithme. Dans le contexte de l'optimisation ME, (Johnson et McGarry, 2003) utilise une matrice diagonale  $\mathbf{M}_k$  définie à partir des éléments diagonaux du Hessien :

$$\mathbf{M}_k = [\text{Diag}(\text{diag}(\mathbf{K}^t \mathbf{K})) + \lambda \text{Diag}(\mathbf{x}_k)^{-1}]^{-1} \quad (9.38)$$

La méthode de préconditionnement que nous proposons se base sur la décomposition en valeurs singulières de  $\mathbf{K}$ . D'après (9.10), la SVD de  $\mathbf{K}$  est donnée par  $\mathbf{K} = \mathbf{U}^t \mathbf{\Sigma} \mathbf{V}$ , avec  $\mathbf{U} = \mathbf{U}_1 \otimes \mathbf{U}_2$ ,  $\mathbf{V} = \mathbf{V}_1 \otimes \mathbf{V}_2$ ,  $\mathbf{\Sigma} = \mathbf{\Sigma}_1 \otimes \mathbf{\Sigma}_2$ ,  $\mathbf{U}_i^t \mathbf{\Sigma}_i \mathbf{V}_i$  étant les SVDs de  $\mathbf{K}_i$ ,  $i = 1, 2$  (Laub, 2004, Chap.13), (Van Loan, 2000). Définissons

$$\mathbf{M}_k = [\tilde{\mathbf{V}} \tilde{\mathbf{\Sigma}}^2 \tilde{\mathbf{V}}^t + \lambda \text{Diag}(\mathbf{x}_k)^{-1}]^{-1} \quad (9.39)$$

où  $\tilde{\mathbf{V}}$  et  $\tilde{\mathbf{\Sigma}}$  correspondent aux versions tronquées de  $\mathbf{V}$  et  $\mathbf{\Sigma}$ . Si l'on n'effectue pas de troncature,  $\tilde{\mathbf{V}} = \mathbf{V}$  et  $\tilde{\mathbf{\Sigma}} = \mathbf{\Sigma}$ , et  $\mathbf{M}_k$  s'identifie alors avec le Hessien de  $F$  en  $\mathbf{s}_k$ . Dans le cas contraire,  $\mathbf{M}_k$  constitue une approximation de rang faible de  $\mathbf{H}_k$ . La troncature de la SVD de  $\mathbf{K}$  est obtenue par une technique inspirée de (Venkataramanan et al., 2002) : On calcule tout d'abord les TSVD de  $\mathbf{K}_1$  et  $\mathbf{K}_2$  de rangs respectifs  $v_1$  et  $v_2$  puis on définit  $\tilde{\mathbf{V}}$  et  $\tilde{\mathbf{\Sigma}}$  par

$$\tilde{\mathbf{V}} = \tilde{\mathbf{V}}_1 \otimes \tilde{\mathbf{V}}_2, \quad (9.40)$$

$$\tilde{\mathbf{\Sigma}} = \tilde{\mathbf{\Sigma}}_1 \otimes \tilde{\mathbf{\Sigma}}_2. \quad (9.41)$$

L'approximation de  $\mathbf{K}$  qui résulte de cette technique est légèrement différente de la TSVD de  $\mathbf{K}$ . En effet, bien que  $\tilde{\mathbf{\Sigma}}_1$  et  $\tilde{\mathbf{\Sigma}}_2$  contiennent chacune les plus grandes valeurs singulières de  $\mathbf{\Sigma}_1$  et  $\mathbf{\Sigma}_2$ ,  $\tilde{\mathbf{\Sigma}}$  ne contient pas nécessairement les plus grandes valeurs singulières de  $\mathbf{\Sigma}$ . Par conséquent, notre approximation peut être sous-optimale par rapport à la TSVD, cette dernière étant optimale au sens des moindres carrés (Eckart et Young, 1936), mais le fait d'obtenir des expressions factorisées pour les matrices  $\tilde{\mathbf{V}}$  et  $\tilde{\mathbf{\Sigma}}$  est essentiel pour conserver un coût de calcul réduit.

Le choix des paramètres  $v_1$ ,  $v_2$  influe sur les performances de l'algorithme comme nous l'avons montré dans (Chouzenoux et al., 2010b,c). Augmenter  $v_i$  améliore les performances de l'algorithme en terme de nombre d'itérations car le préconditionneur est plus proche de l'inverse du Hessien. Par contre, le coût de calcul par itération augmente. Le réglage optimal dépend de la répartition des valeurs propres des matrices  $\mathbf{K}_i$ . Sur les données traitées dans (Chouzenoux et al., 2010b,c), un bon compromis est réalisé pour  $4 \leq v_i \leq 8$ .

L'intégration du préconditionneur dans l'algorithme de gradient conjugué est décrite dans l'algorithme 9.1, noté GCP.

**Entrées :**  $\mathbf{g}_k, \mathbf{H}_k, \mathbf{M}_k$   
**Sorties :** Solution approchée  $\mathbf{d}_k$  de (9.32)

$\mathbf{u}_0 \leftarrow \mathbf{0}$   
 $\mathbf{r}_0 \leftarrow -\mathbf{g}_k - \mathbf{H}_k \mathbf{u}_0$   
 $\mathbf{p}_0 \leftarrow \mathbf{M}_k \mathbf{r}_0$

**tantque** (9.33) n'est pas vérifié **faire**

$\theta_i \leftarrow (\mathbf{r}_i^t \mathbf{M}_k \mathbf{r}_i) / (\mathbf{p}_i^t \mathbf{H}_k \mathbf{p}_i)$   
 $\mathbf{u}_{i+1} \leftarrow \mathbf{u}_i + \theta_i \mathbf{p}_i$   
 $\mathbf{r}_{i+1} \leftarrow \mathbf{r}_i - \theta_i \mathbf{H}_k \mathbf{p}_i$   
 $\beta_i \leftarrow (\mathbf{r}_{i+1}^t \mathbf{M}_k \mathbf{r}_{i+1}) / (\mathbf{r}_i^t \mathbf{M}_k \mathbf{r}_i)$   
 $\mathbf{p}_{i+1} \leftarrow \mathbf{M}_k \mathbf{r}_{i+1} + \beta_i \mathbf{p}_i$   
 $\mathbf{d}_k \leftarrow \mathbf{u}_{i+1}$

**fin tantque**

Algorithme 9.1 – Algorithme de gradient conjugué préconditionné (GCP)

### 9.4.1.3 Critères d'arrêt

La résolution de (9.26) est effectuée par l'algorithme de Newton tronqué 9.2. Nous utilisons le critère d'arrêt global (Nocedal et Wright, 1999) :

$$\|\mathbf{g}_k\|_\infty < 10^{-8}(1 + |F(\mathbf{x}_k)|). \quad (9.42)$$

Le critère d'arrêt des itérations de GCP dépend du paramètre de tolérance  $\eta$ . Ce paramètre influence sur le nombre de sous-itérations GCP nécessaires pour déterminer la direction de Newton tronquée. Le réglage optimal de  $\eta$  résulte d'un compromis entre les nombres de sous itérations de GCP et d'itérations globales. Nous fixerons par la suite  $\eta = 10^{-4}$  car ce réglage fournit les meilleures performances en terme de temps de calcul dans les expérimentations menées dans (Chouzenoux et al., 2010b,c).

**Entrées :** Valeur initiale  $\mathbf{x}_0 \succeq 0$ , paramètres  $v_1, v_2, J, \lambda, \gamma$  et précisions  $\epsilon, \eta$ .  
**Sorties :** Résolution de (9.26)  
 Calculer les TSVDs de  $\mathbf{K}_1$  et de  $\mathbf{K}_2$  de rangs respectifs  $v_1$  et  $v_2$ .  
**tantque** (9.42) n'est pas vérifié **faire**  
     Calculer  $\mathbf{g}_k, \mathbf{M}_k$  et  $\mathbf{H}_k$ .  
     Calculer  $\mathbf{d}_k$  par l'algorithme 9.1.  
     Déterminer  $\alpha_k$  après  $J$  itérations de (9.34).  
     Mettre à jour  $\mathbf{x}_k$  selon (9.31).  
**fin tantque**

Algorithme 9.2 – Algorithme TN pour le problème ME

## 9.4.2 Réduction du coût de calcul

Nous montrons dans cette section que le coût de calcul de chaque itération de l'algorithme de reconstruction 9.2 peut être réduit significativement en exploitant la forme factorisée de la matrice d'observation.

Trois opérations majeures sont effectuées à chaque itération : le calcul du gradient  $\mathbf{g}_k = \nabla F(\mathbf{x}_k)$  et les produits de  $\mathbf{M}_k$  et  $\mathbf{H}_k$  avec un vecteur. Ces trois quantités peuvent être calculées à l'aide d'opérations peu coûteuses ne nécessitant pas d'explicitement la matrice  $\mathbf{K}$  grâce à la propriété suivante du produit de Kronecker (Van Loan, 2000) :

$$(\mathbf{K}_1 \otimes \mathbf{K}_2)\mathbf{x} = \mathbf{K}_1 \mathbf{X} \mathbf{K}_2^T, \quad \text{si } \mathbf{x} = \text{vect}(\mathbf{X}). \quad (9.43)$$

### 9.4.2.1 Gradient

D'après (9.43), le gradient du critère se calcule sans expliciter la matrice  $\mathbf{K}$  :

$$\mathbf{g}_k = -\text{vect}(\mathbf{K}_1^t(\mathbf{Y} - \mathbf{K}_1 \mathbf{X}_k \mathbf{K}_2^t)\mathbf{K}_2) + \lambda(1 + \log \mathbf{x}_k). \quad (9.44)$$

### 9.4.2.2 Hessien

Pour la même raison, tout produit entre la matrice Hessienne et un vecteur  $\mathbf{w} = \text{vect}(\mathbf{W})$  s'obtient par :

$$\mathbf{H}_k \mathbf{w} = \text{vect}(\mathbf{K}_1^t \mathbf{K}_1 \mathbf{W} \mathbf{K}_2^t \mathbf{K}_2) + \lambda(\mathbf{w} ./ \mathbf{x}_k), \quad (9.45)$$

où ./ correspond à la division terme à terme.

### 9.4.2.3 Préconditionneur

Notons tout d'abord que le preconditionneur s'écrit aussi

$$\mathbf{M}_k = \mathbf{A}_k - \mathbf{A}_k \tilde{\mathbf{V}} (\tilde{\Sigma}^{-2} + \tilde{\mathbf{V}}^t \mathbf{A}_k \tilde{\mathbf{V}})^{-1} \tilde{\mathbf{V}}^t \mathbf{A}_k, \quad (9.46)$$

avec  $\mathbf{A}_k = \lambda^{-1} \text{Diag}(\mathbf{x}_k)$ , d'après le lemme d'inversion matriciel. De plus, les composantes de la matrice  $\mathbf{P} = \tilde{\mathbf{V}}^t \mathbf{A}_k \tilde{\mathbf{V}} \in \mathbb{R}^{v_1 \times v_2}$  s'écrivent sous la forme factorisée :

$$P_{ij} = \frac{1}{\lambda} \sum_{m=1}^{n_1} \sum_{n=1}^{n_2} (X_k)_{mn} (\tilde{V}_1)_{ma} (\tilde{V}_2)_{nb} (\tilde{V}_1)_{mc} (\tilde{V}_2)_{nd},$$

avec  $(a, b)$  et  $(c, d)$  les indices bi-dimensionnels correspondant respectivement aux indices scalaires  $i$  et  $j$ . Par conséquent, le produit  $\mathbf{M}_k \mathbf{w}$  s'obtient avec un coût de calcul raisonnable en utilisant l'expression

$$\begin{aligned} \mathbf{M}_k \mathbf{w} &= \mathbf{b}_k - \mathbf{A}_k \tilde{\mathbf{V}} (\tilde{\Sigma}^{-2} + \mathbf{P})^{-1} \tilde{\mathbf{V}}^t \mathbf{b}_k, \\ &= \mathbf{b}_k - \mathbf{A}_k \text{vect} (\tilde{\mathbf{V}}_1 \mathbf{Q}_k \tilde{\mathbf{V}}_2^t), \end{aligned} \quad (9.47)$$

où  $\mathbf{b}_k = \mathbf{A}_k \mathbf{w}$ ,  $\mathbf{q}_k = (\tilde{\Sigma}^{-2} + \mathbf{P})^{-1} \text{vect} (\tilde{\mathbf{V}}_1^t \mathbf{B}_k \tilde{\mathbf{V}}_2)$  et  $\mathbf{Q}_k = \text{mat}(\mathbf{q}_k)$ ,  $\mathbf{B}_k = \text{mat}(\mathbf{b}_k)$ .

## 9.5 Application à la reconstruction de données réelles

### 9.5.1 Réglage du coefficient de régularisation

Lors de la reconstruction de spectres issus de données synthétiques, une stratégie naturelle de réglage de  $\lambda$  consiste à minimiser l'erreur entre le spectre simulé et le spectre reconstruit. Cependant, dans un contexte expérimental, il est nécessaire de trouver une façon de régler automatiquement ce paramètre. Pour ce faire, de nombreuses méthodes existent (Galatsanos et Katsaggelos, 1992), telles que la validation croisée généralisée (Reeves, 1990; Reeves et Mersereau, 1990), la méthode de la courbe en L (Hansen et O'Leary, 1993; Hansen, 1992) ou la méthode du Chi-carré (ou Chi-deux) (Trussell, 1983).

Cette dernière méthode est couramment employée pour l'inversion d'intégrales de Fredholm (Butler *et al.*, 1981; Mariette *et al.*, 1996; Venkataramanan *et al.*, 2002) et pour l'optimisation ME (Pichon et Thiébaud, 1998; Skilling et Bryan, 1984). Etant données des mesures  $\mathbf{Y}$  et une estimation de l'écart type du bruit  $\hat{\sigma}$ , le Chi-carré correspond à la différence entre la valeur observée et la valeur attendue :

$$\chi^2 = \|\mathbf{K}_1 \mathbf{X} \mathbf{K}_2^t - \mathbf{Y}\|_F^2 / \hat{\sigma}^2 \quad (9.48)$$

La variable  $\chi^2$  suit une distribution normale de moyenne et de variance

$$\text{Moyenne}(\chi^2) = m_1 m_2, \quad \text{Variance}(\chi^2) = 2m_1 m_2. \quad (9.49)$$

Par conséquent, une méthode classique pour fixer  $\lambda$  est de contraindre  $\chi^2$  à sa valeur moyenne attendue

$$\chi_{\text{cible}}^2 = m_1 m_2. \quad (9.50)$$

Cependant, la contrainte (9.50) conduit souvent à une solution sur-régularisée (Galatsanos et Katsaggelos, 1992; Pichon et Thiébaud, 1998; Rice, 1986) et un meilleur résultat est obtenu en imposant

$$\chi_{\text{cible}}^2 = m_1 m_2 - \sqrt{2m_1 m_2}. \quad (9.51)$$

Dans les applications où le niveau de bruit est important ou l'estimation de  $\sigma$  imprécise, le test du Chi-carré peut être impossible à vérifier car  $\chi^2$  peut être toujours strictement supérieur à  $\chi_{\text{cible}}^2$  même pour une valeur de  $\lambda$  très petite. Dans ce cas, la méthode de la courbe en S est préconisée

(Song *et al.*, 2002) : il s'agit de choisir  $\lambda$  tel que la décroissance de la courbe en S, définie comme les variations du  $\chi^2$  en fonction de  $\lambda$  (Fig. 9.6), se stabilise :

$$\frac{\partial \log_{10} \chi^2}{\partial \log_{10} \lambda} \ll \epsilon, \quad (9.52)$$

avec  $\epsilon$  un seuil très petit à spécifier. Nous proposons de combiner les deux stratégies pour déterminer  $\lambda$ , comme le décrit l'algorithme 9.3. Notons que les minimisations (9.53) peuvent être effectuées avec un coût de calcul très faible, en initialisant l'algorithme 9.2 par la solution obtenue au précédent  $\lambda$ .

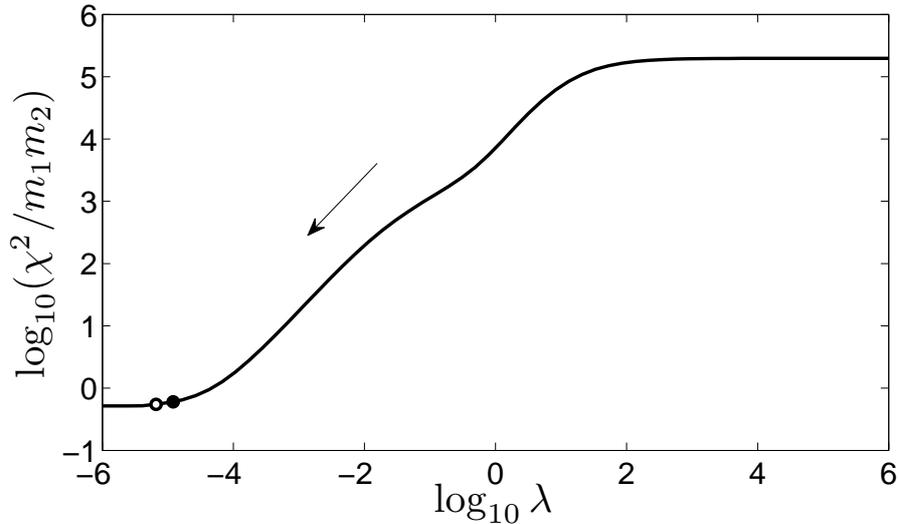


FIGURE 9.6 – Courbe en S obtenue lors de la reconstruction ME du jeu de données synthétique A dans (Chouzenoux *et al.*, 2010b). La vérification du test du Chi-carré (9.51) et du test de la courbe en S (9.52) sont illustrées respectivement par des cercles noirs et blancs. D'après l'algorithme 9.3, le paramètre  $\lambda$  retenu est celui vérifiant le test du Chi-carré.

<p><b>Entrées :</b> <math>x_0 \succeq 0</math>, <math>\lambda_0</math> et paramètre de relaxation <math>\theta \in (0, 1)</math></p> <p><b>Sorties :</b> Reconstruction ME avec réglage de <math>\lambda</math> basé sur le principe du Chi-carré</p> <p><b>tantque</b> (9.51) et (9.52) ne sont pas vérifiées <b>faire</b></p> <p style="padding-left: 2em;">Avec l'algorithme 9.2, calculer</p> $\hat{\mathbf{X}} = \arg \min_{\mathbf{X}} J(\mathbf{X}) + \lambda_n R(\mathbf{X}). \quad (9.53)$ <p style="padding-left: 2em;">Calculer <math>\chi^2(\hat{\mathbf{X}})</math> avec (9.48).</p> <p style="padding-left: 2em;"><math>\lambda_{n+1} \leftarrow \theta \lambda_n</math></p> <p><b>fin tantque</b></p>
--

Algorithme 9.3 – Méthode de la courbe en S pour l'estimation du paramètre de régularisation

### 9.5.2 Réglage du paramètre d'angle de bascule

Lors d'une expérience RMN, l'angle de bascule  $\Phi$  réellement engendré par le champ  $B_1$  peut différer de sa valeur désirée  $\hat{\Phi}$  (Song *et al.*, 2002). Par conséquent, le paramètre  $\gamma$  du modèle d'observation est entaché d'une erreur. Celle-ci a une influence importante sur les résultats de reconstruction, indépendamment de l'algorithme d'optimisation utilisé (Chouzenoux *et al.*, 2010b).

Une procédure de prétraitement des données est proposée dans (Song *et al.*, 2002) afin que celles-ci ne dépendent plus de  $\gamma$ . Elle consiste à soustraire aux données une mesure obtenue pour  $\tau_1$  très grand (typiquement,  $\tau_1 = 5 T_{1\max}$ ) :

$$Y_{\text{cor}}(\tau_1, \tau_2) = Y(\tau_1, \tau_2, \Phi) - Y(\tau_1 \rightarrow \infty, \tau_2, \Phi). \quad (9.54)$$

D'après l'expression du noyau  $k_1(\tau_1, T_1)$ , les mesures corrigées sont liées au spectre par le nouveau modèle intégral

$$Y_{\text{cor}}(\tau_1, \tau_2) = \iint e^{-\tau_1/T_1} X(T_1, T_2) e^{-\tau_2/T_2} dT_1 dT_2 \quad (9.55)$$

indépendant de  $\Phi$ . Un inconvénient de cette procédure est qu'elle ne permet pas de connaître la valeur réelle de l'angle  $\Phi$ .

Dans cette section, nous présentons une stratégie d'estimation de ce paramètre qui ne nécessite pas de modifier le modèle ni de prétraiter les données. Nous définissons tout d'abord comme *spectre de référence* le spectre mono-dimensionnel  $X(T_2)$  reconstruit à partir de données  $Y(\tau_1 \rightarrow \infty, \tau_2, \Phi)$ . Une propriété remarquable de ce spectre est qu'il ne dépend pas du réglage de  $\Phi$  puisqu'il est issu du modèle

$$Y(\tau_1 \rightarrow \infty, \tau_2, \Phi) = \int X(T_2) e^{-\tau_2/T_2} dT_2. \quad (9.56)$$

Pour un réglage exact de  $\Phi$  dans le modèle d'observation, le spectre de référence issu de l'inversion de (9.56) est identique à celui issu de la marginalisation de  $X(T_1, T_2)$  obtenu par inversion du modèle (9.3). Par conséquent, parmi plusieurs reconstructions 2D avec différentes valeurs de  $\gamma$ , nous retenons celle dont la marginalisation fournit le spectre le plus proche du spectre  $T_2$  de référence, comme l'illustre la figure 9.7. A nouveau, les reconstructions successives peuvent être effectuées avec un coût de calcul minimal, en utilisant la solution obtenue au précédent  $\gamma$  comme initialisation de l'algorithme d'optimisation.

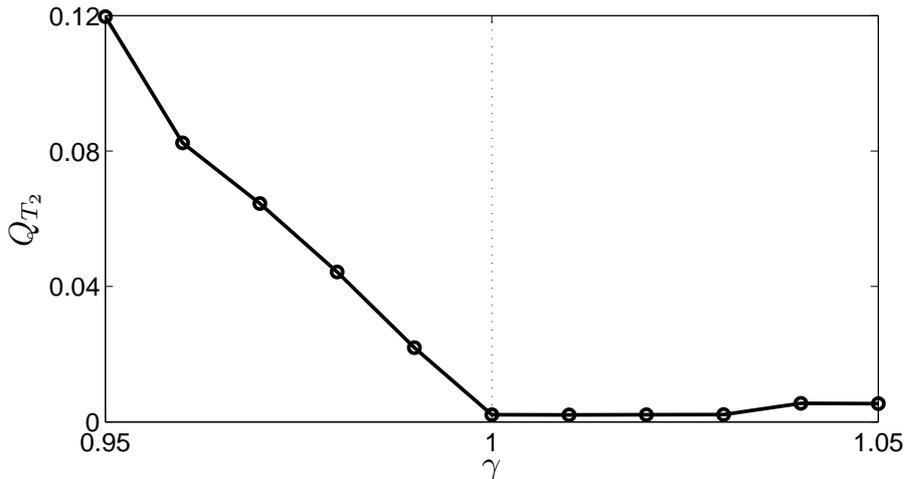


FIGURE 9.7 – Estimation de  $\gamma$  lors de la reconstruction ME du jeu de données synthétique A dans (Chouzenoux *et al.*, 2010b).  $Q_{T_2}$  est la norme  $\ell_2$  de l'erreur entre le spectre  $T_2$  marginalisé et le spectre  $T_2$  de référence. Elle est minimale lorsque  $\gamma$  est égal à sa valeur effective, i.e.  $\gamma = 1$ .

### 9.5.3 Application à l'analyse d'échantillons de pomme par RMN

Nous présentons les résultats d'estimation d'un spectre  $T_1-T_2$  d'un échantillon de pomme. Les mesures ont été effectuées pour  $m_1 = 50$  valeurs de  $\tau_1$ , réparties non uniformément entre 30 ms et 12 s. Dans tous les cas, le moment magnétique résultant est mesuré en  $m_2 = 10000$  valeurs du temps d'écho réparties uniformément entre  $600 \mu\text{s}$  et 8 s. L'algorithme 9.2 a été utilisé pour reconstruire le spectre en  $N_1 = N_2 = 200$  valeurs des temps de relaxation  $T_1$  et  $T_2$ , réparties uniformément entre 25 ms et 3 s. Les paramètres d'optimisation sont choisis afin de minimiser le temps de calcul avant convergence soit  $J = 1$ ,  $\eta = 10^{-4}$  et  $v = 7$ . D'autre part, la stratégie décrite dans l'algorithme 9.3 est utilisée pour le réglage du paramètre de régularisation. Nous présentons à présent les résultats de la procédure d'estimation de l'angle de bascule.

Les données ont été obtenues à partir d'un champ de radio-fréquence d'amplitude et de durée réglées afin de produire un angle de bascule de  $\hat{\Phi} = 90^\circ$  soit  $\hat{\gamma} = 1$ . La figure 9.8 illustre la forme des spectres reconstruits pour différentes valeurs du paramètre  $\gamma$  entre 0.9 et 1. La position et l'amplitude des pics varient fortement en fonction de  $\gamma$ . Par conséquent, une reconstruction exploitable des données nécessite de connaître de façon précise la valeur effective de ce paramètre. Nous utilisons la stratégie d'estimation décrite dans la section 9.5.2. D'après la courbe obtenue en figure 9.9, la valeur retenue correspond à un angle de bascule de  $\Phi \approx 85^\circ$  (i.e.,  $\gamma = 0.92$ ), différent de  $\hat{\Phi}$ .

La figure 9.10 présente le spectre 2D reconstruit pour cette valeur de  $\gamma$ . On remarque que la distribution présente une corrélation entre  $T_1$  et  $T_2$ , par exemple au niveau du pic localisé autour de  $[T_1 = 1.4 \text{ s}, T_2 = 0.9 \text{ s}]$ . Cette information est très utile pour l'analyse de la structure moléculaire de l'échantillon (Hürlimann et Venkataramanan, 2002). En ce qui concerne les performances de l'algorithme de reconstruction, le critère d'arrêt est atteint après 59 s et 67 itérations pour la valeur de  $\lambda = 1.3 \cdot 10^{-4}$ . La similarité entre les spectres 1D  $T_1$  et  $T_2$  reconstruits et les spectres 1D déduits des marginalisations du spectre 2D mise en évidence sur la figure 9.11 illustre la qualité de la reconstruction 2D.

Comparons à présent les spectres issus de la minimisation du critère ME avec ceux fournis par la stratégie de Tikhonov associée à des contraintes de positivités (TIK<sup>+</sup>), décrite en section 9.3.1. L'algorithme BRD est associé à la stratégie de compression décrite dans (Venkataramanan et al., 2002) avec un taux de compression fixé à  $\tilde{m}_i = 10$ . De plus, la stratégie décrite dans l'algorithme 9.3 est utilisée pour régler le paramètre de régularisation. Le spectre 2D obtenu et ses marginales sont présentés en figures 9.12 et 9.13. L'algorithme TIK<sup>+</sup> nécessite 11 s et 14 itérations pour converger, avec  $\lambda = 2 \cdot 10^{-5}$ .

Les spectres 2D issus des méthodes ME et TIK<sup>+</sup> fournissent une mesure d'attachement aux données similaire :

$$\|\mathbf{K}_1 \mathbf{X} \mathbf{K}_2 - \mathbf{Y}\|^2 / \|\mathbf{Y}\| \approx 98\%. \quad (9.57)$$

et présentent des pics aux mêmes couples  $(T_1, T_2)$ . Cependant, leur forme est différente en terme de régularité et d'amplitude. Les pics du spectre ME sont plus lisses et ont une amplitude plus élevée. De plus, ils mènent à des marginales 1D plus proches des spectres 1D obtenus par inversion.

## 9.6 Conclusion

La reconstruction d'un spectre  $T_1$ - $T_2$  en RMN nécessite l'inversion numérique d'une transformée de Laplace 2D, connue pour être un problème inverse mal posé. Dans ce chapitre, nous avons présenté une méthode de reconstruction efficace basée sur une pénalisation par maximum d'entropie et une optimisation par l'algorithme de Newton tronqué. Une difficulté importante réside dans la grande taille de la matrice intervenant dans le modèle direct. Plutôt que de réduire cette matrice par l'étape de compression proposée dans (Venkataramanan *et al.*, 2002), nous avons conservé sa forme exacte pour un coût de calcul raisonnable grâce à une implémentation de l'algorithme utilisant la séparabilité du noyau de convolution. La pénalisation entropique introduisant une barrière dans le critère, la recherche de pas MMLQ a été employée. Enfin, la convergence de l'algorithme a été accélérée grâce à un préconditionneur basé sur les TSVDs des noyaux de convolution.

L'applicabilité de l'algorithme a été testée sur des données simulées et réelles. De plus, nous avons comparé les résultats avec ceux obtenus par la méthode de Tikhonov contrainte de (Venkataramanan *et al.*, 2002). Les deux méthodes fournissent des spectres similaires. La méthode de Tikhonov contrainte est plus rapide mais elle nécessite une réduction de la taille du problème par compression des données passant par le réglage de deux paramètres.

Le traitement de données réelles nécessite d'élaborer une stratégie de réglage du coefficient de régularisation. Nous avons proposé une méthode itérative basée sur une hybridation entre le test du Chi carré et le test de la courbe en S. Les résultats présentés dans (Chouzenoux *et al.*, 2010b) se sont révélés satisfaisants à la fois sur les jeux de données simulés que sur les spectres réels. Une perspective est de comparer cette procédure avec les différentes méthodes présentées dans (Galatsanos et Katsaggelos, 1992).

La reconstruction des données réelles nous a permis de souligner l'importance d'un réglage adéquat du paramètre d'angle de bascule dans le modèle. A notre connaissance, il s'agit d'un problème rarement évoqué dans la littérature de RMN. Nous avons proposé dans ce chapitre une stratégie d'estimation de l'angle de bascule qui s'avère efficace sur nos jeux de données. Une perspective envisagée à court terme consiste à valider cette méthode et à la comparer avec la méthode de prétraitement proposée par Song *et coll.*

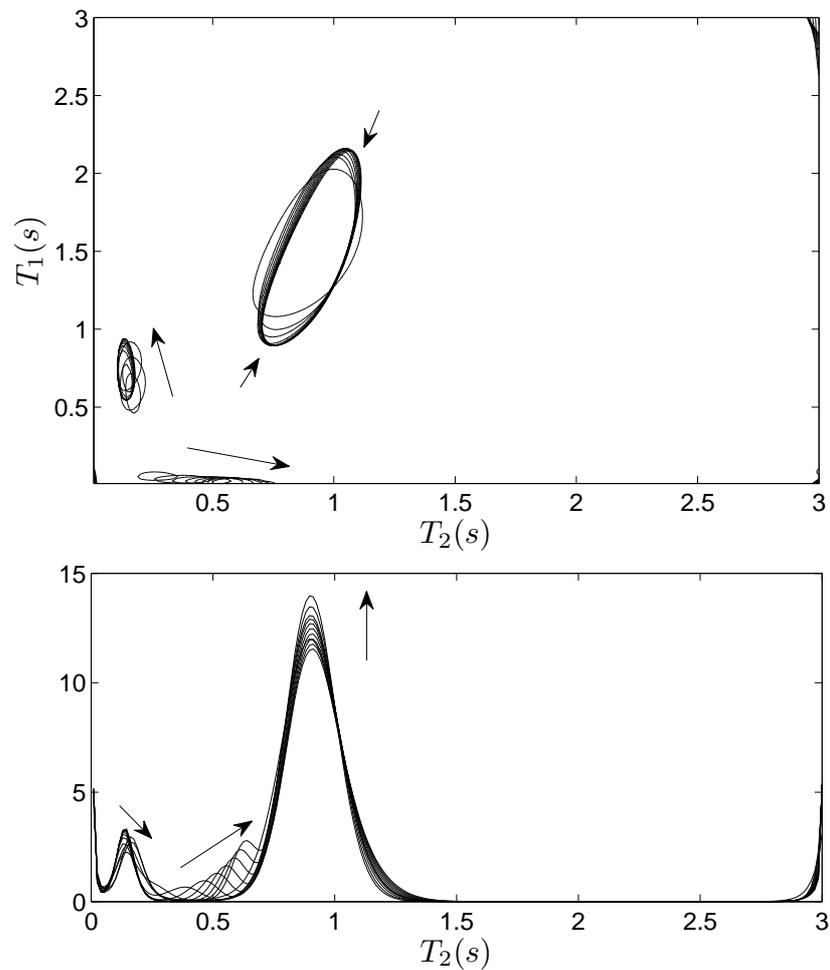


FIGURE 9.8 – Reconstruction  $T_1$ - $T_2$  (haut) à partir des données réelles et reconstruction  $T_2$  déduite par marginalisation (bas), pour différentes valeurs du paramètre  $\gamma$  dans  $[0.9, 1]$ . L'effet d'une augmentation de  $\gamma$  sur la position et l'amplitude des pics est illustré par les flèches.

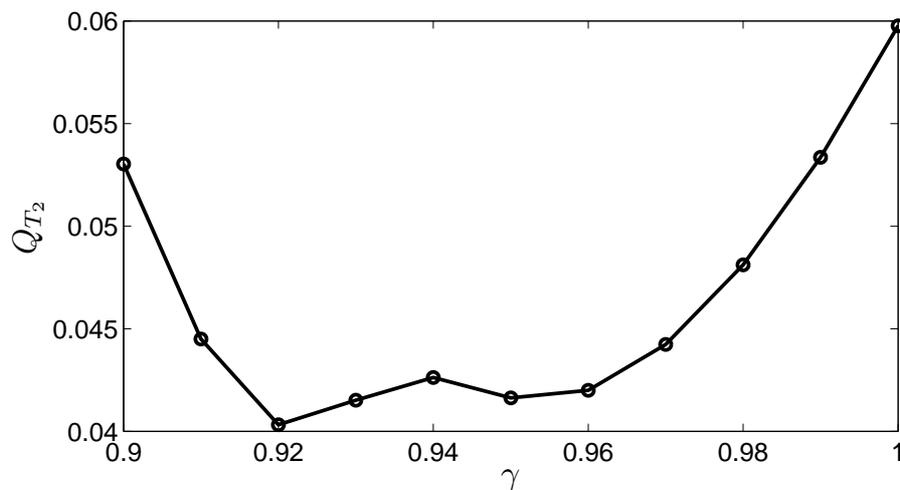


FIGURE 9.9 – Estimation de  $\gamma$  lors de la reconstruction d'un spectre issue de données réelles. L'erreur minimale entre le spectre  $T_2$  marginalisé et le spectre  $T_2$  de référence est obtenue pour  $\gamma = 0.92$  ce qui correspond à l'angle de bascule  $\Phi = 85^\circ 24'$  différent de  $\hat{\Phi} = 90^\circ$ .

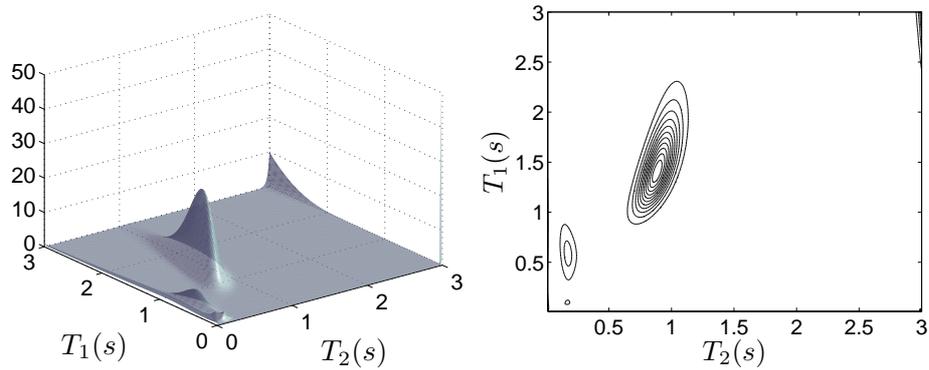


FIGURE 9.10 – Spectre RMN 2D reconstruit par la méthode ME à partir des données réelles.

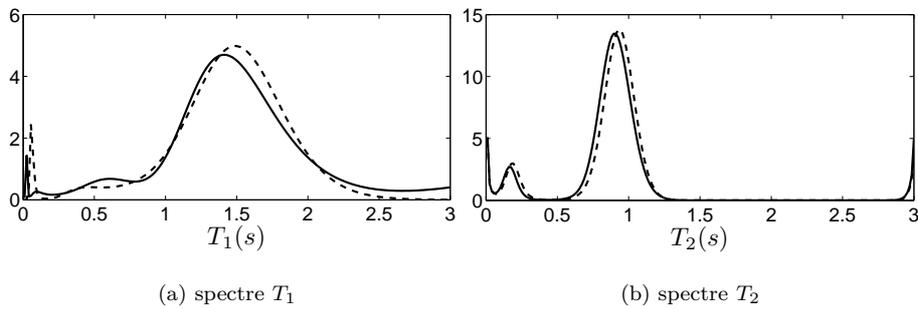


FIGURE 9.11 – Distributions 1D obtenues par marginalisation du spectre 2D reconstruit par la méthode ME (trait plein) ou par inversion (trait pointillé).

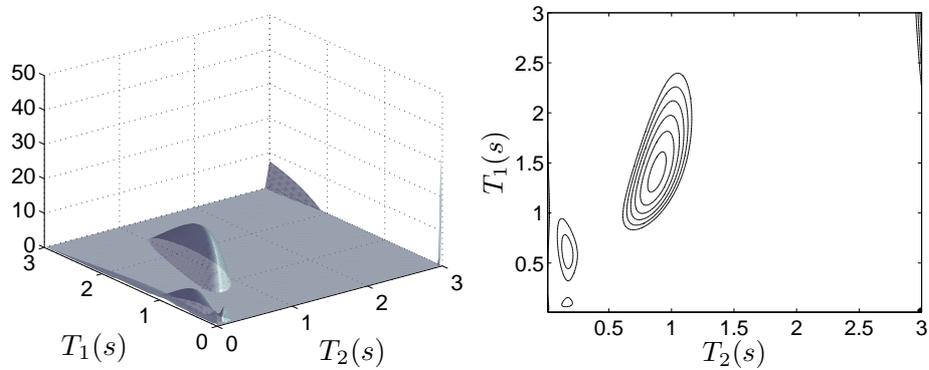


FIGURE 9.12 – Spectre RMN 2D reconstruit par la méthode TIK<sup>+</sup> à partir des données réelles.

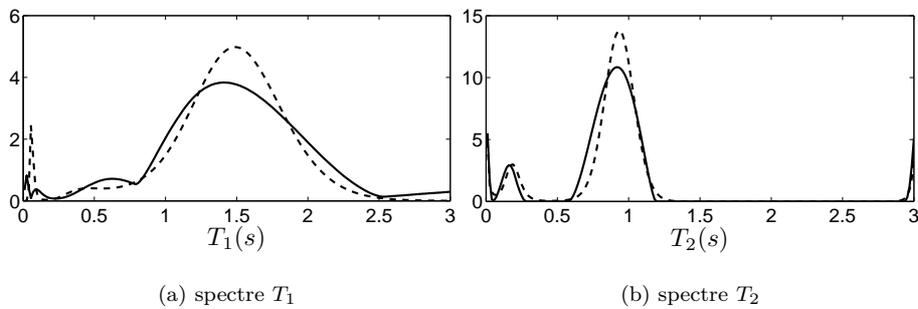


FIGURE 9.13 – Distributions 1D obtenues par marginalisation du spectre 2D reconstruit par la méthode TIK<sup>+</sup> (trait plein) ou par inversion (trait pointillé).

# Chapitre 10

## Conclusions et perspectives

Dans le cadre de ce travail de thèse, nous nous sommes intéressés à des questions algorithmiques liées à la mise en œuvre de l’approche pénalisée pour la résolution de problèmes inverses linéaires de grande taille. Ce travail de thèse nous a amené à nous intéresser de près aux stratégies de pas pour l’optimisation itérative de critères différentiables. Nous nous sommes tout d’abord intéressés à la recherche de pas par Majoration-Minoration quadratique. Nous avons ensuite proposé une nouvelle méthode de pas adaptée à la minimisation des critères à barrières. Enfin, nous avons développé un algorithme efficace de reconstruction de spectres RMN bi-dimensionnels par maximum d’entropie.

Nous concluons maintenant brièvement sur ces trois aspects et soulignons quelques perspectives qui nous paraissent intéressantes à court et à moyen terme.

### Recherche de pas par Majoration-Minoration quadratique

La stratégie de pas par Majoration-Minoration quadratique MMQ 1D développée dans la thèse de Christian Labat (Labat, 2006) pour les algorithmes de gradient conjugué non linéaire (GCNL) se distingue des méthodes standards par sa grande simplicité et ses nombreuses propriétés de convergence. Nous avons montré dans le chapitre 4 que ces propriétés s’étendaient aux cas des algorithmes de descente à gradient relié, regroupant les méthodes de Newton, quasi-Newton et Newton tronqué.

Les algorithmes de descente font en fait partie d’une classe d’algorithmes plus générale qui est celle des algorithmes de sous-espace. Dans ceux-ci, l’étape de calcul du pas nécessite d’explorer le critère à l’intérieur d’un sous-espace de faible dimension. Notre contribution est de définir une nouvelle méthode de pas pour les algorithmes de sous-espace, correspondant à une généralisation multi-dimensionnelle de la méthode MMQ 1D notée MMQ  $r$ D. Nous avons établi dans le chapitre 4 la convergence de ce schéma sous des conditions peu restrictives sur le sous-espace.

En associant le schéma de pas MMQ 2D au sous-espace de mémoire de gradient (MG), nous obtenons un algorithme similaire au GCNL et même identique à celui-ci dans le cas d’un critère quadratique. Nous avons illustré expérimentalement dans le chapitre 5 la pertinence de l’algorithme MG + MMQ 2D par rapport au schéma de descente GCNL + MMQ 1D sur un problème de déconvolution d’images. L’intérêt de l’algorithme MG + MMQ 2D est qu’il permet de s’affranchir du choix de la formule de conjugaison qui a des répercussions importantes sur les performances de l’algorithme GCNL (Hager et Zhang, 2006b). De plus, sa convergence est assurée lorsqu’un préconditionneur variable est utilisé. A l’opposé, la convergence de l’algorithme GCNL associé à un préconditionneur variable reste un sujet ouvert (Al-Baali et Fletcher, 1996; Pytlak, 2009).

**Lien avec les algorithmes de régions de confiance.** Revenons à la stratégie de pas multi-dimensionnel développée par (Shi et Xu, 2009). Dans cette dernière, le calcul du pas résulte de minimisations de l’approximation quadratique  $q_k(\mathbf{s})$  du critère  $f(\mathbf{s}) = F(\mathbf{x}_k + \mathbf{D}_k \mathbf{s})$  à l’intérieur

d'une région de confiance de taille  $\Delta_k$  conditionnée par la valeur du résidu

$$r_k = \frac{f(\mathbf{0}) - f(\mathbf{s}_k)}{f(\mathbf{0}) - q_k(\mathbf{s}_k)}. \quad (10.1)$$

Le pas est accepté lorsque le test  $r_k \geq \mu$  est vérifié, avec  $\mu \in (0, 1)$ . Sinon, la taille de la région est réduite. Supposons que  $q_k(\cdot)$  soit définie comme une approximation majorante quadratique de  $f$  en  $\mathbf{0}$ . Alors le minimiseur  $\mathbf{s}_k^*$  de  $q_k(\cdot)$  vérifie

$$f(\mathbf{0}) = q_k(\mathbf{0}) \geq q_k(\mathbf{s}_k^*) \geq f(\mathbf{s}_k^*). \quad (10.2)$$

Si  $\Delta_k$  est supérieur à  $\|\mathbf{D}_k \mathbf{s}_k^*\|$ , alors  $\mathbf{s}_k^*$  s'identifie avec la solution contrainte  $\mathbf{s}_k$  de (4.30). De plus, la propriété de majoration implique que le résidu  $r_k$  est supérieur à 1, ce qui entraîne l'acceptation du pas quelque soit  $\mu$ . Par conséquent, lorsque l'approximation quadratique  $q_k$  est choisie afin de majorer le critère, le pas obtenu par la stratégie de régions de confiance de (Shi et Xu, 2009) s'identifie à un pas MMQ rD contraint.

Il serait intéressant de voir si un tel lien existe entre stratégies de régions de confiance et de Majoration-Minoration appliquées à la minimisation du critère global  $F(\mathbf{x})$ . Une question importante serait de savoir si les bonnes propriétés des algorithmes de régions de confiance, notamment en ce qui concerne la minimisation des critères non convexes ou très mal conditionnés, se transmettent aux algorithmes MM.

**Recherche de pas et optimisation sous contraintes.** Les algorithmes de descente itérative ou plus généralement de sous-espace sont dédiés à l'optimisation non contrainte. L'utilisation de directions admissibles dans l'algorithme de descente permet toutefois de prendre en compte simplement des contraintes convexes. Les stratégies de pas MMQ et MMLQ 1D peuvent être utilisées dans ce contexte à condition de restreindre la valeur du pas dans l'intervalle  $(0, 1]$ . La convergence est assurée si la direction admissible est à gradient reliée (Bertsekas, 1999, Prop.2.2.1). C'est par exemple le cas de la direction de gradient projeté associée à l'orthant positif (Bertsekas, 1999, Sec.2.3), ou de la direction utilisée dans les algorithmes de décomposition du gradient ('gradient splitting') (Lanteri et al., 2002, 2001).

Une accélération significative de l'algorithme de directions admissibles s'obtient en réalisant une recherche de pas le long du domaine contraint (Ni et Yuan, 1997; Pytlak, 1998; Xiao et Zengxin, 2007). Cela s'exprime comme la minimisation approchée d'une fonction scalaire différentiable par morceaux. Une perspective serait de développer une stratégie de pas MM permettant de traiter ce problème non différentiable. Cela passerait par exemple par l'utilisation de fonctions majorantes ne vérifiant pas partout la condition de tangence. Une autre stratégie pourrait être de relaxer la valeur du pas afin de ne jamais arriver en un point de non différentiabilité. Une méthode similaire est utilisée dans (Fuchs, 2007) dans le cadre de la minimisation d'un critère  $\ell_2 - \ell_1$  par un algorithme semi-quadratique.

La prise en compte de contraintes dans l'algorithme de sous-espace est discutée dans (Conn et al., 1994a; Yuan, 2007). La principale difficulté réside dans la minimisation contrainte du critère dans le sous-espace. Celle-ci peut être réalisée par l'algorithme de Newton projeté (Conn et al., 1994a), par exemple celui décrit dans (Bertsekas, 1981). Lorsque le sous-espace de Quasi-Newton est employé, il en résulte un algorithme très similaire à l'algorithme L-BFGS-B de (Byrd et al., 1995). Yuan et coll. proposent d'utiliser un algorithme de régions de confiance contraint tel que celui décrit dans (Conn et al., 1988). Là encore, il serait intéressant d'étudier les propriétés de la méthode lorsqu'une approximation majorante est utilisée pour l'approximation quadratique dans la région de confiance.

**Recherche de pas et optimisation non différentiable.** Dans ce travail de thèse, nous nous sommes placés dans le cadre de l'optimisation d'un critère différentiable. Cependant, dans certaines applications du traitement du signal et de l'image, l'utilisation d'un critère non différentiable est plus adaptée. C'est le cas par exemple lorsque l'on cherche à restaurer des représentations

parcimonieuses, dans un objectif de compression de données (Chen *et al.*, 1998). L'*a priori* de parcimonie est introduit de façon efficace à travers l'utilisation d'une norme  $\ell_1$  dans la fonction de pénalisation, donnant lieu à un critère pénalisé non différentiable. La non-différentiabilité du critère nécessite de recourir à des stratégies d'optimisation spécifiques (Carlavan *et al.*, 2010; Combettes et Pesquet, 2010) basées pour la plupart sur l'utilisation de l'opérateur proximal (Rockafellar, 1976). Citons par exemple l'algorithme FISTA (Beck et Teboulle, 2009), et les algorithmes de 'forward-backward splitting' (Combettes et Wajs, 2005) et de Douglas-Rachford (Combettes et Pesquet, 2007). L'implémentation de ces deux derniers algorithmes nécessite le réglage à chaque itération d'un paramètre de relaxation. En pratique, ce paramètre est maintenu à une valeur constante choisie afin de vérifier des conditions suffisantes de convergence (Carlavan *et al.*, 2010). Cependant, la question se pose de savoir si un réglage plus élaboré, basé sur une méthode de recherche de pas, entrainerait une accélération de l'algorithme. Là encore, une stratégie de pas MM adaptée à la non différentiabilité du critère pourrait être employée.

## Recherche de pas par Majoration-Minoration log-quadratique

La deuxième partie du travail de thèse est consacrée à la minimisation de critères contenant une barrière. Nous montrons dans le chapitre 6 que les stratégies de pas standards, ainsi que la recherche de pas MMQ 1D sont inefficaces pour résoudre ces problèmes. Nous avons alors considéré dans le chapitre 7 des fonctions majorantes log-quadratiques menant à la procédure MMLQ 1D. Une part importante de nos efforts a porté sur l'établissement des preuves de convergence des algorithmes de descente itérative associés au pas MMLQ 1D. Cette méthode fournit une alternative aux stratégies d'interpolation log-quadratiques utilisées dans la littérature, qui se démarque par sa simplicité d'implémentation. Dans le chapitre 8, nous avons illustré l'efficacité du pas MMLQ 1D pour résoudre des problèmes quadratiques contraints par la méthode de points intérieurs primale. Dans le cas de la méthode primale-duale, notre conclusion est qu'une simple stratégie de restriction du pas dans le domaine contraint est préférable, ce qui était déjà noté dans (Roos *et al.*, 2006, Sec.20.6.2). Des extensions de la recherche de pas MMLQ 1D peuvent être envisagées.

**Hypothèses sur la forme de la barrière.** Les résultats du chapitre 7 sont obtenus sous l'hypothèse que le terme de barrière s'écrit

$$B(\mathbf{x}) = \sum_{i=1}^m \psi_i(\mathbf{c}_i^T \mathbf{x} + \rho_i) \quad (10.3)$$

où  $\psi_i$  est une barrière scalaire qui vérifie l'inégalité  $-\frac{2}{u} \ddot{\psi}_i(u) \leq \ddot{\psi}_i(u) \leq 0$ . Une perspective intéressante consiste à généraliser ses résultats à d'autres formes de fonctions barrières, incluant par exemple des contraintes non linéaires.

L'hypothèse sur la forme de  $B$  nous permet de définir dans le théorème 7.1 une fonction majorante du critère le long d'une direction, et de construire la recherche de pas MMLQ 1D. Cependant, il n'y a pas de façon unique de définir une fonction majorante log-quadratique du critère. En particulier, nous souhaitons souligner le fait que, lorsque les contraintes sont linéaires, les propriétés de convergence du chapitre 7 sont vérifiées pour toute majorante log-quadratique de la forme (7.7) de courbure bornée i.e., vérifiant le lemme 7.5.

**Contraintes quadratiques.** Il est possible de construire une majoration d'un critère contenant la barrière logarithmique

$$B(\mathbf{x}) = \sum_{i=1}^m \psi_i(C_i(\mathbf{x})) \quad (10.4)$$

associée aux contraintes quadratiques

$$C_i(\mathbf{x}) = -\frac{1}{2} \mathbf{x}^T \mathbf{C}_i \mathbf{x} + \mathbf{c}_i^T \mathbf{x} + \rho_i, i = 1, \dots, m \quad (10.5)$$

avec  $C_i$  des matrices définies positives de  $\mathbb{R}^{n \times n}$ . Pour tout  $\mathbf{x}$  tel que  $C_i(\mathbf{x}) > 0$ , la fonction  $\alpha \mapsto C_i(\mathbf{x} + \alpha \mathbf{d})$  est concave, quadratique et positive en  $\alpha = 0$ . Par conséquent, elle s'annule en deux valeurs finies  $(\alpha_-^i, \alpha_+^i)$ , correspondant aux racines d'un polynôme du second ordre. Donc elle se factorise en

$$C_i(\mathbf{x} + \alpha \mathbf{d}) = -q_i(\alpha - \alpha_-^i)(\alpha - \alpha_+^i), \quad q_i > 0. \quad (10.6)$$

Il suffit alors d'appliquer la technique de construction de la majorante sur les deux termes de la fonction  $B$  :

$$B(\mathbf{x} + \alpha \mathbf{d}) = \sum_i -\log(-q_i(\alpha - \alpha_+^i)) - \log(\alpha - \alpha_-^i) \quad (10.7)$$

en positionnant les barrières en  $\alpha_- = \max_i \alpha_-^i$  et  $\alpha_+ = \min_i \alpha_+^i$ . Notons que notre analyse de convergence ne permet pas de couvrir cette technique de majoration car l'hypothèse sur la linéarité des contraintes intervient dans la démonstration du lemme 7.6.

**Barrière non convexe.** La stratégie que nous avons proposée se généralise également au cas d'une fonction

$$B(\mathbf{x}) = \sum_i a_i \psi_i(C_i(\mathbf{x})), \quad C_i(\mathbf{x}) = \mathbf{c}_i^T \mathbf{x} + \rho_i \quad (10.8)$$

où les  $a_i$  sont des réels non nécessairement positifs, conduisant à des termes concaves. La construction de la majorante s'obtient en utilisant une propriété bien connue des fonctions concaves, à savoir qu'elles sont majorables par leur approximation du premier ordre (i.e. linéaire). Les termes convexes sont majorés avec la méthode existante. La convergence de la recherche de pas résultant de cette nouvelle majorante doit cependant être étudiée, en particulier la vérification du lemme 7.5. Cette formulation plus générale permet par exemple de traiter le cas de la fonction de pénalisation suivante

$$R(\mathbf{x}) = \sum_i x_i \log\left(\frac{x_i}{[\mathbf{D}\mathbf{x}]_i}\right) + [\mathbf{D}\mathbf{x}]_i - x_i, \quad (10.9)$$

utilisée dans (O'Sullivan, 1995) pour la reconstruction d'images en tomographie.

**Barrière auto-concordante.** Dans le chapitre 6, nous avons souligné que la recherche de pas développée par Yurii Nesterov pour les fonctions à barrières auto-concordantes était similaire à une stratégie MM. En effet, le pas de Nesterov résulte de la minimisation d'une fonction log-quadratique majorante du critère le long de la direction de descente. Cependant, alors que nous imposons la position de la barrière de la majorante MMLQ à la valeur  $\alpha_+$  délimitant le domaine contraint, la singularité de la majorante de Nesterov est située en  $\bar{\alpha} < \alpha_+$ . La question serait alors de savoir si nos résultats de convergence s'étendent à cette méthode de pas. L'intérêt serait d'obtenir une stratégie de pas MMLQ pour les formes de barrières suivantes, vérifiant l'hypothèse d'auto-concordance (Nesterov et Nemirovski, 1994) :

$$\begin{aligned} B(\mathbf{x}) &= \sum_i (\mathbf{c}_i^T \mathbf{x} + \rho_i)^{-1} \\ B(\mathbf{x}) &= -\sum_i \log(-\frac{1}{2} \mathbf{x}^T \mathbf{C}_i \mathbf{x} + \mathbf{c}_i^T \mathbf{x} + \rho_i) \quad \text{avec } \mathbf{C}_i \text{ définies positives} \\ B(\mathbf{x}) &= -\sum_i \log(\rho_i^2 - \|\mathbf{x}\|^2) \\ B(\mathbf{x}) &= -\log \det \mathbf{C}(\mathbf{x}) \quad \text{avec } \mathbf{C}(\cdot) \text{ matrice de } \mathbb{R}^{m \times m} \end{aligned}$$

**Majorante log-quadratique multi-dimensionnelle.** La construction du pas MMQ se base sur l'existence de fonctions majorantes du critère global  $F(\mathbf{x})$ . De ces majorantes globales, on déduit des majorantes du critère sur un espace restreint (droite supportant la direction de descente, ou espace engendré par les directions de recherche) afin de construire la recherche de pas. A l'opposé, le pas MMLQ est issu d'une majorante log-quadratique construite en considérant uniquement le critère le long d'une direction. La question se pose de savoir si cette majorante s'écrit comme la restriction scalaire d'une majorante globale du critère. L'intérêt serait de pouvoir ensuite généraliser la recherche de pas MMLQ 1D en recherche de pas MMLQ multi-dimensionnelle pour l'algorithme de sous-espace. On pourrait également envisager une minimisation du critère avec une stratégie MM log-quadratique directement dans le domaine  $\mathbb{R}^N$ .

---

## Résonance Magnétique Nucléaire bi-dimensionnelle

Le dernier chapitre de la thèse est consacré au développement d'une méthode de résolution du problème inverse qui se pose en RMN 2D. Pour pallier le caractère mal posé du problème, nous proposons une stratégie basée sur l'inversion pénalisée par le maximum d'entropie (ME). Le problème d'optimisation qui en résulte est résolu par l'algorithme itératif de Newton tronqué, associé à un préconditionneur par TSVD et à la recherche de pas MMLQ. La prise en compte de la séparabilité du modèle dans l'implémentation permet de réduire le coût de calcul de l'algorithme sans passer par une étape de compression des données. Le réglage du coefficient de régularisation est basé sur le test du Chi carré. De plus, le paramètre de l'angle de bascule est estimé à partir des données acquises lors de l'expérimentation RMN afin de limiter les erreurs liées au modèle direct. Les performances de la méthode sont illustrées à travers des reconstructions de spectres réels issus de l'agro-alimentaire.

**Autres algorithmes.** Une première question se pose de savoir s'il existe des algorithmes itératifs plus performants que l'algorithme de Newton tronqué pour minimiser le critère pénalisé par maximum d'entropie. L'algorithme de Bryan et Skilling se montre plus rapide que notre méthode en terme de décroissance du critère sur les premières itérations. Cependant, les itérés qu'il génère restent ensuite bloquées en un point qui ne minimise pas le critère (Chouzenoux *et al.*, 2009b). Une perspective intéressante serait de proposer un algorithme de sous-espace pour le problème ME, présentant de bonnes performances aussi bien à l'initialisation qu'à l'approche de la solution optimale.

Les algorithmes de la famille ART (ex : SMART, MART) présentent l'intérêt d'être très simples à implémenter grâce à leur forme multiplicative. Cependant, leur formulation actuelle permet une résolution sous contraintes d'égalité, menant à des solutions de mauvaises qualités lorsque le problème inverse est mal posé. Une perspective serait de proposer un algorithme multiplicatif permettant de résoudre le problème de maximum d'entropie sous sa forme pénalisée. Une approche par Majoration-Minoration séparable (De Pierro, 1993; Lee et Seung, 2001) ou bien par décomposition du gradient (Lanteri *et al.*, 2002, 2001) pourrait par exemple être envisagée.

**Estimation de l'angle de bascule.** La stratégie que nous avons proposée pour estimer la valeur de l'angle de bascule se base sur une approximation de l'expression du noyau  $k_1$  après un temps de mesure  $\tau_1$  supposé 'grand'. On en déduit un spectre  $T_2$  de référence qui nous permet de retrouver l'angle exact du basculement réalisé lors de la mesure. Une perspective à court terme serait d'analyser la sensibilité de notre méthode à la valeur de  $\tau_1$  utilisée pour reconstruire le spectre de référence. Il faudrait également comparer les spectres obtenus avec ceux issus de données prétraitées suivant la stratégie de (Song *et al.*, 2002).

**Noyaux non séparables.** Les stratégies de préconditionnement et de réduction de la complexité de l'algorithme sont basées sur la séparabilité du noyau présent dans le modèle intégral. Dans certaines applications RMN, le noyau n'est pas séparable (Hürlimann et Venkataramanan, 2002). Une perspective serait alors de traiter ces problèmes en approchant le noyau de convolution par une somme de noyaux séparables en se basant sur une décomposition de matrices en produits de Kronecker (Kamm et Nagy, 1998; Van Loan et Pitsianis, 1993).



# Annexes



# Annexe A

## Démonstrations du chapitre 7

### A.1 Propriétés de la récurrence MM

Nous considérons dans cette section  $\mathbf{x} \in \mathcal{C}$  et  $\mathbf{d}$  une direction de descente. Tout d'abord, nous établissons deux résultats préliminaires issus de l'expression de la majorante et de son minimiseur.

#### A.1.1 Résultats préliminaires

**Lemme A.1.** Soit  $j \in \{0, \dots, J-1\}$ . Si  $\dot{f}(\alpha^j) \leq 0$ , alors :

$$f(\alpha^j) - f(\alpha^{j+1}) + \frac{1}{2}(\alpha^{j+1} - \alpha^j)\dot{f}(\alpha^j) \geq 0 \quad (\text{A.1})$$

*Démonstration.* La démonstration est triviale pour  $\dot{f}(\alpha^j) = 0$ . Nous supposons par la suite que  $\dot{f}(\alpha^j) < 0$  de telle sorte que  $\alpha_+ > \alpha^{j+1} > \alpha^j$ . Soit la fonction scalaire

$$\xi(u) = -\log(1-u) - u. \quad (\text{A.2})$$

Une analyse rapide de  $\xi$  montre que

$$\frac{\xi(u)}{u\xi(u)} \leq \frac{1}{2}, \quad \forall u \in (0,1) \quad (\text{A.3})$$

En posant  $u = \frac{\alpha - \alpha^j}{\alpha_+ - \alpha^j}$  dans (A.3) et en notant  $\varphi(\alpha) = \xi(u)$ , on obtient

$$\frac{\varphi(\alpha)}{(\alpha - \alpha^j)\dot{\varphi}(\alpha)} \leq \frac{1}{2}, \quad \forall \alpha \in (\alpha^j; \alpha_+). \quad (\text{A.4})$$

De plus, définissons les fonctions scalaires

$$Q(\alpha) = \frac{1}{2}m^j(\alpha - \alpha^j)^2, \quad (\text{A.5})$$

$$\tau(\alpha) = Q(\alpha) + \gamma^j(\alpha_+ - \alpha^j)\varphi(\alpha). \quad (\text{A.6})$$

La fonction  $Q$  vérifie

$$Q(\alpha) = \frac{1}{2}(\alpha - \alpha^j)\dot{Q}(\alpha). \quad (\text{A.7})$$

De plus, la fonction majorante s'écrit

$$h(\alpha, \alpha^j) = f(\alpha^j) + (\alpha - \alpha^j)\dot{f}(\alpha^j) + \tau(\alpha), \forall \alpha \in [\alpha^j, \alpha_+] \quad (\text{A.8})$$

et, d'après (A.4) et (A.7),

$$\frac{\tau(\alpha)}{(\alpha - \alpha^j)\dot{\tau}(\alpha)} \leq \frac{1}{2}, \forall \alpha \in (\alpha^j; \alpha_+). \quad (\text{A.9})$$

$h(\cdot, \alpha^j)$  est une fonction tangente majorante de  $f$  donc

$$h(\alpha, \alpha^j) - f(\alpha) = f(\alpha^j) - f(\alpha) + (\alpha - \alpha^j)\dot{f}(\alpha^j) + \tau(\alpha) \geq 0 \quad (\text{A.10})$$

En posant  $\alpha = \alpha^{j+1} > \alpha^j$  dans (A.9) et (A.10), on obtient

$$f(\alpha^j) - f(\alpha^{j+1}) + (\alpha^{j+1} - \alpha^j)\dot{f}(\alpha^j) + \frac{1}{2}(\alpha^{j+1} - \alpha^j)\dot{\tau}(\alpha^{j+1}) \geq 0 \quad (\text{A.11})$$

D'où le résultat, en utilisant

$$\dot{\tau}(\alpha^{j+1}) = \dot{h}^j(\alpha^{j+1}, \alpha^j) - \dot{f}(\alpha^j) \quad (\text{A.12})$$

$$= -\dot{f}(\alpha^j) \quad (\text{A.13})$$

□

**Lemme A.2.** Soit  $j \in \{0, \dots, J-1\}$ . Si  $\dot{f}(\alpha^j) \leq 0$ , alors le minimiseur de  $h(\cdot, \alpha^j)$ , noté  $\alpha^{j+1}$ , vérifie

$$-\frac{q_3}{q_2} \leq \alpha^{j+1} - \alpha^j \leq -\frac{2q_3}{q_2}. \quad (\text{A.14})$$

où  $q_1, q_2$  et  $q_3$  sont donnés en (7.51).

*Démonstration.* Si  $\dot{f}(\alpha^j) \leq 0$ , alors  $\alpha^{j+1}$  s'écrit

$$\alpha^{j+1} = \alpha^j - \frac{2q_3}{q_2 + \sqrt{q_2^2 - 4q_1q_3}} \quad (\text{A.15})$$

avec

$$\begin{cases} q_1 = -m^j \\ q_2 = \gamma^j - \dot{f}(\alpha^j) + m^j(\bar{\alpha}^j - \alpha^j) \\ q_3 = (\bar{\alpha}^j - \alpha^j)\dot{f}(\alpha^j) \end{cases} \quad (\text{A.16})$$

où les valeurs  $(m^j, \gamma^j, \bar{\alpha}^j)$  sont données par

$$\begin{cases} \bar{\alpha}^j = \alpha_+ \\ m^j = m_p^j + \mu \sum_{i|\delta_i > 0} \phi_i(\alpha^j) \\ \gamma^j = \mu(\alpha_+ - \alpha^j) \sum_{i|\delta_i < 0} \phi_i(\alpha^j) \end{cases} \quad (\text{A.17})$$

avec  $\phi_i(\alpha) = \delta_i^2 \ddot{\psi}_i(\theta_i + \alpha \delta_i)$ , fonction à valeurs positives. On en déduit le signe des coefficients  $q_i$  :

$$\begin{aligned} q_1 &< 0 \\ q_2 &> 0 \\ q_3 &> 0 \end{aligned} \quad (\text{A.18})$$

D'où la relation (A.14). □

### A.1.2 Démonstration du lemme 7.5

Montrons que la courbure de la fonction majorante est bornée. Posons  $j \in \{0, \dots, J-1\}$  et le pas  $\alpha^j$  obtenu par la récurrence MM. D'après (7.46) et (7.47), la dérivée seconde de  $h(\cdot, \alpha^j)$  est

$$\ddot{h}^j(\alpha, \alpha^j) = \begin{cases} m_p^j + \mu \sum_{i|\delta_i < 0} \phi_i(\alpha^j) + \mu \sum_{i|\delta_i > 0} \phi_i(\alpha^j) \frac{(\alpha_- - \alpha^j)^2}{(\alpha_- - \alpha)^2} & \forall \alpha \in (\alpha_-, \alpha^j] \\ m_p^j + \mu \sum_{i|\delta_i > 0} \phi_i(\alpha^j) + \mu \sum_{i|\delta_i < 0} \phi_i(\alpha^j) \frac{(\alpha_+ - \alpha^j)^2}{(\alpha_+ - \alpha)^2} & \forall \alpha \in [\alpha^j, \alpha_+) \end{cases}$$

En particulier,

$$\ddot{h}(\alpha^j, \alpha^j) = m_p^j + \mu \sum_i \phi_i(\alpha^j) \quad (\text{A.19})$$

$$= m_p^j + \mu \ddot{b}(\alpha^j). \quad (\text{A.20})$$

La dérivée seconde de  $b$  en  $\alpha^j$  s'écrit

$$\ddot{b}(\alpha^j) = \mathbf{d}^T \nabla^2 B(\mathbf{x} + \alpha^j \mathbf{d}) \mathbf{d},$$

et

$$m_p^j = \mathbf{d}^T \mathbf{A}(\mathbf{x} + \alpha^j \mathbf{d}) \mathbf{d}.$$

De plus,  $\mathbf{x} + \alpha^j \mathbf{d} \in \mathcal{V}$ . Par conséquent, il suffit de montrer que l'ensemble  $\{\mathbf{A}(\mathbf{x}) + \mu \nabla^2 B(\mathbf{x}) | \mathbf{x} \in \mathcal{V}\}$  est à spectre borné strictement positif.

Soit  $\mathbf{x} \in \mathcal{V}$ .

$$\nabla^2 B(\mathbf{x}) = \mathbf{C}^T \text{diag}(\tau_i C_i(\mathbf{x})^{-t_i}) \mathbf{C} \quad (\text{A.21})$$

avec

$$(\tau_i, t_i) = \begin{cases} (2, \kappa_i) & \text{if } \phi_i(u) = -\kappa_i \log u \\ (1, 1) & \text{if } \phi_i(u) = u \log u \\ (-r^2 + r, 2 - r) & \text{if } \phi_i = -u^r \end{cases}$$

et  $\mathbf{C} = [\mathbf{c}_1, \dots, \mathbf{c}_m]^T$ .  $\mathbf{x} \in \mathcal{V}$  donc (7.55) et (7.54) entraînent

$$\mathbf{d}^T \mathbf{T}(E_0) \mathbf{d} \leq \mathbf{d}^T \nabla^2 B(\mathbf{x}) \mathbf{d} \leq \mathbf{d}^T \mathbf{T}(\epsilon_0) \mathbf{d} \quad (\text{A.22})$$

avec  $\mathbf{T}(m) = \mathbf{A}^T \text{diag}(\tau_i m^{-t_i}) \mathbf{A}$ . La matrice  $\mathbf{T}(m)$  est symétrique et a un spectre borné positif de bornes  $(\nu_{\min}^T(m), \nu_{\max}^T(m))$ . En outre, d'après l'hypothèse 7.1,  $\mathbf{A}(\mathbf{x})$  a un spectre borné strictement positif de bornes  $(\nu_{\min}^A, \nu_{\max}^A)$ . Par conséquent, le lemme 7.5 est vérifié avec  $\nu_{\min} = \nu_{\min}^A + \mu \nu_{\min}^T(E_0) > 0$  et  $\nu_{\max} = \nu_{\max}^A + \mu \nu_{\max}^T(\epsilon_0)$ .

### A.1.3 Démonstration du lemme 7.6

Montrons que le pas  $\alpha^1$ , obtenu après une itération de la récurrence MM est borné. Le vecteur  $\mathbf{d}$  est une direction de descente, donc  $\dot{f}(0) < 0$  et  $h(\cdot, 0)$  présente une barrière en  $\bar{\alpha}^0 = \alpha_+$ .

Si  $\alpha_+ = +\infty$  alors  $h(\cdot, 0)$  est une fonction quadratique de courbure  $m^0$ . Son minimiseur est  $\alpha^1 = -\dot{f}(0)/m^0$  et d'après le lemme 7.5,

$$\frac{-\mathbf{g}^T \mathbf{d}}{\nu_{\max} \|\mathbf{d}\|^2} \leq \alpha^1 \leq \frac{-\mathbf{g}^T \mathbf{d}}{\nu_{\min} \|\mathbf{d}\|^2} \quad (\text{A.23})$$

Si  $\alpha_+ < +\infty$ , d'après le lemme A.2 on a

$$\frac{-\mathbf{g}^T \mathbf{d}}{\frac{\gamma^0}{\alpha_+} - \frac{\mathbf{g}^T \mathbf{d}}{\alpha_+} + m^0} \leq \alpha^1 \leq \frac{-2\mathbf{g}^T \mathbf{d}}{\frac{\gamma^0}{\alpha_+} - \frac{\mathbf{g}^T \mathbf{d}}{\alpha_+} + m^0} \quad (\text{A.24})$$

Le lemme 7.5 et la positivité de  $-\mathbf{g}^T \mathbf{d}$  entraînent

$$\nu_{\min} \|\mathbf{d}\|^2 \leq \frac{\gamma^0}{\alpha_+} - \frac{\mathbf{g}^T \mathbf{d}}{\alpha_+} + m^0 \quad (\text{A.25})$$

D'autre part, en posant  $\iota = \arg \max_i -\mathbf{c}_i^T \mathbf{d}$ , nous obtenons de (7.54) que

$$\alpha^+ \geq \frac{\epsilon_0}{|\mathbf{c}_\iota^T \mathbf{d}|}. \quad (\text{A.26})$$

Par conséquent, d'après l'inégalité de Cauchy-Schwartz inequality et (7.53),

$$\begin{aligned} \frac{-\mathbf{g}^T \mathbf{d}}{\alpha^+} = \frac{|\mathbf{g}^T \mathbf{d}|}{\alpha^+} &\leq |\mathbf{g}^T \mathbf{d}| \cdot |\mathbf{c}_\iota^T \mathbf{d}| \frac{1}{\epsilon_0} \\ &\leq \|\mathbf{g}\| \|\mathbf{c}_\iota\| \|\mathbf{d}\|^2 \frac{1}{\epsilon_0} \\ &\leq \frac{\eta C_{\max}}{\epsilon_0} \|\mathbf{d}\|^2 \end{aligned} \quad (\text{A.27})$$

où  $C_{\max} = \max_i \|\mathbf{c}_i\| > 0$ . De plus, le lemme 7.5 implique l'existence de  $\nu_{\max}$  tel que

$$m^0 + \frac{\gamma^0}{\alpha^+} \leq \nu_{\max} \|\mathbf{d}\|^2 \quad (\text{A.28})$$

Les relations (A.25), (A.27) et (A.28) permettent de démontrer le lemme 7.6 avec

$$\begin{aligned} \nu &= \nu_{\max} + \eta C_{\max} / \epsilon_0, \\ \nu' &= \nu_{\min} / 2. \end{aligned}$$

#### A.1.4 Démonstration de la propriété 7.1

La vérification de (7.63) est directe pour  $j = 1$ , avec  $c_{\max}^1 = 1$ . Démontrons que (7.63) est vérifiée pour  $j > 1$ .

Si  $\dot{f}(\alpha^j) < 0$ , alors  $\bar{\alpha}^j = \alpha^+$  et on peut déduire du lemme A.2 que

$$\begin{aligned} \alpha^{j+1} - \alpha^j &\leq \frac{-2\dot{f}(\alpha^j)}{(\gamma^j - \dot{f}(\alpha^j)) / (\alpha^+ - \alpha^j) + m^j} \\ &\leq \frac{-2\dot{f}(\alpha^j)}{\gamma^j / (\alpha^+ - \alpha^j) + m^j} \end{aligned} \quad (\text{A.29})$$

D'après le lemme 7.5 :

$$\|\mathbf{d}\|^2 \geq (\gamma^0 / \alpha^+ + m^0) / \nu_{\max} \quad (\text{A.30})$$

et

$$\gamma^j / (\alpha^+ - \alpha^j) + m^j \geq \nu_{\min} \|\mathbf{d}\|^2 \quad (\text{A.31})$$

soit

$$\gamma^j / (\alpha^+ - \alpha^j) + m^j \geq \nu_{\min} (\gamma^0 / \alpha^+ + m^0) / \nu_{\max} > 0 \quad (\text{A.32})$$

Donc, d'après (A.29) :

$$\alpha^{j+1} \leq \alpha^j + |\dot{f}(\alpha^j)| \frac{2\nu_{\max}}{(\gamma^0 / \alpha^+ + m^0)\nu_{\min}} \quad (\text{A.33})$$

Si  $\dot{f}(\alpha^j) \geq 0$ ,  $\alpha^{j+1}$  est inférieur à  $\alpha^j$  et (A.33) est toujours vérifiée.

D'après l'hypothèse 7.2,  $\nabla F$  est Lipschitz, d'où

$$|\dot{f}(\alpha^j) - \dot{f}(0)| \leq L \|\mathbf{d}\|^2 \alpha^j \quad (\text{A.34})$$

De plus, on a la relation

$$|\dot{f}(\alpha^j)| \leq |\dot{f}(\alpha^j) - \dot{f}(0)| + |\dot{f}(0)|, \quad (\text{A.35})$$

et  $\dot{f}(0) < 0$ , soit

$$|\dot{f}(\alpha^j)| \leq L \alpha^j \|\mathbf{d}\|^2 - \dot{f}(0). \quad (\text{A.36})$$

D'après le lemme 7.6 et (A.30) :

$$\begin{aligned} -\dot{f}(0) &\leq \alpha^1 \nu \|\mathbf{d}\|^2 \\ &\leq \alpha^1 \nu (m^0 + \gamma_0/\alpha_+)/\nu_{\min} \end{aligned} \quad (\text{A.37})$$

Les relations (A.30)- (A.37) entraînent que

$$\alpha^{j+1} \leq \alpha^j + \frac{2\nu_{\max}}{(m^0 + \gamma_0/\alpha_+)\nu_{\min}} \left[ L\alpha^j \left( \frac{m^0 + \gamma_0/\alpha_+}{\nu_{\min}} \right) + \alpha^1 \frac{\nu}{\nu_{\min}} (m^0 + \gamma_0/\alpha_+) \right] \quad (\text{A.38})$$

Soit

$$\alpha^{j+1} \leq \alpha^j \left( 1 + \frac{2\nu_{\max}L}{\nu_{\min}^2} \right) + 2\alpha^1 \frac{\nu_{\max}\nu}{\nu_{\min}^2} \quad (\text{A.39})$$

Cela définit récursivement la suite  $\{c_{\max}^j\}$  avec

$$c_{\max}^{j+1} = c_{\max}^j \left( 1 + 2\frac{\nu_{\max}L}{\nu_{\min}^2} \right) + 2\frac{\nu\nu_{\max}}{\nu_{\min}^2} \quad (\text{A.40})$$

En posant  $c_{\max}^1 = 1$ , (7.64) est le terme général de la série.

### A.1.5 Démonstration de la propriété 7.2

Nous devons démontrer que le pas  $\alpha^j$  vérifie

$$\alpha^j \geq \alpha_{\min} \quad (\text{A.41})$$

avec

$$\alpha_{\min} = c_{\min}\alpha^1, \quad c_{\min} = \frac{\sqrt{1 + 2L/\nu_{\min}} - 1}{2L/\nu_{\min}}. \quad (\text{A.42})$$

Définissons  $\phi$  la fonction quadratique concave

$$\phi(\alpha) = f(0) + \alpha\dot{f}(0) + m\frac{\alpha^2}{2}$$

où

$$m = -L(m^0 + \gamma^0/\alpha_+)/\nu_{\min}. \quad (\text{A.43})$$

On a  $\phi(0) = f(0)$  et  $\dot{\phi}(0) = \dot{f}(0) < 0$ , donc  $\phi$  est décroissante sur  $\mathbb{R}^+$ . Soit  $\alpha \in [0, \alpha^j]$ , alors  $\mathbf{x} + \alpha\mathbf{d} \in \mathcal{V}$ . D'après l'hypothèse 7.2,

$$|\dot{f}(\alpha) - \dot{f}(0)| \leq \|\mathbf{d}\|^2 L|\alpha| \quad (\text{A.44})$$

et le lemme 7.5 implique

$$|\dot{f}(\alpha) - \dot{f}(0)| \leq L\alpha(m^0 + \gamma^0/\alpha_+)/\nu_{\min}$$

Donc on obtient

$$|\dot{f}(\alpha)| \leq L\alpha(m^0 + \gamma^0/\alpha_+)/\nu_{\min} - \dot{f}(0)$$

Soit

$$\dot{\phi}(\alpha) \leq \dot{f}(\alpha), \quad \forall \alpha \in [0, \alpha^j] \quad (\text{A.45})$$

l'intégration de (A.45) entre 0 et  $\alpha^j$  donne

$$\phi(\alpha^j) \leq f(\alpha^j) \quad (\text{A.46})$$

D'autre part, la valeur de  $\phi$  en  $\alpha_{\min} = c_{\min}\alpha^1$  s'écrit

$$\phi(\alpha_{\min}) = f(0) + C\alpha^1\dot{f}(0)$$

avec

$$C = c_{\min} - c_{\min}^2 L \alpha^1 \frac{m^0 + \gamma^0 / \alpha_+}{2\dot{f}(0)\nu_{\min}}.$$

D'après (A.29)

$$\alpha^1 \leq \frac{-2\dot{f}(0)}{m^0 + \gamma^0 / \alpha_+},$$

d'où

$$C \leq c_{\min} + c_{\min}^2 \frac{L}{\nu_{\min}} = \frac{1}{2},$$

d'après l'expression de  $c_{\min}$ . La fonction  $\phi$  étant décroissante sur  $\mathbb{R}^+$ , on obtient

$$\phi(\alpha_{\min}) \geq f(0) + \frac{1}{2}\alpha^1 \dot{f}(0) \geq f(\alpha^1), \quad (\text{A.47})$$

où la deuxième inégalité est la première condition de Wolfe (7.67) en  $j = 1$ .

Le pas  $\alpha^j$  est strictement positif pour tout  $j \geq 1$ . Supposons l'existence d'un indice  $j$  tel que  $\alpha^j < \alpha_{\min}$ . D'après (A.46) et étant donné que  $\phi$  est décroissante sur  $\mathbb{R}^+$ ,

$$f(\alpha^j) \geq \phi(\alpha^j) > \phi(\alpha_{\min}) \geq f(\alpha^1),$$

ce qui entre en contradiction avec la décroissance de  $\{f(\alpha^j)\}$ . Par conséquent, (A.41) est vérifiée, ainsi que (7.65), d'après le lemme 7.6.

### A.1.6 Démonstration de la propriété 7.3

La propriété 7.3 se déduit directement du lemme A.1 et de la propriété 7.1. Si  $j = 1$ , la condition d'Armijo (7.67) est vérifiée d'après le lemme A.1, puisqu'elle s'identifie avec (A.1) lorsque  $j = 0$ , avec  $c_{\max}^1 = 1$ . Si  $j > 1$ , (7.67) est vérifiée par récurrence, d'après la propriété 7.1.

## A.2 Convergence de l'algorithme de conjugué non linéaire

Nous établissons dans cette section la preuve de convergence du théorème 7.4. Tout d'abord, nous établissons deux lemmes préliminaires, à partir de résultats de la démonstration présentés en section 7.2.

### A.2.1 Résultats préliminaires

**Lemme A.3.** *Soit  $j \in \{0, \dots, J-1\}$ . Pour tout  $\alpha \in [\alpha^j, \alpha_+)$ ,  $\dot{h}_b^j(\alpha, \alpha^j)$  est une fonction tangente majorante de  $\dot{b}(\alpha)$  en  $\alpha^j$ .*

*Démonstration.* La fonction  $\dot{\phi}_i$  est concave. Donc pour tout  $\alpha^j$ ,

$$\ddot{h}_{b,1}(\alpha, \alpha^j) = \ddot{b}_1(\alpha^j) \geq \ddot{b}_1(\alpha), \forall \alpha \in [\alpha^j, \alpha_+) \quad (\text{A.48})$$

Par conséquent, la fonction  $\dot{h}_{b,1}(\alpha, \alpha^j) - \dot{b}_1(\alpha)$  est croissante sur  $[\alpha^j; \alpha_+)$ . De plus, elle s'annule en  $\alpha^j$ , d'où

$$\dot{h}_{b,1}(\alpha, \alpha^j) \geq \dot{b}_1(\alpha), \forall \alpha \in [\alpha^j, \alpha_+). \quad (\text{A.49})$$

La relation (7.40) entraîne alors le résultat.  $\square$

**Lemme A.4.** *Si pour tout  $j \in \{0, \dots, J-1\}$ , (7.75) est vérifiée, alors*

$$\dot{f}(\alpha^j) \leq 0, \forall j \in \{0, \dots, J-1\} \quad (\text{A.50})$$

*et la suite  $\{\alpha^j\}$  est croissante.*

*Démonstration.* L'inégalité (7.75) s'écrit aussi

$$\mathbf{d}^T \nabla P(\mathbf{x} + \alpha \mathbf{d}) \leq \mathbf{d}^T \nabla P(\mathbf{x} + \alpha^j \mathbf{d}) + (\alpha - \alpha^j) \mathbf{d}^T \mathbf{A}(\mathbf{x} + \alpha^j \mathbf{d}) \mathbf{d}, \quad \forall \alpha \geq \alpha^j, \quad (\text{A.51})$$

Soit, en utilisant les notations de la section 7.2,

$$\dot{p}(\alpha) \leq \dot{h}_p(\alpha, \alpha^j), \quad \forall \alpha \geq \alpha^j. \quad (\text{A.52})$$

D'après le lemme A.3, (A.52) implique que pour tout  $j \in \{0, \dots, J-1\}$ ,

$$\dot{f}(\alpha) \leq \dot{h}(\alpha, \alpha^j), \quad \forall \alpha \geq \alpha^j. \quad (\text{A.53})$$

L'inégalité (A.50) est vérifiée en  $j = 0$  car  $\mathbf{d}$  est une direction de descente. Donc  $\alpha^1 \geq 0$  d'après (7.56). Posons  $j \in \{0, \dots, J-1\}$  et supposons que  $\dot{f}(\alpha^j) \leq 0$ . Par conséquent, d'après (7.56),  $\alpha^{j+1} \geq \alpha^j$ . En utilisant (A.53) avec  $\alpha = \alpha^{j+1}$ , on obtient

$$\dot{f}(\alpha^{j+1}) \leq \dot{h}(\alpha^{j+1}, \alpha^j) \quad (\text{A.54})$$

De plus,  $\alpha^{j+1}$  est le minimiseur de  $h^j(\cdot, \alpha^j)$  donc  $\dot{h}(\alpha^{j+1}, \alpha^j) = 0$ , d'où le résultat par récurrence immédiate sur  $j$   $\square$

## A.2.2 Démonstration du théorème 7.4

Nous pouvons désormais établir la convergence des algorithmes de gradient conjugué non linéaire PRP+, FR et DY. Nous allons démontrer par récurrence sur  $k$  que  $\mathbf{d}_k$  est une direction de descente suffisante pour  $F$ , i.e., il existe  $\eta > 0$  tel que

$$\mathbf{g}_k^T \mathbf{d}_k \leq -\eta \|\mathbf{g}_k\|^2. \quad (\text{A.55})$$

D'après (Hager et Zhang, 2006b, Th. 4.2, Th. 5.1), l'association de la condition (A.55) avec la condition de Zoutendijk garantit la convergence des algorithmes de gradient conjugué FR, DY et PRP+.

Soit  $\mathbf{x}_k \in \mathcal{V}$  et  $\mathbf{d}_k$  une direction de descente suffisante. Montrons que  $\mathbf{d}^{k+1}$  est une direction de descente suffisante. D'après le lemme A.4, (7.75) implique que  $\dot{f}(\alpha^j) < 0$  pour tout  $j$ . Donc  $\mathbf{g}_{k+1}^T \mathbf{d}_k \leq 0$ . D'après la récurrence (4.101),

$$\mathbf{g}_{k+1}^T \mathbf{c}_{k+1} = -\|\mathbf{g}_{k+1}\|^2 + \beta_{k+1} \mathbf{g}_{k+1}^T \mathbf{d}_k$$

Les coefficients de conjugaison des méthodes FR et PRP+ vérifient

$$\beta_k^{\text{FR}} = \frac{\|\mathbf{g}_{k+1}\|^2}{\|\mathbf{g}_k\|^2} \geq 0 \quad (\text{A.56})$$

$$\beta_k^{\text{PRP+}} = \max(\beta_k^{\text{PRP}}, 0) \geq 0 \quad (\text{A.57})$$

Par conséquent,

$$\mathbf{g}_{k+1} \mathbf{c}_{k+1} \leq -\|\mathbf{g}_{k+1}\|^2, \quad (\text{A.58})$$

et  $\mathbf{d}^{k+1} = \mathbf{c}^{k+1}$  est une direction de descente suffisante. Le coefficient de conjugaison de la méthode DY donné par

$$\beta_k^{\text{DY}} = \frac{\|\mathbf{g}_{k+1}\|^2}{\mathbf{d}_k^T (\mathbf{g}_{k+1} - \mathbf{g}_k)}$$

est du signe de  $\mathbf{d}_k^T (\mathbf{g}_{k+1} - \mathbf{g}_k)$ . Si  $F$  est convexe, la récurrence (4.101) implique

$$|\mathbf{g}_{k+1}^T \mathbf{d}_k| \leq |\mathbf{g}_k^T \mathbf{d}_k|. \quad (\text{A.59})$$

Comme  $\mathbf{d}_k$  est une direction de descente,  $\beta_k^{\text{DY}} \geq 0$ , donc  $\mathbf{d}^{k+1} = \mathbf{c}^{k+1}$  est une direction de descente suffisante. Donc, la condition (A.55) est vérifiée pour tout  $k$  pour les méthodes de FR, DY et PRP+. Par conséquent, la convergence des algorithmes de gradient conjugué FR, DY et PRP+ est assurée.



# Bibliographie

- Al-Baali, M. et Fletcher, R. : An efficient line search for nonlinear least squares. *Journal of Optimization Theory and Applications*, 48(3):377–359, 1986.
- Al-Baali, M. et Fletcher, R. : On the order of convergence of preconditioned nonlinear conjugate gradient methods. *SIAM Journal on Scientific Computing*, 17(3):658–665, 1996.
- Ali, S. M. et Silvey, S. D. : A general class of coefficients of divergence of one distribution from another. *Journal of the Royal Statistical Society B*, 28:131–142, 1966.
- Allain, M. : *Approche pénalisée en tomographie hélicoïdale en vue de l'application à la conception d'une prothèse personnalisée du genou*. Thèse de doctorat, Ecole Polytechnique MontréalCanada, décembre 2002. <http://tel.archives-ouvertes.fr/docs/00/04/57/49/PDF/tel-00003756.pdf>.
- Allain, M., Idier, J. et Goussard, Y. : On global and local convergence of half-quadratic algorithms. *IEEE Transactions on Image Processing*, 15(5):1130–1142, 2006.
- Armand, P., Gilbert, J. C. et Jan-Jégou, S. : A feasible BFGS interior point algorithm for solving strongly convex minimization problems. *SIAM Journal on Optimization*, 11:199–222, 2000.
- Armijo, L. : Minimization of functions having Lipschitz continuous first partial derivatives. *Pacific Journal of Mathematics*, 16:1–3, 1966.
- Aubert, G. et Kornprobst, P. : *Mathematical problems in images processing*. Springer-Verlag, Berlin, 2002.
- Aujol, J.-F. et Chambolle, A. : Dual norms and image decomposition models. *International Journal of Computer Vision*, 63(1):85–104, 2005.
- Auslender, A. et Haddou, M. : An interior-proximal method for convex linearly constrained problems and its extension to variational inequalities. *Mathematical Programming*, 71(1):77–100, 1995.
- Beck, A. et Teboulle, M. : A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2(1):183–202, janvier 2009.
- Belge, M., Kilmer, M. et Miller, E. : Wavelet domain image restoration with adaptive edge-preserving regularization. *IEEE Transactions on Image Processing*, 9(4):597–608, avril 2000.
- Bellavia, S. : Inexact interior-point method. *Journal of Optimization Theory and Applications*, 96(1):109–121, 1998.
- Ben-Tal, A., Ben-Israel, A. et Teboulle, M. : Certainty equivalents and information measures : duality and extremal principles. *Journal of Mathematical Analysis and Applications*, 157:211–236, 1991.
- Ben-Tal, A., Charnes, A. et Teboulle, M. : Entropic means. *Journal of Mathematical Analysis and Applications*, 139:537–551, 1989.

- Bertero, M. et Boccacci, P. : *Introduction to Inverse Problems in Imaging*. IOP Publishing Ltd., London, Royaume-Uni, 1998.
- Bertero, M., Boccacci, P., Desidera, G. et Vicidomini, G. : Image deblurring with poisson data : from cells to galaxies. *Inverse Problems*, 25, 2009.
- Bertsekas, D. P. : Projected Newton methods for optimization problems with simple constraints. *IEEE Conference on Decision and Control including the Symposium on Adaptive Processes*, 20:762–767, décembre 1981.
- Bertsekas, D. P. : *Constrained Optimization and Lagrange Multiplier Methods*. Athena Scientific, Belmont, MA, USA, 2ème édn, 1996.
- Bertsekas, D. P. : *Nonlinear Programming*. Athena Scientific, Belmont, MA, USA, 2ème édn, 1999.
- Bertsekas, D. P. : *Convex analysis and optimization*. Athena Scientific, Belmont, MA, USA, 1ère édn, 2003.
- Birgin, E. G., Martínez, J. M. et Raydan, M. : Nonmonotone spectral projected gradient methods on convex sets. *SIAM Journal on Optimization*, 10(4):1196–1211, 2000.
- Boccacci, P. et Bertero, M. : Image-restoration methods : basics and algorithms. *In Confocal and Two-Photon Microscopy : Foundations, Applications, and Advances*, p. 253–269. Wiley-Liss, New York, NY, USA, 2002.
- Böhning, D. et Lindsay, B. G. : Monotonicity of quadratic-approximation algorithms. *Annals of the Institute of Statistical Mathematics*, 40(4):641–663, 1988.
- Bonettini, S., Zanella, R. et Zanni, L. : A scaled gradient projection method for constrained image deblurring. *Inverse Problems*, 25(1):015002+, 2009.
- Bonnans, J., Gilbert, J., Lemaréchal, C. et Sagastizèbal, C. : *Numerical Optimization. Theoretical and Practical Aspects*. Springer-Verlag, New York, NY, USA, 2ème édn, 2006.
- Borgia, G. C., Brown, R. J. S. et Fantazzini, P. : Uniform-penalty inversion of multiexponential decay data. *Journal of Magnetic Resonance*, 132(1):65 – 77, 1998.
- Bose, C., Cui, Z., Deng, X., Han, Y., Li, Q., Piché, R., Venkataramanan, L., Wang, Q. et Zhou, L. : Inversion of NMR data. Rap. tech., Schlumberger Limited, 2003. <http://www.maths-in-industry.org/miis/185/1/schlumberger.pdf>.
- Bouman, C. et Sauer, K. : A unified approach to statistical tomography using coordinate descent optimization. *IEEE Transactions on Image Processing*, 5(3):480–492, 1996.
- Bouman, C. et Sauer, K. D. : A generalized Gaussian image model for edge-preserving MAP estimation. *IEEE Transactions on Image Processing*, 2(3):296–310, juillet 1993.
- Boyd, S. et Vandenberghe, L. : *Convex Optimization*. Cambridge University Press, New York, 1ère édn, 2004.
- Bregman, L. M. : The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. *USSR Computational Mathematics and Mathematical Physics*, 7:200–217, 1967.
- Bretthorst, G. L., Hutton, W. C., Garbow, J. R. et Ackerman, J. J. H. : Exponential parameter estimation (in NMR) using bayesian probability theory. *Concepts in Magnetic Resonance Part A*, 27A:55–63, 2005.

- Butler, J. P., Reeds, J. A. et Dawson, S. V. : Estimating solutions of first kind integral equations with nonnegative constraints and optimal smoothing. *SIAM Journal on Numerical Analysis*, 18 (3):381–397, juin 1981.
- Byrd, R. H., Lu, P., Nocedal, J. et Zhu, C. : A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16:1190 – 1208, septembre 1995.
- Byrd, R. H., Schnabel, R. B. et Shultz, G. A. : A trust region algorithm for nonlinearly constrained optimization. *SIAM Journal on Numerical Analysis*, 24(5):1152–1170, octobre 1987.
- Byrne, C. L. : Iterative image reconstruction algorithms based on cross-entropy minimization. *IEEE Transactions on Image Processing*, 2:96–103, 1993.
- Byrne, C. L. : Block-iterative methods for image reconstruction from projections. *IEEE Transactions on Image Processing*, 5(5):792–794, mai 1996.
- Byrne, C. L., Fitzgerald, R. M., Fiddy, M. A., Hall, T. J. et Darling, A. M. : Image restoration and resolution enhancement. *Journal of the Optical Society of America*, 73:1481–1487, novembre 1983.
- Canet, D., Boubel, J.-C. et Canet-Soulas, E. : *La RMN. Concepts, méthodes et applications*. Dunod, Paris, 2ème éd., 2002.
- Cantrell, J. : Relation between the memory gradient method and the Fletcher-Reeves method. *Journal of Optimization Theory and Applications*, 4(1):67–71, 1969.
- Cao, Y., Eggermont, P. P. B. et Terebey, S. : Cross Burg entropy maximization and its application to ringing suppression in image reconstruction. *IEEE Transactions on Image Processing*, 8 (2):286–292, 1999.
- Carlavan, M., Weiss, P., Blanc-Féraud, L. et Zerubia, J. : Régularité et parcimonie pour les problèmes inverses en imagerie : algorithmes et comparaisons. *Traitement du Signal*, à paraître, 2010.
- Censor, Y. : Finite series-expansion reconstruction methods. *Proceedings of the IEEE*, 71(3):409–419, 1983.
- Censor, Y. et Lent, A. : An iterative row-action method for interval convex programming. *Journal of Optimization Theory and Applications*, 34(3):321–353, 1981.
- Chambolle, A., DeVore, R., Lee, N. et Lucier, B. J. : Nonlinear wavelet image processing : Variational problems, compression, and noise removal through wavelet shrinkage. *IEEE Transactions on Image Processing*, 7:320–335, mars 1998.
- Chan, T. F. et Mulet, P. : On the convergence of the lagged diffusivity fixed point method in total variation image restoration. *SIAM Journal on Numerical Analysis*, 36(2):354–367, 1999.
- Charbonnier, P., Blanc-Féraud, L., Aubert, G. et Barlaud, M. : Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions on Image Processing*, 6:298–311, 1997.
- Chaâri, L., Pustelnik, N., Chaux, C. et Pesquet, J.-C. : Solving inverse problems with overcomplete transforms and convex optimization techniques. In *SPIE 2009 Wavelets XIII*, San Diego, Chili, août 2009.
- Chaux, C., Combettes, P. L., Pesquet, J.-C. et Wajs, V. R. : A variational formulation for frame-based inverse problems. *Inverse Problems*, 23:1495–1518, juin 2007.
- Chen, K. : *Matrix preconditioning techniques and applications*. Cambridge University Press, Cambridge, Royaume-Uni, 1ère éd., 2005.

- Chen, S., Donoho, D. et Saunders, M. : Atomic decompositions by basis pursuit. *SIAM Journal on Scientific Computing*, 20:30–61, 1998.
- Chen, X. et Sun, J. : Global convergence of two-parameter family of conjugate gradient methods without line search. *Journal of Computational and Applied Mathematics*, 146:37–45, 2002.
- Chiang, Y.-W., Borbat, P. P. et Freed, J. H. : Maximum entropy : A complement to Tikhonov regularization for determination of pair distance distributions by pulsed ESR. *Journal of Magnetic Resonance*, 177(2):184–196, décembre 2005.
- Chouzenoux, E., Idier, J. et Moussaoui, S. : A Majorize-Minimize strategy for subspace optimization applied to image restoration. Rap. tech., IRCCyN, septembre 2010a. <http://hal.archives-ouvertes.fr/hal-00516585>.
- Chouzenoux, E., Moussaoui, S. et Idier, J. : A majorize-minimize line search algorithm for barrier function optimization. In *Proc. 17th EUSIPCO*, p. 1379–1393, Glasgow, Royaume-Uni, août 2009a.
- Chouzenoux, E., Moussaoui, S., Idier, J. et Mariette, F. : Reconstruction d’un spectre RMN 2D par maximum d’entropie. In *Actes 22e coll. GRETSI*, Dijon, septembre 2009b.
- Chouzenoux, E., Moussaoui, S., Idier, J. et Mariette, F. : Efficient maximum entropy reconstruction of nuclear magnetic resonance T1-T2 spectra. *IEEE Transactions on Signal Processing*, à paraître, 2010b.
- Chouzenoux, E., Moussaoui, S., Idier, J. et Mariette, F. : Optimization of a maximum entropy criterion for 2D nuclear magnetic resonance reconstruction. In *Proc. 35th IEEE ICASSP*, Dallas, TX, USA, mai 2010c.
- Combettes, P. L. et Pesquet, J.-C. : A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery. *IEEE Journal of Selected Topics in Signal Processing*, 1(4):564–574, décembre 2007.
- Combettes, P. L. et Pesquet, J.-C. : Proximal splitting methods in signal processing. In Bauschke, H. H., Burachik, R., Combettes, P. L., Elser, V., Luke, D. R. et Wolkowicz, H., édés : *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*. Springer-Verlag, New York, NY, USA, 2010.
- Combettes, P. L. et Wajs, V. R. : Signal recovery by proximal forward-backward splitting. *Multiscale Modeling and Simulation*, 4(4):1168–1200, 2005.
- Conn, A. R., Gould, N., Sartenaer, A. et Toint, P. L. : On iterated-subspace minimization methods for nonlinear optimization. Rap. tech. 94-069, Rutherford Appleton Laboratory, Oxfordshire, Royaume-Uni, mai 1994a. <ftp://130.246.8.32/pub/reports/cgstRAL94069.ps.Z>.
- Conn, A. R., Gould, N. et Toint, P. L. : A note on using alternative second-order models for the subproblems arising in barrier function methods for minimization. *Numerische Mathematik*, 68:17–33, 1994b.
- Conn, A. R., Gould, N. I. M. et Toint, P. L. : Global convergence of a class of trust region algorithms for optimization with simple bounds. *SIAM Journal on Numerical Analysis*, 25(2):433–460, 1988.
- Conn, R. A., Gould, N. I. M. et Toint, P. L. : *Trust-Region Methods*. MPS-SIAM Series on Optimization. Society for Industrial Mathematics, 1987.
- Cragg, E. E. et Levy, A. V. : Study on a supermemory gradient method for the minimization of functions. *Journal of Optimization Theory and Applications*, 4(3):191–205, 1969.

- Csiszàr, I. : Eine informationstheoretische ungleichung und ihre anwendung auf den beweis der ergodizitaet von markoffschen ketten. *Publication of the Mathematical Institute of the Hungarian Academy of Science*, 8:85–108, 1963.
- Csiszàr, I. : Information-type measures of difference of probability distributions and indirect observations. *Studia Scientiarum Mathematicarum Hungarica*, 2:299–318, 1967.
- Csiszàr, I. : Information measures : A critical survey. In *Prague Conference on Information Theory, Academia Prague*, vol. A, p. 73–86, 1974.
- Csiszàr, I. : Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems. *Annals of Statistics*, 19(4):2032–2066, 1991.
- Dai, Y. : On the nonmonotone line search. *Journal of Optimization Theory and Applications*, 112(2):315–330, 2002.
- Dai, Y. et Yuan, Y. : A nonlinear conjugate gradient method with a strong global convergence property. *SIAM Journal on Optimization*, 10:177–182, 1999.
- Dai, Y. et Yuan, Y. : A class of globally convergent conjugate gradient methods. *Science in China*, 46(2):251–261, 2002.
- Dai, Y. et Yuan, Y. : A three-parameter family of nonlinear conjugate gradient methods. *Mathematics of Computation*, 70:1155–1167, 2001.
- Daniel, J. W. : The conjugate gradient for linear and nonlinear operator equations. *SIAM Journal on Numerical Analysis*, 4:10–26, 1967.
- Danielle, G. J. et Gull, S. F. : Maximum entropy algorithm applied to image enhancement. *Proceedings of the IEEE*, 127(5):170–172, septembre 1980.
- Daube-Witherspoon, M. E. et Muehllehner, G. : An iterative image space reconstruction algorithm suitable for volume ECT. *IEEE Transactions on Medical Imaging*, 5(2):61–66, 1986.
- Daubechies, I., Defrise, M. et De Mol, C. : An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*, 57(11):1413–1457, 2004.
- De Pierro, A. : On the relation between the ISRA and the EM algorithm for positron emission tomography. *IEEE Transactions on Medical Imaging*, 12(2):328–333, juin 1993.
- De Pierro, A. : A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography. *IEEE Transactions on Medical Imaging*, 14(1):132–137, mars 1995.
- Dembo, R. S. et Steihaug, T. : Truncated-Newton methods algorithms for large scale unconstrained optimization. *Mathematical Programming*, 26:190–212, 1983.
- Dembo, R., Eisenstat, S. C. et Steihaug, S. : Inexact Newton methods. *SIAM Journal on Numerical Analysis*, 19(2):400–408, avril 1982.
- Demoment, G. : Image reconstruction and restoration : Overview of common estimation structure and problems. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37(12):2024–2036, décembre 1989.
- Dempster, A. P., Laird, N. M. et Rubin, D. B. : Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society B*, 39:1–38, 1977.
- Dennis, J. E. et Moré, J. J. : Quasi-Newton methods, motivation and theory. *SIAM Review*, 19:46–89, 1977.

- Dennis, JR., J. E. et Schnabel, R. B. : *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Classics in Applied Mathematics 16. Soc for Industrial & Applied Math, 1996.
- Donoho, D. : Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.
- Donoho, D., Johnstone, I. M., Stern, A. S. et Hoch, J. C. : Does the maximum entropy method improve sensitivity. *In National Academy of Sciences of the United States of America*, vol. 87, p. 5066–5068, 1990.
- Doyle, M. : *A Barrier Algorithm for Large Nonlinear Optimization Problems*. thèse de doctorat, University of Stanford, 2003. [www.stanford.edu/group/SOL/dissertations/maureenthesis.pdf](http://www.stanford.edu/group/SOL/dissertations/maureenthesis.pdf).
- Dusaussoy, N. J. et Abdou, I. E. : The extended MENT algorithm : a maximum entropy type algorithm using prior knowledge for computerized tomography. *IEEE Transactions on Signal Processing*, 39(5):1164–1180, 1995.
- Eckart, C. et Young, G. : The approximation of one matrix by another of lower rank. *Psychometrika*, 1(3):211–218, 1936.
- Eggermont, P. P. B. : Multiplicative iterative algorithms for convex programming. *Linear Algebra and its Applications*, 130:25–42, 1990.
- Eggermont, P. P. B. : Maximum entropy regularization for Fredholm integral equations of the first kind. *SIAM Journal on Matrix Analysis and Applications*, 24(6):1557–1576, 1993.
- Elad, M., Matalon, B. et Zibulevsky, M. : Coordinate and subspace optimization methods for linear least squares with non-quadratic regularization. *Applied and Computational Harmonic Analysis*, 23:346–367, 2006.
- Elad, M., Milanfar, P. et Rubinstein, R. : Analysis versus synthesis in signal priors. *Inverse Problems*, 23(3):947–968, 2007.
- Elfving, T. : On some methods for entropy maximization and matrix scaling. *Linear Algebra and its Applications*, 34:321–339, décembre 1980.
- English, A. E., Whittall, K. P., Joy, M. L. G. et Henkelman, R. M. : Quantitative two-dimensional time correlation relaxometry. *Magnetic Resonance in Medicine*, 22:425–434, 1991.
- Erdogan, H. et Fessler, J. : Monotonic algorithms for transmission tomography. *IEEE Transactions on Medical Imaging*, 18(9):801–814, septembre 1999.
- Ernst, R. R., Bodenhausen, G. et Wokaun, A. : *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*. International Series of Monographs on Chemistry. Oxford University Press, Oxford, NY, USA, 2ème édn, 1997.
- Fessler, J. A. et Booth, S. D. : Conjugate-gradient preconditioning methods for shift-variant PET image reconstruction. *IEEE Transactions on Image Processing*, 8(5):688–699, mai 1999.
- Fessler, J. et Erdogan, H. : A paraboloidal surrogates algorithm for convergent penalized-likelihood emission image reconstruction. *IEEE Nuclear Science Symposium*, 2:1132–1135, 1998.
- Fiacco, A. V. et McCormick, G. P. : The sequential unconstrained minimization technique (SUMT) without parameters. *Operations Research*, 15(5):820–827, septembre 1967.
- Figueiredo, M., Nowak, R. et Wright, S. : Gradient projection for sparse reconstruction : Application to compressed sensing and other inverse problems. *IEEE Journal of Selected Topics in Signal Processing*, 1(4):586–597, décembre 2007.

- Fletcher, R. : *Practical methods of optimization*. Wiley-Interscience, New York, NY, USA, 2ème éd., 1987.
- Fletcher, R. : On the Barzilai-Borwein method. *In Optimization and Control with Applications*, vol. 96, p. 235–256. Springer, New York, NY, USA, juin 2006.
- Fletcher, R. et Reeves, C. M. : Function minimization by conjugate gradients. *Comp. J.*, 7:149–157, 1964.
- Forsgren, A., Gill, P. et Wright, M. : Interior methods for nonlinear optimization. *SIAM Review*, 44(4):525–597, 2002.
- Frieden, B. : Image enhancement and restoration. *In Picture Processing and Digital Filtering*, vol. 6 de *Topics in Applied Physics*, p. 177–248. Springer-Verlag, New York, NY, USA, 1975.
- Fuchs, J.-J. : Convergence of a sparse representations algorithm applicable to real or complex data. *IEEE Journal of Selected Topics in Signal Processing*, 1(4):598–605, décembre 2007.
- Galatsanos, N. et Katsaggelos, A. : Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation. *IEEE Transactions on Image Processing*, 1(3):322–336, juillet 1992.
- Gallegos, D. P. et Smith, D. M. : A NMR technique for the analysis of pore structure : Determination of continuous pore size distributions. *Journal of Colloid and Interface Science*, 122(1):143–153, 1988.
- Geman, D. et Reynolds, G. : Constrained restoration and the recovery of discontinuities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(3):367–383, mars 1992.
- Geman, D. et Yang, C. : Nonlinear image recovery with half-quadratic regularization. *IEEE Transactions on Image Processing*, 4(7):932–946, juillet 1995.
- Geman, S. et Geman, D. : Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6(6):721–741, novembre 1984.
- Geman, S. et McClure, D. : Statistical methods for tomographic image reconstruction. *In Proceedings of the 46th Session of the ICI, Bulletin of the ICI*, vol. 52, p. 5–21, 1987.
- Gilbert, J. C. et Nocedal, J. : Global convergence properties of conjugate gradient methods for optimization. *SIAM Journal on Optimization*, 2(1):21–42, 1992.
- Gilles, L., Vogel, C. et Bardsley, J. : Computational methods for a large-scale inverse problem arising in atmospheric optics. *Inverse Problems*, 18(1):237–252, 2002.
- Glineur, F. : Self-concordant functions in structured convex optimization. Rap. tech. 7, Faculté Polytechnique, MonsBelgique, octobre 2000. <http://www.core.ucl.ac.be/~glineur/Papers/Image0007.ps.gz>.
- Goldstein, A. et Price, J. : An effective algorithm for minimization. *Numerische Mathematik*, 10:184–189, 1967.
- Golub, G. H. et Matt, V. U. : Quadratically constrained least squares and quadratic problems. *Numerische Mathematik*, 59(1):561–580, 1991.
- Golub, G. H., Heath, M. et Wahba, G. : Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21(2):215–223, mai 1979.
- Golub, G. H. et Van Loan, C. F. : *Matrix computations*. John Hopkins University Press, Baltimore, 3ème éd., 1996.

- Gonzaga, C. C. et Carlos, L. A. : A primal affine-scaling algorithm for constrained convex programs. Rap. tech. ES-238/90, Federal University of Rio de Janeiro Brésil, 1990. revised in 2002, [http://www.optimization-online.org/DB\\_FILE/2002/09/531.ps](http://www.optimization-online.org/DB_FILE/2002/09/531.ps).
- Gordon, R. and Herman, G. T. : Reconstruction of pictures from their projections. *Communications of the ACM*, 14(12):759–768, 1971.
- Gorodnitsky, I. F. : An extension of an interior-point method for entropy minimization. *Proceedings of the IEEE International Conference on Acoustic, Speech and Signal Processing*, 3:1697–1700, 1999.
- Gould, N., Orban, D. et Toint, P. : Numerical methods for large-scale nonlinear optimization. *Acta Numerica*, 14:299–361, 2005.
- Grippo, L., Lampariello, F. et Lucide, S. : A nonmonotone line search technique for Newton’s method. *SIAM Journal on Numerical Analysis*, 23(4):707–716, août 1986.
- Gull, S. F. et Skilling, J. : Maximum entropy method in image processing. *Proceedings of the IEEE*, 131:646–659, octobre 1984.
- Hadamard, J. : Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton Univ. Bull.*, 13:49–52, 1902.
- Hager, W. W. et Zhang, H. : Recent advances in bound constrained optimization. In *System Modeling and Optimization*, vol. 199 de *IFIP International Federation for Information Processing*, p. 67–82. Springer, 2006a.
- Hager, W. W. et Zhang, H. : A survey of nonlinear conjugate gradient methods. *Pacific Journal of Optimization*, 2(1):35–58, janvier 2006b.
- Hansen, P. C. : *Rank-deficient and discrete ill-posed problems : numerical aspects of linear inversion*. SIAM, Philadelphia, PA, USA, 1998.
- Hansen, P. C. et O’Leary, D. P. : The use of the L-curve in the regularization of discrete ill-posed problems. *SIAM Journal on Scientific Computing*, 14(6):1497–1503, novembre 1993.
- Hansen, P. : Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Review*, 34:561–580, 1992.
- Hertog, D. D. : *Interior Point Approach to Linear, Quadratic and Convex Programming*. Mathematics and its Applications 277. Kluwer Academic Publishers, 1994.
- Hestenes, M. R. et Stiefel, E. L. : Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards*, 49(6):409–436, 1952.
- Holland, J. H. : *Adaptation in natural and artificial systems*. MIT Press, Cambridge, MA, USA, 1992.
- Huber, P. J. : *Robust Statistics*. John Wiley, New York, NY, USA, 1981.
- Hunter, D. R. et L., K. : A tutorial on MM algorithms. *The American Statistician*, 58(1):30–37, février 2004.
- Hürlimann, M. D. et Venkataramanan, L. : Quantitative measurement of two-dimensional distribution functions of diffusion and relaxation in grossly inhomogeneous fields. *Journal of Magnetic Resonance*, 157:31–42, 2002.
- Idier, J. : Introduction. In Idier, J., éd. : *Bayesian Approach to Inverse Problems*, p. 15–22. ISTE Ltd and John Wiley & Sons Inc, avril 2008.

- Iusem, A. N. : On the convergence properties of the projected gradient method for convex optimization. *Computational and Applied Mathematics*, 22(1):37–52, 2003.
- Iusem, A. N., Svaiter, B. F. et Teboulle, M. : Entropy-like proximal methods in convex programming. *Mathematics of Operations Research*, 19:790–814, 1994.
- Jacobson, M. et Fessler, J. : An expanded theoretical treatment of iteration-dependent majorize-minimize algorithms. *IEEE Transactions on Image Processing*, 16(10):2411–2422, octobre 2007.
- Jansson, P. : *Deconvolution of Images and Spectra*. Academic Press, San Diego, California, USA, 2ème édn, 1997.
- Jarre, F. : Interior-point methods via self-concordance or relative Lipschitz condition. *Optimization Methods and Software*, 5(1), 1995.
- Jaynes, E. T. : On the rationale of maximum-entropy methods. *Proceedings of the IEEE*, 70:939–952, 1982.
- Ji, H., Huang, M., Moore, J. et Manton, J. : A globally convergent conjugate gradient method for minimizing self-concordant functions with application to constrained optimization problems. *In American Control Conference*, p. 540–545, juillet 2007.
- Jiang, L., Byrd, R. H., Eskow, E. et Schnabel, R. B. : A preconditioned L-BFGS algorithm with application to molecular energy minimization. Rap. tech. CU-CS-982-04, University of Colorado, Boulder, Colorado, USA, novembre 2004.
- Johnson, C. A. et McGarry, D. : Maximum entropy reconstruction methods in electron paramagnetic resonance imaging. *Annals of Operations Research*, 119:101–118, 2003.
- Johnson, C. A., Seidel, J. et Sofer, A. : Interior-point methodology for 3-D PET reconstruction. *IEEE Transactions on Medical Imaging*, 19(4), avril 2000.
- Kamm, J. et Nagy, J. G. : Kronecker product and SVD approximations in image restoration. *Linear Algebra and its Applications*, 284(1-3):177–192, novembre 1998.
- Kaufman, L. : Implementing and accelerating the EM algorithm for positron emission tomography. *IEEE Transactions on Medical Imaging*, 6(1), mars 1987.
- Khalfan, H., Byrd, R. et Schnabel, R. B. : A theoretical and experimental study of the symmetric rank-one update. *SIAM Journal on Optimization*, 3(1):1–24, 1993.
- Khurd, P., Xing, Y., Hsiao, I. T. et Gindi, G. : Fast preconditioned conjugate gradient reconstruction for 2D SPECT. *In IEEE Nuclear Science Symposium and Medical Imaging Conference*, p. 741–745, 2002.
- Labat, C. : *Algorithmes d’optimisation de critères pénalisés pour la restauration d’images. Application à la déconvolution de trains d’impulsions en imagerie ultrasonore*. Thèse de doctorat, Ecole Centrale Nantes, décembre 2006. <http://tel.archives-ouvertes.fr/tel-00132861>.
- Labat, C. et Idier, J. : Convergence of truncated half-quadratic and Newton algorithms, with application to image restoration. Rap. tech., IRCCyN, juin 2007. <http://www.irccyn.ec-nantes.fr/~idier/pub/labat07b.pdf>.
- Labat, C. et Idier, J. : Convergence of conjugate gradient methods with a closed-form stepsize formula. *Journal of Optimization Theory and Applications*, 136(1):43–60, janvier 2008.
- Lam, E. Y. et Goodman, J. W. : Iterative statistical approach to blind image deconvolution. *Journal of the Optical Society of America (A)*, 17(7):1177–1184, 2000.

- Lamanna, R. : On the inversion of multicomponent NMR relaxation and diffusion decays in heterogeneous systems. *Concepts in Magnetic Resonance Part A*, 26A(2):78–90, 2005.
- Lange, K. : Convergence of EM image reconstruction algorithms with gibbs smoothing. *IEEE Transactions on Medical Imaging*, 9(4):439–446, 1990.
- Lange, K. : An adaptive barrier method for convex programming. *Methods and Applications of Analysis*, 1(4):392–402, 1994.
- Lange, K., Bahn, M. et Little, R. : A theoretical study of some maximum likelihood algorithms for emission and transmission tomography. *IEEE Transactions on Image Processing*, 6(2):106–114, juin 1987.
- Lange, K. et Fessler, J. A. : Globally convergent algorithms for maximum a posteriori transmission tomography. *IEEE Transactions on Image Processing*, 4(10):1430–1438, octobre 1995.
- Lanteri, H., Roche, M. et Aime, C. : Penalized maximum likelihood image restoration with positivity constraints : multiplicative algorithms. *Inverse Problems*, 18:1397–1419, 2002.
- Lanteri, H., Roche, M., Cuevas, O. et Aime, C. : A general method to devise maximum-likelihood signal restoration multiplicative algorithms with non-negativity constraints. *Signal Processing*, 81:945–974, 2001.
- Laub, A. J. : *Matrix Analysis For Scientists And Engineers*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2004.
- Laue, E. D., Skilling, J., Staunton, J., Sibisi, S. et Brereton, R. G. : Maximum entropy method in nuclear magnetic resonance spectroscopy. *Journal of Magnetic Resonance*, 62(3):437–452, 1985.
- Le Besnerais, G., Bercher, J.-F. et Demoment, G. : A new look at entropy for solving linear inverse problems. *IEEE Transactions on Information Theory*, 45(5):1565–1578, 1999.
- Lee, D. D. et Seung, H. S. : Algorithms for non-negative matrix factorization. In *Advances in Neural and Information Processing Systems*, vol. 13, p. 556–562, 2001.
- Lent, A. : The primal-dual algorithm as a constraint-set-manipulation device. *Mathematical Programming*, 50:343–357, 1991.
- Lewitt, R. M. et Matej, S. : Overview of methods for image reconstruction from projections in emission computed tomography. *Proceedings of the IEE*, 91(1):1588–1611, 2003.
- Lin, C. : *A Null-Space Primal-Dual Algorithm for NonLinear Network Optimization*. thèse de doctorat, University of Stanford, 2002.
- Liu, D. C. et Nocedal, J. : On the limited memory BFGS method for large scale optimization. *Mathematical Programming*, 45(3):503–528, 1989.
- Liu, J., Liu, H. et Zheng, Y. : A new supermemory gradient method without line search for unconstrained optimization. In *The Sixth International Symposium on Neural Networks*, vol. 56, p. 641–647. Springer Berlin, HeidelbergAllemagne, 2009.
- Liu, Y. et Storey, C. : Efficient generalized conjugate gradient algorithm, part 1 : Theory. *Journal of Optimization Theory and Applications*, 69:129–137, 1991.
- Lu, L.-Z., Ng, M. K. et Lin, F.-R. : Approximation BFGS methods for nonlinear image restoration. *Journal of Computational and Applied Mathematics*, 226(1):84–91, avril 2009.
- Lucy, L. B. : An iterative technique for the rectification of observed distributions. *Astronomical Journal*, 79:745–754, 1974.

- Mallat, G. et Zhang, Z. : Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41(12):3397–3415, 1993.
- Mariette, F. : Relaxation RMN et IRM : un couplage indispensable pour l'étude des produits alimentaires. *Compte-rendus de l'académie des sciences, Chimie* 7:221–232, 2004.
- Mariette, F., Guillement, J. P., Tellier, C. et Marchal, P. : Continuous relaxation time distribution decomposition by MEM. *Signal Treatment and Signal Analysis in NMR*, p. 218–234, 1996.
- Maros, I. et Mészáros, C. : A repository of convex quadratic programming problems. *Optimization Methods and Software*, 11(1-4):671–681, 1999.
- Miele, A. et Cantrell, J. W. : Study on a memory gradient method for the minimization of functions. *Journal of Optimization Theory and Applications*, 3(6):459–470, 1969.
- Mohammad-Djafari, A. et Demoment, G. : Maximum entropy image reconstruction in X-ray and diffraction tomography. *IEEE Transactions on Medical Imaging*, 7(4):345–354, 1988a.
- Mohammad-Djafari, A. et Demoment, G. : Utilisation de l'entropie dans les problèmes de restauration et de reconstruction d'images. *Traitement du Signal*, 5(4):235–248, 1988b.
- Moré, J. J. et Sorensen, D. C. : Computing a trust region step. *SIAM Journal on Statistical Scientific Computing*, 4:553–572, 1983.
- Moré, J. J. et Thuente, D. J. : Line search algorithms with guaranteed sufficient decrease. *ACM Transactions on Mathematical Software*, 20(3):286–307, 1994.
- Moré, J. J. et Toraldo, G. : Algorithms for bound constrained quadratic programming problems. *Numerische Mathematik*, 55:377–400, 1989.
- Mumcuoglu, E. et Leahy, R. : A gradient projection conjugate gradient algorithm for bayesian 'pet' reconstruction. In *IEEE Nuclear Science Symposium and Medical Imaging Conference*, vol. 3, p. 1212–1216, Norfolk, VA, USA, octobre 1994.
- Murray, W. : Analytical expressions for the eigenvalues and eigenvectors of the Hessian matrices of barrier and penalty functions. *Journal of Optimization Theory and Applications*, 7(3):189–196, 1971.
- Murray, W. et Wright, M. H. : Line search procedures for the logarithmic barrier function. *SIAM Journal on Optimization*, 4(2):229–246, 1994.
- Naranayan, A., Hartman, J. S. et Bain, A. D. : Characterizing nonexponential spin-lattice relaxation in solid-state NMR by fitting to the stretched exponential. *Journal of Magnetic Resonance*, 112:58–65, 1995.
- Narkiss, G. et Zibulevsky, M. : Sequential subspace optimization method for large-scale unconstrained problems. Rap. tech. 559, Israel Institute of Technology, octobre 2005.
- Narushima, Y. et Yabe, H. : Global convergence of a memory gradient method for unconstrained optimization. *Computational Optimization and Applications*, 35(3):325–346, juin 2006.
- Nash, S. G. : A survey of truncated-Newton methods. *Journal of Computational and Applied Mathematics*, 124:45–59, 2000.
- Nash, S. G. et Sofer, A. : A barrier method for large-scale constrained optimization. *ORSA Journal on Computing*, 5(41), 1993.
- Nash, S. G. et Sofer, A. : On the complexity of a practical interior-point method. *SIAM Journal on Optimization*, 8(3):833–849, 1998.

- Nashed, M. Z. : Operator-theoretic and computational approaches to ill-posed problems with applications to antenna theory. *IEEE Transactions on Antennas and Propagation*, 29:220–231, 1981.
- Nemirovski, A. : Orth-method for smooth convex optimization. *Izvestia AN SSSR, Transl. : Eng. Cybern. Soviet J. Comput. Syst. Sci.*, 2, 1982.
- Nesterov, Y. : *Introductory lectures on convex optimization*. Num. 87 de Applied Optimization. Kluwer Academics, Norwell, Massachusetts, USA, 2003.
- Nesterov, Y. et Nemirovski, A. : *Interior point polynomial algorithms in convex programming*. Num. 13 de Studies in Applied and Numerical Mathematics. SIAM, Philadelphia, Pennsylvania, USA, 1994.
- Nesterov, Y. E. : Smooth minimization of non-smooth functions. *Mathematical Programming*, 103(1):127–152, 2005.
- Ng, M. K., Chan, R. H. et Tang, W.-C. : A fast algorithm for deblurring models with Neumann boundary conditions. *SIAM Journal on Scientific Computing*, 21(3):851–866, 1999.
- Ni, Q. et Yuan, Y. : A subspace limited memory quasi-Newton algorithm for large-scale nonlinear bound constrained optimization. *Mathematics of Computation*, 66(220):1509–1520, 1997.
- Nikolova, M. : Minimizers of cost-functions involving nonsmooth data-fidelity terms. application to the processing of outliers. *SIAM Journal on Numerical Analysis*, 40(3):965–994, 2002.
- Nikolova, M. : Weakly constrained minimization : application to the estimation of images and signals involving constant regions. *Journal of Mathematical Imaging and Vision*, 21:155–175, 2004.
- Nikolova, M., Idier, J. et Mohammad-Djafari, A. : Inversion of large-support ill-posed linear operators using a piecewise Gaussian MRF. *IEEE Transactions on Image Processing*, 7(4):571–585, 1998.
- Nikolova, M. et Ng, M. K. : Analysis of half-quadratic minimization methods for signal and image recovery. *SIAM Journal on Scientific Computing*, 27:937–966, 2005.
- Nityananda, R. et Narayan, R. : Maximum entropy image reconstruction—a practical non-information-theoretic approach. *Journal of Astrophysics and Astronomy*, 3(4), décembre 1982.
- Nocedal, J. et Wright, S. J. : *Numerical Optimization*. Springer-Verlag, New York, NY, USA, 1999.
- Ortega, J. M. et Rheinboldt, W. C. : *Iterative solution of nonlinear equations in several variables*. Academic Press, New York, USA, 1970.
- O’Sullivan, J. : Roughness penalties on finite domains. *IEEE Transactions on Image Processing*, 4(9), septembre 1995.
- Pichon, C. et Thiébaud, E. : Non-parametric reconstruction of distribution functions from observed galactic discs. *Monthly Notices of the Royal Astronomical Society*, 301(2):419–434, 1998.
- Polak, E. et Ribière, G. : Note sur la convergence de directions conjuguées. *Revue française d’informatique et de recherche opérationnelle*, 3(1):35–43, 1969.
- Polyak, B. T. : The conjugate gradient method in extreme problems. *USSR Computational Mathematics and Mathematical Physics*, 9:94–112, 1969.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T. et Flannery, B. P. : *Numerical Recipes : The Art of Scientific Computing*. Cambridge Univ. Press, New York, 3ème édn, 1992.

- Purcell, E. M. : Research in nuclear magnetism. *Nobel Lecture*, décembre 1952.
- Pytlak, R. : An efficient algorithm for large-scale nonlinear programming problems with simple bounds on the variables. *SIAM Journal on Optimization*, 8(2):532–560, 1998.
- Pytlak, R. : *Conjugate Gradient Algorithms in Nonconvex Optimization*. Springer Berlin Heidelberg, 1ère éd., 2009.
- Raydan, M. : The Barzilai and Borwein gradient method for the large scale unconstrained minimization problem. *SIAM Journal on Optimization*, 7(1):26–33, février 1997.
- Reeves, S. J. : *A cross-validation approach to image restoration and blur identification*. Thèse de doctorat, Georgia Institute of Technology, Atlanta, USA, août 1990. <ftp://ftp.eng.auburn.edu/pub/sjreeves/thesis.pdf>.
- Reeves, S. J. et Mersereau, R. M. : Optimal estimation of the regularization parameters and stabilizing functional for regularized image restoration. *Optical Engineering*, 29:446–454, mai 1990.
- Rice, J. A. : Choice of smoothing parameter in deconvolution problems. In *Function Estimates*, vol. 59 de *Contemporary Mathematics*, p. 137–151. American Mathematical Society, 1986.
- Richardson, W. H. : Bayesian-based iterative method of image restoration. *Journal of the Optical Society of America*, 62(1):55–59, 1972.
- Rivera, M. et Marroquin, J. : Efficient half-quadratic regularization with granularity control. *Image and Vision Computing*, 21(4):345–357, avril 2003.
- Rockafellar, R. T. : *Convex analysis*. Princeton University Press, Princeton, NJ, USA, 1ère éd., 1970.
- Rockafellar, R. T. : A dual approach to solving nonlinear programming problems by unconstrained minimization. *Mathematical Programming*, 5:354–373, 1973.
- Rockafellar, R. T. : Monotone operators and the proximal point algorithm. *SIAM Journal on Control*, 14:877–898, 1976.
- Roos, C., Terlaky, T. et Vial, J. : *Interior Point Methods for Linear Optimization*. Springer-Verlag, New York, NY, USA, 2ème éd., 2006.
- Rudin, L. I., Osher, S. et Fatemi, E. : Nonlinear total variation based noise removal algorithms. *Physical Review D*, 60:259–268, 1992.
- Saad, Y. : *Iterative methods for sparse linear systems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2003.
- Segalat, P. : *Méthodes de points intérieurs et de Quasi-Newton*. thèse de doctorat, Université de Limoges, 2002. <http://tel.archives-ouvertes.fr/tel-00005478/en/>.
- Shi, Y. : Solving linear systems involved in constrained optimization. *Linear Algebra and its Applications*, 229:175–189, 1995.
- Shi, Z.-J. : Convergence of line search methods for unconstrained optimization. *Applied Mathematics and Computations*, 157:393–405, 2004.
- Shi, Z.-J. et Guo, J. : A new algorithm of nonlinear conjugate gradient method with strong convergence. *Computational and Applied Mathematics*, 27(1):93–106, 2008.
- Shi, Z.-J. et Shen, J. : A new super-memory gradient method with curve search rule. *Applied Mathematics and Computations*, 170:1–16, 2005.

- Shi, Z.-J. et Shen, J. : Convergence of nonmonotone line search method. *Journal of Computational and Applied Mathematics*, 193:397–412, 2006.
- Shi, Z.-J. et Shen, J. : Convergence of supermemory gradient method. *Applied Mathematics and Computations*, 24(1-2):367–376, 2007.
- Shi, Z.-J. et Xu, Z. : The convergence of subspace trust region methods. *Journal of Computational and Applied Mathematics*, 231(1):365–377, 2009.
- Skilling, J. et Bryan, R. K. : Maximum entropy image reconstruction : General algorithm. *Monthly Notices of the Royal Astronomical Society*, 211:111–124, 1984.
- Snyder, D. L., Hammoud, A. M. et White, R. L. : Image recovery from data acquired with a charge-coupled-device camera. *Journal of the Optical Society of America (A)*, 10(5):1014–1023, 1993.
- Song, Y. Q., Venkataramanan, L., Hürlimann, M. D., Flaum, M., Frulla, P. et Straley, C. : T1-T2 correlation spectra obtained using a fast two-dimensional Laplace inversion. *Journal of Magnetic Resonance*, 154:261–268, 2002.
- Sotthivirat, S. et Fessler, J. : Image recovery using partitioned-separable paraboloidal surrogate coordinate ascent algorithms. *IEEE Transactions on Image Processing*, 11(3):306–317, mars 2002.
- Soulez, F. : *Une approche “problèmes inverses” pour la reconstruction de données multi-dimensionnelles par méthodes d’optimisation*. Thèse de doctorat, Université de Saint-Etienne, décembre 2008. [http://tel.archives-ouvertes.fr/docs/00/47/69/16/PDF/these\\_ferreol.pdf](http://tel.archives-ouvertes.fr/docs/00/47/69/16/PDF/these_ferreol.pdf).
- Sternin, E. : Use of inverse theory algorithms in the analysis of biomembrane NMR data. *In Methods in Membrane Lipids*, vol. 400, p. 103–125. Humana Press, 2008.
- Sun, J. et Zhang, J. : Global convergence of conjugate gradient methods without line search. *Annals of Operations Research*, 103:161–173, mars 2001.
- Sun, W. et Yuan, Y.-X. : *Optimization Theory and Methods*, vol. 1 de *Springer Optimization and Its Applications*. 2006.
- Taneja, I. J. et Krumar, P. : Relative information of type  $s$ , Csiszàr’s  $f$ -divergence, and information inequalities. *Information Science*, 166:105–125, 2004.
- Teboulle, M. : Convergence of proximal-like algorithms. *SIAM Journal on Optimization*, 7(4):1069–1083, 1997.
- Tikhonov, A. : Regularization of incorrectly posed problems. *Soviet. Math. Dokl.*, 4:1624–1627, 1963.
- Tikhonov, A. et Arsenin, V. : *Solutions of Ill-Posed Problems*. Winston, Washington, DC, USA, 1977.
- Tikhonov, A. et Arsénine, V. : *Méthodes de résolution de problèmes mal posés*. Éditions MIR, Moscou, Russie, 1976.
- Titterton, D. M. : Common structure of smoothing techniques in statistics. *International Statistical Review*, 53(2):141–170, 1985a.
- Titterton, D. M. : General structure of regularization procedures in image restoration. *Astronomy and Astrophysics*, 144:381–387, 1985b.

- Tropp, J. : Just relax : convex programming methods for identifying sparse signals. *IEEE Transactions on Information Theory*, 51:1030–1051, 2006.
- Trummer, M. R. : SMART - an algorithm for reconstructing pictures from projections. *Journal of Applied Mathematics and Physics*, 34:746–753, 1983.
- Trussell, H. J. : Convergence criteria for iterative restoration methods. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 31(1):129–136, 1983.
- Tseng, P. : An analysis of the EM algorithm and entropy-like proximal point methods. *Mathematics of Operations Research*, 29(1):27–44, 2004.
- Tseng, P. et Bertsekas, D. P. : On the convergence of the exponential multiplier method for convex programming. *Mathematical Programming*, 60:1–19, 1993.
- Van de Ven, F. J. M. : *Multidimensional NMR in Liquids. Basic Principles and Experimental Methods*. Wiley-VCH, New York, NY, USA, 1995.
- Van Loan, C. F. : The ubiquitous Kronecker product. *Journal of Computational and Applied Mathematics*, 123:85–100, 2000.
- Van Loan, C. et Pitsianis, N. : *Approximation with Kronecker products*, p. 293–314. Kluwer Publications, 1993.
- Vardi, Y., Shepp, L. A. et Kaufman, L. : A statistical model for positron emission tomography. *Journal of American Statistical Association*, 80(389):8–37, 1985.
- Venkataramanan, L., Song, Y. Q. et Hürlimann, M. D. : Solving Fredholm integrals of the first kind with tensor product structure in 2 and 2.5 dimensions. *IEEE Transactions on Signal Processing*, 50(5):1017–1026, 2002.
- Wang, Y. et Zhao, X. : An iterative method for maximum entropy regularization reconstruction in MRI. *International Journal of Imaging Systems and Technology*, 10(6):427–431, 2000.
- Wang, Z., Wen, Z. et Yuan, Y. : A subspace trust region method for large scale unconstrained optimization. In *Numerical linear algebra and optimization*, p. 264–274. Science Press, Marrickville, SydneyAustralie, 2004.
- Weiss, P., Blanc-Féraud, L. et Aubert, G. : Efficient schemes for total variation minimization under constraints in image processing. *SIAM Journal on Scientific Computing*, 31(3):2047–2080, 2010.
- Willis, M. : Algebraic reconstruction algorithms for remote sensing image enhancement. Mémoire de D.E.A., Brigham Young University, Utah, USA, août 2000. [http://www.mers.byu.edu/docs/thesis/msthesis\\_willism\\_lib.pdf](http://www.mers.byu.edu/docs/thesis/msthesis_willism_lib.pdf).
- Willis, M., Jeffes, B. et Long, D. : Maximum entropy image restoration revisited. *Proceedings of the International Conference on Image Processing*, 1:89–92, 2000.
- Wolfe, M. et Viazminsky, C. : Supermemory descent methods for unconstrained minimization. *Journal of Optimization Theory and Applications*, 18(4):455–468, 1976.
- Wolfe, P. : Convergence conditions for ascent methods. *SIAM Review*, 11:226–235, 1969.
- Wright, M. H. : Interior methods for constrained optimization. In *Acta Numerica 1992*, p. 341–407. Cambridge University Press, 1991.
- Wright, M. H. : Some properties of the Hessian of the logarithmic barrier function. *Mathematical Programming*, 67(2):265–295, 1994.
- Wright, M. H. : Why a pure primal Newton barrier step may be infeasible. *SIAM Journal on Optimization*, 5:1–12, 1995.

- Wright, M. H. : Ill-conditioning and computational error in interior methods for nonlinear programming. *SIAM Journal on Optimization*, 9(1):84–111, 1998.
- Wright, S. J. : *Primal-dual interior-point methods*. SIAM, Philadelphia, PA, USA, 1ère édn, 1997.
- Xiao, Y. et Zengxin, W. : A new subspace limited memory BFGS algorithm for large-scale bound constrained optimization. *Applied Mathematics and Computations*, 185:350–359, 2007.
- Yarlagadda, R., Bednar, J. B. et Watt, T. L. : Fast algorithms for  $\ell_p$  deconvolution. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 33(1):174–182, février 1985.
- Yu, Z. : Global convergence of a memory gradient method without line search. *Journal of Applied Mathematics and Computing*, 26(1-2):545–553, février 2008.
- Yuan, Y. : Subspace techniques for nonlinear optimization. *In Some Topics in Industrial and Applied Mathematics*, vol. 8, p. 206–218. Series on Concrete and Applicable Mathematics, 2007.
- Yuan, Y.-X. et Stoer, J. : A subspace study on conjugate gradient algorithms. *Zeitschrift für angewandte Mathematik und Mechanik*, 75(1):69–77, 1995.
- Zeng, G. L. : Image reconstruction : A tutorial. *Computerized Medical Imaging and Graphics*, 25(2):97–103, 2001.
- Zhang, H. et Hager, W. : A nonmonotone line search technique and its application to unconstrained optimization. *SIAM Journal on Optimization*, 14(4):1043–1056, 2004.
- Zheng, J., Saquib, S. S., Sauer, K. et Bouman, C. A. : Parallelizable bayesian tomography algorithms with rapid, guarenteed convergence. *SIAM Journal on Optimization*, 14(4):1043–1056, 2004.
- Zibulevsky, M. : *Penalty / Barrier multiplier methods for large-scale nonlinear and semidefinite programming*. Thèse de doctorat, Technion - Israël Institute of Technology, septembre 1996. <http://iew3.technion.ac.il/~mcib/fulthes.ps.gz>.
- Zibulevsky, M. : SESOP-TN : Combining sequential subspace optimization with truncated Newton method. Rap. tech., Israel Institute of Technology, septembre 2008. [http://www.optimization-online.org/DB\\_FILE/2008/09/2098.pdf](http://www.optimization-online.org/DB_FILE/2008/09/2098.pdf).
- Zibulevsky, M. et Elad, M. :  $\ell_2 - \ell_1$  optimization in signal and image processing. *IEEE Signal Processing Magazine*, 27(3):76–88, mai 2010.



**Titre :** *Recherche de pas par Majoration-Minoration. Application à la résolution de problèmes inverses.*

**Résumé :** La solution des problèmes inverses en traitement du signal et de l'image est souvent définie comme le minimiseur d'un critère pénalisé qui prend en compte conjointement les observations et les informations préalables. Ce travail de thèse s'intéresse à la minimisation des critères pénalisés différentiables. Nous discutons plus précisément de la mise en œuvre algorithmique de l'étape de recherche de pas dans l'algorithme de descente itérative. Les travaux de thèse de Christian Labat (Labat, 2006) ont mené à l'élaboration de la stratégie de pas par Majoration-Minoration quadratique (MMQ 1D). Cette stratégie se démarque des méthodes de pas standards par sa simplicité d'implémentation et ses propriétés de convergence lorsqu'elle est associée à l'algorithme du gradient conjugué non linéaire (GCNL). Nous étendons ces propriétés à la famille des algorithmes à gradient relié. Nous montrons de plus que l'approche MMQ 1D s'étend en une stratégie de pas multi-dimensionnelle MMQ  $r$ D assurant la convergence d'algorithmes de sous-espace. Nous illustrons expérimentalement en déconvolution d'image que l'algorithme de super mémoire de gradient SMG + MMQ 2D est préférable à l'algorithme de gradient conjugué non linéaire GCNL + MMQ 1D.

Lorsque le critère pénalisé contient une barrière, c'est-à-dire une fonction dont le gradient est non borné, la procédure de pas MMQ est inapplicable. Nous développons une stratégie de pas tenant compte de la singularité de la barrière à travers des approximations majorantes quadratiques augmentées d'un terme logarithmique. La recherche de pas résultante, notée MMLQ 1D, est simple à mettre en œuvre et garantit la convergence des algorithmes standards de descente itérative. Nous montrons expérimentalement que la méthode MMLQ 1D accroît les performances de l'algorithme de point intérieur primal pour la programmation quadratique. Nous appliquons enfin cette approche à la reconstruction de spectres RMN bi-dimensionnels par maximum d'entropie.

**Mots-clés :** problèmes inverses, critère pénalisé, optimisation, convergence, descente itérative, recherche de pas, majoration-minoration, sous-espace, barrière logarithmique, distance de Kullback, maximum d'entropie, points intérieurs, résonance magnétique nucléaire.

**Title :** *Majorize-Minimize algorithms for stepsize determination. Application to the resolution of inverse problems.*

**Abstract :** The solution to inverse problems encountered in signal and image processing is often defined as the minimizer of a penalized criterion that accounts simultaneously for the data and the prior. This thesis deals with the minimization of differentiable criteria. We focus more specifically on algorithmic strategies for stepsize determination in iterative descent optimization methods. In Christian Labat's thesis (Labat, 2006), a simple and efficient quadratic majorize-minimize line search method has been proposed (MMQ 1D). This method benefits from strong convergence results when the nonlinear conjugate gradient (NLCG) direction is used. In this thesis, we extend those properties to gradient related algorithms. Moreover, we prove that MMQ 1D generalizes into a multi-dimensional search MMQ  $r$ D that ensures the convergence of several subspace algorithms. We experimentally show that the super memory gradient algorithm SMG + MMQ 2D performs better than NLCG + MMQ 1D algorithms, for an image deconvolution test problem.

MMQ strategies exclude the case of barrier criteria *i.e.*, functions with unbounded gradient. In this thesis, we propose an original MM line search algorithm for barrier functions, based on a quadratic majorizing function augmented with a logarithmic term. This leads to a simple line search MMLQ 1D that ensures the convergence of several classical descent optimization strategies. We experimentally show that the proposal scheme improves the performances of primal interior point algorithm, for a set of quadratic programming test problems. Finally, we apply the MMLQ 1D approach to maximum entropy reconstruction of bi-dimensional NMR spectra.

**Keywords :** inverse problems, penalised criterion, optimization, convergence, iterative descent, line search, majorization-minimization, subspace, logarithmic barrier, Kullback distance, maximum entropy, interior point, nuclear magnetic resonance.

