Algorithmes pour la résolution de problèmes d'optimisation et de minimax

Bernard Martinet

▶ **To cite this version:**

Bernard Martinet. Algorithmes pour la résolution de problèmes d'optimisation et de minimax. Modélisation et simulation. Université Joseph-Fourier - Grenoble I, 1972. tel-00284255

HAL Id: tel-00284255
https://theses.hal.science/tel-00284255

Submitted on 2 Jun 2008

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.

N° d'ordre

T H E S E

présentée à

L'UNIVERSITE SCIENTIFIQUE ET MEDICALE DE GRENOBLE

pour obtenir

LE GRADE DE DOCTEUR ES-SCIENCES

par

Bernard MARTINET

ALGORITHMES POUR LA RESOLUTION DE PROBLEMES D'OPTIMISATION ET DE MINIMAX

Thèse soutenue le 24 avril 1972 devant la Commission d'Examen :

MM. J. KUNTZMANN Président

A. AUSLENDER Examinateur
N. GASTINEL Examinateur
P.J. LAURENT Examinateur
R. TEMAM Examinateur

UNIVERSITE SCIENTIFIQUE

ET MEDICALE DE GRENOBLE

LISTE DES PROFESSEURS

Président : Monsieur Michel SOUTIF
Vice-Président : Monsieur Gabriel CAU

PROFESSEURS TITULAIRES

MM.	ANGLES D'AURIAC Paul	Mécanique des fluides
	ARNAUD Georges	Clinique des maladies infectieuses
	ARNAUD Paul	Chimie
	AYANT Yves	Physique approfondie
Mme	BARBIER Marie-Jeanne	Electrochimie
MM.	BARBIER Jean-Claude	Physique expérimentale
	BARBIER Reynold	Géologie appliquée
	BARJON Robert	Physique nucléaire
	BARNOUD Fernand	Biosynthèse de la cellulose
	BARRA Jean-René	Statistiques
	BARRIE Joseph	Clinique chirurgicale
	BENOIT Jean	Radioélectricité
	BESSON Jean	Electrochimie
	BEZES Henri	Chirurgie générale
	BLAMBERT Maurice	Mathématiques Pures
	BOLLIET Louis	Informatique (IUT B)
	BONNET Georges	Electrotechnique
	BONNET Jean-Louis	Clinique ophtalmologique
	BONNET-EYMARD Joseph	Pathologie médicale
	BONNIER Etienne	Electrochimie Electrométallurgie
	BOUCHERLE André	Chimie et Toxicologie
	BOUCHEZ Robert	Physique nucléaire
	BRAVARD Yves	Géographie
	BRISSONNEAU Pierre	Physique du Solide
	BUYLE-BODIN Maurice	Electronique
	CABANAC Jean	Pathologie chirurgicale
	CABANEL Guy	Clinique rhumatologique et hydrologie
	CALAS François	Anatomie
	CARRAZ Gilbert	Biologie animale et pharmacodynamie
	CAU Gabriel	Médecine légale et Toxicologie
	CAUQUIS Georges	Chimie organique
	CHABAUTY Claude	Mathématiques Pures
	CHARACHON Robert	Oto-Rhino-Laryngologie
	CHATEAU Robert	Thérapeutique
	CHENE Marcel	Chimie papetière
	COEUR André	Pharmacie chimique
	CONTAMIN Robert	Clinique gynécologique
	COUDERC Pierre	Anatomie Pathologique
	CRAYA Antoine	Mécanique
Mme	DEBELMAS Anne-Marie	Matière médicale
MM.	DEBELMAS Jacques	Géologie générale
	DEGRANGE Charles	Zoologie
	DESSAUX Georges	Physiologie animale
	DODU Jacques	Mécanique appliquée
	DREYFUS Bernard	Thermodynamique
	DUCROS Pierre	Cristallographie
	DUGOIS Pierre	Clinique de Dermatologie et Syphiligraphie
	FAU René	Clinique neuro-psychiatrique
	FELICI Noël	Electrostatique
	GAGNAIRE Didier	Chimie physique
	GALLISSOT François	Mathématiques Pures
	GALVANI Octave	Mathématiques Pures

MM.	GASTINEL Noël	Analyse numérique
	GERBER Robert	Mathématiques Pures
	GIRAUD Pierre	Géologie
	KLEIN Joseph	Mathématiques Pures
Mme	KOFLER Lucie	Botanique et Physiologie végétale
MM.	KOSZUL Jean-Louis	Mathématiques Pures
	KRAVTCHENKO Julien	Mécanique
	KUNTZMANN Jean	Mathématiques Appliquées
	LACAZE Albert	Thermodynamique
	LACHARME Jean	Biologie végétale
	LATREILLE René	Chirurgie générale
	LATURAZE Jean	Biochimie pharmaceutique
	LAURENT Pierre	Mathématiques Appliquées
	LEDRU Jean	Clinique médicale B
	LLIBOUTRY Louis	Géophysique
	LOUP Jean	Géographie
Mle	LUTZ Elisabeth	Mathématiques Pures
	MALGRANGE Bernard	Mathématiques Pures
	MALINAS Yves	Clinique obstétricale
	MARTIN-NOEL Pierre	Séméiologie médicale
	MASSEPORT Jean	Géographie
	MAZARE Yves	Clinique médicale A
	MICHEL Robert	Minéralogie et Pétrographie
	MOURIQUAND Claude	Histologie
	MOUSSA André	Chimie nucléaire
	NEEL Louis	Physique du Solide
	OZENDA Paul	Botanique
	PAUTHENET René	Electrotechnique
	PAYAN Jean-Jacques	Mathématiques Pures
	PEBAY-PEYROULA Jean-Claude	Physique
	PERRET René	Servomécanismes
	PILLET Emile	Physique industrielle
	RASSAT André	Chimie systématique
	RENARD Michel	Thermodynamique
	REULOS René	Physique industrielle
	RINALDI Renaud	Physique
	ROGET Jean	Clinique de pédiatrie et de puériculture
	SANTON Lucien	Mécanique
	SEIGNEURIN Raymond	Microbiologie et Hygiène
	SENGEL Philippe	Zoologie
	SILBERT Robert	Mécanique des fluides
	SOUTIF Michel	Physique générale
	TANCHE Maurice	Physiologie
	TRAYNARD Philippe	Chimie générale
	VAILLAND François	Zoologie
	VAUQUOIS Bernard	Calcul électronique
Mme	VERAIN Alice	Pharmacie galénique
M.	VERAIN André	Physique
Mme	VEYRET Germaine	Géographie
MM.	VEYRET Paul	Géographie
	VIGNAIS Pierre	Biochimie médicale
	YOCCOZ Jean	Physique nucléaire théorique

PROFESSEURS ASSOCIES

MM.	BULLEMER Bernhard	Physique
	RADHAKRISHNA Pidatala	Thermodynamique

PROFESSEURS SANS CHAIRE

MM.	AUBERT Guy	Physique
	BEAUDOING André	Pédiatrie
	BERTRANDIAS Jean-Paul	Mathématiques Appliquées
	BIARES Jean-Pierre	Mécanique
	BONNETAIN Lucien	Chimie minérale
Mme	BONNIER Jane	Chimie générale
MM.	CARLIER Georges	Biologie végétale
	COHEN Joseph	Electrotechnique
	COUMES André	Radioélectricité
	DEPASSEL Roger	Mécanique des Fluides
	DEPORTES Charles	Chimie minérale
	DESRE Pierre	Métallurgie
	DOLIQUE Jean-Michel	Physique des Plasmas
	GAUTHIER Yves	Sciences biologiques
	GEINDRE Michel	Electroradiologie
	GIDON Paul	Géologie et Minéralogie
	GLENAT René	Chimie organique
	HACQUES Gérard	Calcul numérique
	JANIN Bernard	Géographie
Mme	KAHANE Josette	Physique
MM.	MULLER Jean-Michel	Thérapeutique
	PERRIAUX Jean-Jacques	Géologie et minéralogie
	POULOUJADOFF Michel	Electrotechnique
	REBECQ Jacques	Biologie (CUS)
	REVOL Michel	Urologie
	REYMOND Jean-Charles	Chirurgie générale
	ROBERT André	Chimie papetière
	SARRAZIN Roger	Anatomie et chirurgie
	SARROT-REYNAULD Jean	Géologie
	SIBILLE Robert	Construction Mécanique
	SIROT Louis	Chirurgie générale
Mme	SOUTIF Jeanne	Physique générale
M.	VALENTIN Jacques	Physique nucléaire

MAITRES DE CONFERENCES ET MAITRES DE CONFERENCES AGREGES

Mle	AGNIUS-DELORD Claudine	Physique pharmaceutique
	ALARY Josette	Chimie analytique
MM.	AMBLARD Pierre	Dermatologie
	AMBROISE-THOMAS Pierre	Parasitologie
	ARMAND Yves	Chimie
	BEGUIN Claude	Chimie organique
	BELORIZKY Elie	Physique
	BENZAKEN Claude	Mathématiques Appliquées
Mme	BERTRANDIAS Françoise	Mathématiques Pures
MM.	BLIMAN Samuel	Electronique (EIE)
	BLOCH Daniel	Electrotechnique
Mme	BOUCHE Liane	Mathématiques (CUS)
MM.	BOUCHET Yves	Anatomie
	BOUSSARD Jean-Claude	Mathématiques Appliquées
	BOUVARD Maurice	Mécanique des Fluides
	BRIERE Georges	Physique expérimentale
	BRODEAU François	Mathématiques (IUT B)
	BRUGEL Lucien	Energétique
	BUISSON Roger	Physique
	BUTEL Jean	Orthopédie
	CHAMBAZ Edmond	Biochimie médicale
	CHAMPETIER Jean	Anatomie et organogénèse

MM.	CHIAVERINA Jean	Biologie appliquée (EFP)
	CHIBON Pierre	Biologie animale
	COHEN-ADDAD Jean-Pierre	Spectrométrie physique
	COLOMB Maurice	Biochimie médicale
	CONTE René	Physique
	CROUZET Guy	Radiologie
	DURAND Francis	Métallurgie
	DUSSAUD René	Mathématiques (CUS)
Mme	LTERRADOSSI Jacqueline	Physiologie
MM.	FAURE Jacques	Médecine légale
	GAVEND Michel	Pharmacologie
	GENSAC Pierre	Botanique
	GERMAIN Jean-Pierre	Mécanique
	GIDON Maurice	Géologie
	GRIFFITHS Michaël	Mathématiques Appliquées
	GROULADE Joseph	Biochimie médicale
	HOLLARD Daniel	Hématologie
	HUGONOT Robert	Hygiène et Médecine préventive
	IDELMAN Simon	Physiologie animale
	IVANES Marcel	Electricité
	JALBERT Pierre	Histologie
	JOLY Jean-René	Mathématiques Pures
	JOUBERT Jean-Claude	Physique du Solide
	JULLIEN Pierre	Mathématiques Pures
	KAHANE André	Physique générale
	KUHN Gérard	Physique
Mme	LAJZEROWICZ Jeannine	Physique
MM.	LAJZEROWICZ Joseph	Physique
	LANCIA Roland	Physique atomique
	LE JUNTER Noël	Electronique
	LEROY Philippe	Mathématiques
	LOISEAUX Jean-Marie	Physique Nucléaire
	LONGEQUEUE Jean-Pierre	Physique Nucléaire
	LUU DUC Cuong	Chimie Organique
	MACHE Régis	Physiologie végétale
	MAGNIN Robert	Hygiène et Médecine préventive
	MARECHAL Jean	Mécanique
	MARTIN-BOUYER Michel	Chimie (CUS)
	MAYNARD Roger	Physique du Solide
	MICOUD Max	Maladies infectieuses
	MOREAU René	Hydraulique (INP)
	NEGRE Robert	Mécanique
	PARAMELLE Bernard	Pneumologie
	PECCOUD François	Analyse (IUT B)
	PEFFEN René	Métallurgie
	PELMONT Jean	Physiologie animale
	PERRET Jean	Neurologie
	PERRIN Louis	Pathologie expérimentale
	PFISTER Jean-Claude	Physique du Solide
	PHELIP Xavier	Rhumatologie
Mle	PIERY Yvette	Biologie animale
MM.	RACHAIL Michel	Médecine interne
	RACINET Claude	Gynécologie et obstétrique
	RICHARD Lucien	Botanique
Mme	RINAUDO Marguerite	Chimie macromoléculaire
MM.	ROMIER Guy	Mathématiques (IUT B)
	ROUGEMONT (DE) Jacques	Neuro-Chirurgie
	STIEGLITZ Paul	Anesthésiologie

MM. STOEBNER Pierre	Anatomie pathologique
VAN CUTSEM Bernard	Mathématiques Appliquées
VEILLON Gérard	Mathématiques Appliquées (INP)
VIALON Pierre	Géologie
VOOG Robert	Médecine interne
VROUSSOS Constantin	Radiologie
ZADWORNY François	Electronique

MAITRES DE CONFERENCES ASSOCIES

MM. BOUDOURIS Georges	Radioélectricité
CHEEKE John	Thermodynamique
GOLDSCHMIDT Hubert	Mathématiques
YACOUD Mahmoud	Médecine légale

CHARGES DE FONCTIONS DE MATIRES DE CONFERENCES

Mme BERIEL Hélène	Physiologie
Mme RENAUDET Jacqueline	Microbiologie

Fait le 8 MARS 1972.

Cette thèse a été préparée au département d'informatique du Centre d'Etudes Nucléaires de Grenoble et au Laboratoire de Mathématiques Appliquées de l'Université de Grenoble.

Je tiens à exprimer ma profonde gratitude à Monsieur KUNTZMANN qui me fait l'honneur de présider le jury.

Que Monsieur LAURENT qui a dirigé ce travail et qui par son aide incessante, ses encouragements et ses conseils amicaux m'a permis de mener à bien cette entreprise, trouve ici l'expression de ma plus vive reconnaissance.

Monsieur GASTINEL m'a initié à l'analyse numérique ; je tiens à lui témoigner toute ma gratitude pour l'intérêt qu'il a porté à ce travail, pour ses critiques et ses suggestions.

Je remercie très vivement Messieurs AUSLENDER et TEMAM qui ont accepté de faire partie du jury. L'aide amicale et l'esprit de collaboration de Monsieur AUSLENDER m'ont été très précieux.

J'adresse mes sincères remerciements à Messieurs AMOUYAL et TERRIER qui m'ont donné la possibilité d'entreprendre et de mener à bien ce travail.

Je tiens enfin à remercier chaleureusement Mademoiselle PAYERNE dont j'ai pu apprécié l'efficacité et la compétence.

TABLE DES MATIERES

pages

- INTRODUCTION -

- CHAPITRE I : Méthodes de résolution de problèmes avec contraintes par approximation de l'ensemble admissible. Cas général.

(1.1.) Introduction..... I-1
(1.2.) Description de l'ensemble admissible par une famille d'inégalités..... I-3
(1.3.) Algorithme fondamental. Généralisation des méthodes de "coupe" (cut-off methods)..... I-5
(1.4.) Exemples de sous problèmes..... I-7
(1.5.) Convergence de l'algorithme (1.3.)..... I-11
(1.6.) Autres hypothèses permettant de justifier l'algorithme (1.3.)..... I-16
(1.7.) Des variantes de l'algorithme (1.3.) permettant de simplifier les fonctionnelles g^n I-18
(1.8.) Exemples de problèmes..... I-21
(1.9.) Cas de problèmes sur un espace produit..... I-27
Bibliographie..... I-33

- CHAPITRE II : Méthodes de résolution de problèmes d'optimisation avec contraintes par approximation extérieure de l'ensemble admissible.

(2.1.) Introduction..... II-1
Première partie - Cas non convexe..... II-5
(2.2.) Algorithme fondamental. Généralisation des méthodes de coupe et de pénalisation extérieure..... II-5
(2.3.) Autres hypothèses permettant de justifier l'algorithme (2.2.)..... II-12

(2.4.) Quelques variantes de l'algorithme (2.2.) permettant de simplifier la fonctionnelle φ_n II-16

(2.5.) Algorithmes pour lesquels le problème principal est simplifié..... II-21

Deuxième partie - Cas d'un problème convexe où la fonctionnelle convexe est uniformément convexe..... II-26

(2.6.) Algorithme..... II-26

(2.7.) Applications..... II-40

Bibliographie..... II-46

- CHAPITRE III : Méthodes de résolution de problèmes minimax - Applications à l'optimisation.

(3.1.) Introduction..... III-1

(3.2.) Généralités sur les problèmes minimax..... III-7

Première partie - Méthodes de résolution de problèmes minimax..... III-12

(3.3.) Méthodes de type accumulation..... III-13

(3.4.) Méthodes sans accumulation dans le cas où L(x,.) est strictement concave..... III-37

(3.5.) Méthodes sans accumulation dans un cas plus général..... III-45

(3.6.) Problèmes sur les espaces produits..... III-49

(3.7.) Procédés de montée obtenus par symétrie..... III-50

Deuxième partie - Quelques exemples d'applications aux problèmes d'optimisation..... III-51

(3.8.) Problèmes d'optimisation convexe..... III-51

(3.9.) Problèmes d'optimisation convexe avec contraintes inégalités..... III-55

(3.10) Problèmes d'optimisation non convexe avec contraintes..... III-59

Bibliographie..... III-61

- CHAPITRE IV : Minimisation d'une fonctionnelle sur un espace produit par une méthode de relaxation.

(4.1.) Introduction..... IV-1
(4.2.) Notations. Hypothèses. Généralités..... IV-3
(4.3.) Algorithmes de décomposition. Généralités..... IV-4
(4.4.) Méthodes de descente associées à la relaxation. Etude générale..... IV-5
(4.5.) Méthodes de descente associées à la relaxation. Exemples..... IV-7
(4.6.) Autres méthodes associées à la relaxation. Etude générale..... IV-14
(4.7.) Autres méthodes associées à la relaxation. Exemples..... IV-18
Bibliographie..... IV-25

- CHAPITRE V : Recherche d'un point fixe d'une application pseudo-contractante. Cas de l'application prox. V-1

(5.1.) Introduction..... V-1
(5.2. Généralités..... V-3
(5.3.) Recherche d'une solution d'une inéquation variationnelle par régularisation..... V-6
(5.4.) Cas d'un problème d'optimisation convexe. Recherche d'un point fixe de l'application prox. V-10
(5.5.) Cas d'une application composée de plusieurs applications pseudo-contractantes..... V-13
(5.6.) Recherche de points fixes d'une application pseudo-contractante par des itérations approchées. V-17
Bibliographie..... V-25

INTRODUCTION

On présente dans ce travail une étude mathématique d'algorithmes pour la résolution approchée de problèmes d'optimisation et de certains autres problèmes présentant des caractéristiques voisines : minimax, points fixes etc ... . Nous allons passer en revue les principaux sujets abordés dans chacun des chapitres. Mis à part le chapitre II qui suppose connus les paragraphes (1.2.), (1.3.), (1.4.) du chapitre I, les autres chapitres sont, à quelques points de détail près, indépendants.

Etant donnés un espace E , un sous ensemble de E , noté C, défini par des contraintes de type inégalité et une fonctionnelle f définie sur E , on considère le problème d'optimisation qui consiste à chercher l'élément de C pour lequel f atteint son minimum. Les méthodes d'accumulation des contraintes (méthodes de coupe ou en anglais cut-off methods, la méthode des plans sécants de KELLEY) sont bien connues pour la résolution approchée de tels problèmes d'optimisation avec contraintes. Elles consistent à remplacer le problème initial par une suite de problèmes P_n approchés, plus simples à mettre en oeuvre : trouver $x_n \in K_n$ pour lequel f atteint son minimum sur K_n . Les ensembles K_n sont définis par une partie des contraintes qui décrivent C : ils sont donc plus simples à mettre en oeuvre que C et peuvent être considérés comme des approximations extérieures de C . On passe de K_n à K_{n+1} en ajoutant une (ou plusieurs) contrainte non satisfaite en x^n .

Le chapitre I donne une théorie générale de ces procédés qui peut s'appliquer à des problèmes avec contraintes plus généraux que les problèmes d'optimisation , par exemple : points fixes d'application, optimisation vectorielle, point de selle d'une fonction convexe concave, point d'équilibre d'un jeu à n joueurs etc ... .

Le chapitre II présente une classe de procédés qui constitue une synthèse des méthodes d'accumulation des contraintes et de pénalisation extérieure en optimisation. Celle-ci conduit à plusieurs procédés originaux. On étudie d'une part le cas d'un problème d'optimisation non convexe, d'autre part le cas où C est défini par une intersection de convexes et où la fonctionnelle f est uniformément convexe. Sous ces dernières conditions on envisage des procédés remarquablement simples : on peut en particulier faire l'élimination des contraintes non actives.

Le chapitre III est consacré aux méthodes de résolution de problèmes minimax. On sait qu'un problème d'optimisation peut, sous des conditions très générales (ROCKAFELLAR), s'y ramener et il apparait très fructueux de donner une interprétation lagrangienne des méthodes d'optimisation. C'est dans ce but que nous présentons une classe d'algorithmes permettant de résoudre les problèmes de minimax dans des situations assez variées. Ils s'appuient sur des idées de base simples et analogues à celles des procédés des chapitres I et II : ce sont des procédés où à chaque itération on résoud un problème approchant le problème initial et on améliore cette approximation compte tenu du résultat trouvé. On étudie ensuite, très succintement quelques exemples d'application des méthodes précédentes à des problèmes d'optimisation. On décrit plusieurs algorithmes originaux, en particulier un algorithme de descente par sous gradient, pour minimiser une fonction non dérivable. On retrouve ainsi des méthodes connues telles que : méthodes de descente de type

gradient dans le cas sans contrainte, méthode de type FRANCK et WOLFE, méthodes "duales" (AUSLENDER, FALK), méthode des colonnes (DANTZIG), méthodes par approximation extérieure de l'ensemble admissible du chapitre II etc ... . Le formalisme introduit dans ce chapitre se révèle très prometteur tant du point de vue théorique que pratique : il recouvre une très large classe de problèmes et il apparait très constructif. Les résultats obtenus ne sont qu'une première étape et il reste de nombreuses questions ouvertes quant à l'élaboration d'une théorie plus structurée et quant à la mise en oeuvre et l'efficacité des méthodes nouvelles qu'elle nous suggère.

Le chapitre IV est relatif aux méthodes de décomposition pour minimiser une fonctionnelle convexe f sur un espace produit. Si la fonctionnelle f est convexe et quadratique, un procédé bien connu pour résoudre un tel problème est la méthode de GAUSS-SEIDEL qui se généralise aisément si f n'est pas quadratique. Dans cette méthode on doit à chaque itération résoudre des problèmes de minimisation partielle. Ceux-ci peuvent être couteux et on montre ici qu'il suffit d'effectuer sur ces problèmes intermédiaires une seule itération d'une méthode d'optimisation, telle que la méthode de descente, méthode de pénalisation extérieure etc ... . Les algorithmes originaux ainsi obtenus nous semblent très utiles dans les applications. D'autre part l'idée qui consiste dans un procédé de décomposition donné, à remplacer un problème de minimisation dont l'exécution demande en général un nombre infini d'opérations par un processus comprenant un nombre fini et connu d'opérations, nous semble susceptible d'être étendue à bien d'autres domaines.

Le chapitre V est consacré à la convergence d'algorithmes de recherche d'un point fixe d'un opérateur dans un espace de Hilbert. Le principal exemple qui est proposé est la détermination d'un point fixe de l'application prox liée à une fonctionnelle convexe (notion introduite par J.J. MOREAU). On en déduit un procédé nouveau de régularisation pour un problème d'optimisation convexe et on étend ces résultats aux inéquations variationnelles.

NOTATIONS ET DEFINITIONS

Nous donnons brièvement quelques notations et définitions qui seront utilisées dans la suite et qui d'ailleurs seront le plus souvent rappelées dans le texte. On pourra consulter sur ces questions les livres de ROCKAFELLAR,(Convex Analysis (1970)) et LAURENT (Approximation et Optimisation, Hermann (1972)).

Soit E un espace de Banach réel et C un sous ensemble de E . On désigne par $\|\ \|$ la norme dans E . Si E' est le dual topologique de E on désigne par $< x,y >$ (ou par (x,y) dans les chapitres III et IV) la valeur de la fonctionnelle $y \in E'$ au point $x \in E$.

L'ensemble C est appelé convexe si pour tout $x \in C$, tout $y \in C$ et tout $\lambda \in [0,1]$ on a : $\lambda x + (1-\lambda)\, y \in C$. On note conv (C) l'enveloppe convexe de C c'est-à-dire l'intersection de tous les ensembles convexes contenant C . On note cône (C) l'enveloppe conique de C c'est-à-dire l'intersection de tous les cônes de sommet l'origine qui contiennent C.

On appelle fonction indicatrice de C la fonction χ_C : $E \to [0,+\infty]$ telle que :

$$\chi_C(x) = \begin{cases} 0 & \text{si} \quad x \in C \\ +\infty & \text{sinon.} \end{cases}$$

On appelle fonction (ou fonctionnelle) définie sur C une application de C dans $\mathbb{R}$. Dans certains cas qui sont précisés on généralise cette définition en considérant $f : E \to \bar{\mathbb{R}}$ (où $\bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$) ; c'est ce que nous ferons dans la suite de ce paragraphe.

On appelle épigraphe de f noté : epi f l'ensemble défini par :

$$\text{epi } f = \{(x,\mu) \mid \mu \in \mathbb{R} \ , \ x \in E, \mu \geq f(x)\} \quad .$$

Nous dirons que f est convexe si son épigraphe est un sous ensemble convexe de $E \times \mathbb{R}$.

Nous considérerons en général des fonctions convexes propres c'est-à-dire que pour tout $x \in E$ on a $f(x) > -\infty$ et il existe $z \in E$ tel que $f(z) < +\infty$.

Si on considère $f : C \to \mathbb{R}$ où C est convexe on peut définir la convexité de f par la propriété classique suivante :

$$\forall x \in C \quad \forall y \in C \quad \forall \lambda \in [0,1] \quad \text{on a :}$$

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) .$$

On dit que f est concave si $-f$ est convexe.

Etant donné une fonction $g : E \to \bar{\mathbb{R}}$ on appelle enveloppe convexe de g notée : conv g la plus grande fonction convexe majorée par g .

Nous dirons qu'une fonction f est semi-continue inférieurement (s.c.i.) si pour tout $\lambda \in R$ l'ensemble $\{x \in E \mid f(x) \leq \lambda\}$ est fermé.

On dit que $y \in E'$ est un sous grandient de f au point $x_o \in E$ si $f(x_o)$ est fini et si on a :

$$f(x) \geq f(x_o) + \langle y_o, x-x_o \rangle \qquad \forall x \in E .$$

L'ensemble des sous gradients de f en x_o sera appelé sous différentiel de f en x_o et sera noté : $\partial f(x_o)$.

La dérivée directionnelle de f en x_o pour la direction $y \in E$ est définie par :

$$f'(x_o;y) = \lim_{\substack{\lambda \to 0 \\ \lambda > 0}} \frac{f(x_o+\lambda y) - f(x_o)}{\lambda}$$

ERRATA

Page 6 — 6ème ligne à partir du bas :

Au lieu de : " On dit que $y \in E'$ est ... "

lire : " On dit que $y_o \in E'$ est ... "

Page II-13 — 7ème ligne à partir du haut :

Au lieu de : " ... est continue en 0^+ "

lire : " ... est continue en 0 ".

Page II-15 — 4ème ligne à partir du haut :

Au lieu de : " $\lim_{\substack{n\to\infty \\ n\in S}} \varepsilon_n = 0^+$ "

lire : " $\lim_{\substack{n\to\infty \\ n\in S}} \varepsilon_n = 0$ " .

Page II-21 — 2ème ligne à partir du bas :

Au lieu de : " Il existe $D \in \mathbb{R}$ tel que ... "

lire : " Il existe $D \geq 0$ tel que ... "

Page II-23 — 10ème ligne à partir du bas :

Au lieu de : ... (où $d(.,c)$ désigne la distance à C)

lire : ... (où $d(.,C)$ désigne la distance à C)

Page III-7 — 6ème ligne et 3ème ligne à partir du bas :

Au lieu de : " ... le triplet (X,Y,L) a une valeur "

lire : " ... le triplet (X,Y,L) a une valeur finie ".

Page III-16 — 6ème ligne à partir du bas :

Au lieu de : " ... une fonctin $L_n : X \times Y \to \mathbb{R}$ continue "

lire : " ... une fonction $L_n : X_n \times Y \to \mathbb{R}$ continue ".

Page III-17 $2^{ème}$ ligne à partir du haut :

Au lieu de : " ... une fonction $L_{n+1} : X \times Y \to \mathbb{R}$ continue "

lire : " ... une fonction $L_{n+1} : X_{n+1} \times Y \to \mathbb{R}$ continue ".

$4^{ème}$ ligne à partir du haut :

Au lieu de :

" $L_n(x,y) \geq L_{n+1}(x,y) \geq L(x,y) \quad \forall\, x \in X \quad \forall\, y \in Y$ "

lire :

" $L_n(x,y) \geq L_{n+1}(x,y) \geq L(x,y) \quad \forall\, x \in X_n \quad \forall\, y \in Y$

$L_{n+1}(x,y) \geq L(x,y) \quad \forall\, x \in X_{n+1} \quad \forall\, y \in Y$ ".

Page III-20 $8^{ème}$ ligne à partir du haut :

Au lieu de : " $\lim_{n\to\infty} (C^n, \overline{x^n}, x^n) = 0$ ... "

lire : " $\lim_{n\to\infty} (C^n, \overline{x^n} - x^n) = 0$ " .

$1^{ère}$ ligne à partir du bas :

Au lieu de : (3) $\lim_{n\to\infty} (C^n, \overline{x^n}, x^n) = 0$

lire : (3) $\lim_{n\to\infty} (C^n, \overline{x^n} - x^n) = 0$.

Page III-51 $10^{ème}$ ligne à partir du bas :

Au lieu de : (3.8.) PROBLEME D'OPTIMISATION CONVEXE (*)

lire : (3.8.) PROBLEMES D'OPTIMISATION CONVEXE (*) .

Page V-5 $8^{ème}$ ligne à partir du bas :

Au lieu de : " ... ait deux point d'adhérence distincts"

lire : " ... ait deux points d'adhérence faible distincts ".

Page V-10 $10^{ème}$ ligne à partir du haut :

Au lieu de : " ($\mathcal{P}$) "trouver $\bar{x} \in C$ tel que ... "

lire : " ($\mathcal{P}$) "trouver $\bar{x} \in H$ tel que ... " .

CHAPITRE 1

METHODES DE RESOLUTION DE PROBLEMES AVEC CONTRAINTES PAR APPROXIMATION EXTERIEURE DE L'ENSEMBLE ADMISSIBLE. CAS GENERAL

(1.1.) INTRODUCTION.

Soient :

- un espace de Banach E et $\mathcal{P}(E)$ l'ensemble des parties de E ,
- un sous ensemble de E noté C (ensemble admissible) ,
- une famille $\mathcal{F}$ de parties de E telles que :

$$C \in \mathcal{F} \quad ; \quad A \supseteq C \qquad \forall A \in \mathcal{F} \ ,$$

- une application Γ de $\mathcal{F}$ dans $\mathcal{P}(E)$ telle qu'on ait pour tout $A \in \mathcal{F}$:

$$\Gamma A \subseteq A \quad ; \quad \Gamma A \neq \emptyset \quad \text{et} \quad \Gamma A \cap C \subseteq \Gamma C \ . \qquad (*)$$

Nous nous proposons dans ce chapitre d'étudier des procédés de résolution approchés pour le problème P suivant :

(P) "trouver $\bar{x} \in \Gamma C$" .

Parmi ces problèmes, on peut signaler tout particulièrement la recherche d'un élément de C pour lequel le minimum de la fonctionnelle f sur C est atteint. Ce problème est bien connu et de nombreux travaux y ont été consacrés. On définit alors pour toute partie A de $\mathcal{F}$ l'application Γ par :

$$\Gamma A = \{x \in A \,|\, f(x) = \min (f(y) \,|\, y \in A)\} \ .$$

(*) Dans tout ce chapitre l'image de A par Γ sera notée ΓA au lieu de la notation usuelle $\Gamma(A)$.

L'idée générale des méthodes étudiées dans ce chapitre pour résoudre P est de remplacer ce problème par une suite de problèmes P_n plus simples.

De façon plus précise, on résoud à la $n^{ème}$ itération le problème principal qui consiste à trouver un élément x^n de ΓK_n. Ici l'ensemble K_n appartient à $\mathcal{F}$ et il est défini "plus simplement" que C. Si l'élément x^n appartient à C, le processus est arrêté car compte tenu des propriétés de Γ, l'élément x^n est une solution du problème P. Dans le cas contraire, on effectue le sous problème qui consiste à générer, connaissant l'élément x^n, un nouvel ensemble K_{n+1} contenant C qui "approche mieux" l'ensemble C que K_n et on passe à l'itération suivante.

Nous opérons donc par approximation extérieure* de l'ensemble admissible. On peut remarquer que la suite des ensembles K_n n'est pas déterminée à l'avance (comme c'est le cas pour la méthode de RITZ-GALERKIN par exemple), mais le nouvel ensemble K_{n+1} est construit compte tenu de l'élément x^n. Ceci est intéressant car on a un procédé adaptatif, l'ensemble K_n n'approche C que "là où c'est nécessaire". Par construction les approximations successives x^n n'appartiennent pas en général à C (si $x^n \in C$, le processus est arrêté car x^n est solutinn du problème P).

Ces méthodes par approximation extérieure de l'ensemble admissible sont bien connues pour les problèmes d'optimisation avec contraintes : on les appelle méthodes d'accumulation des contraintes, méthodes de relaxation, méthodes du plan sécant (cutting plane method), méthodesde coupe (cut-off methods) etc... .

La littérature sur ce sujet est très abondante. On trouvera dans l'introduction du chapitre II une bibliographie du sujet, nous contentant simplement de signaler ici les articles de CHENEY-GOLSTEIN [5] et de KELLEY [9] pour leur importance historique. En fait, nous n'aborderons ici que très sommairement les problèmes d'optimisation; étant donné leur importance et leurs propriétés particulières, ils seront étudiés en détail au chapitre II.

---

* Les ensembles K_n contiennent C et peuvent être considérés comme des approximations de C.

Notre but est d'étendre les méthodes par approximation extérieure dans les deux directions suivantes :

- donner une classe générale de problèmes qui peuvent être résolus par de tels procédés (cf. ZANGWILL [13]) ;
- étudier des variantes de l'algorithme général intéressantes en pratique .

Nous considérons un formalisme qui peut s'appliquer à de très nombreux problèmes P : problèmes de points fixes d'applications continues ou d'applications faiblement contractantes, optimisation vectorielle, point de selle d'une fonction concave-convexe, point d'équilibre d'un jeu à n joueurs, équilibre économique, etc... .

Quant aux algorithmes, on considère des variantes dans le choix des sous problèmes. Le but de ces considérations étant de proposer un procédé efficace mais le plus "économique" possible.

Dans toute la suite, l'ensemble C sera défini par des contraintes inégalités ; c'est la façon usuelle de définir des ensembles pour de nombreux problèmes pratiques et elle est très générale. Cette description de C permet d'introduire d'une façon naturelle la famille $\mathcal{F}$ de parties considérée plus haut : l'ensemble C est considéré comme l'intersection des ensembles de la famille $\mathcal{F}$.

(1.2.) DESCRIPTION DE L'ENSEMBLE ADMISSIBLE PAR UNE FAMILLE D'INEGALITES.

Nous supposerons dans la suite que l'ensemble admissible C est défini par :

$$(1) \qquad C = \{x \in X \mid g(v,x) \leq 0 \qquad \forall\, v \in V\}$$

avec les hypothèses suivantes :

H1 : X est une partie compacte de $\mathbb{R}^q$.
L'ensemble C est non vide. La fonctionnelle g définie sur $V \times X$ est telle que la famille $g(v,.)$ soit équi-uniformément continue sur X c'est-à-dire que :

"pour tout $\varepsilon > 0$, il existe un voisinage $\mathcal{U}$ de 0 dans $\mathbb{R}^p$ tel que pour tout $x,y \in X$: $(x-y) \in \mathcal{U}$ et tout $v \in V$ on ait :

$|g(v,x)-g(v,y)| < \varepsilon$."

Nous ne faisons pas d'hypothèse explicite sur l'ensemble V. On remarque d'abord qu'à toute représentation de C, telle que (1), on peut faire correspondre une infinité d'autres représentation du même type. Cette représentation d'un ensemble est très fréquente dans les applications : ou bien elle est donnée à priori ou on peut s'y ramener facilement comme nous allons le voir sur quelques exemples.

(1.2.1.) Représentation de KELLEY.

Supposons que l'ensemble C soit défini par :

$$C = \{x \in X | h(x) \leq 0\}$$

où h est une fonctionnelle définie sur $\mathbb{R}^q$, convexe, non identiquement égale à $+\infty$ et qui est continue sur un ouvert contenant X. Une propriété classique des fonctionnelles convexes permet d'affirmer que (cf. LAURENT [10] ou ROCKAFELLAR [12]) :

a) Pour tout $s \in X$, l'ensemble $\partial h(s)$ (sous différentiel de h en s) est une partie de $\mathbb{R}^q$ convexe, non vide et compacte.

b) Pour tout $x,s \in X$ et tout $p \in \partial h(s)$ on a :
$h(x) \geq h(s) + < p,x-s >$.

Et pour tout $x \in X$, on a :

$$h(x) = \max (h(s) + < p,x-s > | s \in X, p \in \partial h(s))$$

On peut donc écrire que :

(2) $$C = \{x \in X | h(s) + < p,x-s > \leq 0 \ \forall s \in X \ ; \ \forall p \in \partial h(s)\} .$$

Cette forme est analogue à (1).
En [9], KELLEY utilise cette remarque dans le cas où h est une fonction dérivable.

(1.2.2.) Représentation d'un ensemble convexe et fermé.

Soit C un ensemble convexe fermé de $\mathbb{R}^q$.

On a évidemment :

$$(3) \qquad C = \{x \in \mathbb{R}^q \mid \langle x,v \rangle - \sup_{z \in C} \langle z,v \rangle \leq 0 \quad \forall\, v \in V\}$$

où V est un ensemble absorbant de $\mathbb{R}^q$ c'est-à-dire tel qu'à tout élément $y \in \mathbb{R}^q$ il correspond $\alpha > 0$ tel que $\alpha.y \in V$.

En (3), on décrit l'ensemble C comme intersection des demi espaces fermés qui sont d'appui à C .

(1.3.) Algorithme fondamental. Généralisation des "méthodes de coupe" (cut off methods).

(1.3.1.) Notations et hypothèses.

Considérons l'ensemble C défini en (1.2.) par :

$$C = \{x \in X \mid g(v,x) \quad 0 \quad \forall\, v \in V\}$$

et qui satisfait aux hypothèses H1.

On attache à cette description de l'ensemble C la famille $\mathcal{F}$ des parties de X suivante :

A toute fonctionnelle $h : V \times X \to [-\infty,+\infty[$ telle que $h(v,x) \leq g(v,x)$ $\forall\, v \in V$ $\forall\, x \in X$ et telle que $h(v,.)$ soit semi-continue inférieurement $\forall\, v \in V$, on fait correspondre l'ensemble A de $\mathcal{F}$ défini par :

$$A = \{x \in X \mid h(v,x) \leq 0 \quad \forall\, v \in V\}.$$

H2. Soit une application Γ qui à tout compact A de X fait correspondre le compact ΓA tel que $\Gamma A \subseteq A$ et $\Gamma A \neq \emptyset$ si A est non vide. Nous supposons de plus la propriété topologique suivante :

Pour toute suite $\{K_n\}$ d'éléments de $\mathcal{F}$ on a :

$$\limsup_{n\to\infty} (\Gamma K_n) \cap C \subseteq \Gamma C .$$

(Etant donnée une suite $\{Q_n\}$ de parties de X, on note :

$\limsup_{n\to\infty} Q_n = \{x \in X | \lim_{k\to\infty} x^{n_k} = x$ pour une suite $x^{n_k} \in Q_{n_k}$ et

$\{Q_{n_k}\}$ est une sous suite de la suite $\{Q_n\}$ } .

Le problème P que nous nous proposons de résoudre de façon approchée est le suivant.

(P) "trouver un élément $\bar{x}$ de l'ensemble ΓC".

(1.3.2.) Description de l'algorithme.

Il s'agit en fait d'une méthode itérative qui comprend à chaque itération n :

un problème principal qui consiste à chercher un élément de ΓK_n (K_n est un élément de $\mathcal{F}$) et un sous problème où compte tenu de la solution x^n trouvée au problème principal on détermine l'ensemble K_{n+1}, nouvelle approximation extérieure de C.

Précisons ces étapes . Au début de la $n^{ème}$ itération, on connait une "approximation inférieure" g^n de g telle que :

$$(1) \quad \begin{cases} g^n : V \times X \to [-\infty, +\infty[\\ g^n(v,.) \text{ est semi-continue inférieurement pour tout } v \in V \\ g^n(v,x) \leq g(v,x) \text{ pour tout } x \in X \text{ et tout } v \in V . \end{cases}$$

Il y correspond une approximation extérieure de C notée K_n :

$$K_n = \{x \in X | g^n(v,x) \leq 0 \quad \forall\, v \in V\} .$$

Le problème principal consiste à déterminer un élément x^n de ΓK_n . Si cet élément appartient à C , alors grace à l'hypothèse H2 il appartient aussi à ΓC et c'est une solution du problème P ; on est donc conduit à arrêter l'algorithme. Dans le cas contraire, on passe à l'exécution du sous-problème : chercher un élément v^n de V qui correspond à une contrainte "mal satisfaite" en x^n (voir (1.4.2.)) puis déterminer la fonctionnelle g^{n+1} possédant les propriétés (1) et telle que :

$$\begin{cases} g^{n+1}(v^n,x) = g(v^n,x) \quad \text{pour tout} \quad x \in X \\ g(v,x) \geq g^{n+1}(v,x) \geq g^n(v,x) \quad \forall\, x \in X \quad \text{et} \quad \forall\, v \in V \ . \end{cases}$$

(1.4.) EXEMPLES DE SOUS PROBLEMES.

Plusieurs sous problèmes permettent de justifier l'algorithme décrit ci-dessus. Nous nous contentons d'en citer quelques uns qui sont utilisés fréquemment dans les applications.

(1.4.1.) Choix de la fonctionnelle g^n .

Un choix pratique très courant consiste à prendre :

$$g^n(v,x) = \begin{cases} g(v,x) \quad \text{pour tout} \quad v \in V_n \quad \text{et tout} \quad x \in X \\ -\infty \qquad \text{si} \quad v \notin V_n \ . \end{cases}$$

Le sous ensemble V_n de V définit ainsi la fonctionnelle g^n . Connaissant V_n et v^n , on choisira V_{n+1} tel que :

$$V \supseteq V_{n+1} \supseteq V_n \cup \{v^n\} \ .$$

Exemple 1 :

Les ensembles V_n sont finis et on passe de V_n à V_{n+1} par :

$$V_{n+1} = V_n \cup \{v^n\} \ .$$

Exemple 2 :

Si V est une partie convexe d'un espace vectoriel on prend :

$$V_{n+1} = \text{conv}\ (V_n \cup \{v^n\}\ ). \quad (*)$$

Exemple 3 :

Si V est un cône d'un espace vectoriel :

$$V_{n+1} = \text{cône}\ (V_n \cup \{v^n\}). \quad (*)$$

(1.4.2.) Choix de l'élément v^n .

Il s'agit en principe d'une problème d'optimisation. En voici trois exemples :

(1.4.2.1.) L'ensemble V est fini.

C'est le cas le plus simple. Il suffit de prendre $v^n \in V$ tel que :

$$g(v^n, x^n) > 0 .$$

(i.e. trouver une contrainte qui n'est pas satisfaite en x^n) .

(1.4.2.2.) Cas général.

On choisit $v^n \in V$ tel que :

$$(1) \quad \begin{cases} g(v^n, x^n) > 0 \\ g(v^n, x^n) \geq \theta \ \sup\ (g(v, x^n) \mid v \in V) - \rho_n \end{cases}$$

où sont donnés le nombre $\theta \in]0,1]$ et la suite de réels $\{\rho_n\}$ telle que $\rho_n \geq 0$ et $\lim_{n\to\infty} \rho_n = 0$. Le sup considéré doit être fini. On peut vérifier qu'il en est bien ainsi pour les hypothèses du (1.3.1.)

(*) Etant donné un ensemble A on note conv (A) l'enveloppe convexe de A et cône (A) l'enveloppe conique de A .

Nous allons expliciter trois cas particuliers de cette méthode où ce choix a une signification simple.

a) Une généralisation de la méthode de KELLEY.

Supposons que l'ensemble C est défini comme en (1.2.1.) par :

$$C = \{x \in X \mid h(x) \leq 0\}$$

où la fonctionnelle h est convexe, définie et continue sur $\mathbb{R}^q$.
On sait que (voir (1.2.1.) :

$$(2) \qquad C = \{x \in X \mid h(s) + < p, x-s > \leq 0 \quad \forall\, s \in X \quad \forall\, p \in \partial h(s)\} .$$

On montre (cf. ROCKAFELLAR [12] Théorème 24.7) qu'il existe $d > 0$ tel que : $\|p\| \leq d \quad \forall\, s \in X \quad \forall\, p \in \partial g(s)$.

Si on pose : $g(v,x) = h(s) + < p, x-s >$ où $v = (s,p)$, on a pour tout v et tout $x,y \in X$:

$|g(v,x) - g(v,y)| = |<p,x-y>| \leq \|p\|.\|x-y\| \leq d\|x-y\|$ ce qui montre que l'hypothèse H1 est satisfaite.

Nous allons voir maintenant un sous problème particulièrement simple attaché à la définition (2) de C . Etant donné x^n on choisit $s^n = x^n$ et p^n un élément quelconque de $\partial h(x^n)$. Ce choix correspond à celui de (1) avec $\theta = 1$ et $\rho_n = 0$ car, puisque h est convexe, on sait que :

$$h(x^n) \geq h(s) + < p, x^n - s > \quad \text{pour tout} \quad s \in X \quad \text{et tout} \quad p \in \partial h(s).$$

En [9] KELLEY a utilisé cette technique pour résoudre un problème d'optimisation convexe dans le cas où h est dérivable.

b) Une méthode de projection.

Soit C un ensemble convexe, fermé, borné de $\mathbb{R}^q$. Comme en (1.2.2.), on peut le représenter par :

$$C = \{x \in \mathbb{R}^q \mid < x,v > - \sup_{z \in C} < z,v > \leq 0 \quad \forall\, v : \|v\| = 1\} .$$

On vérifie facilement l'hypothèse H1 : pour tout $v \in V$ et $x,y \in X$ on a :

$$|g(v,x)-g(v,y)| = |< x-y,v >| \leq ||v|| \cdot ||x-y|| \leq ||x-y|| .$$

Etant donné le point x^n , le sous problème peut consister à chercher l'élément v^n pour lequel l'optimum suivant est atteint :

$$\max (<x^n,v> - \sup_{z \in C} < z,v > | ||v|| = 1).$$

On a : $$v^n = \frac{x^n - P_c(x^n)}{||x^n - Pc(x^n)||} .$$

(Où $P_c(x^n)$ est la projection de x^n sur C).

c) Méthode de KAPLAN généralisée.

On considère le même ensemble C que ci-dessus.
On suppose de plus que l'intérieur de C n'est pas vide et qu'on connait un élément x^* de cet intérieur. On considère le sous-problème suivant : Etant donné x^n qui n'appartient pas à C , on détermine l'élément y^n à l'intersection du segment $[x^*,x^n]$ avec la frontière de C . On choisit ensuite v^n ($||v^n|| = 1$) dans le cône normal à C en y^n .*

On peut vérifier facilement que ce choix de v^n satisfait à

(1) avec $\rho_n = 0$ et $\theta = \inf \left(\frac{\pi}{||x^*-y||} \middle| y \in \text{Front } (C) \right) > 0$

où r est un nombre strictement positif fixé tel que la boule ouverte de centre x^* et de rayon r soit contenue dans l'intérieur de C .

Ce procédé est utilisé par KAPLAN en [8] pour la résolution approchée d'un problème d'optimisation convexe dans R^n où la fonctionnelle à minimiser est linéaire.

---

* On appelle cône normal à C en $z \in C$ l'ensemble :
$\{y \in \mathbb{R}^q | < x-z,y > \leq 0 \quad \forall x \in C\}$.

(1.4.2.3.) Cas où V est convexe et séquentiellement compact.

Nous ferons dans ce paragraphe les hypothèses supplémentaires suivantes :

L'ensemble V est convexe et séquentiellement compact.
Pour tout $x \in X$ la fonctionnelle $g(.,x)$ est concave et possède une dérivée de Fréchet : $\nabla_v g(.,x)$ bornée sur tout borné de $V \times X$.

On choisit la fonctionnelle g^n comme en (1.4.1.) :

$$g^n(v,x) = \begin{cases} g(v,x) & \text{si} \quad v \in V_n \quad \text{et} \quad x \in X \\ -\infty & \text{sinon,} \end{cases}$$

où V_n est une partie convexe et fermée de V .

Le sous problème de l'itération n est le suivant :
On détermine $\tilde{v}^n \in V_n$ tel que :

(i) $\max (g(v,x^n) \mid v \in V_n) = g(\tilde{v}^n, x^n)$.

On pose $C^n = \nabla_v g(\tilde{v}^n, x^n)$. On cherche $v^n \in V$ tel que :

(ii) $\max (< C^n, v > \mid v \in V) = < C^n, v^n >$.

On prend ensuite $V_{n+1} = \text{conv} (V_n \cup \{v^n\})$. On est donc conduit à résoudre deux problèmes d'optimisation qui sont relativement simples. On pourrait bien entendu se contenter d'en chercher des solutions approchées.

(1.5.) CONVERGENCE DE L'ALGORITHME (1.3.).

Nous allons démontrer un théorème qui montre que l'algorithme (1.3.) permet en un certain sens de trouver des solutions approchées du problème P .

(1.5.1.) Théorème.

Soit une suite $\{x^n\}$ générée par un procédé du (1.3.).
Les hypothèses du (1.3.1.) entrainent que toute valeur d'adhérence de cette suite est solution du problème P .

Démonstration :

Nous avons à considérer les trois sous problèmes étudiés en (1.4.2.). Nous éliminons le cas trivial où la suite $\{x^n\}$ est finie.

a) Cas du sous problème (1.4.2.1.).

Evident.

b) Cas du sous problème (1.4.2.2.).

Par construction on sait que pour tout n et tout $m > n$ on a :

$$x^m \in \Gamma K_m \subseteq K_m \subseteq K_{n+1} .$$

Donc :

$$g^{n+1}(v,x^m) \leq 0 \qquad \text{pour tout } v \in V .$$

D'où on tire par construction de la fonctionnelle g^{n+1} que :

(i) $\quad g(v^n,x^m) \leq 0 .$

Puisque la suite $\{x^n\}$ appartient à l'ensemble compact X , elle possède une sous suite infinie $\{x^n\}_{n\in S}$ qui converge vers un élément $\tilde{x}$ de X donc qui est une suite de Cauchy.

Par conséquent on déduit des hypothèses H1 que :

(ii) $$\lim_{\substack{n\to\infty \\ m>n \\ m,n\in S}} |g(v^n,x^n) - g(v^n,x^m)| = 0$$

Grace à (i) et du fait que : $g^n(v^n,x^n) > 0 \ \forall\, n$, on tire que :

$$|g(v^n,x^n) - g(v^n,x^m)| \geq g(v^n,x^n) > 0 .$$

D'où on déduit grace à (ii) que :

$$\lim_{\substack{n\to\infty \\ n\in S}} g(v^n,x^n) = 0 .$$

Du choix de v^n considéré il vient :

$$\limsup_{\substack{n\to\infty \\ n\in S}} g(v,x^n) \leq 0 \qquad \text{pour tout } v \in V .$$

Ceci prouve (semi continuité inférieure de $g(v,.)$) que :

$$g(v,\tilde{x}) \leq 0 \qquad \text{pour tout } v \in V .$$

Donc, puisque $\tilde{x} \in X$ alors $\tilde{x} \in C$.

On termine la démonstration en utilisant l'hypothèse H2 qui permet d'affirmer que dans ces conditions le point $\tilde{x}$ appartient à ΓC .

c) Cas du sous problème (1.4.2.3.).

On fait dans ce cas les hypothèses supplémentaires du (1.4.2.3.). Soit des indices n,m : $m > n$. Par définition de v^m on a :

$$<c^m,v^m> \leq < c^m,v > \qquad \text{pour tout } v \in V .$$

Puisque $\tilde{v}^m \in V_m \subseteq V$, il vient :

$$(1) \qquad < c^m,v^m-\tilde{v}^m > \geq 0 .$$

Par définition de $\tilde{v}^m$, on tire :

$$g(\tilde{v}^m,x^m) \geq g(v,x^m) \qquad \text{pour tout } v \in V_m .$$

Et, puisque V_m est un ensemble convexe et que la fonctionnelle $g(.,x^m)$ est concave, il vient :

$$<c^m,v-\tilde{v}^m> \leq 0 \qquad \text{pour tout } v \in V_m .$$

Or $v^n \in V_m$ (par construction), donc :

(2) $\quad < c^m, v^n - \hat{v}^m > \leq 0$.

Puisque V est séquentiellement compact, il existe une sous suite infinie $\{v^n\}_{n \in S}$ convergente, donc de Cauchy. Par conséquent, puisque les éléments c^m sont bornés :

$$\lim_{\substack{n \to \infty \\ m > n ; m,n \in S}} < c^m, v^n - v^m > = 0 .$$

D'où on tire grâce à (1) et (2) que :

$$\lim_{\substack{m \to \infty \\ m \in S}} < c^m, v^m - \hat{v}^m > = 0 .$$

On montre (par un raisonnement par l'absurde) que ceci est encore vrai pour toute la suite :

(3) $$\lim_{n \to \infty} < c^n, v^n - \hat{v}^n > = 0 .$$

Puisque la fonctionnelle $g(.,x)$ est concave on peut écrire :

$$g(v,x^n) \leq g(\hat{v}^n, x^n) + < c^n, v - \hat{v}^n > \quad \text{pour tout } v \in V$$

donc :

$$g(v,x^n) \leq g(\hat{v}^n, x^n) + < c^n, v^n - \hat{v}^n > \quad \text{pour tout } v \in V .$$

Par construction on a : $g(\hat{v}^n, x^n) \leq 0$. D'où on tire que, grâce à (3) :

$$\limsup_{n \to \infty} g(v,x^n) \leq 0 \quad \text{pour tout } v \in V .$$

On termine la démonstration du théorème comme en b).

C.Q.F.D.

(1.5.2.) Remarques sur l'hypothèse de compacité de X .

Nous remarquons que l'hypothèse de compacité de X (voir H1) est assez limitative pour de nombreux cas pratiques. En fait cette hypothèse peut être écartée dans les cas suivants :

a) au début de l'algorithme, on connait une fonctionnelle g^o telle que l'ensemble :

$$K_o = \{x \in X \mid g^o(v,x) \leq 0 \quad \forall\, v \in V\} \quad \text{soit compact.}$$

Ceci implique bien entendu que l'ensemble C est compact.

b) On peut enfermer à priori dans un ensemble borné connu la solution cherchée.

(1.5.3.) Cas où l'ensemble C est vide (X non vide).

Il est facile de montrer par un raisonnement par l'absurde simple qu'il ne peut exister une suite infinie d'itérations n pour lesquelles l'ensemble K_n serait non vide.

(1.5.4.)

On pourrait étendre aisément les résultats précédents à des espaces plus généraux. On remplacerait H1 par les hypothèses suivantes :

{ X est une partie séquentiellement compacte d'un espace topologique séparé E . L'ensemble C est non vide et la famille $g(v,.)$ est équi-uniformément continue sur X .

En fait cette généralisation ne semble présenter qu'un intérêt pratique relativemnt limité. Signalons cependant qu'en [1] 10.2, AUSLENDER donne un tel exemple dans le cas d'espaces de Banach réflexifs.

(1.6.) AUTRES HYPOTHESES PERMETTANT DE JUSTIFIER L'ALGORITHME (1.3.)

(1.6.1.) Hypothèses.

Les hypothèses que nous introduisons dans ce paragraphe font intervenir la compacité séquentielle de V et la continuité des fonctionnelles $g(.,x)$. Elles permettent donc plus aisément d'aborder le cas où X est une partie d'un espace de dimension infinie.

H3 : L'ensemble V est une partie séquentiellement compacte de l'espace vectoriel topologique séparé F. La famille de fonctionnelles $g(.,x)$ est équi-uniformément continue c'est-à-dire : pour tout $\varepsilon > 0$, il existe un voisinage $\mathcal{U}$ de 0 dans F tel que pour tout $u,v \in V$: $(u-v) \in \mathcal{U}$ et pour tout $x \in X$ on ait :

$$|g(u,x) - g(v,x)| < \varepsilon.$$

L'ensemble X est une partie convexe, fermée, bornée et non vide d'un espace de Banach réflexif E.

Pour tout $v \in V$ la fonctionnelle $g(v,.)$ est semi-continue inférieurement et convexe.

L'ensemble C défini par :

$$C = \{x \in X | g(v,x) \leq 0 \quad \forall\, v \in V\}$$

est non vide.

H4 : Cette hypothèse concerne l'application Γ.
A tout ensemble A convexe fermé borné non vide de E elle fait correspondre l'ensemble non vide ΓA tel que : $\Gamma A \subseteq A$.
Pour toute suite d'ensembles $\{K_n\}$ convexes fermés bornés contenant C on a :

$$w.\lim_{n\to\infty} \sup\, (\Gamma K_n) \cap C \subseteq \Gamma C .$$

On rappelle que :

$$\text{w. lim sup}_{n\to\infty} Q_n = \{x \in E \mid \text{. Il existe une suite } \{x^{n_k}\}$$

telle que $x^{n_k} \in Q_{n_k}$ et $\text{w. lim}_{k\to\infty} x^{n_k} = x$ où $\{Q_{n_k}\}$ est une sous suite de $\{Q_n\}$ } .

(w.lim x^n désigne ici la limite faible de la suite $\{x^n\}$).

Nous allons montrer un théorème qui est l'analogue du théorème (1.5.1.), et qui justifie dans ce cadre d'hypothèses l'algorithme (1.3.).

(1.6.2.) Théorème.

Si les hypothèses ci-dessus : H3 et H4 sont vérifiées, alors toute valeur d'adhérence faible de toute suite générée par l'algorithme (1.3.) où on choisit v^n par le procédé (1.4.2.2.), est solution du problème P .

Démonstration :

Comme dans le théorème (1.5.1.) on peut montrer que : pour tout n et tout $m > n$ on a :

$$g(v^n, x^m) \leq 0 .$$

Et par construction : $g(v^m, x^m) > 0$.
Grâce à l'hypothèse H3, on déduit comme dans la démonstration du théorème (1.5.1.) que :

$$\text{(i)} \quad \lim_{n\to\infty} g(v^n, x^n) = 0 .$$

D'où, grâce au choix de v^n :

$$\limsup_{n\to\infty} g(v, x^n) \leq 0 \quad \text{pour tout } v \in V .$$

Ce qui compte tenu des propriétés de faible compacité de V et de faible semi continuité inférieure de $g(v,.)$, entraine que toute valeur d'adhérence faible $\tilde{x}$ de la suite $\{x^n\}$ appartient à X et satisfait aux inégalités : $g(v,\tilde{x}) \leq 0$ pour tout $v \in V$. Par conséquent, $\tilde{x}$ appartient à C et donc grâce à l'hypothèse H4 est une solution du problème P .

C.Q.F.D.

(1.7.) DES VARIANTES DE L'ALGORITHME (1.3.) PERMETTANT DE SIMPLIFIER LES FONCTIONNELLES g^n .

Dans l'algorithme (1.3.) on doit au cours du sous problème de la $n^{ème}$ itération déterminer la fonctionnelle g^{n+1} telle que :

$$g^{n+1}(v,x) \geq g^n(v,x) \qquad \forall\, v \in V \qquad \forall\, x \in X .$$

Ainsi les fonctionnelles g^n deviennent elles de plus en plus "complexes" quand l'indice n augmente car elles doivent garder toute l'information des itérations antérieures. Nous proposons dans ce paragraphe deux variantes de cet algorithme qui évitent partiellement cet inconvénient. Nous nous contenterons de justifier ces modifications dans le cas des hypothèses H1 et H2.

(1.7.1.) Première variante.

Le principe de ce procédé consiste à simplifier l'expression de g^{n+1} pour une sous suite d'itérés $\{x^n\}_{n \in S}$ pour laquelle on sait, dans le cas où cette sous suite est infinie, que :

$$\lim_{\substack{n\to\infty \\ n\in S}} g(v^n,x^n) = 0 .$$

Plus précisémment, on se donne à priori, une suite arbitraire de réels $\{\varepsilon_n\}$ tels que :

$$\varepsilon_n > 0 \text{ pour tout } n \quad \text{et} \quad \lim_{n\to\infty} \varepsilon_n = 0 .$$

A l'itération n on détermine $x^n \in \Gamma K_n$ (problème principal) puis (sous problème) un élément v^n de V comme en (1.4.2.2.).
Si $g(v^n,x^n) \geq \varepsilon_n$, on détermine g^{n+1} comme dans l'algorithme (1.3.).
Sinon on cherche x^{n+1} élément de $\Gamma(K_n \cap \{x \mid g(v^n,x) \leq 0\})$ puis la fonctionnelle g^{n+1} : $V \times X \to [-\infty,+\infty[$ telle que (ces conditions définissent alors une fonction g^{n+1} plus simple que dans le cas précédent) :

$$\begin{cases} g^{n+1}(v,.) \text{ est semi continue inférieurement pour tout } v \in V \\ g^{n+1}(v,x) = g(v,x) \quad \forall x \in X \quad \forall v \in \{v \in V | g(v,x^{n+1}) = 0\} \\ g^{n+1}(v,x) \leq g(v,x) \quad \forall x \in X \quad \forall v \in V \end{cases}$$

On pose $S = \{n | g(v^n,x^n) - \varepsilon_n \leq 0\}$.

On démontre sans difficulté le lemme suivant :

LEMME.

Soit la suite $\{x^n\}$ infinie construite par l'algorithme ci-dessus. Les hypothèses H1 et H2 entrainent qu'on a l'alternative suivante :

a) Le cardinal de S est fini : à partir d'un rang N on applique l'algorithme du (1.3.) d'où les mêmes propriétés.

b) Le cardinal de S est infini : toute valeur d'adhérence de la sous suite $\{x^n\}_{n \in S}$ est solution du problème P .

(1.7.2.) Deuxième variante.

Pour simplifier nous supposons dans ce paragraphe qu'on choisit g^n tel que :

$$g^n(v,x) = \begin{cases} g(v,x) & \forall v \in V_n \quad \forall x \in X . \\ -\infty & \text{sinon.} \end{cases}$$

L'ensemble V_n est un sous ensemble fini de V tel que :

$$V_n \subseteq \bigcup_{i=1}^{(n-1)} \{v^i\} ,$$

où les éléments v^i sont choisis comme en (1.4.2.2.).

On se donne à priori une suite de réels $\{a_n\}$ non décroissants et bornés supérieurement :

A la $m^{ème}$ itération on choisit alors V_{m+1} tel que :

$$V_{m+1} = V_m \cup \{v^m\} - A_m .$$

où $A_m = \{n | v^n \in V_m \ , \ g(v^n,x^m) < 0 \ , \ a_m - a_n \geq g(v^n,x^n)\}$.

On élimine donc à chaque itération une partie des v^i "non actifs".

LEMME.

Pour toute suite $\{x^n\}$ infinie construite par l'algorithme ci-dessus, les hypothèses H1 et H2 entrainent que toute valeur d'adhérence de cette suite est solution du problème P .

Démonstration :

Soit $A_1 = \{n | v^n \in V_m \quad \forall\, m > n\}$.

On a donc par construction :

$$g(v^n,x^m) \leq 0 \quad \forall\, n \in A_1 \quad \forall\, m > n \ .$$

On en déduit comme en (1.5.1.) que :

(1) $$\lim_{\substack{n\to\infty \\ n \in A_1}} g(v^n,x^n) = 0 \ .$$

Soit A_2 l'ensemble des indices n qui n'appartiennent pas à A_1 . On sait donc que par construction à tout $n \in A_2$ il correspond $m > n$ tel que :

$$a_m - a_n \geq g(v^n,x^n) > 0 \ .$$

D'où on déduit que si l'ensemble A_2 est infini on a, puisque $\{a_n\}$ est une suite de Cauchy,

$$\lim_{\substack{n\to\infty \\ n \in A_2}} g(v^n,x^n) = 0 \ .$$

On termine facilement la démonstration.

C.Q.F.D.

(1.8.) EXEMPLES DE PROBLEMES.

On explicite ici quelques problèmes qui rentrent dans le cadre de cette théorie:

- l'application Γ est fermée
- problèmes d'optimisation dans le cas convexe ou non convexe
- recherche d'un point fixe d'une application
- optimisation vectorielle
- recherche d'un point de selle d'une fonction convexe-concave
- point d'équilibre de NASH d'un jeu à n joueurs
- équilibre économique.

La variété de ces exemples montre l'intérêt pratique de cette étude. Pour chacun d'eux nous considérons un ensemble C défini comme en (1.2.) (ce qui implique une famille $\mathcal{F}$ de parties de X) et une application Γ de $\mathcal{F}$ dans $\mathcal{P}(E)$.
Pour pouvoir appliquer les résultats précédents nous devrons vérifier que l'une ou l'autre des séries d'hypothèses suivantes sont satisfaites :

α) H1 et H2 : On doit vérifier en particulier que pour toute suite infinie $\{x^n\}$ générée par l'algorithme (1.3.) toute valeur d'adhérence qui appartient à C , appartient aussi à ΓC .

β) H2 et H4 : On doit vérifier que pour toute suite infinie $\{x^n\}$ générée par l'algorithme (1.3.) toute valeur d'adhérence faible de cette suite qui appartient à C , appartient aussi à ΓC .

(1.8.1.) Cas d'une application Γ fermée.

On considère un ensemble C vérifiant H1 , une application Γ qui à tout compact non vide A de $\mathbb{R}^q$ fait correspondre le compact non vide ΓA tel que $\Gamma A \subset A$.

Nous considérons l'ensemble $\mathcal{F}$ des parties compactes de $\mathbb{R}^q$ avec la topologie définie par la distance de HAUSDORFF, cf. [7]. Supposons que pour cette topologie l'application Γ soit fermée dans le sens suivant :

"Pour toute suite $\{x^n\} \subset X$ telle que $\lim_{n\to\infty} x^n = \bar{x}$

et toute suite de compacts $\{A_n\}$ telle que $\lim_{n\to\infty} A_n = \bar{A}$

qui vérifient $x^n \in \Gamma A_n$ on a :

$$\bar{x} \in \Gamma\bar{A} \ .$$

On montre alors que l'application Γ vérifie les hypothèses H2. En général la réciproque n'est pas vraie. Malgré son apparente généralité cet exemple est assez restrictif car il implique le comportement de la suite $\{x^n\}$ pour toute suite $\{A_n\}$ de compacts qui convergent vers C, ce qui n'est pas indispensable pour le sujet qui nous intéresse (on ne considère en fait que les ensembles A_n contenant C).

(1.8.2.) Problème d'optimisation.

Nous ne donnerons pas de développements particuliers à cet exemple très important car cette question sera traitée en détails dans le chapitre II.

Soit le problème :

"trouver $\bar{x} \in C$ tel que $f(\bar{x}) = \min (f(x) | x \in C) = m$"

On notera :

$$\Gamma K = \{x \in K | f(x) = \min (f(y) | y \in K)\} \ .$$

L'ensemble ΓC est l'ensemble des solutions du problème considéré.

(1.8.2.1.) Problème d'optimisation non convexe.

On suppose que C satisfait aux hypothèse Hl et que f est une fonctionnelle semi continue inférieurement.

Pour tout ensemble compact K_n contenant C, il est évident que : ΓK_n est non vide.

Soit $\{K_n\}$ la suite d'ensembles contenant C, considérée dans l'algorithme (1.3.).

Le problème principal de l'itération n est le suivant :

"Trouver $x^n \in K_n$ tel que : $f(x^n) = \min (f(x) | x \in K_n) = m_n$"

Puisque $K_n \supseteq K_{n+1} \supseteq C$ on a :

$m \geq m_{n+1} \geq m_n$, donc $\lim_{n\to\infty} f(x^n)$ existe et on a :

(1) $$\lim_{n\to\infty} f(x^n) = f^* \leq m .$$

Soit $\tilde{x}$ un point d'adhérence de $\{x^n\}$: il existe une sous suite $\{x^n\}_{n\in S}$ telle que :

$$\lim_{\substack{n\to\infty \\ n\in S}} x^n = \tilde{x} .$$

Puisque f est s.c.i., il vient :

$$f(\tilde{x}) \leq \lim_{\substack{n\to\infty \\ n\in S}} \inf f(x^n) = \lim_{n\to\infty} f(x^n) = f^* .$$

La relation (1) entraine que :

$$f(\tilde{x}) \leq m .$$

Si on suppose que $\tilde{x} \in C$: $f(\tilde{x}) \geq m$, par conséquent :

$$f(\tilde{x}) = m = f^*$$

ce qui montre que les hypothèses H2 sont satisfaites.

(1.8.2.2.) Problème d'optimisation convexe.

Supposons que l'ensemble C satisfait aux hypothèses H3 et que la fonctionnelle f est convexe et semi continue inférieurement. On montre comme ci-dessus que les hypothèses H4 sont satisfaites.

(1.8.3.) Problème de point fixe (application faiblement contractante).

Soit un ensemble C convexe qui vérifie les hypothèses H3 où on suppose de plus que E est un espace de Hilbert et une application U univoque de X dans X faiblement contractante c'est-à-dire telle que pout tout $x, y \in X$ on a :

$$\|Ux-Uy\| \leq \|x-y\| .$$

A toute partie A convexe fermée bornée non vide de X on fait correspondre l'ensemble ΓA convexe fermé borné non vide (cf. BROWDER et PETRYSHIN [3]) défini par :

$$\Gamma A = \{x \in A \mid Ux = x\} .$$

Pour vérifier que Γ satisfait à H4, il suffit de montrer que pour toute suite $\{x^n\}$ d'éléments tels que :

$$\begin{cases} x^n = Ux^n & \forall\ n \\ \text{w. } \lim_{n\to\infty} x^n = \tilde{x} \in C & \text{(limite faible)} , \end{cases}$$

on a $\tilde{x} = U\tilde{x}$.

On démontre aisément cette affirmation par un raisonnement analogue à celui du lemme a) du théorème (5.2.2.).

(1.8.4.) Problème de point fixe (application continue).

Soit un ensemble C convexe compact qui vérifie les hypothèses H1. On considère une application T de X dans X continue,

Etant donnéela famille $\mathcal{F}$ des parties convexes compactes non vide de X, on définit l'application Γ de $\mathcal{F}$ dans $\mathcal{F}$ par :

$$\Gamma A = \{x \in A \mid x = Tx\} \qquad \forall\, A \in \mathcal{F}.$$

Grace au théorème de BROUWER on sait que si A est non vide ΓA est non vide. On vérifie (c'est une conséquence immédiate de la continuité de T) que l'application Γ satisfait aux hypothèses H2.

(1.8.5.) Optimisation vectorielle.

Cf. DA CUNHA, POLAK [6] .

Soit un ensemble C défini comme en (1.2.) avec les hypothèses H1. On suppose de plus que C est convexe, plus précisémment que :

$$\begin{cases} X \text{ est convexe} \\ g(v,.) \text{ est convexe pour tout } v \in V . \end{cases}$$

Soit f une application de X dans R^p . On notera :

$$L = \{1, 2, \ldots, p\}$$

$$f(x) = (f_1(x), f_2(x), \ldots, f_p(x))$$

où $f_i : X \to \mathbb{R}$ pour $i = 1, 2, \ldots, p$.

Pour tout convexe compact A tel que $X \supseteq A \supseteq C$ on définit A (ceci précise le problème qu'on se pose) par :

$$\Gamma A = \{x \in A \mid \{y \in A \mid \begin{matrix} f_i(y) \leq f_i(x) & \forall\, i \in L \\ \sum_{i \in L} f_i(y) < \sum_{i \in L} f_i(x) \end{matrix} \} = \emptyset\} .$$

Nous supposerons que les fonctionnelles f_i sont semi-continues inférieurement et strictement convexes.

On vérifie facilement que pour tout A convexe compact contenant C l'ensemble ΓA est non vide. Pour que les hypothèses H2 sur Γ soient satisfaites il nous reste à montrer le lemme suivant :

Lemme :

Pour toute suite $\{x^n\}$ générée par l'algorithme (1.3.) toute valeur d'adhérence de cette suite qui appartient à C appartient aussi à ΓC .

Démonstration :

Supposons qu'il existe (raisonnement par l'absurde) un point d'adhérence $\tilde{x}$ qui appartienne à C et non à ΓC . Alors il existe $y \in C$ tel que :

$$\begin{cases} f_i(y) \leq f_i(\tilde{x}) \quad \forall\, i \in L\ , \\ \text{il existe } y \in L \text{ tel que } f_j(y) < f_j(\tilde{x})\ . \end{cases}$$

Soit $u = \frac{y+\tilde{x}}{2}$ qui appartient à C car C est convexe.
On a (stricte convexité de f) :

$$(1) \qquad f_i(u) < \frac{1}{2}\,(f_i(y) + f_i(\tilde{x})) < f_i(\tilde{x}) \qquad \forall\, i \in L\ .$$

La sous suite $\{x^n\}_{n \in S}$ qui converge vers $\tilde{x}$ sera donc telle que (semi continuité inférieure de f) :

$$(2) \qquad \begin{cases} \text{Il existe } u \in C \text{ et un entier } N \text{ tel que on ait :} \\ f_i(u) < f_i(x^n) \qquad \forall\, n > N\ ,\ (n \in S)\ , \quad \forall\, i \in L\ . \end{cases}$$

Mais par définition l'élément x^n appartient à ΓK_n , donc :

$$\{x \in K_n \mid f_i(x) \leq f_i(x^n)\ \forall\, i \in L\ ;\ \sum_{i \in L} f_i(x) < \sum_{i \in L} f_i(x^n)\} = \emptyset$$

ce qui, puisque l'ensemble K_n contient C, contredit (2).

C.Q.F.D.

(1.9.) CAS DE PROBLEMES SUR UN ESPACE PRODUIT.

(1.9.1.) Introduction.

On peut étendre sans difficulté les résultats précédents au cas de problèmes sur un espace produit. En vue des exemples que nous traiterons dans ce paragraphe, nous introduisons les hypothèses et notations suivantes :

$$E = \prod_{i=1}^{N} E_i \quad , \qquad X = \prod_{i=1}^{N} X_i \ .$$

Les ensembles X_i et C_i sont des parties convexes fermées bornées non vides de E_i et on a :

$$\begin{cases} C_i = \{x_i \in X_i \,|\, g_i(v,x_i) \leq 0 \quad \forall\, v \in V_i\} \quad i = 1, 2.., N \ . \\ \\ C = \prod_{i=1}^{N} C_i \ . \end{cases}$$

Nous ferons sur les C_i soit les hypothèses H1 soit les hypothèses H3. On suppose toujours que les fonctionnelles $g_i(v,.)$ sont convexes.

On considère la famille $\mathcal{F}$ des ensembles A où :

$$A = \prod_{i=1}^{N} A_i$$

les ensembles A_i étant des parties convexes fermées bornées telles que $X_i \supseteq A_i \supseteq C_i$.

On étend facilement l'algorithme (1.3.) à ces problèmes. Nous allons considérer dans la suite quelques exemples; pour chacun d'eux nous montrons que l'application Γ correspondante satisfait aux hypothèses H2 ou H4.

(1.9.2.) Problème de point de selle.

Cas où $N = 2$. Soit L une fonctionnelle définie sur $X_1 \times X_2$ telle que :

$L(.,x_2)$ est convexe et s.c.i. $\quad \forall\, x_2 \in X_2$

$L(x_1,.)$ est concave et s.c.s. $\quad \forall\, x_1 \in X_1$

Nous ferons sur les ensembles C_1 et C_2 les hypothèses H3. Etant donné un couple (K_1,K_2) de convexes fermés bornés contenant C_1 et C_2 on posera :

$$\Gamma(K_1 \times K_2) = \{(\bar{x}_1,\bar{x}_2) \in K_1 \times K_2 \mid L(\bar{x}_1,x_2) \leq L(\bar{x}_1,\bar{x}_2) \leq L(x_1,\bar{x}_2) \quad \forall\, (x_1,x_2 \in K_1 \times K_2\} .$$

$\Gamma(K_1 \times K_2)$ désigne l'ensemble des points de selle de L sur $K_1 \times K_2$. Cet ensemble n'est pas vide grace au théorème de SION-KAKUTANI (cf; BERGE [2]). Pour vérifier que l'application Γ satisfait à H4 il reste à montrer le résultat suivant :

LEMME :

Soit une suite $\{(x_1^n,x_2^n)\}$ où le couple (x_1^n,x_2^n) est point de selle de L sur $K_{1,n} \times K_{2,n}$.

(Les ensembles convexes fermés bornés $K_{1,n}$ et $K_{2,n}$ contiennent respectivement C_1 et C_2). Toute valeur d'adhérence faible de cette suite qui appartient à $C_1 \times C_2$ appartient aussi à $\Gamma(C_1 \times C_2)$.

Démonstration :

Par définition on a :

$$(1) \quad \begin{cases} (x_1^n,x_2^n) \in K_{1,n} \times K_{2,n} \\ L(x_1^n,x_2) \leq (L(x_1^n,x_2^n) \leq L(x_1,x_2^n) \quad \forall (x_1,x_2) \in K_{1,n} \times K_{2,n} \end{cases} .$$

Soit $(\tilde{x}_1,\tilde{x}_2)$ une valeur d'adhérence faible de cette suite :

$$\lim_{\substack{n\to\infty \\ n\in S}} (x_1^n,x_2^n) = (\tilde{x}_1,\tilde{x}_2) \quad \text{faible.}$$

Supposons que $(\tilde{x}_1,\tilde{x}_2) \in C_1 \times C_2$. On a donc :

$$L(x_1^n,\tilde{x}_2) \leq L(x_1^n,x_2^n) \leq L(\tilde{x}_1,x_2^n) .$$

D'où grâce à la semi continuité inférieure de $L(.,\tilde{x}_2)$ et la semi continuité supérieure de $L(\tilde{x}_1,.)$:

$$\begin{cases} L(\tilde{x}_1,\tilde{x}_2) \leq \liminf\limits_{\substack{n\to\infty \\ n\in S}} L(x_1^n,\tilde{x}_2) \leq \liminf\limits_{\substack{n\to\infty \\ n\in S}} L(x_1^n,x_2^n) . \\ L(\tilde{x}_1,\tilde{x}_2) \geq \limsup\limits_{\substack{n\to\infty \\ n\in S}} L(\tilde{x}_1,x_2^n) \geq \limsup\limits_{\substack{n\to\infty \\ n\in S}} L(x_1^n,x_2^n) . \end{cases}$$

Il vient que :

$$(2) \qquad \lim_{\substack{n\to\infty \\ n\in S}} L(x_1^n,x_2^n) = L(\tilde{x}_1,\tilde{x}_2) .$$

De (1) on tire par passage à la limite que pour tout $(x_1,x_2) \in C_1 \times C_2$ on a :

$$L(\tilde{x}_1,x_2) \leq \liminf_{\substack{n\to\infty \\ n\in S}} L(x_1^n,x_2) \leq \lim_{\substack{n\to\infty \\ n\in S}} L(x_1^n,x_2^n) \leq \limsup_{\substack{n\to\infty \\ n\in S}} L(x_1,x_2^n) \leq L(x_1,\tilde{x}_2)$$

D'où grâce à (2) :

$$L(\tilde{x}_1,x_2) \leq L(\tilde{x}_1,\tilde{x}_2) \leq L(x_1,\tilde{x}_2) \qquad \forall\ (x_1,x_2) \in C_1 \times C_2$$

C.Q.F.D.

(1.9.3.) Détermination d'un point d'équilibre de NASH d'un jeu à N joueurs.

Voir BERGE [2] ou BURGER [4].

Nous considérons des ensembles C_i $i = 1, \ldots, N$ vérifiant les hypothèses H1 et les hypothèses supplémentaires du (1.9.1.) (convexité).

On considère N fonctionnelles f_i définies et continues sur X qui sont convexes séparément par rapport à la $i^{ème}$ variable c'est-à-dire que la fonctionnelle $f_i(x_1,\ldots,x_{i-1},.,x_{i+1},.,x_N)$ est convexe pour tout $x_j \in X_j$ $j = 1, \ldots, N$.
On définit la famille $\mathcal{F}$:

$$\mathcal{F} = \{A = \prod_{i=1}^{N} A_i \quad | \quad \begin{array}{l} A_i \text{ convexe compact} \\ X_i \supseteq A_i \supseteq C_i \qquad i = 1, \ldots, N \end{array} \}$$

et pour tout $A \in \mathcal{F}$ on a :

$$\Gamma A = \{(\bar{x}_1,\ldots,\bar{x}_N) \in A \mid f_i(\bar{x}_1,\ldots,\bar{x}_N) = \min_{x_i \in A_i} f_i(\bar{x}_1,\ldots,\bar{x}_{i-1},x_i,\bar{x}_{i+1},..)\}$$
$$i = 1, 2, \ldots, N$$

L'ensemble ΓA est l'ensemble des points d'équilibre du jeu à N joueurs défini par les fonctionnelles f_i et les ensembles A_i .
Pour tout $A \in \mathcal{F}$ on a : $\Gamma A \neq \emptyset$ (voir BERGE [2]).

Il reste à vérifier que Γ satisfait à H2.
Cela peut se faire en remarquant que ce problème se ramène à la recherche d'un point fixe d'une application continue sur un convexe compact d'un espace de dimension finie (c'est de cette façon qu'on montre les théorèmes d'existence pour ces problèmes). On est ramené au problème (1.8.4.).
On peut aussi le montrer directement (cf. MARTINET [11]).

(1.9.4.) Equilibre économique dans une économie d'échange.

Cf. BURGER [4].

Nous considérons un problème sur $C = \prod_{i=1}^{N} C_i$ où les ensembles C_i satisfont aux hypothèses H1 et où de plus :

$$\begin{cases} \text{les ensembles } X_i \text{ sont des parties convexes de } \mathbb{R}^{n_i}. \\ \text{les fonctionnelles } g_i(v,.) \text{ sont convexes.} \end{cases}$$

On considère :

$$\mathcal{P} = \{P \in \mathbb{R}^m \mid P_j \geq 0 \ , \ \sum_{j=1}^{m} P_j = 1\}$$

et une famille d'ensembles $\mathcal{F}$ définie par :

$$\mathcal{F} = \{A = \prod_{i=1}^{N} A_i \times P \ \Big| \ \begin{array}{l} A_i \text{ partie convexe compacte} \\ X_i \supseteq A_i \supseteq C_i \qquad i = 1, \ldots, N \end{array} \}$$

On définit pour tout $A \in \mathcal{F}$ l'ensemble ΓA par :

$$\Gamma A = \{(\bar{x}_1, \ldots, \bar{x}_N, \bar{P}) \in A \mid \begin{array}{l} \text{les relations (1), (2) et (3) sont} \\ \text{satisfaites} \end{array} \},$$

où les relations (1), (2) et (3) sont les suivantes :

(1) $\quad (\bar{P}, \bar{x}_i) \leq a_i \qquad i = 1, \ldots, N$

(2) $\quad f_i(\bar{x}_i) = \min (f_i(x_i) \mid x_i \in A_i, \ (\bar{P}, x_i) \leq a_i) \qquad i = 1, \ldots, N$

(3) $\quad \sum_{i=1}^{N} \bar{x}_i = \omega \ .$

On donne :

- les réels a_i et les fonctions f_i $\qquad i = 1, 2, \ldots, N$
- l'élément ω de R^m .

Dans les modèles économiques habituels les symboles ci-dessus ont les significations suivantes :

m : nombre de biens

N : nombre d'unités économiques (productrices ou consommatrices)

P : vecteur prix , $P \in \mathbb{R}^m$.

$(x_1,\ldots,x_N)$: vecteurs biens des N unités.

$-f_i$: fonction d'utilité de la $i^{ème}$ unité économique.

C_i : ensemble admissible de la $i^{ème}$ unité économique.

On est conduit à faire les hypothèses suivantes :

a) Pour tout $P \in \mathcal{P}$ il existe $\varepsilon > 0$ tel que, pour tout i $i = 1, 2, \ldots, N$, l'ensemble $\{x_i \in C_i \mid (P,x_i) \leq a_i - \varepsilon\}$ soit non vide.

b) Les fonctionnelles f_i sont convexes et semi continues inférieurement.

c) Pour tout A de $\mathcal{F}$ l'ensemble ΓA est non vide (existence d'un point d'équilibre pour tout élément de $\mathcal{F}$).

Il reste à vérifier que Γ satisfait à H2 . Comme dans l'exemple précédent ceci peut être fait en remarquant que ce problème se ramène à la recherche d'un point fixe d'une application continue sur un convexe compact d'un espace de dimension finie (voir problème (1.8.4.)).

On peut aussi le montrer directement (voir MARTINET [11]).

BIBLIOGRAPHIE

[1] AUSLENDER, A.
Problèmes de minimax via l'analyse convexe et les inégalités variationnelles : théorie et applications.
Cours A.E.A. - Université de Clermont (1971).

[2] BERGE, C.
Théorie générale des jeux à n joueurs.
Gauthiers-Villars (1957).

[3] BROWDER, F.E. ; PETRYSHIN, W.V.
Construction of fixed points of non linear mappings in Hilbert space.
Journal of math. anal. and appl. 20 (1967) 197-228.

[4] BURGER, E.
Introduction to the theory of games.
Prentice-Hall (1963).

[5] CHENEY, GOLDSTEIN.
Newton's method for convex programming and Tchebycheff approximation.
Numerishe mathematik. Vol. 1 (1959) 253-268.

[6] DA CUNHA ; POLAK, E.
Constrained minimization under vector-valued criteria in linear topological spaces.
Mathematical Theory of Control. Edited by BALAKRISNHAN.
Academic press (1967). pp 96-108.

[7] HAUSDORFF, F.
Set Theory
Chelsea publishing company. New-York (1962).

[8] KAPLAN, A.A.
Determination of the extremum of a linear function on a convex set.
Soviet math. 9 (1968) 269-271.

[9] KELLEY, J.E.
The cutting plane method for solving convex programs.
J. SIAM 8 (1960) 703-712.

[10] LAURENT, P.J.
Approximation et Optimisation.
Hermann (1972).

[11] MARTINET, B.
a) La méthode d'accumulation des contraintes. Applications.
Colloque d'Analyse Numérique d'Anglet. Juin 1971.
b) Généralisation des algorithmes d'accumulation des contraintes. Applications.
The fourth conference on probability theory.
BRASOV - 12 - 18 sept. 1971.

[12] ROCKAFELLAR, R.T.
Convex analysis. Princeton Univ. Press (1970).

[13] ZANGWILL
Nonlinear programming. A unified approach
Prentice Hall (1969).

CHAPITRE II

METHODES DE RESOLUTION DE PROBLEMES D'OPTIMISATION AVEC CONTRAINTES PAR APPROXIMATION EXTERIEURE DE L'ENSEMBLE ADMISSIBLE

(2.1.) INTRODUCTION.

Les problèmes d'optimisation sont un cas particulier des problèmes étudiés au chapitre précédent. Mais leur importance tant théorique que pratique et leurs propriétés particulières justifient l'étude que nous allons faire.

Soit un espace de Banach E, une partie C non vide de E et f une fonctionnelle définie sur E.
On considère les "problèmes" suivants :

Problème P :

"trouver $\bar{x} \in C$ tel que $f(\bar{x}) = \min (f(x) | x \in C)$".

Problème Pa :

"étant donné un voisinage V de 0 dans E et un nombre réel $\varepsilon > 0$ on cherche $\hat{x}$ tel que :

$(\hat{x}+V) \cap C \neq \emptyset$, $|f(\hat{x}) - \inf (f(x) | x \in C)| < \varepsilon$ ".

On remarque qu'on peut se poser le problème Pa même si le problème P n'a pas de solution.

Problème P_V :

"trouver le nombre m défini par :

$m = \inf (f(x) | x \in C)$".

Ces différents problèmes peuvent être résolus par des procédés analogues à ceux qui ont été étudiés dans le chapitre I. Ceuxici ont fait l'objet de très nombreuses publications qu'on peut classer en deux catégories :

a) Les méthodes de coupe (cut-off methods) dont les variantes s'appellent :

méthodes d'accumulation des contraintes, méthodes de relaxation, méthodes du plan sécant (cutting plane methods). Ces procédés ont d'abord été utilisés pour la résolution approchée de problèmes de programmation linéaire par DANTZIG [4] qui a beaucoup contribué à leur développement. En 1959 CHENEY et GOLDSTEIN [3] précisent cette technique dans le cas de l'optimisation convexe et l'appliquent à des problèmes d'approximation. En 1960 KELLEY [11] propose la célèbre méthode des plans sécants.
Il existe à l'heure actuelle de très nombreuses publications sur ce sujet, nous n'en citerons que quelques unes :

AUSLENDER [1], EAVES et ZANGWILL [5], KAPLAN [10], GEOFFRION [7], LEVITIN-POLYAK [14], MARTINET [16], TOPKIS [19], ZANGWILL [20].

Ces méthodes sont basées sur le même principe que les procédés du chapitre I.

Dans le cas d'un problème convexe où la fonctionnelle est uniformément convexe on peut définir des variantes très intéressantes de ces procédés. Ces méthodes ont été étudiées en particulier par LAURENT-MARTINET [12], EAVES et ZANGWILL [5], HAUGAZEAU [8], LIONS et TEMAM [15].

b) Les méthodes de pénalités extérieures : Ce sont des procédés très connus des utilisateurs par leur simplicité de mise en oeuvre, ils ont donc donné lieu à de très nombreux travaux. Ces méthodes sont sans doute dues à COURANT. Elles consistent à remplacer le problème P par un problème sans contrainte où on ajoute à la fonctionnelle f un terme de pénalité $\lambda\ p(.)$ où la fonctionnelle p est nulle sur C et strictement positive en dehors et le coefficient λ est strictement positif. Ce problème pénalisé "approche" d'autant mieux le problème P que le nombre λ est grand. On pourra sur cette question consulter FIACCO [6], LEVITIN et POLYAK [14], POLAK [18], CEA [2], etc ... .

Nous présentons dans ce chapitre un formalisme qui englobe ces deux types de procédé. Il s'agit de méthodes itératives où à la $n^{ème}$ itération, on cherche x^n (problème principal) tel que :

$$f(x^n) + \varphi_n(x^n) = \min\ (f(x) + \varphi_n(x) \mid x \in E).$$

La fonctionnelle de pénalisation généralisée φ_n prend ses valeurs dans $[0,+\infty]$ et est nulle sur C ; qualitativement on peut la considérer comme une approximation inférieure de la fonction indicatrice* de C ; on notera qu'elle peut prendre la valeur zéro en un point hors de C ce qui n'est pas le cas d'une pénalité usuelle.

Compte tenu de x^n on détermine ensuite (sous problème) une fonctionnelle φ_{n+1} qui "approche mieux" la fonction indicatrice de C que φ_n. On a par rapport aux méthodes du chapitre I appliquées à l'optimisation les différences suivantes :

a) Les inf considérés ne sont pas forcément atteints, en tout cas on se contente de les résoudre de façon approchée.

b) Les problèmes principaux considérés sont de nature différente selon que φ_n prend ses valeurs dans $[0,+\infty]$ ou dans $[0,+\infty[$.

---

* On appelle fonction indicatrice de C la fonction qui est nulle sur C et prend la valeur $+\infty$ ailleurs.

c) Dans le cas où C est un ensemble convexe et où la fonctionnelle f est uniformément convexe, on a des propriétés particulières qui permettent d'envisager d'autres procédés très intéressants.

d) A un problème d'optimisation avec contraintes donné on peut faire correspondre un lagrangien (non unique). L'interprétation lagrangienne d'un procédé d'optimisation présente beaucoup d'intérêt ; ce point de vue qui dépasse le cadre des méthodes d'approximation extérieure sera étudié dans le chapitre III.

Nous avons divisé ce chapitre en deux grandes parties :

α) **Méthodes de résolution de problèmes non convexes.**

Il s'agit d'une extension des procédés du chapitre I. On considère deux classes d'hypothèses qui correspondent à celles du chapitre précédent.

β) **Méthode de résolution de problèmes convexes où la fonctionnelle f est uniformément convexe.**

Si le problème P admet une solution, celle-ci est unique. On décrit comme en LAURENT-MARTINET [12] et MARTINET [16] des méthodes où à chaque itération la nouvelle fonction de pénalisation φ_{n+1} peut être construite sous des conditions moins strictes que précédemment (en ne faisant intervenir que les paramètres localement actifs) ce qui est un avantage pratique très important. On montre de plus avec des hypothèses convenables que les approximations successives de la solution convergent fortement vers la solution du problème P.

Un exemple important (théorie de l'approximation) de tels problèmes est la recherche de l'élément de plus petite norme d'un ensemble C convexe fermé d'un espace de Hilbert. On étudie des cas particuliers de cet exemple : équations linéaires, recherche d'un point de selle d'une fonction convexe-concave.

PREMIERE PARTIE - CAS NON CONVEXE

(2.2.) ALGORITHME FONDAMENTAL - GENERALISATION DES METHODES DE COUPE ET DE PENALISATION EXTERIEURE.

SERVICE POLYCOPIE
MATHEMATIQUES APPLIQUEES
Université de GRENOBLE

(2.2.1.) Notations et hypothèses.

On donne un ensemble C non vide définicomme en (1.2.) par :

$$C = \{x \in X \mid g(v,x) \leq 0 \quad \forall\, v \in V\}$$

et une fonctionnelle f définie sur X et à valeurs réelles.

Nous ferons les hypothèses suivantes :

H1 :

Même hypothèse qu'en (1.2.).
X est une partie compacte de $\mathbb{R}^q$.
L'ensemble C est non vide. La fonctionnelle g définie sur $V \times X$ est telle que la famille $g(v,.)$ soit équi-uniformément continue sur X c'est-à-dire que :

"pour tout $\varepsilon > 0$, il existe un voisinage $\mathcal{U}$ de 0 dans $\mathbb{R}^p$ tel que pour tout $x,y \in X$: $(x-y) \in \mathcal{U}$ et tout $v \in V$ on ait :

$$|g(v,x) - g(v,y)| < \varepsilon" .$$

H2 :

f est semi continue inférieurement sur X.
Pour tout $a \in \mathbb{R}$ l'ensemble $\{x \in X \mid f(x) \leq a\}$ est compact.

(2.2.2.) Description de l'algorithme.

Il s'agit d'une extension de la méthode du (1.3.)

On se donne à priori une suite de nombres non négatifs : $\{\eta_n\}$ tels que $\lim_{n\to\infty} \eta_n = 0$.

A la première itération on se donne une fonctionnelle φ_1 s.c.i., nulle sur C , à valeurs dans $[0,+\infty]$ telle que :

$$\inf (f(x) + \varphi_1(x) \mid x \in X) = F_1 > -\infty .$$

(Dans le cas des hypothèses du (2.2.1.) on peut choisir évidemment φ_1 telle que

$$\varphi_1(x) = 0 \qquad \text{pour tout } x \in X).$$

A la $n^{ème}$ itération on considère la fonctionnelle f_n définie par :

$$f_n = f + \varphi_n ,$$

où la fonctionnelle φ_n (pénalisation généralisée) est s.c.i., nulle sur C et à valeurs dans $[0,+\infty]$.

On distinguera comme dans le chapitre I deux étapes :

a) Problème principal.

trouver x^n tel que :

$$x^n \in X \ , \ f_n(x^n) \leq a_n + \eta_n$$

où $$a_n = \inf (f_n(x) \mid x \in X).$$

Il s'agit donc en fait de la résolution approchée d'un problème d'optimisation.

b) Sous problème.

Il s'agit de déterminer φ_{n+1} . On distingue deux cas :

α) $x^n \in C$

Si $\eta_n = 0$ alors x^n est solution de $\mathcal{P}$. arrêt des itérations.

Sinon prendre $\varphi_{n+1} = \varphi_n$ et passer à l'itération (n+1).

β) $x^n \notin C$

On choisit $v^n \in V$ comme dans (1.4.2.), puis la fonctionnelle φ_{n+1} s.c.i. telle que :

$$\chi_C \geq \varphi_{n+1} \geq \varphi_n + \ell(g(v^n,.))$$

où ℓ est une fonction d'influence donnée et χ_C la fonction indicatrice de C ;

Définition :

On appelle fonction d'influence ℓ une fonction de $\mathbb{R}$ dans $\mathbb{R}^+$ telle que pour toute suite $\{t_n\}$ de réels on ait :

$$\limsup_{n \to \infty} t_n \leq 0 \quad \text{si} \quad \lim_{n \to \infty} \ell(t_n) = 0 .$$

On prendra souvent :

$$\varphi_{n+1} = \varphi_n + \ell(g(v^n,.)) .$$

(2.2.3.) Choix des fonctionnelles φ_n et ℓ.

On peut imaginer de nombreux choix de φ_n et ℓ .

En voici deux qui correspondent à des procédés simples utilisés effectivement dans les applications.

(2.2.3.1.) Les fonctions φ_n et ℓ sont des fonctions indicatrices d'ensemble.

On prend ℓ telle que : $\ell(u) = \begin{cases} 0 & \text{si} \quad u \leq 0 \\ +\infty & \text{si} \quad u > 0 \end{cases}$.

Supposons pour simplifier qu'au départ on choisisse $\varphi_1(x) = 0 \ \forall \ x$ et qu'on choisisse à chaque itération n si $x^n \notin C$:

$$\varphi_{n+1} = \varphi_n + \ell(g(v^n,.)) .$$

Alors, pour tout n, φ_n est une fonction indicatrice d'un ensemble compact contenant C.

On retrouve les procédés décrits dans le chapitre I et les procédés habituels de coupe.

b) Les fonctionnelles φ_n et ℓ sont des fonctions à valeurs finies.

Pour que les fonctionnelles φ_n soient suffisamment régulières sans être trop compliquées, on choisira souvent :

$$\ell(u) = \begin{cases} 0 & \text{si} \quad u \leq 0 \\ \beta u^2 & \text{si} \quad u > 0 \end{cases} \qquad (\beta > 0 \text{ est donné})$$

et dans le cas où $x^n \notin C$:

$\varphi_{n+1} = \varphi_n + \alpha_n \ell(g(v^n,.))$, où sont donnés les nombres $\alpha_n \geq 1$.

On trouve une généralisation de la méthode des pénalisations extérieures.

(2.2.4.) Propriétés de l'algorithme (2.2.).

THEOREME :

Pour toute suite infinie $\{x^n\}$ générée par l'algorithme ci-dessus les hypothèses du (2.2.1.) entrainent que :

a) toute valeur d'adhérence de la suite $\{x^n\}$ est solution du problème P.

b) $\lim_{n\to\infty} f(x^n) = \lim_{n\to\infty} f_n(x^n) = m = \inf\,(f(x)\,|\,x \in C)$.

Démonstration :

Pour alléger celle-ci nous ne considérerons que le choix de v^n du (1.4.2.2.). Par construction on sait que pour tout n on a :

(1) $$f_n(x^n) \leq \inf\,(f_n(x)\,|\,x \in X) + \eta_n \leq f_n(x^{n+1}) + \eta_n \leq f_{n+1}(x^{n+1}) + \eta_n$$

(2) $$f_n(x^n) \leq m + \eta_n$$

(3) $$f_n(x^n) \geq \inf\,(f(x)\,|\,x \in X) = \alpha > -\infty\ .$$

On déduit de (2) et (3) que la suite $\{f_n(x^n)\}$ possède une sous suite convergente.

Considérons un ensemble infini d'indices S tel que :

$$S = \{n\,|\lim_{n\to\infty} f_n(x^n) = f^*\}\ .$$

Pour tout $n,m \in S$ tels que $m > n$ on a par construction :

(4) $$f_m(x^m) \geq f_n(x^m) \geq f_n(x^n) - \eta_n\ .$$

et

$$f_m(x^m) - f_n(x^m) \geq \ell(g(v^n,x^m))\ .$$

D'où on tire de (4) par définition de S que :

$$\lim_{\substack{n\to\infty \\ n,m\in S \\ m>n}} \ell(g(v^n,x^m)) = 0$$

et grace aux propriétés de ℓ :

$$\lim_{\substack{n\to\infty \\ n,m\in S \\ m>n}} g^+(v^n,x^m) = 0 .$$

Par un raisonnement analogue à celui fait dans la démonstration du théorème (1.5.1.) on montre, compte tenu du fait que f est inf compacte sur X et que $g(v,.)$ est équi-uniformément continue, que :

$$\lim_{\substack{n\to\infty \\ n\in S}} g(v^n,x^n) = 0 .$$

En considérant le choix (voir (1.4.2.2.)) de v^n il vient :

$$\limsup_{\substack{n\to\infty \\ n\in S}} g(v,x^n) \leq 0 \qquad \forall\, v \in V .$$

D'où on tire grâce à la semi continuité inférieure de $g(v,.)$ que pour <u>toute</u> valeur d'adhérence $\tilde{x}$ de la suite $\{x^n\}_{n\in S}$ on a :

$$g(v,\tilde{x}) \leq 0 \qquad \forall\, v \in V .$$

D'où on tire puisque $\tilde{x} \in X$ que :

(5) $\qquad \tilde{x} \in C .$

On a donc grâce à la semi continuité inférieure de f :

$$m \leq f(\tilde{x}) \leq \liminf_{\substack{n\to\infty \\ n\in S}} f(x^n) \leq \limsup_{\substack{n\to\infty \\ n\in S}} f(x^n) \leq \lim_{\substack{n\to\infty \\ n\in S}} f_n(x^n) = f^* .$$

Or de (2), il vient :

$$f^* \leq m \ .$$

Par conséquent : $f^* = m$;

et :

$$\lim_{\substack{n\to\infty \\ n\in S}} f(x^n) = \lim_{\substack{n\to\infty \\ n\in S}} f_n(x^n) = m \ .$$

Comme cette limite est indépendante de S , on montre facilement par un raisonnement par l'absurde que :

$$\lim_{n\to\infty} f(x^n) = \lim_{n\to\infty} f_n(x^n) = m \ .$$

Compte tenu de (5) le théorème est démontré.

C.Q.F.D.

(2.2.5.) Cas où l'ensemble C est vide.

Supposons dans ce paragraphe que l'ensemble C soit vide et que : "il existe $\alpha > 0$ tel que pour tout $x \in X$ il existe $v \in V$ tel que $g(v,x) > \alpha$ " .

En gardant par ailleurs les mêmes hypothèses que ci-dessus on peut montrer que pour toute suite $\{x^n\}$ générée par l'algorithme du (2.2.) on a l'alternative suivante :

a) A un certain rang N on a :

$$\inf (f_N(x) \mid x \in X) = + \infty \ .$$

Les itérations ultérieures ne peuvent donc plus être définies, ce qui nous amène à stopper le procédé.

b) Pour tout n les nombres $\inf (f_n(x) | x \in X)$ sont finis.
On génère dans ce cas une suite $\{x^n\}$ infinie telle que :

$$\lim_{n\to\infty} f_n(x^n) \neq +\infty \quad .$$

(2.3.) AUTRES HYPOTHESES QUI PERMETTENT DE JUSTIFIER L'ALGORITHME (2.2.).

(2.3.1.) Hypothèses.

Soit E un espace de Banach, X une partie non vide de E et C une partie de X définie par :

$$C = \{x \in X | g(v,x) \leq 0 \quad \forall\, v \in V\} \quad .$$

On fait de plus les hypothèses :

H3 :

L'ensemble V est une partie séquentiellement compacte de l'espace vectoriel topologique séparé F . La famille de fonctionnelles $g(.,x)$ est équi-uniformément continue, c'est-à-dire :

"pour tout $\varepsilon > 0$, il existe un voisinage $\mathcal{U}$ de 0 dans F tel que pour tout $u,v \in V$: $(u-v) \in \mathcal{U}$ et pour tout $x \in X$, on ait :

$$|g(u,x) - g(v,x)| < \varepsilon \quad ".$$

L'ensemble C est non vide.

H4 :

Définissons la fonction h : $\mathbb{R} \to \bar{\mathbb{R}}$ qui à tout $u \in \mathbb{R}$ fait correspondre :

$$h(u) = \inf (f(x) | x \in X \ , \ g(v,x) \leq u \quad \forall\, v \in V) \ .$$

On suppose que $h(0)$ est fini, qu'il existe un réel $\varepsilon > 0$ tel que $h(\varepsilon)$ soit fini et que h soit semi continue inférieurement en 0. Cette dernière propriété signifie que pour la perturbation considérée le problème est inf-stable (voir LAURENT [13]). En fait cela entraine compte tenu de la non-croissance de h que cette fonction est définie sur $[0,\varepsilon]$ et qu'elle est continue en 0^+.

Cette hypothèse étant peu explicite, il est intéressant de la préciser sur des exemples :

a) Exemple 1 : Considérons les hypothèses du (2.2.1.).

Il est évident que $h(0)$ et $h(\varepsilon)$ sont finis $\forall\ \varepsilon > 0$.

Soit une suite $\{u_n\}$ de nombres ≥ 0 qui tendent vers zéro. Soit $x^n \in X$ une solution du problème :

$$\min\ (f(x) \mid g(v,x) \leq u_n \quad \forall\ v \in V,\ x \in X)\ .$$

Pour toute valeur d'adhérence $\tilde{x}$ de la suite $\{x^n\}$ on montre que :

$$g(v,\tilde{x}) \leq 0 \quad \forall\ v \in V\ ,\ \tilde{x} \in X \quad \text{donc} \quad \tilde{x} \in C\ .$$

Et par conséquent (f est s.c.i.) :

$$h(0) \leq f(\tilde{x}) \leq \liminf f(x^n) = \liminf h(u^n)$$

ce qui montre que h est s.c.i. en 0.

b) Exemple 2 : On suppose en plus de H3 que :

E est un espace de Banach réflexif

X est une partie bornée fermée et convexe

f est s.c.i. et convexe

$g(v,.)$ est s.c.i. et convexe.

On montre comme ci-dessus en considérant la topologie faible sur E que ces hypothèses entrainent H4.

c) Exemple 3 : On suppose en plus de H3 que :

E est un espace vectoriel topologique.

X est une partie convexe de E

f est une fonctionnelle convexe

$g(v,.)$ est convexe $\forall\, v \in V$

la fonction h est définie sur un ouvert contenant 0 .

On en déduit facilement que h est convexe donc continue en 0. Dans ce cas l'optimum n'est pas en général atteint.

(2.3.2.) Théorème :

Pour toute suite $\{x^n\}$ générée par l'algorithme (2.2.2.) les hypothèses H3 et H4 ci-dessus entrainent que :

a) $\limsup_{n\to\infty} g(v,x^n) \leq 0 \quad \forall\, v \in V$

b) $\lim_{n\to\infty} f(x^n) = \lim_{n\to\infty} f_n(x^n) = m$.

Démonstration :

Comme dans la démonstration du théorème (2.2.4.) on considère un ensemble d'indices S (il existe car la suite $f_n(x^n)$ est bornée) :

$$S = \{n \mid \lim_{n\to\infty} f_n(x^n) = f^*\}$$

pour lequel on montre que (en utilisant H3) :

$$\lim_{\substack{n\to\infty \\ n\in S}} g(v^n,x^n) = 0 .$$

Soit par définition de v^n :

(1) $$\limsup_{\substack{n\to\infty \\ n\in S}} g(v,x^n) \leq 0 \qquad \forall\, v \in V .$$

Posons :

$$\varepsilon_n = \max_{v \in V} g(v, x^n) \geq 0$$

Grace à (1) il vient :

$$\lim_{\substack{n \to \infty \\ n \in S}} \varepsilon_n = 0^+ .$$

Donc :

$$h(\varepsilon_n) \leq f(x^n) \leq f_n(x^n) \leq m + \eta_n$$

et par passage à la limite on a compte tenu de H4 :

$$m = h(0) = \lim_{\substack{n \to \infty \\ n \in S}} f(x^n) = \lim_{\substack{n \to \infty \\ n \in S}} f_n(x^n) .$$

Comme cette limite est indépendante de la suite S et que la suite $\{f_n(x^n)\}$ est bornée on en déduit que :

$$\lim_{n \to \infty} f(x^n) = \lim_{n \to \infty} f_n(x^n) = m$$

ce qui compte tenu de (1) montre le théorème.

C.Q.F.D.

COROLLAIRE 1 :

Si en plus des hypothèses précédentes on suppose que E est un espace vectoriel topologique séparé, que les fonctionnelles f et $g(v,.)$ sont semi-continues inférieurement et que l'ensemble X est séquentiellement compact, alors toute valeur d'adhérence de $\{x^n\}$ est solution du problème P.

COROLLAIRE 2 :

Si on suppose que E est un espace de Banach réflexif que f est convexe et s.c.i., que $g(v,.)$ est convexe et s.c.i. pour tout $v \in V$ et que X est un ensemble convexe fermé et borné alors toute valeur d'adhérence faible de $\{x^n\}$ est solution du problème P .

(2.4.) QUELQUES VARIANTES DE L'ALGORITHME (2.2.) PERMETTANT DE SIMPLIFIER LA FONCTIONNELLE φ_n .

Les fonctionnelles φ_n considérées dans l'algorithme (2.2.) sont de plus en plus complexes quand n croît car on impose que pour tout n on ait :

(i) $\varphi_{n+1} \geq \varphi_n + \ell(g(v^n,.))$.

Il est intéressant de trouver des variantes de l'algorithme (2.2.) où pour certaines itérations on n'impose pas aux fonctions φ_n de satisfaire à la relation (i). Ces procédés correspondent aux techniques que nous avons étudiées en (1.7.2.) et (1.7.3.):

pour certaines itérations on peut éliminer en partie ou totalité les contraintes non actives.

On pourra consulter aussi sur ce sujet EAVES-ZANGWILL [5] . Dans le cas où l'ensemble des contraintes est fini, GEOFFRION [7] définit une technique simple d'élimination des contraintes non actives ; nous n'examinerons pas ce cas ici.

(2.4.1.) Algorithme.

(Cf. Algorithme (1.7.2.)

a) Description.

Nous considérerons que les problèmes principaux sont résolus exatement. Plus précisément à l'itération n on considère $x^n \in X$ tel que :

$$f_n(x^n) = \min\,(f_n(x) \mid x \in X)$$

et $v^n \in V$ tel que :

$$g(v^n, x^n) \geq \theta \sup\,(g(v, x^n) \mid v \in V) - \rho_n$$

(où sont donnés $\theta \in]0,1[$ et la suite $\{\rho_n\}$: $\rho_n \geq 0$; $\lim_{n\to\infty} \rho_n = 0$).

On définit :

$$e(x^n) = \begin{cases} h(g(v^n, x^n)) & \text{si } x^n \notin C \\ 0 & \text{sinon.} \end{cases}$$

où h est une fonction d'influence donnée.

On pose :

$$q_{n+1} = \begin{cases} f_n & \text{si } x^n \in C \\ f_n + \ell(g(v^n, .)) & \text{sinon} \end{cases}$$

où $f_n = f + \varphi_n$.

On détermine ensuite x^{n+1} tel que :

$$x^{n+1} \in X \ , \ q_{n+1}(x^{n+1}) = \min\,(q_{n+1}(x) \mid x \in X)$$

Soit :

$$d_n = q_{n+1}(x^{n+1}) - f_n(x^n) - e(x^n) \ .$$

On cherche alors la fonctionnelle φ_{n+1} en envisageant deux cas :

PREMIER CAS : $d_n \geq 0$

On choisit la fonctionnelle φ_{n+1} semi-continue inférieurement qui applique E dans $[0,+\infty]$ telle que :

$$\varphi_{n+1}(x) = 0 \qquad \forall\, x \in C$$

$$q_{n+1}(x^{n+1}) \leq f_{n+1}(x^{n+1}) = \min(f_{n+1}(x) \mid x \in X)$$

On a alors plusieurs choix possibles pour φ_{n+1} (voir (2.6.4.)) dont le plus simple à mettre en oeuvre est le suivant :

$$\varphi_{n+1} = f + \sum_{i \in L_{n+1}} \ell(g(v^i,.))$$

où $L_{n+1} = \{i \in L_n \cup \{n\} \mid g(v^i,x^{n+1}) = 0\}$.

DEUXIEME CAS : $d_n < 0$

On choisit comme en (2.2.) $\varphi_{n+1} = \varphi_n + \ell(g(v^n,.))$. On ne peut dans ce cas simplifier l'expression de φ_{n+1} .

b) Propriétés de cet algorithme.

Soit $S = \{n \mid d_n \geq 0\}$.

On a le résultat suivant.

LEMME :

Si le cardinal de S est nul ou fini, à partir d'un certain rang la méthode se ramène à l'algorithme (2.2.). Si le cardinal de S est infini, toute valeur d'adhérence de la suite $\{x^n\}_{n \in S}$ est solution du problème P et on a :

$$\lim_{n\to\infty} f_n(x^n) = \lim_{\substack{n\to\infty \\ n\in S}} f(x^n) = m .$$

Démonstration :

Contentons-nous de montrer la deuxième partie du lemme. On sait que :

$$m \geq f_{n+1}(x^{n+1}) \geq f_n(x^n) .$$

Donc la suite $\{f_n(x^n)\}$ possède une limite :

(1) $$\lim_{n\to\infty} f_n(x^n) \leq m .$$

D'où on tire par définition de d_n et de l'ensemble S que :

$$\lim_{\substack{n\to\infty \\ n\in S}} e(x^n) = 0 .$$

Et grâce aux propriétés de $e(.)$, il vient :

$$\limsup_{\substack{n\to\infty \\ n\in S}} g(v,x^n) \leq 0 \qquad \forall\, v \in V .$$

Donc toute valeur d'adhérence $\tilde{x}$ de $\{x^n\}_{n\in S}$ appartient à C et il vient :

$$m = f(\tilde{x}) \leq \liminf_{\substack{n\to\infty \\ n\in S}} f(x^n) .$$

C.Q.F.D.

(2.4.2.) Algorithme.

Il s'agit d'une variante analogue à celle du (1.7.3.) dans le contexte d'un problème d'optimisation. On élimine à chaque itération une partie des contraintes non actives.

a) Description :

On choisit

$$\begin{cases} \varphi_1 = 0 \\ \cdots\cdots \\ \varphi_n = \sum_{i\in L_n} \ell(g(v^i,.)) \end{cases}$$

où $L_n \subseteq \{2,3,\ldots,(n-1)\}$.

Précisons comment est définie la fonctionnelle φ_{n+1} :

On détermine l'élément x^n comme en (2.2.), v^n comme en (2.4.1.) puis $L_{n+1} = \{n\} \cup L_n - H_n$

avec $H_n = \{i < n \,|\, g(v^i,x^n) < 0 \ ; \ f_n(x^n) - f_i(x^i) \leq e(x^i)\}$.

b) Propriétés de l'algorithme.

LEMME :

Soit $\{x^n\}$ une suite infinie construite par l'algorithme ci-dessus. Les hypothèses : H1 et H2 (ou H3 et H4) entrainent que :

- toute valeur d'adhérence de la suite $\{x^n\}$ est solution du problème P .
- $\lim_{n\to\infty} f(x^n) = \lim_{n\to\infty} f_n(x^n) = m$.

Démonstration :

Soit $S = \{n \,|\, \lim_{n\to\infty} f_n(x^n) = f^*\}$ et A_1 le sous ensemble de S dont les éléments n sont tels que le terme $\ell(g(v^n,.))$ ne sera plus éliminé ultérieurement. Si l'ensemble A_1 est infini on a comme pour l'algorithme (2.2.) :

$$(1) \qquad \lim_{\substack{n\to\infty \\ n \in A_1}} g(v^n,x^n) = 0$$

Soit A_2 le complémentaire de A_1 par rapport à S . Supposons que l'ensemble A_2 soit infini. On montre par définition de H_m que :

$$\lim_{\substack{n\to\infty \\ n \in A_2}} g(v^n,x^n) = 0 \ .$$

D'où on tire grâce à (1) que :

$$\lim_{\substack{n\to\infty \\ n \in S}} g(v^n,x^n) = 0$$

et

$$\begin{cases} \limsup_{\substack{n \to \infty \\ n \in S}} g(v, x^n) \leq 0 \qquad \forall\, v \in V \\ \lim_{\substack{n \to \infty \\ n \in S}} f_n(x^n) = \lim_{\substack{n \to \infty \\ n \in S}} f(x^n) = m \,. \end{cases}$$

Ces résultats sont vrais pour toute suite S d'où le lemme.

C.Q.F.D.

(2.5.) ALGORITHMES POUR LESQUELS LE PROBLEME PRINCIPAL EST SIMPLIFIE.

Un inconvénient majeur des procédés précédents est que la résolution du problème principal : un problème d'optimisation à résoudre de façon approchée, peut être couteuse et en tout cas le critère d'approximation du problème principal est mal défini pour les besoins pratiques. D'où l'idée de considérer des problèmes principaux où l'on n'exécuterait qu'une itération d'une méthode de minimisation. L'étude systématique de tels procédés n'est pas faite ici. Nous nous contentons de donner deux exemples.

(2.5.1.) Méthode de type gradient projeté.

En ce qui concerne la méthode habituelle du gradient projeté on pourra consulter par exemple LEVITIN et POLYAK [14] .

a) Considérons un problème P avec les hypothèses H1 et H2. On suppose de plus que X est convexe et que la fonctionnelle $g(v,.)$ est convexe pour tout $v \in V$.

La fonctionnelle f est convexe et possède une dérivée de FRECHET f' qui satisfait à une condition de Lipchitz :

il existe $D \in \mathbb{R}$ tel que pour tout $x,y \in E$ on ait :

$$\|f'(x) - f'(y)\| \leq D \, \|x-y\| \,.$$

b) Description de l'algorithme.

Pour simplifier on considère l'adaptation d'une méthode de coupe (les fonctions φ_n sont des fonctions indicatrices). Considérons deux nombres réels $a_1, a_2 > 0$ tels que $\frac{2}{D+2a_2} \geq a_1$. On construit une suite $\{x^n\}$ d'éléments de X à partir d'un élément $x^0 \in X$.

La $n^{ème}$ itération comprend les opérations suivantes :

PROBLEME PRINCIPAL :

On détermine x^n par :

$$x^n = P_{K_n}(x^{n-1} - \alpha_{n-1} f'(x^{n-1}))$$

où :
- $a_1 \leq \alpha_n \leq \frac{2}{D+2a_2}$
- P_{K_n} est l'opérateur projection sur K_n.

SOUS PROBLEME 1

On détermine, comme dans l'algorithme (2.2.), $v^n \in V$ puis l'ensemble K_{n+1} convexe fermé borné contenant C.

c) Propriétés de l'algorithme.

LEMME :

Soit $\{x^n\}$ une suite construite par l'algorithme ci-dessus ; les hypothèses de a) entrainent que :

α) toute valeur d'adhérence de $\{x^n\}$ est solution du problème P

β) $\lim_{n\to\infty} f(x^n) = m$.

Démonstration :

α) On a pour tout n :

$$f(x^{n+1})-f(x^n) = \langle f'(x^n), x^{n+1}-x^n \rangle + \int_0^1 \langle f'(x^n+\tau(x^{n+1}-x^n))-f'(x^n), x^{n+1}-x^n \rangle \, d\tau$$

$$f(x^{n+1})-f(x^n) \leq \langle f'(x^n), x^{n+1}-x^n \rangle + \frac{D}{2} \|x^{n+1}-x^n\|^2$$ puisque f' satisfait à une condition de Lipchitz

$$f(x^{n+1})-f(x^n) \leq -\frac{1}{\alpha_n} \langle x^n-\alpha_n f'(x^n)-x^{n+1}, x^{n+1}-x^n \rangle + (\frac{D}{2} - \frac{1}{\alpha_n}) \|x^{n+1}-x^n\|^2 .$$

Donc :

$$f(x^{n+1})-f(x^n) \leq -\frac{1}{\alpha_n} \langle x^n-\alpha_n f'(x^n)-x^{n+1}, x^{n+1}-P_{K_{n+1}}(x^n) \rangle -$$

(1)
$$-\frac{1}{\alpha_n} \langle x^n-\alpha_n f'(x^n)-x^{n+1}, P_{K_{n+1}}(x^n)-x^n \rangle + (\frac{D}{2} - \frac{1}{\alpha_n}) \|x^{n+1}-x^n\|^2 .$$

Par construction de x^n et grâce aux hypothèses faites sur C on montre (comme dans le théorème (1.5.1.) que :

(2) $\lim_{n\to\infty} d(x^n,C) = 0$ (où $d(.,c)$ désigne la distance à C)

et comme on a :

$C \subseteq K_{n+1}$ il vient :

$$\lim_{n\to\infty} \|P_{K_{n+1}}(x^n) - x^n\| = 0 .$$

Par conséquent, du fait qu'on a $\alpha_n \geq a_1 > 0$ et que la suite $\{x^n\}$ est bornée, il vient :

(3) $$\lim_{n\to\infty} \frac{1}{\alpha_n} \langle x^n - \alpha_n f'(x^n)-x^{n+1}, P_{K_{n+1}}(x^n)-x^n \rangle = 0 .$$

On sait de plus grâce à la propriété de la projection que :

$$\langle x^n - \alpha_n f'(x^n)-x^{n+1}, x^{n+1}-P_{K_{n+1}}(x^n) \rangle \geq 0 .$$

D'où on tire grâce à (1) et (2) que :

$$(4) \qquad f(x^{n+1}) - f(x^n) \leq - a_2 \, \|x^{n+1} - x^n\|^2 + \varepsilon_n$$

où les nombres ε_n sont tels que : $\lim_{n\to\infty} \varepsilon_n = 0$.

β) Soit la suite $S = \{n \mid f(x^{n+1}) > f(x^n)\}$.

On montre que si le cardinal de S est infini on a :

$$\lim_{n\to\infty} f(x^n) = m \ .$$

Grâce à (4), il vient :

$$(5) \qquad \lim_{\substack{n\to\infty \\ n\in S}} \|x^{n+1} - x^n\| = 0 \ .$$

Grâce à (2) et à l'uniforme continuité de f il vient :

$$f(x^n) - m \geq - \rho_n \qquad \text{où} \qquad \lim_{n\to\infty} \rho_n = 0 \ .$$

Soit $\bar{x} \in C$ et tel que $f(\bar{x}) = m$. On a :

$$- \rho_n \leq f(x^n) - m \leq < f'(x^n), x^n - \bar{x} > \quad = < f'(x^n), x^n - x^{n+1} + x^{n+1} - \bar{x} >$$

$$= < f'(x^n), x^n - x^{n+1} > + \frac{1}{\alpha_n} < x^n - x^{n+1} - \alpha_n f'(x^n), \bar{x} - x^{n+1} >$$

$$- \frac{1}{\alpha_n} < x^n - x^{n+1}, \bar{x} - x^{n+1} >$$

$$\leq \quad < f'(x^n), x^n - x^{n+1} > - \frac{1}{\alpha_n} < x^n - x^{n+1}, \bar{x} - x^{n+1} > \ .$$

Grâce aux propriétés de projection sur le convexe K_{n+1} .

Donc :

$$- \rho_n \leq . f(x^n) - m \leq < f'(x^n), x^n - x^{n+1} > - \frac{1}{\alpha_n} < x^n - x^{n+1}, \bar{x} - x^{n+1} > \ .$$

Par passage à la limite il vient, compte tenu de (5) :

$$(6) \qquad \lim_{\substack{n\to\infty \\ n\in S}} f(x^n) = m \ .$$

Si le complémentaire de S est fini ou vide on en déduit :

$$(7) \qquad \lim_{n\to\infty} f(x^n) = m \ .$$

Sinon, à tout $n \notin S$ on fait correspondre :

$i(n) \in S$: $i(n) = \max (\ell \mid \ell < n \ , \ \ell \in S)$

$j(n) \in S$: $j(n) = \min (\ell \mid \ell > n \ , \ \ell \in S)$.

D'après (6) on sait que :

$$(8) \qquad \lim_{n\to\infty} f(x^{i(n)}) = \lim_{n\to\infty} f(x^{j(n)}) = m .$$

Or par construction :

$$f(x^{i(n)}) + \varepsilon_n \geq f(x^n) \geq f(x^{j(n)}) .$$

D'où on tire de (8) par passage à la limite que :

$$\lim_{n\to\infty} f(x^n) = m .$$

γ) On montre que si S est vide ou fini alors $\lim_{n\to\infty} f(x^n) = m$.

A partir d'un certain rang, on a :

$$f(x^{n+1}) \leq f(x^n) .$$

Donc :

$$\lim_{n\to\infty} f(x^n) = f^* \geq m .$$

$$\lim_{n\to\infty} f(x^{n+1}) - f(x^n) = 0$$

et grâce à (4) :

$$\lim_{n\to\infty} \|x^{n+1} - x^n\| = 0 .$$

On termine la démonstration comme plus haut.

C.Q.F.D.

(2.5.2.) Méthodes de type relaxation pour minimiser une fonctionnelle sur un espace produit.

Voir (4.7.3.) et (4.7.4.).

DEUXIEME PARTIE

CAS D'UN PROBLEME CONVEXE OÙ LA FONCTIONNELLE À MINIMISER EST UNIFORMEMENT CONVEXE.

En 1969, EAVES-SANGWILL [5] et LAURENT-MARTINET [12] remarquent que si dans un problème d'optimisation la fonctionnelle f est uniformément convexe alors on peut utiliser des méthodes où à chaque itération il n'est pas nécessaire de conserver dans φ_n les paramètres non actifs (ceci a été fait partiellemnt dans le paragraphe (2.4.)). Il s'agit d'un avantage pratique très appréciable. On peut aussi considérer de nouvelles méthodes quant aux choix de φ_{n+1}. C'est le cas de l'"éclatement des contraintes" trouvée par HAUGAZEAU [8] et LIONS-TEMAM [15].

(2.6.) ALGORITHME.

(2.6.1.) Hypothèses et notations.

Nous ferons maintenant les hypothèses suivantes :

H5 : E est un espace de Banach. X un sous ensemble convexe de E.

H6 : Pour tout $v \in V$ la fonctionnelle $g(v,.)$ définie sur X est convexe.

H7 : La famille $g(v,.)$ est équi uniformément continue c'est-à-dire que :

pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que pour tout $x,y \in X$: $\|x-y\| \leq \eta$ on a :

$$|g(v,x) - g(v,y)| < \varepsilon \text{ , pour tout } v \in V .$$

Par exemple les fonctionnelles $g(v,.)$ peuvent être uniformément lipchitziennes.

H8 : La fonctionnelle f est uniformément convexe sur X c'est-à-dire qu'elle est convexe et que :

"Il existe une fonction $\delta : \mathbb{R}^+ \to \mathbb{R}^+$, vérifiant $\delta(t) \to 0$ si et seulement si $t \to 0$ telle que pour tout $x,y \in X$ on ait :

$$f(\frac{x+y}{2}) \leq \frac{1}{2}(f(x) + f(y)) - \delta(\|x-y\|)" .$$

Voici deux exemples où est vérifiée cette hypothèse :

a) $f(x) = a(x) + (Ax,x)$,

où $x \in H$ un espace de Hilbert
A est une application linéaire continue auto-adjointe définie positive .
a est une fonctionnelle convexe.

b) f est strictement convexe et continue sur l'ensemble convexe et compact X . (Voir (3.4.)).

H4 : Si on définit la fonction h de $\mathbb{R}$ dans $\bar{\mathbb{R}}$ par :

$$h(u) = \inf (f(x) | x \in X, g(v,x) \leq u \quad \forall v \in V).$$

On suppose que : (voir (2.3.1.)) :

$m = h(0)$ est fini,
il existe $\varepsilon > 0$ tel que $h(\varepsilon)$ soit fini,
la fonction h est s.c.i. en zéro.

(2.6.2.) Description de l'algorithme.

Il s'agit essentiellement de l'algorithme décrit en (2.2.) (nous conservons les mêmes notations) où on dispose d'un choix plus large de fonctions φ_{n+1} .

Soit x^n et $v^n \in V$ définis comme en (2.2.) et :

$$q_{n+1} = f_n + \ell\,(g(v^n,.))$$

où ℓ est une fonction d'influence convexe.

On détermine $x^{n+1} \in X$ tel que :

$$q_{n+1}(x^{n+1}) \leq \inf\,(q_{n+1}(x) \mid x \in X) + \eta_{n+1}$$

puis la fonctionnelle φ_{n+1} convexe et s.c.i. telle que :

$$\begin{cases} \varphi_{n+1} : X \to [0,+\infty] \\ \varphi_{n+1}(x) = 0 \qquad \forall\, x \in C \\ q_{n+1}(x^{n+1}) \leq f_{n+1}(x^{n+1}) \leq \inf\,(f_{n+1}(x) \mid x \quad X) + \eta_{n+1} \end{cases} .$$

Nous préciserons plus loin divers choix de φ_{n+1} , nous contentant de remarquer que les relations ci-dessus permettent un très grand choix de telles fonctionnelles.

(2.6.3.) Propriétés de l'algorithme.

On a le théorème suivant :

THEOREME :

Toute suite générée par l'algorithme du (2.6.2.) a, compte tenu des hypothèses du (2.6.1.), les propriétés suivantes :

α) $\limsup_{n \to \infty} g(v,x^n) \leq 0 \qquad \forall\, v \in V$

β) $\lim_{n\to\infty} f_n(x^n) = \lim_{n\to\infty} f(x^n) = m$.

Démonstration :

Celle-ci est assez longue et nous l'avons divisée en plusieurs lemmes. Nous poserons dans toute la suite :

$$\begin{cases} S = \{n \mid f_{n+1}(x^{n+1}) < f_n(x^n)\} \\ a_n = \inf (f_n(x) \mid x \in X) \end{cases}$$

LEMME a :

Si le cardinal de S est infini on a :

$$\lim_{\substack{n\to\infty \\ n\in S}} \|x^{n+1}-x^n\| = 0 .$$

Démonstration :

Par construction on a pour tout n :

$$f_n(x^n) \leq a_n + \eta_n \leq f_n(x^{n+1}) + \eta_n \leq q_{n+1}(x^{n+1}) + \eta_n \leq f_{n+1}(x^{n+1}) + \eta_n$$

donc :

$$f_n(x^n) \leq f_{n+1}(x^{n+1}) + \eta_n$$

et par conséquent, par définition de S il vient :

$$\lim_{\substack{n\to\infty \\ n\in S}} (f_n(x^n) - f_{n+1}(x^{n+1})) = 0 \qquad (1)$$

Grâce à l'uniforme convexité de f (H8) et à la convexité de φ_n on a :

$$\forall\, x,y \in X,\ \forall\, n \qquad f_n\left(\frac{x+y}{2}\right) \leq \frac{1}{2}(f_n(x) + f_n(y)) - \delta(\|x-y\|)$$

Donc :

$$f_n\left(\frac{x^{n+1}+x^n}{2}\right) \leq \frac{1}{2}(f_n(x^n) + f_n(x^{n+1})) - \delta(\|x^{n+1}-x^n\|) .$$

Soit par construction :

$$f_n(x^n) - \eta_n \leq f_n\left(\frac{x^{n+1}+x^n}{2}\right) \leq f_{n+1}(x^{n+1}) + \frac{\eta_n}{2} - \delta\left(\|x^{n+1}-x^n\|\right)$$

et

$$\delta(\|x^{n+1}-x^n\|) \leq f_{n+1}(x^{n+1}) - f_n(x^n) + \frac{3\eta_n}{2}$$

ce qui prouve le lemme grâce à (1) et aux propriétés de δ .

C.Q.F.D.

LEMME b :

Si le cardinal de S est infini on a :

$$\begin{cases} \limsup_{\substack{n\to\infty \\ n\in S}} g(v,x^n) \leq 0 \qquad \forall\, v \in V \\ \lim_{\substack{n\to\infty \\ n\in S}} f(x^n) = \lim_{\substack{n\to\infty \\ n\in S}} f_n(x^n) = m \ . \end{cases}$$

Démonstration :

Par construction on a :

$$f_{n+1}(x^{n+1}) \geq f_n(x^{n+1}) + \ell(g(v^n,x^{n+1})) \geq f_n(x^n) + \ell(g(v^n,x^{n+1})) - \eta_n$$

Grâce à (1), il vient :

$$(2) \qquad \limsup_{\substack{n\to\infty \\ n\in S}} g(v^n,x^{n+1}) \leq 0 \ .$$

On déduit du lemme a et de l'hypothèse H7 que :

$$(3) \qquad \lim_{\substack{n\to\infty \\ n\in S}} (g(v^n,x^n) - g(v^n,x^{n+1})) = 0 \ .$$

Le choix de v^n implique que :

$$g(v^n,x^n) > 0 .$$

Donc il vient de (2) et (3) que :

$$\lim_{\substack{n\to\infty \\ n\in S}} g(v^n,x^n) = 0 .$$

D'où (choix de v^n) :

$$\text{(4)} \qquad \limsup_{\substack{n\to\infty \\ n\in S}} g(v,x^n) \leq 0 \qquad \forall\, v \in V$$

Posons :

$$\varepsilon_n = \sup_{v\in V} g(v,x^n) > 0 .$$

On sait grâce à (4) que :

$$\lim_{\substack{n\to\infty \\ n\in S}} \varepsilon_n = 0 .$$

On a :

$$\text{(5)} \qquad f(x^n) \geq \inf \left(f(x) \,\middle|\, x \in X,\ g(v,x) \leq \varepsilon_n \quad \forall\, v \in V\right) = h(\varepsilon_n) .$$

D'où par passage à la limite grâce à H4 :

$$\text{(6)} \qquad \liminf_{\substack{n\to\infty \\ n\in s}} f(x^n) \geq m .$$

Or, par construction on a :

$$f(x^n) \leq f_n(x^n) \leq m + \eta_n$$

Donc :

$$\limsup f(x^n) \leq \limsup f_n(x^n) \leq m \ .$$

D'où on tire grâce à (6) que :

$$\lim_{\substack{n\to\infty \\ n\notin S}} f(x^n) = \lim_{\substack{n\to\infty \\ n\notin S}} f_n(x^n) = m \ .$$

C.Q.F.D.

LEMME c :

Si le cardinal de S est infini alors :

$$\begin{cases} \lim_{n\to\infty} f(x^n) = \lim_{n\to\infty} f_n(x^n) = m \ . \\ \\ \limsup g(v,x^n) \leq 0 \qquad \forall\, v \in V \ . \end{cases}$$

Démonstration :

Si le complémentaire de S est vide ou fini la conclusion se déduit immédiatement du LEMME b . Supposons donc que le complémentaire de S ait un cardinal infini.

A tout $n \notin S$ assez grand on peut faire correspondre les indices $i(n)$ et $j(n)$ de la façon suivante :

$j(n) \in S$: $j(n) = \min (\ell \mid \ell > n \ , \ \ell \in S)$

$i(n) \in S$: $i(n) = \max (\ell \mid \ell < n \ , \ \ell \in S)$

On a par définition de S :

$$(7) \qquad f_{i(n)}(x^{i(n)}) - \eta_{i(n)} \leq f_n(x^n) \leq f_{j(n)}(x^{j(n)}) \ .$$

Or (LEMME b) :

$$\lim_{n\to\infty} f_{i(n)}(x^{i(n)}) = \lim_{n\to\infty} f_{j(n)}(x^{j(n)}) = m.$$

D'où grâce à (7) :

$$\lim_{n\to\infty} f_n(x^n) = m \ .$$

On termine facilement la démonstration comme dans les lemmes a et b.

C.Q.F.D.

LEMME d :

Si l'ensemble S est vide ou fini alors :

$$\begin{cases} \lim_{n\to\infty} f(x^n) = \lim_{n\to\infty} f_n(x^n) = m \\ \\ \limsup g(v,x^n) \leq 0 \qquad \forall\, v \in V \ . \end{cases}$$

Démonstration :

A partir d'un n assez grand on a :

$$f_{n+1}(x^{n+1}) \geq f_n(x^n) \ .$$

Donc :

$$\lim_{n\to\infty} f_n(x^n) = f^* \leq m \ .$$

On termine la démonstration comme dans les lemmes a et b.

C.Q.F.D.

Le théorème est démontré. Les hypothèses utilisées n'impliquent pas l'existence d'un minimum. Le corollaire qui suit précise le résultat grâce à de nouvelles hypothèses.

COROLLAIRE :

Si en plus des hypothèses du (2.6.1.) on suppose que :

- E est un espace de Banach réflexif
- X est un ensemble fermé et pour tout $a \in \mathbb{R}$ l'ensemble
$$\{x \in X \mid f(x) \leq a\}$$
est borné.
- la fonctionnelle f est s.c.i.
- la fonctionnelle g(v,.) est s.c.i. pour tout $v \in V$

alors :

Il existe un élément $\bar{x}$ unique tel que :

$$\bar{x} \in C \ , \ f(\bar{x}) = \min (f(x) \mid x \in C)$$

et toute suite $\{x^n\}$ générée par l'algorithme du (2.6.2.) converge fortement vers $\bar{x}$;

Démonstration :

Pour tout $z \in C$, l'ensemble $\{x \in C \mid f(x) \leq f(z)\}$ est convexe, fermé et borné donc faiblement compact. Par conséquent la fonctionnelle f qui est faiblement s.c.i. atteint son minimum sur C . L'unicité de celui-ci est une conséquence immédiate de la stricte convexité de f .

Grâce au théorème précédent on a:

$$\lim_{n \to \infty} f_n(x^n) = f_n(\bar{x}) = f(\bar{x}) = m \ .$$

Par définition de x^n , du fait que X est convexe et que la fonctionnelle f est uniformément convexe, on a :

$$f_n(x^n) - \eta_n \leq f_n\left(\frac{\bar{x}+x^n}{2}\right) \leq \frac{1}{2}\left(f_n(x^n) + f_n(\bar{x})\right) - \delta(\|x^n - \bar{x}\|)$$

Donc :

$$\delta(\|x^n-\bar{x}\|) \leq \eta_n + \frac{1}{2}(f_n(\bar{x}) - f_n(x^n))$$

ce qui montre que :

$$\lim_{n\to\infty} \|x^n-\bar{x}\| = 0 .$$

C.Q.F.D.

(2.6.4.) Choix des fonctionnelles φ_n et ℓ.

Nous donnons maintenant des choix pratiques de φ_n et ℓ qui sont compatibles avec les conditions que nous avons rencontrées en (2.6.2.).
Pour simplifier, nous supposerons dans ce paragraphe que :

a) Les hypothèses du corollaire précédent sont vérifiées ; celles-ci assurent l'existence d'une solution unique pour le problème P .

b) Dans l'algorithme considéré, à chaque itération on résoud les problèmes principaux exactement : plus précisément à l'itération n on détermine l'élément x^n unique tel que (on prend $\eta_n = 0$) :

$$x^n \in X,\ f_n(x^n) = \min (f_n(x) \mid x \in X) .$$

En ce qui concerne le choix des fonctionnelles φ_n et ℓ nous envisageons deux classes.

(2.6.4.1.) Les fonctionnelles φ_n et ℓ sont des fonctionnelles indicatrices.

On retrouve alors les méthodes de coupe (cut-off methods) sous leur forme usuelle. (Voir LAURENT-MARTINET [12]).
On choisit ℓ tel que :

$$\ell(u) = \begin{cases} 0 & \text{si} \quad u \leq 0 \\ +\infty & \text{si} \quad u > 0 \end{cases}$$

et

$\varphi_n = \chi_{K_n}$ (fonction indicatrice de l'ensemble convexe fermé K_n

i.e. $\varphi_n(x) = \begin{cases} 0 & \text{si} \quad x \notin K_n \\ +\infty & \text{sinon.} \end{cases}$) .

A l'itération n , on connait le convexe fermé K_n et l'élément x^n tels que :

$$(1) \quad \begin{cases} X \supseteq K_n \supseteq C \\ x^n \in K_n \\ f(x^n) = \min(f(x) \mid x \in K_n) \end{cases} .$$

On choisit $v^n \in V$ (si $x^n \in C$ alors $x^n = \bar{x}$ et le processus est terminé) comme dans l'algorithme général ce qui permet de déterminer l'élément x^{n+1} unique tel que :

$$(2) \quad x^{n+1} \in K_n,\ g(v^n, x^{n+1}) \leq 0 \ ,\ f(x^{n+1}) = \min\ (f(x) \mid x \in K_n,\ g(v^n, x) \leq 0)$$

Puis on détermine l'ensemble K_{n+1} convexe, fermé tel que :

$$\begin{cases} X \supseteq K_{n+1} \supseteq C \\ x^{n+1} \in K_{n+1} \\ f(x^{n+1}) = \min\ (f(x) \mid x \in K_{n+1}) \end{cases} .$$

Ces conditions nous donnent de multiples K_{n+1} acceptables :

K_{n+1} est un ensemble convexe fermé tel que :

$$\text{conv}\,(\{x^{n+1}\} \cup C) \subseteq K_{n+1} \subseteq \{x \in X \mid f'(x^{n+1} ; x-x^{n+1}) > 0\}$$

où :

- conv A désigne l'enveloppe convexe de l'ensemble A .
- $f'(x;y) = \lim_{\substack{\lambda \to 0 \\ \lambda > 0}} (f(x+\lambda y)-f(x))/\lambda$ est la dérivée directionnelle de f en x dans la direction y .

Voici quelques exemples de choix des K_n :

a) Accumulation des contraintes

$$K_{n+1} = K_n \cap \{x \mid g(v^n,x) \leq 0\} \quad .$$

Ce procédé a déjà été signalé en (2.2.3.) a).
Il pourra être avantageusement remplacé par le suivant.:

b) Accumulation des contraintes avec élimination des contraintes non actives.

Cf. EAVES-ZANGWILL [5] , TOPKIS [19] , LAURENT-MARTINET [12] .

Posons :

$$K_n = \{x \in X \mid g(v,s) \leq 0 \quad \forall\, v \in V_n\} \quad .$$

On choisit pour définir K_{n+1} l'ensemble V_{n+1} tel que :

$$V_{n+1} = \{v^n\} \cup \{v \in V_n \mid g(v,x^{n+1}) = 0\} \quad .$$

c) "Eclatement des contraintes".

Cf. LAURENT-MARTINET [12] , HAUGAZEAU [8] .

On suppose de plus ici que la fonctionnelle f possède un sous différentiel ∂f (voir (*)) non vide pour tout $x \in X$.

(*) On rappelle que le sous différentiel $\partial f(x)$ de la fonctionnelle f en x est défini par :

$$\partial f(x) = \{z \in E \mid f(y) \geq f(x) + \langle z, y-x \rangle \quad \forall\, y \in E\} \quad .$$

On choisit :

$$K_n = \{x \in X \mid \langle c^n, x-x^n \rangle \geq 0\}$$

où c^n est un élément de l'ensemble $\partial f(x^n)$: sous différentiel de f en x^n .

On détermine comme plus haut v^n et x^{n+1} et on choisit alors K_{n+1} tel que :

$$K_{n+1} = \{x \in X \mid \langle c^{n+1}, x-x^{n+1} \rangle \geq 0\}$$

où $c^{n+1} \in \partial f(x^{n+1})$.

Remarque :

Dans le cas où la fonctionnelle f possède une dérivée de Frechet ∇f pour tout $x \in X$ on sait que :

$$\partial f(x^n) = \{\nabla f(x^n)\} .$$

L'ensemble K_n est donc l'intersection d'un demi espace fermé avec l'ensemble X , ce qui donne une interprétation géométrique simple du procédé.

(2.6.4.2.) Les fonctionnelles φ_n et ℓ prennent des valeurs finies.

On obtient dans ce cas des généralisations de la méthode habituelle de pénalisation extérieure.

On prend par exemple :

$$\ell(u) = \begin{cases} 0 & \text{si} \quad u \leq 0 \\ k\, u^2 & \text{sinon} \end{cases}$$

(le nombre réel $k > 0$ est donné).

a) Méthode d'accumulation des contraintes (voir (2.2.3.) b).

A partir d'une fonctionnelle initiale donnée φ_1 (on prend souvent $\varphi_1 \equiv 0$) on fait à chaque itération :

$$\varphi_{n+1} = \varphi_n + \ell(g(v^n,.)).$$

b) Accumulation des contraintes avec élimination des contraintes non actives.

On a :

$$\varphi_n = \varphi_1 + \sum_{m \in L_n} \ell(g(v^m,.))$$

où $L_n \subseteq \{2,\ldots,(n-1)\}$.

On détermine $x^{n+1} \in X$ solution unique de :

$$\min (f_n(x) + \ell(g(v^n,x)) \mid x \in X) .$$

On prendra alors pour définir φ_{n+1} l'ensemble L_{n+1} défini par :

$$L_{n+1} = \{k \in L_n \cup \{n\} \mid g(v^k,x^{n+1}) = 0\} .$$

c) Un procédé de type "éclatement des contraintes".

On choisit pour tout n la fonctionnelle φ_n définie par :

$$\varphi_n(x) = (a_n + \langle b^n, x-x^n \rangle)^+ .$$

où :

$$\begin{cases} a_n \in \mathbb{R} , \ b^n \in E' \text{ l'espace dual de } E \\ \langle , \rangle \text{ est la dualité entre } E' \text{ et } E . \end{cases}$$

L'élément x^{n+1} est déterminé par :

$$x^{n+1} \in X , \ f_n(x^{n+1}) + \ell(g(v^n,x^{n+1})) = \min (f_n(x) + \ell(g(v^n,x)) \mid x \in X)$$

Il s'agit d'un problème d'optimisation qui, par définition de φ_n, s'applique à une fonctionnelle non différentiable. On peut le résoudre aisément en se ramenant à un problème d'pptimisation avec contraintes.

On définit ensuite la fonctionnelle φ_{n+1} par :

$$\varphi_{n+1}(x) = (a_{n+1} + < b^{n+1}, x-x^{n+1} >)^+$$

où :

$$\begin{cases} a_{n+1} = \ell(g(v^n, x^{n+1})) + (a_n + < b^n, x^{n+1}-x^n >)^+ \\ b^{n+1} \in \partial f(x^{n+1}) . \end{cases}$$

Si nous supposons qu'il existe un élément z de X où la fonctionnelle f est continue, on peut montrer que ce choix de φ_n satisfait aux conditions du (2.6.2.).

(2.7.) APPLICATIONS :

(2.7.1.)

On donne dans ce paragraphe quelques exemples de problèmes qui peuvent être traités par les algorithmes du (2.6.). Ces problèmes consistent à chercher la projection de 0 sur l'ensemble C défini par des inégalités. Plus précisément, on suppose que E est un espace de Hilbert et on cherche l'élément $\bar{x}$ unique tel que :

$$(M) \qquad \bar{x} \in C \ , \quad \|\bar{x}\| = \min (\|x\| \ x \in C)$$

où

$$C = \{x \in X \mid g(v,x) \leq 0 \quad \forall \ v \in V\} .$$

X est une partie convexe fermée de E .

Ce problème est un cas particulier important de ceux qui ont été définis en (2.6.). Nous allons préciser certains procédés du (2.6.) sur quelques cas particuliers.

(2.7.2.) Résolution d'égalités linéaires.

Soit :

$$C = \{x \in X | < \ell_v, x > - C_v = 0 \quad \forall\, v \in V\}$$

où à tout $v \in V$ correspondent les éléments $\ell_v \in E'$ et $C_v \in \mathbb{R}$. Supposons qu'il existe $d \in R$ tel que :

$$\|\ell_v\| \leq d \qquad \forall\, v \in V .$$

On vérifie facilement que dans ce cas, toutes les hypothèses du (2.6.1.) sont vérifiées on peut donc résoudre le problème (M) correspondant par un algorithme de (2.6.).

Nous utiliserons par exemple la méthode d'éclatement des contraintes du (2.6.4.1.). Précisons cet algorithme.
A la $n^{ème}$ itération on connait $x^n \in X$ et

$$K_n = \{x \in X | < x^n, x - x^n > \geq 0\}$$

on choisit $v^n \in V$ tel que * :

$$\begin{cases} |< \ell_{v^n}, x^n > - C_{v^n}| > 0 \\ |< \ell_{v^n}, x^n > - C_{v^n}| \geq \theta \sup_{v \in V} |< \ell_v, x > - C_v)| - \rho_n \end{cases}$$

où sont donnés le nombre $\theta \in]0,1]$ et la suite des nombres non négatifs $\{\rho_n\}$ telle que $\lim_{n \to \infty} \rho_n = 0$.

---

* Si l'ensemble V est fini il suffit que v^n soit tel que :

$$v^n \in V , \quad |< \ell_{v^n}, x^n > - C_{v^n}| > 0 .$$

L'élément x^{n+1} est la solution unique du problème :

$$\min \left(\|x\|^2 \mid x \in X,\ < x^n, x-x^n > \geq 0,\ < \ell_{v_n}, x > - C_{v_n} = 0 \right)$$

lequel problème est relativement simple à résoudre.

Cet algorithme peut être utilisé dans le cas d'un système linéaire dans un espace de dimension finie. Il conduit alors à un procédé original ayant quelque analogie avec la méthode de KACMARZ [9] et qui converge vers la solution de plus petite norme du système si celui-ci a au moins une solution. Ce procédé semble intéressant pour les systèmes mal conditionnés.

(2.7.3.) Résolution approchée d'un système linéaire.

Soit :

$$(2) \qquad C = \{x \in X \mid \|A(v)x - b(v)\| - \varepsilon \leq 0 \quad \forall v \in V\}$$

où on donne :

le nombre $\varepsilon > 0$
l'application b de V dans l'espace de Banach F
les applications $A(v)$ linéaires et uniformément bornées de X dans F
$(\exists D \in \mathbb{R}, \forall v \in V, \forall x \in X \quad \|A(v)x\| \leq D\|x\|)$.

Ceci entraine grâce à l'inégalité triangulaire :

: $\forall x,y \in X \quad \forall v \in V$:

$$\left| \|A(v)x - b(v)\| - \|A(v)y - b(v)\| \right| \leq D \|x-y\| .$$

On peut donc appliquer les algorithmes du (2.6.) à la résolution du problème (M) correspondant. Cette technique nous semble intéressante pour certains problèmes (équations intégrales de première espèce) où l'ensemble C défini par (2) est non vide pour tout $\varepsilon > 0$ mais peut être vide si $\varepsilon = 0$, elle s'apparente aux procédés habituels de régularisation.

(2.7.4.) Recherche d'un col d'une fonction convexe-concave.

Cf. MARTINET [17] .

Soit L une fonctionnelle définie sur $X \times Y$.

On suppose que :

X , Y sont des ensembles convexes, bornés, fermés non vides des espaces de Hilbert $H1$ et $H2$.
Les fonctionnelles $L(.,y)$ sont convexes pour tout $y \in Y$,
et les fonctionnelles $L(x,.)$ sont concaves pour tout $x \in S$.
La fonctionnelle L est uniformément continue sur $X \times Y$.

On définit l'ensemble M des cols de L par :

$$M = \{(\bar{x},\bar{y}) \in X \times Y \mid L(x,\bar{y}) \geq L(\bar{x},\bar{y}) \geq L(\bar{x},y) \quad \forall\ (x,y) \in X \times Y\} .$$

Grâce aux hypothèses faites cet ensemble n'est pas vide.
Soit :

$$H = R \times H1 \times H2$$
$z = (a,x,y)$ un élément de H .

L'espace H est un espace de Hilbert si on y définit le produit scalaire :

$$< z_1,z_2 >_H = a_1a_2 + < x_1,x_2 >_{H1} + < y_1,y_2 >_{H2} .$$

Posons :

$$A = \mathbb{R} \times X \times Y \subset H$$

et considérons l'ensemble C non vide :

$$C = \{(\bar{a},\bar{x},\bar{y}) \in A \;\Big|\; \begin{array}{l} L(\bar{x},y) \leq \bar{a} \quad \forall\ y \in Y \\ -L(x,\bar{y}) \leq -\bar{a} \quad \forall\ x \in X \end{array} \} .$$

SERVICE POLYCOPIE
MATHÉMATIQUES APPLIQUÉES
Université de GRENOBLE

Nous nous proposons de trouver la solution du problème :

$$(3) \qquad \min\ (|\bar{a}|^2 + \|\bar{x}\|^2 + \|\bar{y}\|^2 \mid (\bar{a},\bar{x},\bar{y}) \in C) .$$

On peut vérifier que les hypothèses du (2.6.1.) sont satisfaites.

Précisons comme exemple des méthodes du (2.6.) la méthodes d'éclatement des contraintes du (2.6.4.1.) pour résoudre (3).

On construit une suite $\{a^n,x^n,y^n\}$ de solutions approchées du problème (3). Montrons comment passer de (a^n,x^n,y^n) à $(a^{n+1},x^{n+1},y^{n+1})$.

On détermine (c'est la façon la plus simple du point de vue théorique) :

$$p_n = \max\ (L(x^n,y)-a^n \mid y \in Y) \qquad \text{optimum atteint en } \tilde{y}^n$$

$$q_n = \max\ (-L(x,y^n)+a^n \mid x \in X) \qquad \text{optimum atteint en } \tilde{x}^n$$

(les hypothèses faites montrent que ces optimums sont atteints).
On considère alors deux cas :

a) $\max\ (p_n,q_n) \leq 0$

Alors (a^n,x^n,y^n) est solution de (3). On arrête le procédé.

b) $\max\ (p_n,q_n) > 0$

$$\text{Soit } h_n(a,x,y) = \begin{cases} L(x,y^n)-a & \text{si } p_n \geq q_n \\ -L(x^n,y)+a & \text{si } p_n < q_n \end{cases}$$

On détermine le triplet $(a^{n+1},x^{n+1},y^{n+1})$ solution unique du problème :

$$\min\left(|a|^2+\|x\|^2+\|y\|^2 \;\middle|\; \begin{array}{l} a^n(a-a^n)+\langle x^n,x-x^n\rangle+\langle y^n,y-y^n\rangle \geq 0 \\ h_n(a,x,y) \leq 0 \\ (a,x,y) \in A \end{array}\right)$$

Ce dernier problème a deux contraintes inégalités ; la première est linéaire, la seconde est toujours active. Il sera opportun d'utiliser cette méthode si ce dernier problème est assez facile à résoudre.

Remarque :

Si les ensembles X et Y ne sont pas bornés et que des hypothèses supplémentaires assurent l'existence d'un point de selle, la méthode précédente n'est plus définie, on peut cependant la modifier pour l'adapter à ce cas.

BIBLIOGRAPHIE

[1] AUSLENDER A.
Méthodes numériques pour la résolution de problèmes d'optimisation avec contraintes.
Thèse, faculté des Sciences de GRENOBLE, (1970).

[2] CEA, J.
Optimisation. Théorie et Applications.
Dunod, (1971).

[3] CHENEY - GOLDSTEIN.
Newton's method for convex programming and Tchebycheff approximation.
Numerishe mathematik, vol.1 (1959) 253-268.

[4] DANTZIG C.B., FUKERSON, D.R. and JOHNSON, S.
Solution of a large scale travelling salesman problem.
Operations research. vol.2 n°4 (1954) 393-410.

[5] EAVES, C. et ZANGWILL, W.I.
Generalized cutting plane algorithms.
Working paper n° 274.
Center for research in management science.
University of California. Berkeley. July (1969)

[6] FIACCO, A.V.
Doctoral dissertation.
Northwertern university. Evanston. Illinois. Juin 1967.

[7] GEOFFRION, A.M.
Relaxation and the dual method in mathematical programming. Western management science institute.
Working paper n° 135.
University of California. Los Angeles. March 1968.

[8] HAUGAZEAU, Y.
Sur la minimisation de formes quadratiques avec contraintes.
C.R.A.S. Paris 264 (1967) 322-324.

[9] KACMARZ, S.
Bull. Insti. Acad. polonaise des sciences A (1937) 355-357.

[10] KAPLAN, A.A.
Determination of the extremum of a linear function on a convex set.
Soviet. Math. 9 (1968) 269-271.

[11] KELLEY, J.E.
The cutting plane method for solving convex programs.
J. SIAM 6 (1958) 15-22.

[12] LAURENT, P.J. et MARTINET, B.
Méthodes duales pour le calcul du minimum d'une fonction convexe sur une intersection de convexes.
Colloque d'optimisation de Nice. Juin 1969.

[13] LAURENT, P.J.
Approximation et Optimisation.
Hermann, (1972).

[14] LEVITIN, E.S. et POLYAK, B.T.
Constrained minimization problems.
U.S.S.R. comp. math. and Phy. math. 5 (1966) 1-50.

[15] LIONS, J.L. et TEMAM, R.
Une méthode d'éclatement des contraintes en calcul des variations.
C.R.A.S. Paris 263 (1966) 563-565.

[16] MARTINET, B.
La méthode d'accumulation des contraintes pour la résolution approchée de problèmes avec contraintes.
Colloque d'analyse numérique d'Anglet. Juin 1971.

[17] MARTINET, B.
Une méthodes générale pour la résolutinn des jeux.
Colloque d'analyse numérique de Super-Besse.
Juin 1970.

[18] POLAK, E.
Computational methods in optimization.
Academic press. New-York and London. 1971.

[19] TOPKIS, D.M.
Cutting plane methods without nested constrained sets.
University of California. Berkeley. June 1969.

[20] ZANGWILL, W.I.
Non linear programming. A unified approach.
Prentice Hall. Englewood cliffs.
New-Jersey (1969).

CHAPITRE III

METHODES DE RESOLUTION DE PROBLEMES MINIMAX. APPLICATIONS A L'OPTIMISATION.

(3.1.) INTRODUCTION.

On étudie des méthodes de résolution de problèmes minimax pour les appliquer à des problèmes d'optimisation. Les procédés considérés ne font pas jouer un rôle similaire aux deux variables. Les hypothèses introduites sont les mêmes que celles qu'on utilise dans les théorèmes d'existence.

Dans une première partie nous considérons des méthodes générales pour résoudre un problème de minimax.

Plus précisément, on se donne un triplet (X,Y,L) où :

- X,Y sont des ensembles convexes fermés de $\mathbb{R}^n$ et $\mathbb{R}^m$.
- L est une fonction : $X \times Y \to \mathbb{R}$ convexe-concave* et continue sur $X \times Y$.
- La fonction L possède un point de selle sur $X \times Y$.

Ces hypothèses ne sont données ici qu'à titre indicatif (on pourra considérer parfois le cas où on n'a pas d'hypothèses de convexité et où il n'existe pas de point de selle).
Pour simplifier l'exposé, nous nous sommes limités au cas où les espaces sont de dimension finie.

Au triplet (X,Y,L) on fait correspondre plusieurs "problèmes" :

α) Chercher la valeur du triplet (X,Y,L) c'est-à-dire le nombre v tel que :

$$v = \inf_{x\in X} \sup_{y\in Y} L(x,y) = \sup_{y\in Y} \inf_{x\in X} L(x,y).$$

---

* On dit que L est convexe-concave si $L(.,y)$ est convexe pour tout $y \in Y$ et si $L(x,.)$ est concave pour tout $x \in X$.

β) Problème P : trouver un élément $\tilde{x} \in X$ tel que :

$$\sup_{y \in Y} L(x,y) \geq \sup_{y \in Y} L(\tilde{x},y) \quad , \quad \forall\, x \in X \ .$$

Sous certaines conditions il existe $\tilde{y} \in Y$ tel que ;

$$L(\tilde{x},\tilde{y}) = \sup_{y \in Y} L(\tilde{x},y) \ .$$

Définition :

Par abus de langage, on appellera indifféremment solution du problème P relatif au triplet (X,Y,L) l'élément $\tilde{x}$ ou le couple $(\tilde{x},\tilde{y})$.

On remarquera qu'en général $(\tilde{x},\tilde{y})$ n'est pas nécessairement point de selle de L , ce couple peut d'ailleurs exister sans qu'il existe un point de selle.

γ) Problème Q : Il s'agit de l'analogue du problème P quand on intervertit les variables x et y .

Trouver $\hat{y} \in Y$ tel que :

$$\inf_{x \in X} L(x,\hat{y}) \geq \inf_{x \in X} L(x,y) \qquad \forall\, y \in Y \ .$$

Sous certaines conditions il existe $\hat{x} \in X$ tel que :

$$L(\hat{x},\hat{y}) = \inf_{x \in X} L(x,\hat{y}) \ .$$

Définition :

Par abus de langage on appellera indifféremment solution du problème Q relatif au triplet (X,Y,L) l'élément $\hat{y}$ ou le couple $(\hat{x},\hat{y})$.

δ) Problème S : Recherche du point de selle relatif à (X,Y,L) :

Trouver $(\bar{x},\bar{y}) \in X \times Y$ tel que :

$$L(\bar{x},y) \leq L(\bar{x},\bar{y}) \leq L(x,\bar{y}) \quad , \quad \forall\ (x,y) \in X \times Y \ .$$

Nous établirons plus loin des relations importantes entre les solutions de ces différents problèmes.

Voici le principe des algorithmes étudiés ici.
Etant donné un triplet (X,Y,L) on considère un procédé itératif pour lequel l'itération n comprend :

un problème "principal" qui consiste à résoudre un problème (par exemple le problème Q) associé au triplet (X_n,Y,L) où X_n est une partie de X qui, en ce qui concerne la mise en oeuvre du procédé, est plus simple que X . Ce problème principal peut être considéré comme une approximation du problème initial.

On choisit ensuite (sous-problème) , compte tenu des résultats trouvés au problème principal, un ensemble X_{n+1} contenu dans X et on passe à l'itération suivante.

On a donc un procédé adaptatif qui rappelle les procédés des chapitres précédents. En fait nous verrons qu'on généralise ici les résultats du chapitre II.

Ces méthodes sont souvent des méthodes de descente pour la fonction f définie par :

$$f(x) = \sup_{y \in Y} L(x,y) \quad , \quad \forall\ x \in X \ .$$

Elles ne sont pas symétriques par rapport aux variables x et y . Elles diffèrent donc essentiellement des méthodes symétriques décrites par AUSLENDER [2] , FALK [10] , LEMAIRE [23] chapitre 1, et des procédés du (2.7.4.).

On peut alors construire autour de ce principe général de très nombreuses variantes, compte tenu des remarques suivantes :

a) Si on considère des procédés symétriques des précédents en échangeant les rôles de x et y, on définit alors des méthodes qui dans de nombreux cas sont des procédés de montée pour la fonction g définie par :

$$g(y) = \inf_{x \in X} L(x,y) \quad , \quad \forall\, y \in Y \ .$$

Cette remarque évidente est très importante pour l'application de ces procédés aux problèmes d'optimisation. Compte tenu de la disymétrie des problèmes (P) et (Q), on obtient des algorithmes pratiques complètement différents.

b) Choix des ensembles X_n .

On considère les deux cas suivants :

- Les ensembles X_n (pas forcément convexes) sont obtenus par "accumulation" des points $\overrightarrow{x^i}$ trouvés dans les sous problèmes des itérations précédentes c'est-à-dire qu'on a :

$$X \supseteq X_n \supseteq \{\overrightarrow{x^1},\ldots,\overrightarrow{x^{n-1}}\} \quad , \text{ pour tout } n \ .$$

On retrouve l'idée des procédés du chapitre I et de la première partie du chapitre II. On peut concevoir des variantes où on élimine une partie des points $\overrightarrow{x^i}$ antérieurs .

- Cette technique peut se révéler d'une mise en oeuvre trop lourde d'où l'intérêt de considérer des procédés où les ensembles X_n sont plus "simples" (segment, demi-droite, triangle, etc...). C'est en particulier le cas si la fonction $L(x,.)$ est strictement concave, on obtient alors des algorithmes de mise en oeuvre aisée. (On généralise ainsi les procédés de la deuxième partie du chapitre II). Dans le cas général, des choix d'ensembles X_n plus simples restent possible soit en se ramenant au cas précédent par une perturbation convenable de la fonction L soit plus directement par extension de techniques proposées par LEVITIN [24] pour la minimisation de fonctions non dérivables.

c) Nature des problèmes.

On peut par exemple dans le problème principal résoudre relativement au triplet (X_n, Y, L) , soit le problème P , soit le problème Q , soit encore le problème S .

Le cadre d'hypothèses n'est forcément identique et les algorithmes n'ont pas les mêmes propriétés.

d) Les ensembles X et Y ne sont pas bornés.

Il est relativement facile de définir ces procédés si les ensembles X et Y sont bornés. Cependant, les applications pratiques (cas où L est un lagrangien d'un problème d'optimisation) montrent l'intérêt considérable qu'on a à supprimer cette hypothèse en la remplaçant par des conditions plus faibles. Mais, dans ce cas, les algorithmes doivent être modifiés en conséquence.

Dans une deuxième partie, on considère quelques exemples d'application de ces méthodes générales à des problèmes d'optimisation. Le sujet étant très vaste, on se limite à quelques exemples caractéristiques ; on n'examine ni la mise en oeuvre pratique de ces méthodes, ni leur application à la décomposition et la décentralisation qui sont des motivations majeures en ce domaine. On retrouve cependant de nombreuses méthodes d'optimisation classiques.

Le premier exemple considéré est la minimisation d'une fonction convexe et continue sur un convexe fermé X . On fait correspondre à ce problème le lagrangien (cf. ROCKAFELLAR [27]) :

$$L(x,y) = (x,y) - f^*(y) \qquad (*) \quad (**)$$

où f^* est la fonction conjuguée de f .

(*) On n'est pas forcément dans le cadre des hypothèses ci-dessus : dom f^* peut être non fermé et f^* non continue.

(**) On note (,) le produit scalaire dans $\mathbb{R}^n$.

On considère alors le triplet $(X, \text{dom } f^*, L)$. Le problème P relatif à ce triplet est le problème qui consiste à minimiser f sur X .

On peut définir ainsi des méthodes originales de déscente par sous-gradient pour minimiser une fonction convexe, continue non dérivable. Si cette fonction possède une dérivée continue, on retrouve des méthodes de descente de type gradient. En considérant sur le même lagrangien les méthodes "symétriques" des précédentes, on définit d'autres procédés par approximation de l'épigraphe de f .

Le deuxième exemple considéré est le problème suivant :

$$\min (f(x) \mid x \in X \ , \ h(x) \leq 0) \quad ,$$

où les fonctions $h : \mathbb{R}^n \to \mathbb{R}^m$ et $f : \mathbb{R}^n \to \mathbb{R}$ sont convexes et l'ensemble X est convexe et fermé.

On lui fait correspondre le lagrangien L : (cf. ROCKAFELLAR [27]) :

$$L(x,y) = f(x) + (y,h(x)) \qquad (*)$$

défini pour $x \in X$ et $y \in Y = \{y \in \mathbb{R}^m \mid y \geq 0\}$.

Les méthodes de descente correspondantes sont des généralisations de la méthode des colonnes sur le problème d'optimisation initial. (cf. HUARD [18]).

Les méthodes de montée symétriques des précédentes correspondent aux méthodes duales sur le problème initial (AUSLENDER [2] , FALK [10], CEA [8] , CEA-GLOVINSKI-NEDELEC [7] , ARROW-HURWITZ-UZAWA [1] .

Enfin, un lagrangien particulier permet de retrouver les procédés d'optimisation par approximation extérieure de l'ensemble admissible du chapitre II.

(*) On note $(,)$ le produit scalaire dans $\mathbb{R}^m$.

(3.2.) GENERALITES SUR LES PROBLEMES MINIMAX.

(Cf. ROCKAFELLAR [27] , JOLY-LAURENT [19] , STOER-WITZGALL [32] , LEMAIRE [23]) .

Nous rappelons dans ce paragraphe des théorèmes d'existence que nous utiliserons constamment dans la suite.

Soit un triplet (X,Y,L). Il nous arrivera de faire les hypothèses suivantes :

H1 : $L : X \times Y \to \mathbb{R}$ est une fonction convexe-concave.

X et Y sont des ensembles convexes fermés de $\mathbb{R}^n$ et $\mathbb{R}^m$.

L est continue sur $X \times Y$.

Il existe $\tilde{x} \in X$ tel que $\sup_{y \in Y} L(\tilde{x},y) < +\infty$

et il existe $\tilde{y} \in Y$ tel que $\inf_{x \in X} L(x,\tilde{y}) > -\infty$.

On aura aussi à utiliser les hypothèses suivantes :

H2 : Il existe une partie D de Y et un réel a tels que :

l'ensemble $\{x \in X \mid \sup_{y \in D} L(x,y) \leq a\}$ soit non vide et borné.

H3 : Il existe une partie C de X et un réel b tels que :

l'ensemble $\{y \in Y \mid \inf_{x \in C} L(x,y) \geq b\}$ soit non vide et borné.

(3.2.1.) Théorème :

a) Les hypothèses H1 et H2 entrainent que le problème P relatif à (X,Y,L) a un ensemble de solutions qui est une partie compacte et non vide de X et que le triplet (X,Y,L) à une valeur.

b) Les hypothèses H1 et H3 entrainent que le problème Q relatif à (X,Y,L) a un ensemble de solution qui est une partie compacte non vide de Y et que le triplet (X,Y,L) a une valeur.

c) Les hypothèses H1, H2 et H3 entrainent que le triplet (X,Y,L) possède un point de selle.

Démonstration :

Montrons a) . On a :

(1) $$f(.) = \sup_{y \in Y} L(.,y) \geq \sup_{y \in D} L(.,y) .$$

Donc la fonction f est semi-continue inférieurement puisque $L(.,y)$ est continue pour tout y .
De plus on sait grâce à H2 qu'il existe $a \in \mathbb{R}$ tel que :

$$\emptyset \neq \{x \in X \mid f(x) \leq a\} \subseteq \{x \in X \mid \sup_{y \in D} L(x,y) \leq a\} .$$

Comme ce dernier ensemble est borné, on en déduit que l'ensemble des éléments où f atteint son minimum sur X est un compact non vide (solutions du problème P).

La démonstration de l'existence d'une valeur est longue. Nous n'en donnons que les grandes lignes en utilisant des résultats de ROCKAFELLAR [27] .
Posons :

$$F(u,x) = \sup_{y \in Y} L(x,y)-(y,u) \qquad \forall x \in X \quad \forall u \in \mathbb{R}^m$$

$$k(u) = \inf_{x \in X} F(u,x) \qquad \forall u \in \mathbb{R}^m$$

$$f(x) = F(0,x)$$

$$g(y) = \inf_{x \in X} L(x,y) \qquad \forall y \in Y .$$

Le problème P consiste à chercher $\bar{x} \in X$ tel que :

$$f(\bar{x}) = \inf (f(x) \mid x \in X) = \alpha \text{ nombre fini.}$$

Le problème Q consiste à chercher $\bar{y} \in Y$ tel que :

$$g(\bar{y}) = \sup (g(y) \mid y \in Y) = \beta .$$

Evidemment $k(0) = \alpha$. Le théorème 30-4-(i) de ROCKAFELLAR [27] nous permet de déduire des hypothèses H1, que puisque l'ensemble des solutions de P est compact et non vide, la fonction k est semi-continue inférieurement en 0 . Il vient alors du théorème 30-3 de [27] que :

$$\alpha = \beta .$$

Ceci montre l'existence d'une valeur pour le triplet (X,Y,L). On montre b) d'une façon analogue. Le point c) est une conséquence immédiate de a) et b) si on remarque que toute solution $\bar{x}$ du problème P et toute solution $\bar{y}$ du problème Q sont telles que le couple $(\bar{x},\bar{y})$ est point de selle du triplet (X,Y,L).

C.Q.F.D.

Remarque :

Il nous arrivera dans la suite de faire les hypothèses suivantes qui sont des cas particuliers de H2 :

H2' : L'ensemble X est borné

H2" : Il existe un élément ω de Y et un réel a tels que l'ensemble $\{x \in X | L(x,\omega)\} \leq a\}$ est non vide et borné.

H2'" : Il existe un réel a tel que l'ensemble $\{x \in X | f(x) \leq a\}$ est non vide et borné.

Nous considérerons de même les hypothèses H3, H3" et H'" analogues à H3.

(3.2.2.) **Théorème** :

Soit (X,Y,L) un triplet. Si $\tilde{x}$ est une solution du problème P , $\tilde{y}$ une solution du problème Q et si le triplet a une valeur alors $(\tilde{x},\tilde{y})$ est solution du problème S .

(3.2.3.) **Théorème** :

Soit un triplet (X,Y,L) tel que :

$$\begin{cases} L \text{ possède un point de selle,} \\ L(x,.) \text{ est strictement concave pour tout } x \in X , \\ (\tilde{x},\tilde{y}) \text{ est solution du problème P ,} \end{cases}$$

alors $(\tilde{x},\tilde{y})$ est solution du problème S .

Démonstration :

Puisque $(\tilde{x},\tilde{y})$ est solution du problème P et que le triplet (X,Y,L) possède un point de selle, on sait qu'il existe $\bar{y} \in Y$ tel que $(\tilde{x},\bar{y})$ soit point de selle c'est-à-dire que :

$$(1) \qquad L(\tilde{x},\bar{y}) = \max_{y \in Y} L(\tilde{x},y) .$$

Or par définition de $(\tilde{x},\tilde{y})$ on a :

$$L(\tilde{x},\tilde{y}) = \max_{y \in Y} L(\tilde{x},y) .$$

Ce qui compte tenu de (1) et de la stricte concavité de $L(\tilde{x},.)$ prouve que $\tilde{y} = \bar{y}$.

C.Q.F.D.

Remarque :

On a bien entendu un résultat similaire en échangeant les rôles de x et y .

(3.2.4.) Théorème (cas non convexe) :

Soit un triplet (X,Y,L) qui vérifie les hypothèses :

$$\begin{cases} L : X \times Y \to \mathbb{R} \ , \ L(.,y) \text{ est continue pour tout } y \in Y \\ X \text{ est une partie fermée de } \mathbb{R}^n \ . \\ \text{Il existe } \hat{x} \in X \text{ tel que } \sup_{y \in Y} L(\hat{x},y) < \infty \ . \\ \text{Il existe une partie } D \text{ de } Y \text{ telle que pour tout } a \in \mathbb{R} \\ \text{l'ensemble } \{x \in X \mid \sup_{y \in D} L(X,y) \leq a\} \text{ soit borné.} \end{cases}$$

Dans ces conditions le problème P possède au moins une solution x .

Remarque :

Même résultat pour le problème Q avec les hypothèses analogues aux précédentes (intervertir les rôles de x et y).

PREMIERE PARTIE

METHODES DE RESOLUTION DE PROBLEMES MINIMAX

Nous étudions une large classe de procédés itératifs basés sur des principes simples (voir (3.1.)). Ceci nous conduit à faire une synthèse constructive d'un grand nombre de méthodes de descente (et de montée) en optimisation.
Nous avons divisé ces méthodes en trois classes :

a) Les algorithmes de "type accumulation" dont voici le prototype. (On fait les hypothèses H1 et on suppose que X et Y sont bornés).

A l'itération n on connait une partie convexe compacte X_n de X . On cherche (x^n,y^n) solution du problème S relatif à (X_n,Y,L). Soit $a_n = L(x^n,y^n)$.
On détermine ensuite $\overline{x^n}$ tel que :

$$\overline{x^n} \in X \qquad b_n = L(\overline{x^n},y^n) = \min_{x\in X} L(x,y^n) .$$

Si $a_n = b_n$ le processus est arrêté. Sinon on détermine $X_{n+1} = \text{conv}\,(X_n \cup \{\overline{x^n}\})$ et on passe à l'itération suivante.

b) Les algorithmes "sans accumulation" dans le cas où $L(x,.)$ est strictement concave.

Le prototype de ces procédés est l'analogue du précédent où on choisit $X_n = [x^{n-1}, \overline{x^{n-1}}]$.

c) Les algorithmes sans accumulation dans le cas général.

Ces procédés dérivent d'une méthode exposée par LEVITIN [24] pour la minimisation de fonction non dérivable. Malheureusement, ils paraissent d'une mise en oeuvre difficile.

(3.3.) METHODES DE TYPE ACCUMULATION.

(3.3.1.) Cas où les ensembles X et Y sont bornés.

En plus des hypothèses H1, on suppose que les ensembles X et Y sont bornés. Pour beaucoup d'applications en vue, les ensembles X et Y ne sont pas bornés mais il nous a paru intéressant d'étudier d'abord ce cas en raison de sa simplicité. L'affaiblissement de ces hypothèses nous conduira par la suite à compliquer les procédés de base vus ici. On note v la valeur du triplet (X,Y,L).

(3.3.1.1.) Algorithme.

Cet algorithme permet de résoudre de façon approchée le problème Q relatif à (X,Y,L) . Cf. LEMAIRE [23]

a) Description :

A l'itération n , on connait une partie compacte X_n (pas forcément convexe) de X .

Le problème principal consiste à déterminer y^n solution du problème Q relatif au triplet (X_n,Y,L).
Cet élément existe (cf. (3.2.4.)).
On pose $a_n = \inf_{x \in X_n} L(x,y^n) = \sup_{y \in Y} \inf_{x \in X_n} L(x,y)$.

Sous-problème :

Déterminer $\overline{x^n} \in X$ tel que :

$$b_n = L(\overline{x^n},y^n) = \min_{x \in X} L(x,y^n) .$$

Si $a_n = b_n$ arrêt (on montre que y^n est solution du problème Q relatif à (X,Y,L).)
Sinon, on détermine l'ensemble X_{n+1} compact tel que :

$$X \supseteq X_{n+1} \supseteq X_n \cup \{\overline{x^n}\}$$

et on passe à l'itération suivante. Cette dernière opération explique le mot "accumulation" utilisé dans la dénomination de ces procédés.

b) Théorème :

Si H1 est satisfaite et si X et Y sont bornés, pour toute suite $\{y^n\}$ construite par le procédé ci-dessus on a l'alternative suivante :

α) A une certaine itération N on a $a_N = b_N$ alors y^N est solution du problème Q relatif à (X,Y,L) .

β) La suite $\{y^n\}$ est infinie. Alors :

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = v$$

et toute valeur d'adhérence de $\{y^n\}$ est solution du problème Q relatif à (X,Y,L).

Démonstration :

On a, puisque pour tout n : $X \supseteq X_{n+1} \supseteq X_n$, les inégalités :

(1) $\quad a_n \geq a_{n+1} \geq v \geq b_n$.

Si

$a_N = b_N$, alors $a_N = b_N = v$;

donc :

$$\inf_{x\in X} L(x,y^N) = \sup_{y\in Y} \inf_{x\in X} L(x,y) .$$

Ceci démontre α).

Supposons la suite $\{y^n\}$ infinie. De (1) il vient :

(2) $\quad \lim_{n\to\infty} a_n = a^* \geq v$.

Soit $m > n$. On a, puisque $\overline{x^n} \in X_{n+1} \subseteq X_m$:

(3) $\quad L(\overline{x^n},y^m) \geq \min_{x\in X_m} L(x,y^m) = a_m \geq v \geq L(\overline{x^m},y^m)$.

Comme la suite $\{\overline{x^n}\}$ est bornée, on peut extraire une sous suite convergente $\{\overline{x^n}\}_{n \in S}$. On a donc puisque la suite $\{y^n\}$ est bornée et que L est unfformément continue sur $X \times Y$:

$$\lim_{\substack{n\to\infty \\ m>n \\ m,n \in S}} (L(\overline{x^n},y^m) - L(\overline{x^m},y^m)) = 0 \quad .$$

D'où il vient par un raisonnement par l'absurde utilisant (2) et (3) que :

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = v \quad .$$

On a par construction :

$$b_n \leq L(x,y^n) \qquad \forall\, x \in X \quad .$$

Donc pour toute valeur d'adhérence $\tilde{y}$ de la suite $\{y^n\}$, on a : par passage à la limite :

$$v \leq L(x,\tilde{y}) \qquad \forall\, x \in X \quad .$$

D'où :

$$v \leq \min_{x\in X} L(x,\tilde{y}) \leq \sup_{y\in Y} \inf_{x\in X} L(x,y) = v \quad .$$

Donc :

$$v = \min_{x\in X} L(x,\tilde{y}) \quad ,$$

ce qui prouve que $\tilde{y}$ est solution du problème Q relatif à (X,Y,L).

C.Q.F.D.

c) Corollaire :

Si le problème Q initial a une solution unique, la suite $\{y^n\}$ converge vers cette solution .

d) Remarque :

Soit $x^n \in X_n$: $L(x^n,y^n) = \min_{x \in X_n} L(x,y^n)$.

Ces valeurs de x^n n'ont pas de signification très intéressante, ce ne sont pas en général des approximations du problème P .

e) Remarque :

Le procédé qui consisterait à modifier l'itération n en considérant le problème principal suivant :

"trouver (x^n,y^n) solution du problème P relatif à (X_n,Y,L)" ,

n'a en général aucun intérêt.

f) Remarque :

L'algorithme ci-dessus peut être utilisé dans le cas non convexe. On suppose simplement que les ensembles X et Y sont compacts et que L est continue sur $X \times Y$.
On montre alors que :

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = \sup_{y \in Y} \inf_{x \in X} L(x,y) .$$

Toute valeur d'adhérence de $\{y^n\}$ est solution du problème Q.

g) Remarque :

On peut généraliser le procédé précédent.
A l'itération n on considère une fonction $L_n : X \times Y \to \mathbb{R}$ continue sur $X \times Y$, convexe concave et telle que :

$$L_n(x,y) \geq L(x,y) \qquad \forall x \in X \quad \forall y \in Y .$$

Le problème principal consiste à trouver y^n solution du problème Q relatif au triplet (X,Y,L_n) .
On pose $a_n = \min_{x \in X} L_n(x,y^n)$.

On détermine $\overline{x^n}$ par le même sous-problème que précédemment. Puis, si $a_n \neq b_n$ on construit une fonction $L_{n+1} : X \times Y \to \mathbb{R}$ continue sur $X \times Y$ convexe concave et telle que :

$$\begin{cases} L_n(x,y) \geq L_{n+1}(x,y) \geq L(x,y) & \forall\, x \in X \quad \forall\, y \in Y \\ L_{n+1}(\overline{x^n},y) = L(\overline{x^n},y) & \forall\, y \in Y\ , \end{cases}$$

et on passe à l'itération suivante.

(3.3.1.2.) Algorithme :

Permet de résoudre de façon approchée le problème S relatif à (X,Y,L).

a) Description :

Ce procédé se différencie du précédent par le fait que le problème principal consiste à trouver le couple (x^n,y^n) point de selle du triplet (X_n,Y,L) et que les ensembles X_n sont convexes et compacts.

b) Théorème :

Pour toute suite $\{(x^n,y^n)\}$ construite par le procédé ci-dessus, l'hypothèse H1 et le fait que X et Y sont bornés impliquent l'alternative suivante :

α) A une certaine itération N on a : $a_N = b_N$ alors (x^N,y^N) est solution du problème S relatif au triplet (X,Y,L) .

β) On a un nombre infini d'itérations. Toutes valeurs d'adhérence $\tilde{x}$ et $\tilde{y}$ respectivement des suites $\{x^n\}$ et $\{y^n\}$ sont telles que $(\tilde{x},\tilde{y})$ est point de selle de (X,Y,L).

Démonstration :

Démonstration analogue à celle du théorème du (3.3.1.1.). Il suffit de montrer de plus que toute valeur d'adhérence de $\{x^n\}$, notée $\tilde{x}$, est solution du problème P relatif à (X,Y,L) .

On a :

$$a_n \geq L(x^n, y) \qquad \forall\, y \in Y \ .$$

Par passage à la limite pour une sous suite convenable on a :

$$v \geq L(\tilde{x}, y) \qquad \forall\, y \in Y \ .$$

$$v \geq \max_{y \in Y} L(\tilde{x}, y) \geq \inf_{x \in X} \sup_{y \in Y} L(x,y) = v \ .$$

Donc :

$$v = \max_{y \in Y} L(\tilde{x}, y) \ .$$

C.Q.F.D.

La remarque g) du paragraphe précédent est intéressante. On peut prendre dans ce cas :

$$L_n(.,y) = \text{conv}\,(A(.,y)) \qquad (*)$$

avec
$$A(x,y) = \begin{cases} L(x,y) & \forall\, y \quad x = \overline{x^i} \qquad i = 0, \ldots, (n-1) \\ +\infty & \text{sinon} \ . \end{cases}$$

(3.3.1.3.) Algorithme avec linéarisation du sous-problème.

Nous considérons un algorithme où le sous problème est linéarisé ce qui permet une mise en oeuvre simple. Ce procédé peut être interprété comme une généralisation au cas de fonctions non dérivables de la méthode de FRANK et WOLFE.

a) En plus de H1 et du fait que X et Y sont bornés nous ferons l'hypothèse suivante :

H4 : Pour tout $y \in Y$ la fonction $L(.,y)$ est finie (donc continue car elle est convexe) sur un ouvert convexe $\mathcal{O}$ contenant X .

(*) On note conv (f) l'enveloppe convexe de la fonction f .

Ceci implique que pour tout $x \in X$ et tout $y \in Y$ la fonction $L(.,y)$ possède un sous différentiel (cf. ROCKAFELLAR [27] théorème 23-4) $\partial_x L(x,y)$ non vide.

Le théorème (24.7) de ROCKAFELLAR [27] entraine que :

$$\exists M \in \mathbb{R} \quad \forall x \in X$$
$$\forall \tilde{y} \in Y \text{ tel que } L(x,\tilde{y}) = \sup_{z \in Y} L(x,z)$$
$$\forall C \in \partial_x L(x,\tilde{y}) \text{ on a :}$$

$$\|C\| \leq M .$$

b) Nous décrivons maintenant la $n^{ème}$ itération. Soit (x^n,y^n) le couple obtenu au cours du problème principal, supposons pour fixer les idées, que (x^n,y^n) est solution du problème S relatif à (X_n,Y,L) .

X_n est convexe.

On choisit $C^n \in \partial_x L(x^n,y^n)$ tel que :

$$(C^n, x-x^n) \geq 0 \qquad \forall x \in X_n$$

(un tel élément existe toujours car on a : $L(x^n,y^n) \leq L(x,y^n) \quad \forall x \in X_n$) .

Puis on détermine $\overline{x^n} \in X$ tel que :

$$(C^n,\overline{x^n}) = \min_{x \in X} (C^n,x) .$$

Si $(C^n, x^n - \overline{x^n}) = 0$ on s'arrête, sinon on détermine X_{n+1} convexe compact tel que :

$$X \supseteq X_{n+1} \supseteq X_n \cup \{\overline{x^n}\} ,$$

et on passe à l'itération suivante.

c) Théorème :

Si les hypothèses H1 et H4 sont satisfaites et si les ensembles X et Y sont bornés, pour toute suite (x^n,y^n) générée par l'algorithme ci-dessus, on a l'alternative suivante :

α) Si la suite est finie le dernier couple (x^n,y^n) déterminé est solution du problème S pour le triplet (X,Y,L).

β) Si la suite est infinie on a : $\lim_{n\to\infty} a_n = v$,

$\lim_{n\to\infty} (C^n,\overline{x}^n,x^n) = 0$ et toutes valeurs d'adhérence $\tilde{x}$ et $\tilde{y}$ des suites $\{x^n\}$ et $\{y^n\}$ sont telles que $(\tilde{x},\tilde{y})$ est point de selle du triplet (X,Y,L).

Démonstration :

Montrons la partie β) de ce théorème.

Soit $m > n$. On a :

$$(C^m,\overline{x}^m-x) \leq 0 \quad \forall\, x \in X \quad \text{par définition de } \overline{x}^m ,$$

donc :

$$(1) \qquad (C^m,\overline{x}^m-x^m) \leq 0 .$$

Par construction de C^m il vient :

$$(C^m,x-x^m) \geq 0 \qquad \forall\, x \in X_m .$$

Comme $\overline{x}^n \in X_m$:

$$(2) \qquad (C^m,\overline{x}^n-x^m) \geq 0 .$$

Les suites $\{x^m\}$ et $\{C^m\}$ étant bornées, on montre grâce à (1) et (2) que :

$$(3) \qquad \lim_{n\to\infty} (C^n,\overline{x}^n,x^n) = 0 .$$

Par définition de C^n (convexité de $L(.,y^n)$ il vient :

$$L(x,y^n) \geq L(x^n,y^n) + (C^n, x-x^n) \qquad \forall\, x \in X$$

d'où :

$$\min_{x\in X} L(x,y^n) \geq L(x^n,y^n) + (C^n,\overline{x^n}-x^n)$$

et :

$$(4) \qquad L(x^n,y^n) \geq v \geq \min_{x\in X} L(x,y^n) \geq L(x^n,y^n) + (C^n,\overline{x^n}-x^n) .$$

D'où grâce à (3) :

$$\lim_{n\to\infty} L(x^n,y^n) = v .$$

Soit $\tilde{x}$ et $\tilde{y}$ des valeurs d'adhérence des suites $\{x^n\}$ et $\{y^n\}$. Par passage à la limite en (4) il vient :

$$(5) \qquad v = \min_{x\in X} L(x,\tilde{y}) .$$

Par définition de (x^n,y^n) , on a :

$$L(x^n,y^n) = \max_{y\in Y} L(x^n,y) .$$

D'où par passage à la limite :

$$v = L(\tilde{x},\tilde{y}) = \max_{y\in Y} L(\tilde{x},y) .$$

C.Q.F.D.

d) Remarque :

Si dans l'algorithme précédent (x^n,y^n) est solution du problème Q relatif à (X_n,Y,L) , on montre seulement que toute valeur d'adhérence de $\{y^n\}$ est solution du problème Q relatif à (X,Y,L).

e) Remarque :

Si $L(.,y)$ possède une dérivée bornée sur X : $\nabla_x L(.,y)$ on prend alors $C^n = \nabla_x L(x^n,y^n)$.

f) Remarque :

L'algorithme décrit en b) a une signification très intéressante qui ne fait intervenir que la fonction f .

On rappelle qu'on note :

$$f(.) = \sup_{y \in Y} L(.,y) .$$

On sait (VALADIER [34]) que :

$$\partial f(x^n) = \overline{\text{conv}} \bigcup_{z \in Z_n} \partial_x L(x^n,z)$$

où :

$$Z_n = \{z \in Y \mid L(x^n,z) = \sup_{y \in Y} L(x^n,y) = f(x^n)\} .$$

On peut donc donner une interprétation de l'itération n de la méthode du a) qui ne fait intervenir que la fonction f :

On connait une partie convexe compacte X_n de X .

PROBLEME PRINCIPAL :

On cherche $x^n \in X_n$ tel que $f(x^n) = \min (f(x) \mid x \in X_n)$.

Puis $C^n \in \partial f(x^n)$ tel que :

$$(C^n, x-x^n) \geq 0 \qquad \forall\, x \in X_n .$$

SOUS-PROBLEME :

On détermine $\overline{x^n} \in X$ tel que :

$$(C^n, \overline{x^n}) = \min ((C^n,x) \mid x \in X) .$$

Si $(C^n, \overline{x^n} - x^n) = 0$ arrêt.

Sinon on détermine l'ensemble X_{n+1} convexe compact tel que :

$$X \supseteq X_{n+1} \supseteq X_n \cup \{\overline{x^n}\} .$$

Cette interprétation est intéressante car elle donne un procédé original pour minimiser f sur X . Cet algorithme a des analogies avec l'algorithme de FRANK et WOLFE (cf. CEA [8]) , il est cependant fondamentalement différent car il s'applique au cas où f n'est pas dérivable. Nous reviendrons sur ce point en (3.8.2.).

(3.3.1.4.) Quelques variantes permettant de simplifier les ensembles X_n .

a) Par analogie avec les techniques du (1.7.) et (2.4.), nous pouvons "simplifier" les ensembles X_n qui dans les méthodes précédentes "mémorisent" tous les éléments $\overline{x^i}$ trouvés dans les itérations antérieures. A titre d'exemple, nous décrivons une stratégie qui permet d'éliminer une partie des $\overline{x^i}$ précédents. A l'itération n , on connait :

$$X_{n-1} \subseteq \{\overline{x^0}, \overline{x^1}, \ldots, \overline{x^{n-2}}\} .$$

Soit $A_n = X_{n-1} \cup \{\overline{x^{n-1}}\}$.

On détermine (problème principal) une solution y^n du problème Q relatif au triplet (A_n, Y, L) . On pose $a_n = \inf_{x \in A_n} L(x, y^n)$.

Puis $\overline{x^n} \in X$ tel que :

$$L(\overline{x^n}, y^n) = \min_{x \in X} L(x, y^n) .$$

Etant donné un nombre $K > 0$ arbitraire fixé, on détermine le compact $X_n = A_n - B_n$, où :

$$B_n = \{\overline{x^i} \in A_n \mid L(\overline{x^i}, y^n) > a_n, |a_n - a_i| \geq k |L(\overline{x^n}, y^n) - L(\overline{x^i}, y^n)|\} .$$

b) Pour justifier ce procédé il nous suffit de montrer le :

LEMME :

Les hypothèses H1, H2 et le fait que X et Y sont bornés entraînent que si la suite $\{y^n\}$ construite est infinie on a :

$$\lim_{n\to\infty} a_n = v .$$

Démonstration :

Comme :

$$a_n \geq a_{n+1} \geq v \qquad \forall\, n$$

on a :

(1) $$\lim_{n\to\infty} a_n \geq v$$

(2) $$\lim_{\substack{i\to\infty \\ n>i}} (a_n - a_i) = 0 .$$

Soit $E_1 = \{i \mid \overline{x^i} \in X_n \quad \forall\, n > i\}$ et E_2 son complémentaire.

Si l'ensemble E_2 est de cardinal fini, on montre comme en (3.3.1.1.) que :

(3) $$\lim_{n\to\infty} a_n = v .$$

Si l'ensemble E_2 est de cardinal infini on sait que :

Pour tout $i \in E_2$ il existe $n(i) > i$ tel que :

$$|a_{n(i)} - a_i| \geq K\left|L(\overline{x^i}, y^{n(i)}) - L(\overline{x^{n(i)}}, y^{n(i)})\right| .$$

Ce qui grâce à (2) implique que :

(4) $$\lim_{\substack{i\to\infty \\ i\in E_2}} \left(L(\overline{x^i}, y^{n(i)}) - L(\overline{x^{n(i)}}, y^{n(i)})\right) = 0 .$$

Or on sait que :

$$L(\overline{x^{n(i)}}, y^{n(i)}) \leq v \leq a_{n(i)} < L(\overline{x^{I}}, y^{n(i)}) \qquad \forall\, i \in E_2$$

D'où grâce à (4) :

$$\lim_{\substack{i \to \infty \\ i \in E_2}} a_{n(i)} = v$$

ce qui prouve, compte tenu de (1) que :

$$\lim_{n \to \infty} L(x^n, y^n) = v \ .$$

C.Q.F.D.

(3.3.1.6.)

Dans les paragraphes qui suivent nous examinerons le cas où les ensembles X et Y ne sont pas bornés. Mais bien entendu on pourra pour certains problèmes se ramener à priori au cas des ensembles X et Y bornés et utiliser les méthodes ci-dessus. (C'est le cas par exemple si on sait qu'on peut remplacer X et Y par des ensembles convexes compacts connus sans modifier la solution du problème considéré). Si Y n'est pas borné et si les problèmes principaux ont des solutions bornées pour tout n on peut utiliser les procédés du (3.3.) sans modification.

(3.3.2.) Cas où Y n'est pas borné.

En plus des hypothèses H1, on suppose que X est borné et que :

H3 : Il existe une partie C de X et un réel a tel que l'ensemble $\{y \in Y \mid \inf_{x \in C} L(x,y) \geq a\}$ soit non vide et borné.

On peut adapter sans difficulté théorique les algorithmes du (3.3.1.), il suffit que les problèmes principaux sur (X_n, Y, L) aient une solution. On pourra (cf. théorème (3.2.1;)) pour cela considérer des

ensembles X_n qui contiennent C, Nous les résultats du (3.3.1.) restent alors valables.

(3.3.3.) Cas où X n'est pas borné.

(3.3.3.1.)

En plus des hypothèses H1 nous supposerons ici sauf si le contraire est précisé que Y est borné et que :

H2 : Il existe une partie D de Y et un réel a tel que l'ensemble $\{x \in X \mid \sup_{y \in D} L(x,y) \leq a\}$ soit non vide et borné.

Les algorithmes précédentsne sont plus applicables car les sous problèmes correspondants ne sont plus définis. Nous considérerons trois types de sous-problèmes qui sont adaptés à ce cas.

Nous nous limiterons au cas des procédés où le problème principal consiste à chercher (x^n,y^n) solution du problème S relatif au triplet (X_n,Y,L). On posera $a_n = L(x^n,y^n)$.

(3.3.3.2.) Algorithme.

a) Description :

Soit V un voisinage convexe compact de $0 \in \mathbb{R}^n$.

A l'itératinn n on considère le sous problème suivant :

si on pose $A_n = X \cap (x^n+V)$,

on cherche $\overline{x^n} \in A_n$ (il existe toujours) tel que :

$$\min_{x \in A_n} L(x,y^n) = L(\overline{x^n},y^n) = b_n .$$

Si $a_n = b_n$ arrêt. Sinon on détermine l'ensemble convexe fermé X_{n+1} (pas forcément borné) tel que $X \supseteq X_{n+1} \supseteq X_n \cup \{\overline{x^n}\}$, et on passe à l'itération suivante.

b) Théorème :

Les hypothèses du (3.3.3.1.) entrainent les mêmes conclusions que le théorème du (3.3.1.2.).

Démonstration :

Montrons d'abord que les suites $\{x^n\}$ et $\{\overline{x^n}\}$ sont bornées.

On a :

$$(1) \qquad a_1 \geq a_n = f(x^n) \geq a_{n+1} \geq v \qquad \forall\, n \ .$$

Donc : $x^n \in \{x \in X \mid f(x) \leq a_1\}$. Grâce à H2 cet ensemble est borné, la suite $\{x^n\}$ l'est donc aussi. On en déduit que la suite $\{\overline{x^n}\}$ est bornée car V est borné.

On montre alors comme dans le théorème du (3.3.1.1.) que :

$$(2) \qquad \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = \tilde{a} \geq v \ .$$

Donc pour tout n on a :

$$(3) \qquad a_n = \min_{\substack{x \in X \\ x - x^n \in V}} L(x, y^n) + \varepsilon_n$$

SERVICE POLYCOPIE
MATHÉMATIQUES APPLIQUÉES
Université de GRENOBLE

où :

$$\lim_{n\to\infty} \varepsilon_n = 0 \ .$$

Soit $\tilde{x}$ et $\tilde{y}$ des valeurs d'adhérence des suites $\{x^n\}$ et $\{y^n\}$ (existent toujours car ces suites sont bornées).

On a par passage à la limite en (3) :

$$\lim_{n\to\infty} a_n = \tilde{a} = L(\tilde{x},\tilde{y}) \geq \min_{\substack{x \in X \\ x - \tilde{x} \in V}} L(x,\tilde{y}) \ .$$

Et comme l'intérieur de V contient 0 on a :

$$L(\tilde{x},\tilde{y}) = \min_{x \in X} L(x,\tilde{y}) \leq \sup_{y \in Y} \inf_{x \in X} L(x,y) = v \ .$$

D'où on tire grâce à (2) :

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = v \ .$$

On termine sans difficulté la démonstration.

C.Q.F.D.

Remarque :

On peut aussi avec les hypothèses supplémentaires du (3.3.1.3.) considérer le sous problème suivant :

Déterminer C^n élément de $\partial_x L(x^n,y^n)$ tel que :

$(C^n, x-x^n) \geq 0 \qquad \forall\, x \in X_n$;

déterminer $\overline{x^n} \in A_n = \{x \in X \mid (x-x^n) \in V\}$

tel que $(C^n,\overline{x^n}) = \min\ ((C^n,x) \mid x \in A_n)$.

Cet algorithme est une extension du procédé du (3.3.1.3.).

(3.3.3.3.) Le sous problème est un problème de minimax particulier.

a) A titre d'exemple considérons le sous problème suivant :

trouver $(\overline{x^n},\overline{y^n})$ solution du problème S relatif à (X,Z_n,L) où Z_n est une partie convexe fermée de Y définie par :

$$Z_n = \overline{\text{conv}}\ (D \cup \{y^n\}) \ ,$$

où l'ensemble D est celui qui intervient dans l'hypothèse du (3.3.3.1.). Le théorème (3.2.1.) permet d'affirmer que grâce aux hypothèses faites le couple $(\overline{x^n},\overline{y^n})$ existe.
On pose $b_n = L(\overline{x^n},\overline{y^n})$. Si $a_n = b_n$ arrêt. Sinon on détermine X_{n+1} comme plus haut.

b) Théorème :

Les hypothèses du (3.3.3.1.) entrainent que pour toute suite $\{(x^n,\overline{y^n})\}$ générée par l'algorithme ci-dessus on a l'alternative suivante :

α) Si la suite est finie le dernier couple calculé $(x^N,\overline{y^N})$ est point de selle du problème initial.

β) Si la suite est infinie alors toute valeur d'adhérence $\tilde{x}$ et $\tilde{y}$ respectivement de $\{x^n\}$ et $\{\overline{y^n}\}$ sont telles que $(\tilde{x},\tilde{y})$ est point de selle du problème initial et on a :

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = v .$$

Démonstration :

Montrons le point β). Vérifions d'abord que les suites $\{x^n\}$ et $\{\overline{y^n}\}$ sont bornées. C'est évident pour la suite $\{x^n\}$ car f est inf compacte sur X .

On a pour tout n :

$$v \geq \overline{L(x^n,y^n)} = \sup_{y\in Y} L(\overline{x^n},y) \geq \sup_{y\in D} L(\overline{x^n},y) ,$$

ce qui montre (H2) que la suite $\{\overline{x^n}\}$ est bornée.

Si $m > n$ on a puisque $\overline{x^n} \in X_m$:

$$L(\overline{x^n},y^m) \geq L(x^m,y^m) \geq v \geq L(\overline{x^m},\overline{y^m}) \geq L(\overline{x^m},y^m) .$$

D'où on tire comme dans le théorème du (3.3.1.1.), puisque les suites $\{\overline{x^n}\}$ et $\{\overline{y^n}\}$ sont bornées, que :

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = v .$$

Or par construction on sait que :

$$\begin{cases} L(x^n,y^n) \geq L(x^n,y) & \forall\, y \in Y \\ L(\overline{x^n},\overline{y^n}) \leq L(x,\overline{y^n}) & \forall\, x \in X . \end{cases}$$

On termine sans difficulté la démonstration par passage à la limite dans les inégalités ci-dessus.

C.Q.F.D.

Remarque :

Un cas pratique intéressant est le cas où $D = \{\omega\}$ où ω est un élément connu de Y .

Remarque :

On trouvera dans le tableau du (3.3.4.) d'autres méthodes du même type.

(3.3.3.4.) Cas particulier :

a) En plus des hypothèses H1, H2 et H3 nous supposerons que :

H5 : X est un cône convexe de sommet 0 .
La fonction $L(.,y) - L(0,y)$ est positivement homogène pour tout $y \in Y$.

Ces hypothèses sont vérifiées pour les lagrangiens usuels associés aux problèmes d'optimisation.

On considère dans ce paragraphe un voisinage V convexe compact de $0 \in \mathbb{R}^n$.

b) Algorithme :

Au départ on impose que $0 \in X_1$. Nous décrivons le sous problème de la $n^{ème}$ itération. On détermine l'élément $\overline{x^n}$.

$$\overline{x^n} \in V \cap X \text{ tel que } L(\overline{x^n},y^n) = \min_{x \in V \cap X} L(x,y^n) .$$

Si $L(0,y^n) = L(\overline{x^n},y^n)$ arrêt.
Sinon on détermine l'ensemble X_{n+1} convexe compact tel que :

$$X \supseteq X_{n+1} \supseteq A_n$$

où $$A_n = X_n + [0,\overline{x^n}] .$$

Il est évident que A_n est un ensemble convexe compact contenant 0 tel que :

$$A_n \supseteq X_n .$$

c) Théorème :

Les hypothèses faites en a) entrainent que pour toute suite $\{x^n\}$ générée par l'algorithme ci-dessus on a l'alternative suivante :

α) Si la suite générée est finie, le dernièr couple calculé (x^N,y^N) est point de selle du triplet (X,Y,L) .

β) La suite $\{x^n\}$ est infinie. Alors : $\lim_{n\to\infty} a_n = v$ et toutes valeurs d'adhérence $\tilde{x}$ et $\tilde{y}$ respectivement de $\{x^n\}$ et $\{y^n\}$ sont telles que $(\tilde{x},\tilde{y})$ est point de selle du triplet (X,Y,L).

Démonstration :

Montrons d'abord α) .

On a pour un rang N :

$$(1)\quad \begin{cases} L(x^N,y) \leq L(x^N,y^N) \leq L(x,y^N) & \forall\, x \in X_N \quad \forall\, y \in Y \\ L(x,y^N) \geq L(0,y^N) & \forall\, x \in V \cap X . \end{cases}$$

Donc grâce aux propriétés de V et de L on a :

$$L(x,y^N) \geq L(0,y^N) \qquad \forall\, x \in X .$$

Et comme $0 \in X_N$ il vient :

$$L(0,y^N) \geq L(x^N,y^N) .$$

Donc :

$$L(x,y^N) \geq L(x^N,y^N) \qquad \forall\, x \in X .$$

Ce qui prouve α) .

Montrons maintenant $\beta)$.

La suite $\{\overline{x^n}\}$ est bornée et on montre comme pour le théorème du (3.3.3.1.) que les suites $\{x^n\}$ et $\{y^n\}$ sont bornées.

On a les inégalités :

$$(2) \qquad a_n \geq a_{n+1} \geq v \qquad \forall\, n \ .$$

Soit $m > n$ on a par construction :

$$(3) \qquad L(x^m,y^m) \leq L(x^n+\overline{x^n},y^m) \ .$$

On voit grâce aux propriétés de L que :

$$(4) \qquad L(x^n+\overline{x^n},y^m) \leq L(x^n,y^m) + L(\overline{x^n},y^m) - L(0,y^m) \ .$$

Et comme :

$$L(x^n,y^m) \leq L(x^n,y^n)$$

il vient de (3) et (4) :

$$L(x^m,y^m) - L(x^n,y^n) \leq L(\overline{x^n},y^m) - L(0,y^m) \ .$$

On déduit de (2) que :

$$(5) \qquad \liminf_{\substack{n \to \infty \\ m > n}} L(\overline{x^n},y^m) - L(0,y^m) \geq 0 \ .$$

Comme les suites $\{\overline{x^n}\}$ et $\{y^n\}$ sont bornées et que :

$$L(\overline{x^m},y^m) - L(0,y^m) \leq 0 \qquad \forall\, m \ ,$$

on tire de (5) que :

$$\lim_{n\to\infty} L(\overline{x^n},y^n) - L(0,y^n) = 0 \ .$$

On termine aisément la démonstration.

C.Q.F.D.

Remarque :

On peut adapter sans difficulté ce procédé à un cas non convexe.

On modifie simplement dans l'algorithme ci-dessus le problème principal qui devient : trouver y^n solution du problème Q relatif à (X_n,Y,L).

Cette méthode permet de trouver une solution approchée du problème Q relatif au triplet (X,Y,L).

(3.3.3.4.) Cas où X et Y ne sont pas bornés.

On adaptera aisément les procédés précédents au cas où X n'est pas borné compte tenu des remarques du (3.3.2.).

(3.3.4.) Récapitulation des principaux procédés du (3.3.).

Nous présentons sous forme d'un tableau un résumé des procédés précédents. A l'itération n , on y trouve les problèmes principaux suivants :

a) PP1 trouver (x^n,y^n) point de selle du triplet (X_n,Y,L) (dans ce X_n est supposé convexe)

b) PP2 trouver y^n solution du problème Q relatif au triplet (X_n,Y,L) .

On pose $a_n = L(x^n,y^n)$.

Dans la plupart des sous problèmes on détermine $\overline{x^n} \in X$ puis l'ensemble X_{n+1} compact tel que :

$$X \supseteq X_{n+1} \supseteq X_n \cup \{\overline{x^n}\} \quad .$$

On a les sous problèmes suivants :

SP1 : $\overline{x^n} \in X$ tel que :

$$b_n = L(\overline{x^n},y^n) = \min_{x \in X} L(x,y^n) \quad .$$

SP2 : $\overline{x^n} \in X$ tel que :

$$(c^n, \overline{x^n}) = \min \{(c^n,x) \mid x \in X\}, \text{ où } c^n \in \partial_x L(x^n,y^n)$$

est tel que :

$$(c^n, x-x^n) \geq 0 \quad \forall\, x \in X_n .$$

SP3 : trouver $\overline{x^n} \in A_n$ tel que :

$$b_n = L(\overline{x^n},y^n) = \min_{x \in A_n} L(x,y^n) \quad \text{où} \quad A_n = X \cap (x^n+V)$$

SP4 : trouver $(\overline{x^n},\overline{y^n})$ point de selle du triplet $(X, \text{conv}(\{y^n\} \cup D), L)$

On pose $b_n = L(\overline{x^n},\overline{y^n})$.

SP5 : trouver $\overline{x^n}$ solution du problème P relatif à $(X,\{y^n\} \cup D, L)$.

SP6 : trouver $\overline{x^n} \in V \wedge X$ tel que :

$$L(\overline{x^n},y^n) = \min_{x \in V \wedge X} L(x,y^n) .$$

$$A_n = X_n + [0,\overline{x^n}] .$$

On choisit ici X_{n+1} convexe compact tel que :

$$X \supseteq X_{n+1} \supseteq A_n .$$

Paragraphe	Hypothèses particulières	Problème principal	Sous problème	Test d'arrêt	Propriétés
(3.3.1.1.)	H1, X et Y bornés	PP2	SP1	$a_n = b_n$	$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = v$ adh $\{y^n\}$ solution du problème Q initial
(3.3.2.1.)	L continue sur $X \times Y$ X,Y compacts (non convexes)	PP2	SP1	$a_n = b_n$	$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = \sup_{y\in Y} \inf_{x\in X} L$ adh $\{y^n\}$ solution du problème Q initial
(3.3.1.2.)	H1, X,Y bornés	PP1 X_n convexe	SP1	$a_n = b_n$	si $\tilde{x} \in$ adh $\{x^n\}$, $\tilde{y} \in$ adh $\{y^n\}$ $(\tilde{x},\tilde{y})$ solution du problème S initial
(3.3.1.3.)	H1, H4, X,Y bornés	id.	SP2	$(C^n, x^n - \overline{x^n}) = 0$	id.
(3.3.3.2.)	H1, H2, Y borné	id.	SP3	$a_n = b_n$	id.

.../...

.../...

Paragraphe	Hypothèses particulières	Problème principal	Sous problème	Test d'arrêt	Propriétés
(3.3.3.3.)	H1, H2, Y borné	PP1 X_n convexe	SP4	$a_n = b_n$	si $\tilde{x} \in$ adh $\{x^n\}$, $\tilde{y} \in$ adh $\{\overline{y^n}\}$ $(\tilde{x},\tilde{y})$ solution du problème S initial
	id.	id.	SP5	id.	adh $\{x^n\}$ est solution du problème P initial $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = v$
	id.	PP2	SP4	id.	adh $\{\overline{y^n}\}$ est solution du problème Q initial $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = v$
	id.	id.	SP5	id.	$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = v$
(3.3.3.4.)	H1, H2, H3, H5	PP1 X_n convexe	SP6	$L(0,y^n) = L(\overline{x^n},y^n)$	$\lim_{n\to\infty} a_n = v$ Si $\tilde{x} \in$ adh $\{x^n\}$, $\tilde{y} \in$ adh $\{y^n\}$ $(\tilde{x},\tilde{y})$ solution du problème S initial

(3.4.) METHODES SANS ACCUMULATION DANS LE CAS OU L(x,.) EST STRICTEMENT CONCAVE.

En plus des hypothèses H1 nous supposerons dans ce paragraphe que :

H6 : Pour tout $x \in X$ la fonction $L(x,.)$ est trictement concave.

Cette nouvelle hypothèse est très importante car elle permet de justifier des procédés analogues aux méthodes (2.5.) où les ensembles X_n considérés dans les problèmes principaux sont beaucoup plus "simples" en ce qui concerne la mise en oeuvre.

(3.4.1.) Théorème :

Soit X_1 et Y_1 des parties convexes compactes non vides de X et Y.

Les hypothèses H1 et H6 entrainent que la fonction $L(x,.)$ est uniformément concave uniformément en x c'est-à-dire que :

il existe une fonction δ de $\mathbb{R}^+$ dans $\mathbb{R}^+$ non décroissante, qui vérifie : $\delta(0) = 0$ et $\delta(\tau) > 0$ si $\tau > 0$ qui est telle que l'on ait :

$$L\left(x,\frac{y_1+y_2}{2}\right) \geq \frac{1}{2}\left(L(x,y_1)+L(x,y_2)\right)+\delta\left(\|y_1-y_2\|\right)$$

$$\forall\, x \in X_1 \qquad \forall\, y_1, y_2 \in Y_1 .$$

(Ici $\|\ \|$ est la norme euclidienne sur $\mathbb{R}^m$) .

Démonstration :

Posons :

$$\alpha(x,y_1,y_2) = L\left(x,\frac{y_1+y_2}{2}\right) - \frac{1}{2}\left(L(x,y_1) + L(x,y_2)\right)$$

et pour $\tau \geq 0$ considérons :

$$(1) \qquad \delta(\tau) = \inf \{\alpha(x,y_1,y_2) \mid x \in X_1; y_1,y_2 \in Y_1; \|y_1-y_2\| \geq \tau\} .$$

On vérifie sans difficulté que l'inf considéré ci-dessus est atteint car on minimise une fonction continue sur un compact.
Il est facile de voir que grâce à la stricte concavité de $L(x,.)$ on a :

$$\alpha(x,y_1,y_2) \begin{cases} > 0 & \text{si} \quad y_1 \neq y_2 \\ = 0 & \text{si} \quad y_1 = y_2 . \end{cases}$$

D'où on déduit de (1) que la fonction δ est non décroissante et que :

$$\delta(\tau) \begin{cases} > 0 & \text{si} \quad \tau > 0 \\ = 0 & \text{si} \quad \tau = 0 . \end{cases}$$

C.Q.F.D.

(3.4.2.) Corollaire :

Avec les conditions du théorème (3.4.1.) ci-dessus la fonction $g(.) = \inf_{x \in X_1} L(x,.)$ est uniformément concave.

Démonstration :

D'après le théorème (3.4.1.) on a :

$$(L(x,\frac{y_1+y_2}{2}) \geq \frac{1}{2}(L(x,y_1)+L(x,y_2))+\delta(\|y_1-y_2\|) \qquad \forall\, x \in X_1 \quad \forall\, y_1,y_2 \in Y_1 .$$

Donc pour tout $y_1,y_2 \in Y_1$ on a :

$$\inf_{x \in X_1} L(x,\frac{y_1+y_2}{2}) \geq \inf_{x \in X_1} \frac{1}{2}(L(x,y_1)+L(x,y_2))+\delta(\|y_1-y_2\|)$$

$$\inf_{x \in X_1} L(x,\frac{y_1+y_2}{2}) \geq \frac{1}{2} \inf_{x \in X_1} L(x,y_1)+\frac{1}{2} \inf_{x \in X_1} L(x,y_2)+\delta(\|y_1-y_2\|)$$

C.Q.F.D.

(3.4.3.) Théorème :

Soit une suite $\{(x^n,y^n)\}$ telle que :

$$\begin{cases} (x^n,y^n) \in X \times Y \qquad \forall\, n \\ L(x^n,y^n) = \sup_{y\in Y} L(x^n,y) = f(x^n) \qquad \forall\, n \\ \lim_{n\to\infty} f(x^n) = v \quad \text{(valeur du triplet } (X,Y,L)\text{)}. \end{cases}$$

Si en plus des hypothèses H1, H2 et H6 on suppose que :

l'ensemble Y est compact et que

la fonction $g(.) = \inf_{x\in X} L(x,.)$ est strictement concave sur son domaine ,

alors la suite $\{y^n\}$ converge vers l'élément unique $\bar{y}$ qui réalise le minimum de g sur Y et toute valeur d'adhérence $\tilde{x}$ de $\{x^n\}$ est telle que $(\tilde{x},\bar{y})$ est point de selle du triplet (X,Y,L).

Démonstration :

La suite $\{f(x^n)\}$ est bornée supérieurement par un nombre M . On a pour tout n :

$$M \geq L(x^n,y^n) \geq \max_{y\in Y} L(x^n,y) .$$

Donc la suite $\{x^n\}$ est bornée grâce à H2.

Soit $\tilde{x}$ et $\tilde{y}$ des valeurs d'adhérence des suites bornées $\{x^n\}$ et $\{y^n\}$.

On a :

$$f(x^n) = L(x^n,y^n) = \sup_{y\in Y} L(x^n,y) .$$

Par passage à la limite sur une sous suite convenable il vient :

$$f(x) \geq f(\tilde{x}) = \sup_{y\in Y} L(\tilde{x},y) = L(\tilde{x},\tilde{y}) = v \qquad \forall\, x \in X$$

ce qui montre que le couple $(\tilde{x},\tilde{y})$ est solution du problème P relatif au triplet (X,Y,L).

Soit $\bar{y}$ l'élément unique où g atteint son maximum sur Y, $\bar{y}$ est donc solution unique du problème Q relatif à (X,Y,L).
Donc $(\tilde{x},\bar{y})$ est point de selle de (X,Y,L) (cf. (3.2.2.)) et on a :

$$v = L(\tilde{x},\tilde{y}) = L(\tilde{x},\bar{y}) = \max_{y\in Y} L(\tilde{x},y) .$$

Mais $L(\tilde{x},.)$ est strictement concave, donc : $\tilde{y} = \bar{y}$.

C.Q.F.D.

Remarque :

Ce théorème montre que toute méthode générant une suite minimisante pour la fonctionnelle f permet d'obtenir une suite qui converge vers la solution unique $\bar{y}$ du problème Q .

Ce théorème est la base des méthodes duales étudiées par AUSLENDER [2] , FALK [10] , [11] pour la résolution de problèmes d'optimisation avec contraintes.

(3.4.4.) Généralités sur les algorithmes.

Nous allons étudier quelques procédés similaires à ceux du (3.3.) pour lesquels on considère à l'itération n :

- un problème principal qui consiste à chercher le point de selle (x^n,y^n) du triplet (X_n,Y,L);
- un sous problème comme en (3.3.).

Les ensembles X_n considérés peuvent être nettement "plus simples" que ceux qu'on considère en (3.3.), on prendra par exemple si on suppose l'hypothèse H3 satisfaite :

$$X_n = \overline{\text{conv}}\,([\overline{x^{n-1}},x^{n-1}] \cup C) \quad (C \text{ ensemble considéré dans H3})$$
$$X_n = [\overline{x^{n-1}},x^{n-1}] \quad \text{si } Y \text{ est borné.}$$

Nous nous limiterons ici à deux exemples caractéristiques de tels procédés.

(3.4.5.) Algorithme.

(3.4.5.1.)

En plus de l'hypothèse H1 nous supposons que les ensembles X et Y sont bornés. On considère le procédé qui correspond à l'algorithme (3.3.1.2.). A l'itération n on connait les éléments x^{n-1} et $\overline{x^{n-1}}$ de X obtenus à l'itération précédente et on considère :

. Le problème principal :

trouver (x^n,y^n) point de selle relatif au triplet $([x^{n-1},\overline{x^{n-1}}],Y,L)$, poser $a_n = L(x^n,y^n)$.

. Le sous problème :

chercher $\overline{x^n} \in X$ tel que :

$$b_n = L(\overline{x^n},y^n) = \min_{x\in X} L(x,y^n) .$$

Si $a_n = b_n$ on arrête le procédé, sinon on passe à l'itération suivante.

Remarque :

On pourrait aussi dans le problème principal choisir X_n convexe compact tel que :

$$X \supseteq X_n \supseteq [\overline{x^{n-1}},x^{n-1}] .$$

(3.4.5.2.) Théorème :

Les hypothèses H1, H6 et le fait que les ensembles X et Y sont bornés entrainent que pour toutes suites $\{x^n\}$ $\{y^n\}$ générée par l'algorithme ci-dessus on a l'alternative suivante :

a) si la suite est finie le dernier couple calculé (x^N,y^N) est point de selle du problème initial.

b) si la suite est infinie on a :

. $\lim_{n\to\infty} \|y^{n+1}-y^n\| = 0$

. $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = v$.

. Toutes valeurs d'adhérence $\tilde{x}$ et $\tilde{y}$ des suites $\{x^n\}$ et $\{y^n\}$ sont telles que les couples $(\tilde{x},\tilde{y})$ sont des points de selle du problème initial.

Démonstration :

Le point a) est évident. Montrons le point b).

Comme les ensembles X et Y sont bornés, il existe (théorème (3.4.1.)) une fonction δ telle que :

$$(1) \qquad L(x^n, \frac{y^n+y^{n-1}}{2}) \geq \frac{1}{2}(L(x^n,y^n) + L(x^n,y^{n+1})) + \delta(\|y^{n+1}-y^n\|) \quad \forall\, n$$

Or par construction on sait que :

$$L(x^n,y^n) \geq L(x^n, \frac{y^n+y^{n+1}}{2})$$

$$L(x^n,y^{n+1}) \geq L(x^{n+1},y^{n+1})$$

D'où on tire grâce à (1) que :

$$(2) \qquad L(x^n,y^n) - L(x^{n+1},y^{n+1}) \geq 2\ \delta(\|y^{n+1}-y^n\|) .$$

Or la suite $\{a_n\}$ $(a_n = L(x^n,y^n))$ est non croissante et bornée inférieurement donc :

$$\lim_{n\to\infty} (a_n - a_{n+1}) = 0 .$$

D'où on tire de (2) et des propriétés de δ que :

$$(3) \qquad \lim_{n\to\infty} \|y^{n+1}-y^n\| = 0 .$$

Par construction, on a :

$$(4) \qquad L(\overline{x^n}, y^{n+1}) \geq L(x^{n+1}, y^{n+1}) \geq v \geq L(\overline{x^n}, y^n) = b_n \ .$$

Puisque la suite $\{\overline{x^n}\}$ est bornée et que L est uniformément continue sur $X \times Y$ on déduit de (3) que :

$$\lim_{n\to\infty} L(\overline{x^n}, y^{n+1}) - L(\overline{x^n}, y^n) = 0$$

et grâce à (4) :

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = v \ .$$

On termine la démonstration sans difficulté.

C.Q.F.D.

(3.4.6.) Algorithme. (Cf. Algorithme (3.3.1.3.)).

En plus des hypothèses ci-dessus, nous supposerons que :

H7 : Pour tout $y \in Y$ la fonction $L(.,y)$ possède une dérivée $\nabla_x L(.,y)$ qui est une fonction continue sur $X \times Y$.

La $n^{ème}$ itération de l'algorithme comporte le même problème principal que dans l'algorithme (3.4.5.) ; on y détermine le couple (x^n, y^n). Posons $C^n = \nabla_x L(x^n, y^n)$.

Le sous problème consiste à chercher $\overline{x^n} \in X$ tel que :

$$(C^n, \overline{x^n}) = \min ((C^n, x) \mid x \in X).$$

Posons $d_n = (C^n, \overline{x^n} - x^n)$. Si $d_n = 0$ arrêt des itérations, sinon on passe à l'itération suivante.

Nous allons montrer que ce procédé n'est autre que l'algorithme de FRANK et WOLFE (cf. CEA [8]) appliqué à la fonction

$f = \sup_{y \in Y} L(.,y)$.

A l'itération n :

L'élément $\underline{x^n}$ obtenu en résolvant le problème principal vérifie : $x^n \in [x^{n-1}, \overline{x^{n-1}}]$ et $f(x^n) = \min(f(x) | x \in [x^{n-1}, \overline{x^{n-1}}])$.

Le sous problème fait intervenir $C^n = \nabla_x L(x^n,y^n)$.
On sait que (théorème de VALADIER [34]) :

$$\partial f(x^n) = \overline{\text{conv}} \bigcup_{y \in Z_n} \nabla_x L(x^n,y)$$

où

$$Z_n = \{y \in Y | L(x^n,y) = \max_{p \in Y} L(x^n,p)\} .$$

Or Z_n est réduit à un seul élément grâce à H6, donc $\partial f(x^n)$ est lui aussi réduit à un seul élément et on a :

$$\partial f(x^n) = \{\nabla_x L(x^n,y^n)\} .$$

Les hypothèses H6 et H7 entrainent donc que f est dérivable et que $f'(x^n) = \nabla_x L(x^n,y^n)$. On remarque de plus que f' est continue sur X (donc uniformément continue),

On retrouve la méthode de FRANK et WOLFE habituelle.

(3.4.7.) Cas où l'ensemble X n'est pas borné.

Nous sommes amenés à définir des algorithmes qui comprennent le même problème principal qu'en (3.4.4.) et les mêmes sous problèmes que dans (3.3.3.).

Nous ne développerons pas en détails ces procédés car ils n'introduisent donc rien de fondamentalement nouveau, tant pour leur description que pour leurs propriétés.

(3.5.) METHODES SANS ACCUMULATION DANS UN CAS PLUS GENERAL.

Il s'agit de méthodes de descente pour la fonction f basées sur des principes différents des précédents ; elles permettent de résoudre le problème P initial par une suite de problèmes P relatifs aux triplets (X_n,Y,L) où les ensembles X_n sont des segments. En [24] LEVITIN étudie un procédé analogue pour la minimisation de fonctions non différentiables. On gagne beaucoup en clarté en exposant ces méthodes dans le cadre des problèmes minimax.

(3.5.1.) Algorithme.

(3.5.1.1.) Description.

On construit au cours de ce procédé des suites :

$\{x^n\}$, $\{\overline{x^n}\}$, $\{\delta_n\}$ où x^n et $\overline{x^n} \in X$ et $\delta_n > 0$.

Voici comment au cours de la $n^{ème}$ itération on passe de $(x^{n-1},\overline{x^{n-1}},\delta_{n-1})$ à $(x^n,\overline{x^n},\delta_n)$:

PROBLEME PRINCIPAL :

trouver x^n solution du problème P relatif au triplet $([x^{n-1},\overline{x^{n-1}}],Y,L)$ c'est-à-dire que :

$x^n \in [x^{n-1},\overline{x^{n-1}}]$ est tel que :

$$\sup_{y\in Y} L(x,y) \geq \sup_{y\in Y} L(x^n,y) = a_n \qquad \forall\, x \in [x^{n-1},\overline{x^{n-1}}] \ .$$

SOUS PROBLEME :

trouver $\overline{x^n}$ solution du problème P relatif au triplet (X,Y_n,L) où on pose :

$$Y_n = \{y \in Y \mid L(x^n,y) \geq \sup_{z\in Y} L(x^n,z)-\delta_n\} \ .$$

En d'autres termes trouver $\overline{x^n} \in X$ tel que :

$$\sup_{y \in Y_n} L(x,y) \geq \sup_{y \in Y_n} L(\overline{x^n},y) = b_n \qquad \forall\, x \in X \ .$$

On fait ensuite :

$$\delta_{n+1} = \begin{cases} \delta_n & \text{si } a_n - b_n \geq \delta_n \\ \dfrac{\delta_n}{2} & \text{sinon .} \end{cases}$$

Ce sous problème peut être très complexe à mettre en oeuvre car l'ensemble Y_n défini plus haut n'est pas commode à utiliser ; c'est ce qui limite l'intérêt de ce procédé.

(3.5.1.2.) Théorème :

Si on suppose que Y est compact, X est convexe et compact, L est continue sur $X \times Y$ et $L(.,y)$ est convexe pour tout $y \in Y$, alors toute suite générée par l'algorithme ci-dessus est telle que :

α) $\lim_{n\to\infty} a_n = A = \inf_{x\in X} \sup_{y\in Y} L(x,y)$.

β) Toute valeur d'adhérence de la suite $\{x^n\}$ est solution du problème P initial.

Démonstration :

Montrons d'abord que :

$$(1) \qquad \lim_{n\to\infty} \delta_n = 0 \ .$$

Supposons que ceci ne soit pas vrai alors il existerait $d > 0$ et un entier N tel que pour tout $n \geq N$ on ait :

$$(2) \qquad \begin{cases} \delta_{n+1} = \delta_n \\ a_n - b_n \geq \delta_n = d > 0 \ . \end{cases}$$

Pour tout $\alpha \in [0,1]$ on définit :

$$z^{n-1}(\alpha) = (1-\alpha)\, x^{n-1} + \alpha \overline{x^{n-1}}$$

L'uniforme continuité de L sur $X \times Y$ entraine qu'il existe $\bar{\alpha} > 0$ tel que pour tout $\alpha \in [0,\bar{\alpha}]$ et pour tout n on ait :

$$\max_{y \in Y} L(z^{n-1}(\alpha),y) = \max_{y \in Y_{n-1}} L(z^{n-1}(\alpha),y) .$$

Donc :

$$\max_{y \in Y} L((1-\bar{\alpha})x^{n-1} + \bar{\alpha}\overline{x^{n-1}},y) = \max_{y \in Y_{n-1}} L((1-\bar{\alpha})x^{n-1} + \bar{\alpha}\overline{x^{n-1}},y)$$

$$\leq (1-\bar{\alpha}) \max_{y \in Y_{n-1}} L(x^{n-1},y) + \bar{\alpha} \max_{y \in Y_{n-1}} L(\overline{x^{n-1}},y)$$

$$\leq (1-\bar{\alpha}) a_{n-1} + \bar{\alpha}\, b_{n-1} .$$

Par conséquent :

$$a_n \leq (1-\bar{\alpha})\, a_{n-1} + \bar{\alpha}\, b_{n-1}$$

Grâce à (2) $\forall\, n \geq N+1$ on a :

$$a_n \leq a_{n-1} - \bar{\alpha}\, d .$$

Cette relation implique que la suite $\{a_n\}$ n'est pas bornée inférieurement ce qui contredit le fait que :

$$a_n \geq A \qquad \forall\, n$$

(On rappelle que : $A = \inf_{x \in X} \sup_{y \in Y} L(X,y)$).

L'hypothèse faite est absurde ce qui entraine (1).

Donc, par construction, il existe un ensemble infini d'indices S tels que :

$$(3) \qquad \lim_{\substack{n\to\infty \\ n\in S}} (a_n - b_n) = 0 \ .$$

Or pour tout n on sait que :

$$(4) \qquad a_n \geq a_{n+1} \geq A \geq b_n \ .$$

Il vient de (3) et (4) que :

$$\lim_{n\to\infty} a_n = A \ .$$

On termine aisément la démonstration.

C.Q.F.D.

(3.5.2.) Remarque :

Il serait intéressant d'étudier d'autres procédés comprenant le même problème principal mais d'autres sous problèmes si possibles plus simples à mettre en oeuvre.

(3.5.3.) Remarque :

L'ensemble X n'est plus borné et on suppose par exemple qu'il existe une partie D de Y telle que pour tout $a \in \mathbb{R}$ l'ensemble $\{x \in X \mid \sup_{y\in Y} L(x,y) \leq a\}$ est borné.

On peut adapter facilement le procédé du (3.5.1.1.) en considérant le sous problème suivant :

trouver $\overline{x^n}$ solution du problème P relatif au triplet (X_n, Y_n, L) où :

$$\begin{cases} X_n = X \cap (x^n + V) \\ V \text{ est un voisinage convexe compact de } 0 \in R^n . \\ Y_n \text{ a la même signification que dans (3.5.1.1.).} \end{cases}$$

On choisit δ_{n+1} comme précédemment.

On peut montrer pour cet algorithme modifié que le théorème (3.5.1.2.) reste vrai.

(3.6.) PROBLEMES SUR DES ESPACES PRODUIT :

Ces questions présentent un très grand intérêt car elles sont à la base des techniques de décomposition en optimisation, citons par exemple les articles de GEOFFRION [14] , [15] , [16] , [17] , FALK [10] , [11] , KRONSJÖ [20] , [21] etc ... .

Soit un triplet $(X, Y \times Y_1, \mathcal{L})$

où $\mathcal{L}$ est une fonction : $X \times Y \times Y_1 \to \mathbb{R}$ telle que :

$$\mathcal{L}(x;y,y_1) = a(x,y) + b(x,y_1) \quad \forall x \in X \quad \forall y \in Y \quad \forall y_1 \in Y_1 .$$

On pose :

$$L(x,y) = \sup_{y_1 \in Y_1} \mathcal{L}(x;y,y_1) \qquad \forall x \in X \qquad \forall y \in Y .$$

On peut sous certaines conditions remplacer un problème initial sur $(X, Y \times Y_1, \mathcal{L})$ par un problème sur le triplet (X,Y,L). Ces problèmes peuvent donner lieu à de très nombreux développements algorithmiques qui ne seront pas étudiés ici.

(3.7.) PROCEDES DE MONTEE OBTENUS PAR "SYMETRIE".

Considérons $\tilde{L} : Y \times X \to \mathbb{R}$ tel que

$\tilde{L}(y,x) = -L(x,y) \quad \forall x \in X \quad \forall y \in Y$

Etant donné un triplet T1 : (X,Y,L) , si on le remplace par le triplet T2 : $(Y,X,\tilde{L})$, on voit que le problème P (resp Q) relatif à T1 est identique au problème Q (resp P) relatif à T2 et que les problèmes S relatifs à T1 et T2 sont identiques.

Ceci revient simplement à échanger les rôles de x et y et à changer les inf par des sup et inversement. Cette remarque triviale n'est pas dénuée d'intérêt car elle permet de générer simplement une nouvelle famille de procédés de montée "symétriques" des précédents. Il est inutile de décrire ces nouvelles méthodes. Précisons à titre d'exemple l'algorithme "symétrique" de l'algorithme (3.3.1.1.) :

A l'itération n on connait une partie compacte Y_n de Y .

PROBLEME PRINCIPAL :

trouver x^n solution du problème P relatif au triplet (X,Y_n,L) . $\quad a_n = \sup_{y \in Y_n} L(x^n,y)$.

SOUS PROBLEME :

trouver $\overline{y^n} \in Y$ telqque :

$b_n = L(x^n,\overline{y^n}) = \max (L(x^n,y) \mid y \in Y)$.

Arrêt si $a_n = b_n$, sinon on cherche Y_{n+1} compact tel que : $Y \supseteq Y_{n+1} \supseteq Y_n \cup \{\overline{y^n}\}$ et on passe à l'itération suivante.

DEUXIEME PARTIE

QUELQUES EXEMPLES D'APPLICATIONS AUX PROBLEMES D'OPTIMISATION

Nous ne donnons qu'un aperçu très rapide de quelques applications. Etant donné un problème d'optimisation, on le transforme d'abord en un problème de minimax (qui n'est pas unique), lequel sera résolu par les procédés de la première partie. Il peut être intéressant de donner une interprétation directe de ces méthodes.

(3.8.) PROBLEME D'OPTIMISATION CONVEXE. (*)

(3.8.1.)

Nous faisons d'abord un rappel très bref sur la théorie des fonctions convexes (voir par exemple ROCKAFELLAR [27] ou LAURENT [22]). Soit X une partie convexe fermée de $\mathbb{R}^n$ et f une fonction convexe, continue, à valeurs finies sur $\mathbb{R}^n$, et inf compacte sur X . On lui associe la fonction polaire f^* : $\mathbb{R}^n \to \bar{\mathbb{R}}$ définie par :

$$f^*(y) = \sup_{x \in R^n} ((x,y) - f(x)) .$$

(*) On lui associe un Lagrangien qui selon la terminologie de JOLY-LAURENT [19] correspond à des perturbations horizontales.

Cette fonction f^* est convexe et grâce aux hypothèses faites sur f, semi-continue inférieurement et propre ($\exists\, y \in \mathbb{R}^n : f^*(y)$ est fini, $\nexists\, y \in \mathbb{R}^n : f^*(y) = -\infty$) . Le domaine effectif de f^* noté dom f^* est un ensemble convexe mais qui n'est pas nécessairement fermé ni borné.

Pour tout $x \in \mathbb{R}^n$, le sous différentiel de f noté $\partial f(x)$ est un ensemble convexe, compacte et non vide et de plus sur tout convexe compact B de $\mathbb{R}^n$ l'ensemble

$\{y \in \mathbb{R}^n | y \in \partial f(x)$ avec $x \in B\}$ est compact. (Cf. [27] , théorème 24;7).

La fonction $(x,.) - f^*(.)$ atteint son maximum sur $\mathbb{R}^n$ en un élément y de $\partial f(x)$ et on a :

$$(1) \qquad f(x) = \sup_{y \in \mathbb{R}^n} (x,y) - f^*(y) \ .$$

Ces développements nous conduisent à considérer les problèmes attachés au triplet $(X, \text{dom } f^*, L)$ où :

$$(2) \qquad L(x,y) = (x,y) - f^*(y). \qquad \forall\, x \in X \qquad \forall\, y \in \text{dom } f^* \ .$$

Etant donné L défini ci-dessus, le problème P relatif à $(X, \text{dom } f^*, L)$ consiste à chercher un élément $\bar{x}$ pour lequel f atteint son minimum sur X .
Le problème S consiste à chercher un couple $(\bar{x},\bar{y})$ tel que :

$$\bar{x} \in X \ ; \ (\bar{y}, x-\bar{x}) \geq 0 \quad \forall\, x \in X \ ; \ \bar{y} \in \partial f(\bar{x})$$

si $X = \mathbb{R}^n$ alors $\bar{y} = 0$.

Les hypothèses H1 ne sont pas toutes satisfaites par le triplet $(X, \text{dom } f^*, L)$ mais il nous suffira de vérifier pour que tous les résultats énoncés en première partie concernant les algorithmes restent valables, que le problème S initial a au moins une solution et que les éléments $x^n, y^n, \overline{x^n}$ introduits dans les algorithmes existent et appartiennent à des bornés.

Dans la mise en oeuvre pratique de ces procédés on ne doit pas oublier que la fonction f^* n'est pas explicitée.

(3.8.2.) Une méthode de descente avec accumulation.

Il s'agit du procédé (3.3.3.2.) appliqué à notre problème. Soit V un voisinage convexe compact de $0 \in \mathbb{R}^n$. A l'itération n on connait X_n un ensemble convexe compact de $\mathbb{R}^n$. On a l'itération n :

a) Problème principal :

trouver (x^n, y^n) point de selle du triplet $(X_n, \text{dom } f^*, L)$ c'est-à-dire que :

$$x^n \in X_n \; ; \; (y^n, x-x^n) \geq 0 \quad \forall\, x \in X_n \; ; \; y^n \in \partial f(x^n)$$

(ces relations impliquent que $f(x^n) = \min (f(x) | x \in X_n)$).

b) Sous problème :

trouver $x^n \in (x^n+V) \cap X$ tel que :

$$(y^n, \overline{x^n}) = \min ((y^n, x) | x \in X \; ; \; x-x^n \in V).$$

(le terme $-f^*(y^n)$ est constant donc il n'apparait pas dans le min ci-dessus).

Si $(y^n, x^n - \overline{x^n}) = 0$ arrêt sinon trouver X_{n+1} convexe et compact tel que :

$$X \supseteq X_{n+1} \supseteq X_n \cup \{\overline{x^n}\}$$

et passer à l'itération suivante.

(3.8.3.) Cas où f est dérivable.

Supposons que la fonction f possède une dérivée continue telle que :

$$\lim_{\|x\| \to \infty} \|f'(x)\| = +\infty .$$

On dira comme ROCKAFELLAR [27] que la fonction f est essentiellement lisse. On sait alors (ROCKAFELLAR [27], théorème 26.3) que

la fonction polaire f^* est strictement convexe, donc $L(x,.)$ est strictement concave. On pourra appliquer à ce problème le procédé (3.4.5.). On retrouve en fait la méthode bien connue de FRANK et WOLFE que nous avons déjà rencontrée dans le paragraphe (3.4.6.).

(3.8.4.) Une méthode de montée par accumulation.

Il s'agit du procédé "symétrique" de la méthode du (3.3.1.1.).

A l'itération n, on connait une partie Y_n de Y comprenant un nombre fini d'éléments : $Y_n = \{\overline{y^1},\ldots,\overline{y^{n-1}}\}$ et on connait les nombres $\{f^*(\overline{y^i})\}$ $i = 1, \ldots, n-1$.

On suppose que la fonction : $\sup_{y \in Y_n} (.,y) - f^*(y)$ est minorée sur X.

Cette condition pose un problème de démarrage non évident pour l'algorithme.

On considère alors :

le problème principal :

trouver $x^n \in X$ solution du problème P relatif à (X,Y_n,L).

Soit $a_n = \inf_{x \in X} \sup_{y \in Y_n} L(x,y)$.

le sous problème :

On cherche $\overline{y^n} \in \text{dom } f^*$ tel que :

$$b_n = (x^n,\overline{y^n}) - f^*(\overline{y^n}) = \max \{(x^n,y) - f^*(y) \mid y \in \text{dom } f^*\} .$$

C'est-à-dire on cherche $\overline{y^n} \in \partial f(x^n)$ et on a de plus :

$f^*(\overline{y^n}) = (x^n,\overline{y^n}) - f(x^n)$.

Si $a_n = f(x^n)$ arrêt, sinon on fait $Y_{n+1} = Y_n \cup \{\overline{y^n}\}$ et on passe à l'itération suivante.

(3.9.) PROBLEMES D'OPTIMISATION CONVEXE AVEC CONTRAINTES INEGALITES (*)

(3.9.1.)

Soit le problème d'optimisation :

$$\min (f(x) \mid x \in X \; ; \; g_i(x) \leq 0 \qquad i=1,\ldots,m)$$

où f est une fonction convexe continue définie sur $\mathbb{R}^n$, X une partie convexe compacte de $\mathbb{R}^n$, les fonctions g_i sont convexes et continues. On suppose aussi (hypothèse de qualification) qu'il existe $z \in X$ tel que : $g_i(z) < 0$ pour tout i.

Soit $\quad Y = \{y \in \mathbb{R}^m \mid y_i \geq 0 \qquad i=1,\ldots,m\}$

On associe à ce problème la fonction $L : X \times Y \to R$ définie par :

$$L(x,y) = f(x) + \sum_{i=1}^{m} y_i \, g_i(x) \; .$$

On vérifie sans peine que le triplet (X,Y,L) vérifie les hypothèses H1, H"2 (X borné) et

H3 : Il existe $z \in X$ et $a \in \mathbb{R}$ tels que l'ensemble $\{y \in Y \mid L(z,y) \geq a\}$ soit non vide et borné.

Donc le théorème (3.2.1.) assure que le triplet (X,Y,L) possède un point de selle $(\bar{x},\bar{y})$ tel que $\bar{x}$ est une solution du problème d'optimisation initial et $\bar{y}$ un vecteur des coefficients de KUHN et TUCKER.

(*) On lui associe un Lagrangien qui selon la terminologie de JOLY-LAURENT [19] correspond à des perturbations verticales.

(3.9.2.) Méthodes de descente par accumulation.

(3.9.2.1.) Algorithme.

C'est algorithme (3.3.1.2.) adapté à notre exemple.

A l'itération n on considère un ensemble X_n convexe compact contenant z .

PROBLEME PRINCIPAL :

trouver (x^n,y^n) point de selle du triplet (X_n,Y,L)

c'est-à-dire tel que :

$$\begin{cases} x^n & \text{solution du problème } \min (f(x) \mid x \in X_n, g_i(x) \leq 0 \quad i=1,\dots,m) \\ y^n & \text{vecteur des coefficients de KUHN et TUCKER.} \end{cases}$$

SOUS PROBLEME :

trouver $\overline{x^n} \in X$ tel que :

$$b_n = f(\overline{x^n}) + (y^n, g(\overline{x^n})) = \min (f(x) + (y^n, g(x)) \mid x \in X)$$

Arrêt si $a_n = b_n$, sinon faire $X_{n+1} = \text{conv } (X_n \cup \{\overline{x^n}\})$ et passer à l'itération suivante.

(3.9.2.2.) Méthode des colonnes.

Il s'agit d'une variante du procédé précédent (cf. HUARD [5], GEOFFRION [15]). On en trouve le principe en remarque g) du (3.3.1.1.). C'est une modification du procédé précédent où le problème principal consiste à chercher le couple (x^n,y^n) point de selle du triplet (X_n,Y,L_n) où L_n est une fonction convexe concave continue sur $X_n \times Y$ telle que $L_n(.,y) = \text{conv } A(.,y)$

où $$A(x,y) = \begin{cases} L(x,y) \quad \forall\, y \in Y \quad \forall\, x = \overline{x^i} \quad i=1,\dots,(n-1); x = z \,. \\ +\infty \quad \text{sinon.} \end{cases}$$

(3.9.2.3.) Méthode du barycentre (adaptation de l'algorithme (3.3.1.3.). Cf. HUARD [18])

Même problème principal qu'en (3.9.2.1.).

On suppose de plus que les fonctions f et g_i possèdent des dérivées f' et g_i' continues sur X.

Le sous problème consiste à chercher $\overline{x^n} \in X$ solution de

$$\min \left((f'(x^n) + \sum_{i=1}^{m} y_i^n \, g_i'(x^n), x) \mid x \in X \right) .$$

(3.9.3.) Méthodes de montée par accumulation.

Il s'agit de méthodes symétriques des précédentes qui s'apparentent aux méthodes du chapitre II.

Nous remarquons que l'ensemble Y n'est pas borné.

(3.9.3.1.) Algorithme.

(Correspond au symétrique de (3.3.3.3.)).

A l'itération n on connait un ensemble Y_n fermé de Y.

PROBLEME PRINCIPAL :

On cherche x^n solution du problème P relatif à (X, Y_n, L). Soit $a_n = \sup_{y \in Y_n} L(x^n, y)$.

SOUS PROBLEME :

On cherche $\overline{y^n} \in Y$ tel que $(\overline{x^n}, \overline{y^n})$ soit un point de selle relatif à $([z, x^n], Y, L)$.

Si $a_n = b_n$ arrêt sinon on détermine Y_{n+1} fermé :

$$Y \supseteq Y_{n+1} \supseteq Y_n \cup \{\overline{y^n}\} .$$

(3.9.3.2.) Algorithme.

(Contient le sous problème du symétrique de l'algorithme (3.3.3.4.)).

Y_n est un cône fermé de sommet l'origine.

PROBLEME PRINCIPAL :

On cherche x^n solution du problème P relatif à (X,Y_n,L) c'est-à-dire solution du problème d'optimisation.

$$\min (f(x)\ x \in X\ ;\ (y,g(x)) \leq 0 \quad \forall\ y \in Y_n)$$

SOUS PROBLEME :

Trouver $\overline{y^n} \in V \wedge Y$ (V voisinage convexe compact de 0) tel que :

$$\max ((y,g(x^n))|y \in V \wedge Y) = (\overline{y^n},g(x^n)) .$$

Puis faire :

$$Y_{n+1} = Y_n \cup \{y = \lambda\overline{y^n}|\lambda \geq 0\} .$$

Par exemple :

$$V \wedge Y = \{y \in \mathbb{R}^m | y_i \in [0,1] \quad i=1,\ldots,m\} .$$

(3.9.4.) Méthodes de montée dans le cas où la fonction f est strictement convexe.

En plus des hypothèses du (3.9.1.) on suppose que f est strictement convexe. Nous donnons deux exemples de méthodes applicables dans ce cas.

(3.9.4.1.) Algorithme. (Cf. Algorithme (3.3.3.3.)).

On connait à l'itération n , y^{n-1} et $y^{n-1} \in Y$.

PROBLEME PRINCIPAL :

Trouver (x^n,y^n) point de selle du triplet $(X,[\overline{y^{n-1}},y^{n-1}],L)$.

SOUS PROBLEME :

Trouver $(\overline{x^n},\overline{y^n})$ point de selle du triplet $([z,x^n],Y,L)$.

(3.9.4.2.) Algorithme. (Cf. Algorithme (3.4.5.1.)).

PROBLEME PRINCIPAL :

Trouver (x^n,y^n) point de selle du triplet

$(X,$cône convexe $(\overline{y^{n-1}},y^{n-1}),L)$.

SOUS PROBLEME :

Trouver $\overline{y^n} \in V \wedge Y$ solution de :

$\max ((y,g(x^n))|y \in V \wedge Y)$.

(3.10) PROBLEMES D'OPTIMISATION NON CONVEXE AVEC CONTRAINTES.

Cf. ROODE [30].

(3.10.1.)

Soit le problème :

(1) $\min (f(x)|x \in X ; g_i(x) \leq 0 \quad i = 1,\ldots, m)$

où $\begin{cases} X \text{ est une partie compacte de } R^n \\ f \text{ est continue sur } R^n \\ g_i(x) \text{ sont continues sur } R^n . \end{cases}$

On considère une fonction $\ell : \mathbb{R} \to \bar{\mathbb{R}}^+$ non décroissante, continue en 0 et telle que : $\ell(\tau) = \begin{cases} 0 & \text{si } \tau \leq 0 \\ >0 & \text{sinon.} \end{cases}$

On considère la fonction $L : X \times Y \to R$ définie par :

$$L(x,y) = f(x) + \sum_{i=1}^{m} y_i \ \ell(g_i(x))$$

(l'ensemble Y est défini comme dans (3.9.)).

On voit facilement que le problème P relatif à (X,Y,L) est équivalent au problème (1).
En fait ce triplet permet de retrouver les méthodes extérieures étudiées au chapitre II en considérant des méthodes de montée.
Nous nous contentons de préciser un procédé par accumulation.

(3.10.2.) Algorithme.

A l'itération n on connait

$$Y_n = \{y \in R^m \mid 0 \leq y \leq q^{n-1}\}$$

où $q^{n-1} \in Y$.

Le problème principal consiste à trouver x^n solution du problème P relatif à (X,Y_n,L) soit chercher x^n tel que :

$$f_n(x^n) = \min (f_n(x) \mid x \in X)$$

où $f_n(x) = f(x) + \sum_{i=1}^{m} q_i^{n-1} \ \ell(g_i(x))$.

Le sous problème consiste à trouver $\overline{y^n} \in V \wedge Y$ tel que

$$\max (L(x^n,y) \mid y \in V \wedge Y) = L(x^n,\overline{y^n}).$$

V est défini comme en (3.9.3.2.).
On détermine ensuite :

$$Y_{n+1} = \{y \in \mathbb{R}^m \mid 0 \leq y \leq q^n\}$$

où $q^n = q^{n-1} + \overline{y^n}$.

Si $\sum_{i=1}^{m} y_i^n \ \ell(g_i(x^n)) = 0$ arrêt, sinon on passe à l'itération suivante.

BIBLIOGRAPHIE

[1] ARROW, K.J. ; HURWICZ, L. ; UZAWA, H.
Studies in linear and non linear programming
Stanford university press. (1958).

[2] AUSLENDER, A.
Méthodes numériques pour la résolution de problèmes d'optimisation avec contraintes.
Thèse. Faculté des Sciences de Grenoble. (Juin 1969).

[3] BENDERS, J.F.
Partitioning procedures for solving mixed variables programming problems.

[4] BENSOUSSAN, A.
Saddle points of convex concave functionals.
Differential games and related topics.
Edited by KUHN, N.W. North Holland (1971)
pages 177-200.

[5] BROISE, P ; HUARD, P. ; SENTENAC, J.
Décomposition des programmes mathématiques.
Monographie de recherche opérationnelle.
6. Dunod (1968).

[6] BUTZ, A.R.
Iterative saddle point techniques.
SIAM. App. Math. 15. 3 (1967) 719-725.

[7] CEA, J. ; GLOWINSKI, R. ; NEDELEC, J.C.
Dualité et problèmes aux dérivées partielles.
(A paraître).

[8] CEA, J. Optimisation. Théorie et application.
Dunod (1971).

|9| DEM'JANOV, V.F.
On a minimax problem.
Soviet math. dokl. vol. 10 n° 4 (1969) 828-832.

[10] FALK, J.
An algorithm for separable convex programming
under linear constraints.
Research analysis corporation. Technical paper 148.
Mac Lean, Virginia, march 1965.

[11] FALK, J.
Lagrange multipliers and non linear programming
Journal of math. anal. and appl. 19-1 (1967) 141-159.

[12] FIACCO, A.V. ; MAC CORMICK, G.P.
Non linear programming.
John Wiley and sons. New-York (1968).

[13] GEOFFRION, A.M.
Relaxation and the dual method in mathematical
programming. Western management science institute.
Working paper n° 135. University of California.
Los Angeles (march 1968)

[14] GEOFFRION, A.M.
Generalized Benders decomposition.
Proceeding of the symposium on non linear programming.
Math. research center. University of Wisconsin. Madison.

[15] GEOFFRION, A.M.
Elements of large scale mathematical programming.
Management science. 6. n° 11 (1970) 652-691.

[16] GEOFFRION, A.M.
Duality in non linear programming : a simplified applications oriented development.
SIAM Review vol. 13 n° 1 (1971).

[17] GEOFFRION, A.M.
Primal ressource directive approaches for optimizing non linear decomposable systems.
Management science. (1970) 375-403.

[19] JOLY, J.L. ; LAURENT,P.J.
Stability and duality in convex minimization problems.
R.I.R.O. R2 (1971) 3-42.

[20] KRONSJÖ, T.O.M.
Centralization and decentralization of decision making.
R.I.R.O. vol. 2 n° 10 (1968) 73-114.

[21] KRONSJÖ, T.O.M.
Decomposition of a large block triangular convex program into only master problems.
7th math. program. symp. sept 1970 The Hague.

[22] LAURENT, P.J.
Approximation et optimisation. Hermann (1972).

[23] LEMAIRE, B.
Problèmes minimax et applications au contrôle optimal des systèmes gouvernés par des équations aux dérivées partielles.
Thèse. Paris (1970).

[24] LEVITIN, E.S.
A general minimization method for unsmouth extremal problems.
U.S.S.R. comp. math. and phy. math.
9-4 (1969) 783-806.

[25] MOREAU, J.J.
Fonctionnelles convexes, séminaire
Collège de France. (1967).

[26] MOREAU, J.J.
Théorèmes inf-sup.
C.R.A.S. Paris 258 (1964) 2720-2722.

[27] ROCKAFELLAR, R.T.
Convex analysis. Princeton Univ. press (1970).

[28] ROCKAFELLAR, R.T.
New applications of duality in convex programming
7th math. progr. symp. (septembre 1970). The Hague.

[29] ROCKAFELLAR, R.T.
Saddle points and convex analysis. pages 109-127.

Differential games and related topics.
Edited by KUHN and SZEGÖ.
North Holland (1971).

[30] ROODE, J.D.
Generalized lagrangien functions in mathematical programming.
Thèse. Université de Leiden (Octobre 1968).

[31] SION, M.
On general minimax theorems, Pacific J. math. 8 (1958) 171-176.

[32] STOER ; WITZGALL
Convexity and optimization in finite dimension.
Springer Verlag (1970).

[33] TREMOLIERES, R.
Algorithmes pour la résolution de problèmes posés sous la forme min-max f" . Cas non linéaire.
Rapport n° 6909. IRIA. Rocquencourt (mai 1969).

[34] VALADIER, M.
Contribution à l'analyse convexe.
Thèse. Paris (1970).

[35] VARAIYA, P.
Decomposition of large scale systems.
J. SIAM control vol. 4 n° 1 (1966) 173-178.

CHAPITRE IV

MINIMISATION D'UNE FONCTIONNELLE SUR UN ESPACE PRODUIT PAR UNE METHODE DE RELAXATION.

(4.1.) INTRODUCTION.

Soient B_i $(i = 1, \ldots, m)$ des espaces de Banach réflexifs réels, C_i $(i = 1, \ldots, m)$ des convexes fermés de B_i , $C = \prod_{i=1}^{m} C_i$, f une fonction convexe à valeurs réelles définie sur $B = \prod_{i=1}^{m} B_i$ possédant une dérivée de FRECHET f' .

On se propose de résoudre le problème d'optimisation suivant :

(P) " trouver $\bar{x} \in C$ tel que $f(\bar{x}) = \min (f(x) | x \in C)$ " .

On emploie souvent pour résoudre ce problème des méthodes de décomposition qu'on peut considérer comme des extensions de la méthode bien connue de GAUSS-SEIDEL (voir GASTINEL [8]) dans le cas d'une fonction quadratique : nous les appellerons méthodes de relaxation.

Ces procédés ont été étudiés en particulier par AUSLENDER [1], CEA [5], MARTINET [12],[13] , MIELLOU' [15] , ORTEGA et RHEINBOLDT [16] , SCHECHTER [18] .

Ils consistent à remplacer le problème P par une suite en général infinie de problèmes d'optimisation P_{i_n} de petite taille. Cependant, ces derniers peuvent être très couteux. L'idée développée ici (cf. AUSLENDER - MARTINET [3], [4]) est la suivante : au lieu de résoudre le problème de minimisation P_{i_n} on se contentera d'appliquer à ce problème une seule itération d'une méthode de minimisation.

On associe ainsi divers méthodes d'optimisation classiques à la décomposition. Nous nous contentons ici de considérer la décomposition associée à la méthode de GAUSS-SEIDEL. (On pourrait aussi examiner d'autres modes de décomposition tels que ceux étudiés par AUSLENDER en [1]).

On considère dans ce chapitre deux classes de méthodes ainsi associées à la décomposition. La première classe comprend la plupart des méthodes de descente usuelles :
méthodes de descente de type gradient dans le cas sans contrainte cf. CEA [5] , méthodes de type gradient projeté (LEVITIN et POLYAK [11]), HILDRETH [9] , MARTINET [12]), les méthodes de descente par linéarisation [5] , etc ...

On montre que sous certaines conditions ces méthodes de descente peuvent être associées à la relaxation.

La deuxième classe n'a pas la même unité. Après avoir donné un théorème de convergence on cite comme exemple de tels procédés : les méthodes de type régularisation cf. TICKONOV [19] , les méthodes de type approximation extérieure de l'ensemble admissible (cf. chapitre II) .

(4.2.) NOTATIONS - HYPOTHESES - GENERALITES.

On note $\| \|_i$ la norme dans B_i et $(,)_i$ la dualité entre B_i et son dual B_i' . On note aussi la dualité entre B et B' et la norme dans B respectivement par $(,)$ et $\| \|$ avec :

$$(x,y) = \sum_{i=1}^{m} (x_i,y_i)_i \quad \text{et} \quad \|x\|^2 = \sum_{i=1}^{m} \|x_i\|_i^2$$

pour tout $x = (x_1, \ldots, x_m) \in B$ $\qquad x_i \in B_i$

et tout $y = (y_1, \ldots, y_m) \in B'$ $\qquad y_i \in B_i'$.

En fait si le contexte ne prête pas à confusion nous pourrons omettre dans ces notations les indices inférieurs.

On notera indifféremment les valeurs de f par :

$f(x)$ ou $f(x_1, \ldots, x_m)$

Nous ferons les hypothèses supplémentaires :

H1 : Pour tout $a \in \mathbb{R}$ l'ensemble :

$\{x \in C \mid f(x) \leq a\}$ est borné.

H2 : Pour tout borné D de B il existe une constante $M > 0$ telle que :
$\|f'(x)-f'(y)\| \leq M\|x-y\| \quad \forall\, x, y \in D$.

On se propose de résoudre le problème suivant :

(P) " trouver $\bar{x} \in C$ tel que $f(\bar{x}) = \min (f(x) \mid x \in C) = v$ " .

Les hypothèses faites entraînent que le problème (P) a au moins une solution car la fonctionnelle f est faiblement semi continue inférieurement et pour tout $a \in \mathbb{R}$ l'ensemble $\{x \in C \mid f(x) \leq a\}$ est faiblement compact.

Mais il n'y a pas en général unicité. On notera :

$$M = \{x \in C \mid f(x) = v\} .$$

Puisque la fonctionnelle f est convexe et qu'elle est dérivable, toute solution $\bar{x}$ du problème (P) est telle que pour tout $i = 1, \ldots, m$ et tout $x_i \in C_i$ on ait :

$$(f'_i(\bar{x}) , x_i - \bar{x}_i) \geq 0 .$$

(On note ici f'_i la dérivée partielle de f sur l'espace B_i , et $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_m)$) .

(4.3.) ALGORITHMES DE DECOMPOSITION - GENERALITES.

Les procédés étudiés ici permettent de construire à partir d'un élément x^0 de B une suite $\{x^n\}$ d'éléments de B . Soit :

$$x^n = (x_1^n, \ldots, x_m^n) \quad \text{où} \quad x_i^n \in B_i .$$

Pour $i = 1, 2, \ldots, m$, on note φ_i^n l'application de B_i dans $\mathbb{R}$ définie par :

$$\varphi_i^n(v_i) = f(x_1^{n+1}, \ldots, x_{i-1}^{n+1}, v_i, x_{i+1}^n, \ldots, x_m^n)$$

et $\varphi_i^{n'}$ la dérivée de φ_i^n .

Dans les procédés que nous étudions, on associe à la relaxation m méthodes d'optimisation Q_i , $i = 1, \ldots, m$ de la façon suivante :

on passe de x^n à x^{n+1} en déterminant successivement pour $i = 1, 2, \ldots, m$ l'élément x_i^{n+1} à partir de x_i^n en effectuant une seule itération de la méthode Q_i appliquée au problème de minimisation de φ_i^n sur C_i .

Les procédés considérés permettent de générer une suite $\{x^n\}$ qui est minimisante c'est-à-dire telle que :

$$\lim_{n\to\infty} f(x^n) = v \quad \text{et} \quad \lim_{n\to\infty} d(x^n,C) = 0 \ .$$

(On désigne par $d(x^n,C)$ la distance de x^n à C).

Les hypothèses faites entrainent alors que toute valeur d'adhérence faible de la suite $\{x^n\}$ (il en existe toujours une car la suite est bornée) est solution du problème P .

(4.4.) METHODES DE DESCENTE ASSOCIEES A LA RELAXATION. GENERALITES.

(4.4.1.)

On associe à la relaxation une méthode de descente Q qui sera la même pour chaque composante. On génère une suite $\{x^n\}$ d'éléments de C tels qu'à l'itération n , on ait :

$$\varphi_i^n(x_i^{n+1}) \leq \varphi_i^n(x_i^n) \qquad \forall\ i \ .$$

On a donc :

$$f(x^{n+1}) \leq f(x^n) \ .$$

Si $x^{n+1} = x^n$ on arrête les itérations, sinon on passe à l'itération suivante.

Si la suite $\{x^n\}$ générée est infinie, on a puisque la suite $\{f(x^n)\}$ est bornée inférieurement :

$$\begin{cases} \lim_{n\to\infty} f(x^n) \ \text{existe} \\ \lim_{n\to\infty} f(x^{n+1}) - f(x^n) = 0 \\ \lim_{n\to\infty} \varphi_i^n(x_i^{n+1}) - \varphi_i^n(x_i^n) = 0 \qquad \forall\ i \ . \end{cases}$$

Dans certains cas on sera amené à faire l'hypothèse supplémentaire suivante :

H3 : Pour tout borné convexe D de B dont la projection sur B_i est notée D_i , il existe une application γ de R^+ dans R^+ ne dépendant que de D et f et vérifiant : (*)

$$\lim_{K\to\infty} t_K = 0 \longleftrightarrow \lim_{K\to\infty} (t_K) = 0 ,$$

telle que :

$\forall\, i = 1,2,\ldots,m$, $\forall\, x \in D$, $\forall\, y_i, z_i \in D_i$, $\forall \lambda \in [0,1]$

on ait, si on note :

$$\Psi_{x,i}(y) = f(x_1,\ldots,x_{i-1},y,x_{i+1},\ldots,x_m) ,$$

la relation suivante :

$$\Psi_{x,i}(\lambda y_i+(1-\lambda)z_i) \leq \lambda\Psi_{x,i}(y_i) + (1-\lambda)\Psi_{x,i}(z_i) - \lambda(1-\lambda)\gamma(\|y_i-z_i\|) .$$

On montrera pour ces méthodes le théorème de convergence suivant :

(4.4.2.) Théorème :

La suite $\{f(x^n)\}$ est décroissante.
Si $x^{n+1} = x^n$, alors x^n est solution du problème P .
La suite $\{x^n\}$ est minimisante et ses valeurs d'adhérence faible sont solutions du problème P .

Soit $S_o = \{x \in C \mid f(x) \leq f(x^o)\}$.
L'ensemble S_o est borné grâce à H1 . On note M_o (hypothèse H2) la

---

(*) Une telle application sera appelée fonction d'influence dans la suite du chapitre.

constante positive telle que :

$$\|f'(x) - f'(y)\| \leq M_o \|x-y\| \quad \forall\ x,y \in S_o \ .$$

Remarque :

Dans le paragraphe qui suit on donne quelques méthodes de descente qui peuvent être associées à la relaxation. On n'a pas étudié ici le problème de savoir quelles sont les propriétés générales des procédés de descente qui peuvent être associés à la relaxation, ainsi on présente indépendamment les uns des autres les différents algorithmes.

(4.5.) METHODES DE DESCENTE ASSOCIEES A LA RELAXATION. EXEMPLES.

Pour plus de détails sur ces questions on pourra se reporter à l'article de AUSLENDER-MARTINET [3] . Pour chacun de ces algorithmes on le décrit et on montre que la suite qu'il permet de générer vérifie le théorème (4.4.2.).

(4.5.1.) Exemple. La méthode associée est un procédé de descente de type gradient dans le cas sans contrainte.

On suppose ici que $C = B$. On donne des fonctions d'influence δ_1 et δ_2 .

Décrivons l'itération n . Pour $i=1,\ldots,m$ on détermine :

$$x_i^{n+1} = \begin{cases} x_i^n & \text{si} \quad \varphi_i'^n(x_i^n) = 0 \\ x_i^n - \rho_i^n W_i^n & \text{sinon} \ . \end{cases}$$

On désigne respectivement par W_i^n et ρ_i^n la direction de descente et la longueur de déplacement. On choisit d'abord (cf. CEA [5]) $W_i^n \in B$ vérifiant :

$$\|W_i^n\| = 1 \ .$$

(A) $\quad a_i^n = (\varphi'^n_i(x_i^n),\ W_i^n) > 0$

(B) $\quad x_i^n \in S_o \Rightarrow \|\varphi_i'^n(x_i^n)\| \leq \delta_1(a_i^n)$,

puis on choisit le scalaire ρ_i^n tel que :

(C) $\quad \rho_i^n > 0$

(D) $\quad x_i^n \in S_o \Rightarrow \varphi_i^n(x_i^n) - \varphi_i^n(x_i^{n+1}) \geq \delta_2(a_i^n)$.

Montrons que si on fait les hypothèses H3 et celles du (4.2.), cet algorithme satisfait au théorème (4.4.2.).

Démonstration du théorème (4.4.2.) :

Si $x^n = x^{n+1}$ on a par construction : $\varphi_i'^n(x_i^n) = 0 \quad \forall\ i$ par conséquent : $f'(x^n) = 0$ et x^n est solution du problème P .

Supposons maintenant que la suite $\{x^n\}$ soit infinie. On a :

(E) $\quad \lim_{n\to\infty} \varphi_i^n(x_i^{n+1}) - \varphi_i^n(x_i^n) = 0 \quad \forall\ i$

Grâce à (E), (D) et (B) il vient :

(F) $\quad \lim_{n\to\infty} \|\varphi_i'^n(x_i^n)\| = 0 \qquad \forall\ i$.

D'après l'hypothèse d'uniforme convexité de f (hypothèse H3) il existe une fonction d'influence γ telle que :

$$\gamma(\|x_i^{n+1}-x_i^n\|) + (\varphi_i'^n(x_i^n)\ ,\ x_i^{n+1}-x_i^n) \leq \varphi_i^n(x_i^{n+1}) - \varphi_i^n(x_i^n) .$$

La suite $\{x^n\}$ étant bornée, les relations (E) et (F) entrainent que :

$$\lim_{n\to\infty} \|x^{n+1}-x^n\| = 0 .$$

Et, puisque la suite $\{x^n\}$ est bornée, et que f' est lipchitzienne, il vient de la relation (F) que :

$$\lim_{n\to\infty} \|f'(x^n)\| = 0 .$$

La suite $\{x^n\}$ est donc minimisante puisqu'elle est bornée et que f est convexe.

C.Q.F.D.

Nous n'insisterons pas sur le choix pratique de ρ_i^n et W_i^n qui se fait exactement comme dans le cas usuel d'application des méthodes de type gradient en considérant la fonction φ_i^n . (Cf. CEA [5]).

(4.5.2.) Exemple. La méthode associée est un procédé de type FRANK-WOLFE.

Cf. CEA [5] .

On suppose que C est borné et que f vérifie l'hypothèse H3. Décrivons l'itération n . Pour $i=1,\ldots,m$, on détermine à partir de x_i^n l'élément $z_i^n \in C_i$ tel que :

$$(\varphi_i'^n(x_i^n), z_i^n - x_i^n) = \min(\varphi_i'^n(x_i^n), x_i - x_i^n) \mid x_i \in C_i)$$

puis l'élément $x_i^{n+1} \in [x_i^n, z_i^n]$ tel que :

$$\varphi_i^n(x_i^{n+1}) = \min(\varphi_i^n(x_i^n + \lambda(z_i^n - x_i^n)) \mid \lambda \in [0,1])$$

$$= \varphi_i^n(x_i^n + \lambda_i^n(z_i^n - x_i^n)).$$

Démonstration du théorème (4.4.2.) :

Si $x^n = x^{n+1}$ on a :

$$(f'(x^n), x - x^n) \geq 0 \qquad \forall\, x \in C$$

donc x^n est solution du problème P .

Supposons maintenant que la suite $\{x^n\}$ soit infinie.

Posons $\Psi_i^n(\rho) = \varphi_i^n(x_i^n + \rho(z_i^n - x_i^n))$.

On a :

$$\Psi_i^n(0) - \Psi_i^n(\lambda) \leq \varphi_i^n(x_i^n) - \varphi_i^n(x_i^{n+1}) \qquad \forall\, \lambda \in [0,1]$$

Par conséquent, par passage à la limite en n , il vient :

$$\limsup_{n \to \infty} \frac{\Psi_i^n(0) - \Psi_i^n(\lambda)}{\lambda} \leq 0 \qquad \forall\, \lambda \in [0,1]$$

Comme f' vérifie une condition de Lipchitz sur C , en faisant tendre λ vers 0, il vient :

(A) $$\liminf_{n \to \infty} (\varphi_i'^n(x_i^n),\ z_i^n - x_i^n) \geq 0 .$$

Et par définition de z_i^n :

(B) $$\liminf_{n \to \infty} (\varphi_i'^n(x_i^n),\ x_i - x_i^n) \geq 0 \qquad \forall\, x_i \in C_i .$$

De l'hypothèse H3 on déduit qu'il existe une fonction d'influence γ telle que :

$$\gamma(\|x_i^{n+1} - x_i^n\|) + \lambda_i^n(\varphi_i'^n(x_i^n), z_i^n - x_i^n) \leq \varphi_i^n(x_i^{n+1}) - \varphi_i^n(x_i^n) .$$

De (A) et des propriétés de γ on tire que :

(C) $$\lim_{n \to \infty} \|x_i^{n+1} - x_i^n\| = 0 \qquad \forall\, i .$$

Comme f' satisfait à une condition de Lipchitz, on déduit de (B) et (C) que :

(D) $$\liminf_{n \to \infty} (f'(x^n), x - x^n) \geq 0 \qquad \forall\, x \in C .$$

Or on sait que :

$$f(x) \geq f(x^n) + (f'(x^n), x-x^n) \qquad \forall\, x \in C \ .$$

D'où on tire de (D) que :

$$\lim_{n\to\infty} f(x^n) \leq f(x) \qquad \forall\, x \in C \ .$$

Donc :

$$\lim_{n\to\infty} f(x^n) = v$$

puisque : $x^n \in C \ \forall\, n$.

C.Q.F.D.

Remarque :

Par analogie avec CEA [5] , on peut étendre cet algorithme en considérant d'autres choix du scalaire λ_i^n qui conduisent à des techniques plus efficaces en pratique. (Tabulation par pas etc...).

(4.5.3.) Exemple. La méthode associée est un procédé de gradient projeté.

Cf. LEVITIN-POLYAK [11] , HILDRETH [9] , MARTINET [12] etc... .

On suppose que B est un espace de Hilbert et qu'il existe $Q > 0$ tel que :

$$\|f'(x) - f'(y)\| \leq Q \ \|x-y\| \qquad \forall\, x,y \in B \ .$$

Nous ne ferons pas dans ce paragraphe l'hypothèse H3. On se donne des nombres $\varepsilon_1 > 0$, $\varepsilon_2 > 0$ tels que :

$$\varepsilon_1 < \frac{2}{Q + 2\,\varepsilon_2} \ .$$

On note P_{C_i} l'opérateur de projection sur C_i .
Décrivons la $n^{ème}$ itération du procédé. Pour $i=1,\ldots, m$ on détermine :

$$x_i^{n+1} = P_{C_i}(x_i^n - \alpha_i^n \varphi_i'^n(x_i^n))$$

où

$$\alpha_i^n \in [\alpha_1 , \frac{2}{Q + 2\varepsilon_2}] \quad .$$

Démonstration du théorème (4.4.2.) :

Si $x^{n+1} = x^n$ il vient :

$$x_i^n = P_{C_i}(x_i^n - \alpha_i^n f_i'(x^n)) \qquad \forall\, i$$

et, d'après les propriétés de la projection et le fait que $\alpha_i^n > 0$, on a :

$$(f_i'(x^n), x_i - x_i^n) \geq 0 \qquad \forall\, x_i \in C_i \qquad \forall\, i$$

ce qui entraine que x^n est solution du problème P .

On suppose maintenant que la suite générée $\{x^n\}$ est infinie.
Soit $a_i^n = \varphi_i^n(x_i^{n+1}) - \varphi_i^n(x_i^n)$.
On a :

$$a_i^n = (\varphi_i'^n(x_i^n),\ x_i^{n+1} - x_i^n) +$$

$$\int_0^1 (\varphi_i'^n(x_i^n + \tau(x_i^{n+1} - x_i^n)) - \varphi_i'^n(x_i^n),\ x_i^{n+1} - x_i^n)\, d\tau \ .$$

$$\leq (\varphi_i'^n(x_i^n),\ x_i^{n+1} - x_i^n) + \frac{Q}{2} \|x_i^{n+1} - x_i^n\|^2$$

$$\leq (\frac{Q}{2} - \frac{1}{\alpha_i^n}) \| x_i^{n+1} - x_i^n \|^2 .$$

Donc :

$$(A) \qquad \varphi_i^n(x_i^{n+1}) - \varphi_i^n(x_i^n) \leq - \varepsilon_2 \, \|x_i^{n+1} - x_i^n\|^2$$

Cela entraine que :

$$\varphi_i^n(x_i^{n+1}) \leq \varphi_i^n(x_i^n) \qquad \forall\, n \quad \forall\, i \ .$$

Donc :

$$(B) \qquad \lim_{n \to \infty} \left(\varphi_i^n(x_i^n) - \varphi_i^n(x_i^{n+1}) \right) = 0 \qquad \forall\, i \ .$$

D'où on déduit de (A) que :

$$(C) \qquad \lim_{n \to \infty} \|x^{n+1} - x^n\| = 0 \ .$$

Les propriétés de projection entrainent que :

$$(x_i^{n+1} - x_i^n + \alpha_i^n \varphi_i'^n(x_i^n), \ x_i - x_i^{n+1}) \geq 0 \qquad \forall\, x_i \in C_i \qquad i=1,\ldots,m \ ,$$

d'où on tire que :

$$\frac{(x_i^{n+1} - x_i^n, \ x_i - x_i^{n+1})}{\alpha_i^n} + (\varphi_i'^n(x_i^n) - f_i'(x^{n+1}), \ x_i - x_i^{n+1})$$

$$+ (f_i'(x_i^{n+1}), \ x_i - x_i^{n+1}) \geq 0 \qquad \forall\, x_i \in C_i \qquad \forall\, i \ .$$

Comme f' est uniformément continue sur C, que la suite $\{x^n\}$ est bornée et qu'on a $\alpha_i^n \geq \alpha_1 > 0$, la relation (C) entraine que :

$$\liminf_{n \to \infty} (f'(x^n), x - x^n) \geq 0 \qquad \forall\, x \in C \ .$$

C.Q.F.D.

Remarque :

Dans le cas sans contrainte la méthode associée est une méthode de gradient particulière. L'algorithme obtenu ne nécessite pas de faire l'hypothèse H3. Un cas particulier d'une telle situation se présente tout naturellement lorsque $B_i = \mathbb{R}$; on notera l'importance pratique de ce cas.

(4.5.4.)

On trouvera en [3] d'autres exemples de méthodes de descente associées à la relaxation, par exemple un procédé de linéarisation.

(4.6.) AUTRES METHODES ASSOCIEES A LA RELAXATION. ETUDE GENERALE.

(4.6.1.)

On suppose que B_i est un espace de Hilbert et que C_i est borné. On note $P_{C_i}(x_i)$ la projection de x_i sur C_i et

$$d(x_i,C_i) = \|x_i - P_{C_i}(x_i)\|_i \ .$$

Nous étudions dans ce paragraphe des procédés qui consistent à construire des suites $\{x^n\}$ qui vérifient pour tout i , les deux conditions suivantes :

(1) $$\lim_{n\to\infty} d(x_i^n,C_i) = 0$$

(2) $$\liminf_{n\to\infty} (\varphi_i^{n'}(x_i^{n+1}), z_i^n - x_i^{n+1})_i \geq 0 \quad . \quad \forall\, z_i^n \in C_i \ .$$

Avant de décrire des algorithmes effectifs qui permettent de générer des suites ayant ces propriétés nous donnons un théorème de convergence.

(4.6.2.) Théorème :

Soit une suite $\{x^n\}$ vérifiant les conditions (1) et (2). Les hypothèses du (4.1.) et H1, H2, H3 entrainent qu'une telle suite est minimisante et que toute valeur d'adhérence faible de cette suite est solution de (P).

Démonstration :

Grâce à (2) et à la convexité de φ_i^n on a :

(A) $$\limsup_{n \to \infty} \varphi_i^n(x_i^{n+1}) - \varphi_i^n(P_{C_i}(x_i^n)) \leq 0 \quad \forall\, i$$

De (1) il vient que la suite $\{x^n\}$ est bornée.

On déduit de (1), de (A) et de l'uniforme continuité de φ_i^n qu'il existe une suite $\{\eta_n\}$ de réels positifs tels que :

$$\lim_{n\to\infty} \eta_n = 0 \quad \text{et}$$

(B) $$\varphi_i^n(x_i^{n+1}) \leq \varphi_i^n(x_i^n) + \eta_n \qquad \forall\, n \quad \forall\, i=1,\ldots,m\ .$$

Posons :

$$\begin{cases} \psi_i^n = \varphi_i^n(x_i^{n+1}) - \varphi_i^n(x_i^n) \\ A_n = \{i \mid \psi_i^n \geq 0\} \quad , \quad B_n = \{i \mid \psi_i^n < 0\}\ . \end{cases}$$

Par construction et d'après (B) , on a :

(C) $$f(x^{n+1}) - f(x^n) = \sum_{i \in A_n} |\psi_i^n| - \sum_{i \in B_n} |\psi_i^n|$$

(D) $$|\psi_i^n| \leq \eta_n \qquad \forall\, i \in A_n$$

(E) $$|\psi_i^n| \leq (m-1)\,\eta_n + f(x^n) - f(x^{n+1}) \qquad \forall\, i \in B_n\ .$$

Soit $S = \{n \mid f(x^n) \leq f(x^{n+1})\}$.

Nous distinguons successivement le cas où le cardinal de S est infini de celui où il ne l'est pas.

a) Le cardinal de S est infini.

α) Pour tout $n \in S$ on a d'après (E) et (D) :

(F) $$|\Psi_i^n| \leq (m-1)\,\eta_n \ .$$

Dans ce cas, l'ensemble A_n ne peut être vide alors que B_n peut l'être. De (F) on tire que :

$$\lim_{n\to\infty} \Psi_i^n = 0 \qquad \forall\, i \ .$$

Grâce à l'hypothèse H3, on a :

$$\varphi_i^n(x_i^n) \geq \varphi_i^n(x_i^{n+1}) + (\varphi_i'^n(x_i^{n+1}), x_i^n - x_i^{n+1})_i + \gamma(\|x_i^{n+1} - x_i^n\|_i) \ .$$

Si on pose $C_i^n = (\varphi_i'^n(x_i^{n+1}), x_i^n - P_{C_i}(x_i^n))_i$,

on a , compte tenu de (1) et (2) :

(G) $$-\Psi_i^n \geq \varepsilon_n + C_i^n + \gamma(\|x_i^{n+1} - x_i^n\|_i) \ ,$$

la suite $\{\varepsilon_n\}$ vérifiant :

$$\liminf_{n\to\infty} \varepsilon_n \geq 0 \ .$$

On tire de (1) et (G) que :

(H) $$\lim_{n\to\infty} \|x_i^{n+1} - x_i^n\|_i = 0 \qquad \forall\, i \ .$$

Donc de (2) et du fait que la fonctionnelle $\varphi_i'^n$ est uniformément continue on déduit que :

$$\liminf_{n\to\infty} (f'(x^n), x - x^n) \geq 0 \qquad \forall\, x \in C \ .$$

Grâce à la convexité de f cela entraine que :

$$(I) \qquad f(\bar{x}) \leq \limsup_{\substack{n \to \infty \\ n \in S}} f(x^n) \qquad \forall\, \bar{x} \text{ solution de } (P) .$$

Soit $\tilde{x}$ une valeur d'adhérence faible de la suite $\{x^n\}_{n \in S}$ (il en existe au moins une).
D'après (1) $\tilde{x} \in C$ et comme f est faiblement semi-continue inférieurement, on a :

$$f(\bar{x}) \leq f(\tilde{x}) \leq \liminf_{\substack{n \to \infty \\ n \in S}} f(x^n) .$$

D'où on tire de (I) que :

$$(J) \qquad f(\bar{x}) = \lim_{\substack{n \to \infty \\ n \in S}} f(x^n) .$$

β) Si le complémentaire de S est vide ou de cardinal fini le théorème est démontré. Dans le cas contraire, à tout $n \notin S$ on fait correspondre les nombres $k(n)$ et $j(n)$ tels que :

$$k(n) \in S \quad , \quad k(n) = \min (\ell \mid \ell > n \ , \ \ell \in S)$$

$$j(n) \in S \quad , \quad j(n) = \max (\ell \mid \ell < n \ , \ \ell \in S) .$$

Par définition, on aura :

$$\begin{cases} f(x^n) \geq f(x^{k(n)}) \\ f(x^n) \leq f(x^{j(n)}) + m. \eta_{j(n)} . \end{cases}$$

Or on tire de (J) que :

$$\lim_{n \to \infty} f(x^{k(n)}) = \lim_{n \to \infty} f(x^{j(n)}) = f(\bar{x}) ,$$

d'où on déduit que :

$$\lim_{n \to \infty} f(x^n) = f(\bar{x}) .$$

b) Le cardinal de S est fini :

Dans ce cas, il existe n_o tel que pour tout $n \geq n_o$, on ait :

$$f(x^{n+1}) < f(x^n) .$$

Cela entraine, puisque la suite $\{f(x^n)\}$ est bornée inférieurement (grâce à (1)) que :

$$\lim_{n\to\infty} (f(x^{n+1}) - f(x^n)) = 0 .$$

Des relations (D) et (E) il vient alors que :

$$\lim_{n\to\infty} \Psi_i^n = 0 .$$

La fin de la démonstration se fait comme dans la partie a).

C.Q.F.D.

(4.7.) AUTRES METHODES ASSOCIEES A LA RELAXATION. EXEMPLES.

(4.7.1.)

Nous explicitons maintenant des algorithmes qui permettent de générer les suites $\{x^n\}$ qui vérifient :

(1) $$\lim_{n\to\infty} d(x_i^n, C_i) = 0 .$$

(2) $$\liminf_{n\to\infty} (\varphi_i^{n'}(x_i^{n+1}), z_i^n - x_i^{n+1})_i \geq 0 \qquad \forall\, z_i^n \in C_i .$$

(On remarquera que tous les procédés du (4.5.) ont ces propriétés). Le théorème (4.6.2.) entraine alors la convergence de tels procédés.

Dans les exemples qui suivent on suppose qu'on associe à chaque composante le même procédé Q .

(4.7.2.) On associe à la relaxation une méthode de régularisation.

Soit une suite $\{\eta_n\}$ de réels positifs et tels que :

$$\lim_{n\to\infty} \eta_n = 0 .$$

L'algorithme est le suivant. On trouve x^{n+1} à partir de x^n en déterminant pour $i = 1, \ldots, m$ x_i^{n+1} tel que :

$$x_i^{n+1} \in C_i \quad : \quad \varphi_i^n(x_i^{n+1}) + \eta_n \|x_i^{n+1}\|_i^2 = \min \left(\varphi_i^n(x_i) + \eta_n \|x_i\|_i^2 \mid x_i \in C_i \right) .$$

La relation (1) est satisfaite trivialement puisque $x_i^{n+1} \in C_i$.

D'autre part, les conditions nécessaires et suffisantes d'optimalité entrainent que :

$$(\varphi_i^{n'}(x_i^{n+1}),\ z_i^n - x_i^{n+1})_i + 2\,\eta_n (x_i^{n+1}, z_i^n - x_i^{n+1})_i \geq 0 \qquad \forall\ z_i^n \in C_i .$$

D'où la relation (2) puisque C_i est borné.

(4.7.3.) On associe une méthode de pénalisation extérieure.

On suppose que chaque ensemble C_i est définie par :

$$C_i = \{x_i \in D_i \mid g_i(x_i) \leq 0\} \qquad i = 1, \ldots, m .$$

et que :

a) D_i est un sous ensemble convexe fermé borné de B_i

b) g_i est une fonctionnelle définie sur B_i convexe et continue

c) l'ensemble C_i est défini "correctement" (cf. LEVITIN [11], AUSLENDER [2])

C'est-à-dire que :

pour toute suite $\{x_i^n\} \subset D_i$ et telle que :

$$\limsup_{n \to \infty} g_i(x_i^n) \leq 0$$

on a :

$$\lim_{n\to\infty} d(x_i^n, C_i) = 0 .$$

C'est le cas si compte tenu des hypothèses précédentes, on suppose de plus que D_i est compact ou bien qu'il existe $z_i \in D_i$ tel que :

$$g_i(z_i) < 0 .$$

d) $\alpha = \sup (f(x) \mid x \in D)$ est fini.

On définit alors l'algorithme suivant :

On se donne n suites de réels positifs $\{K_i^n\}$ $i = 1, \ldots, m$ tels que $\lim_{n\to\infty} K_i^n = +\infty$ et des applications ℓ_i de $\mathbb{R}$ dans $\mathbb{R}^+$ non décroissantes, convexes, dérivables et telles que :

$$\ell_i(\tau) = \ell_i'(\tau) = 0 \quad \text{si} \quad \tau \leq 0$$
$$\ell_i(\tau) > 0 \quad \text{si} \quad \tau > 0 .$$

Par exemple :

$$\ell_i(\tau) = (\tau^+)^2 .$$

On pose $\Psi_i^n(x_i) = \varphi_i^n(x_i) + K_i^n \ell_i(g_i(x_i))$.

On construit pour $i = 1, \ldots, m$ l'élément x_i^{n+1} tel que :

$$x_i^{n+1} \in D_i , \quad \Psi_i^n(x_i^{n+1}) = \min (\Psi_i^n(x_i) \mid x_i \in D_i) .$$

Nous devons montrer maintenant que cette méthode vérifie les relations (1) et (2) du (4.7.1.).

α) Par construction, on a :

$$\Psi_i^n(x_i^{n+1}) \leq \varphi_i^n(x_i) \qquad \forall\, x_i \in C_i$$

car $\ell_i(g_i(x_i)) = 0$ si $x_i \in C_i$.

Donc :

$$0 \leq K_i^n \, \ell_i(g_i(x_i^{n+1})) \leq \varphi_i^n(x_i) - \varphi_i^n(x_i^{n+1}) \qquad \forall\, x_i \in C_i$$

et si on pose :

$$\beta = \inf\,(f(X) \,|\, x \in D) > -\infty \,,$$

il vient :

$$0 \leq K_i^n \, \ell_i(g_i(x_i^{n+1})) \leq \alpha - \beta \,.$$

Les propriétés de K_i^n et de ℓ_i entrainent que :

$$\limsup_{n \to \infty} g_i(x_i^{n+1}) \leq 0$$

et comme C_i est défini correctement :

$$\lim_{n \to \infty} d(x_i^n, C_i) = 0 \,.$$

β) On peut montrer (cf. LESCARRET [10] proposition 2) que l'application : $x_i \to \ell_i(g_i(x_i))$ est convexe et sous différentiable. De plus si $\partial g_i(x_i)$ désigne la valeur du sous différentiel de g_i au point x_i, on a la formule :

$$\partial(\ell_i(g_i(x_i))) = \ell_i'(g_i(x_i))\partial(g_i(x_i)) \,.$$

Par définition de x_i^{n+1}, on a :

(A) $$(\varphi_i'^n(x_i^{n+1}), x_i - x_i^{n+1})_i + K_i^n \ell'(g_i(x_i^{n+1}))(d_i^n, x_i - x_i^{n+1})_i \geq 0 \quad \forall\, x_i \in D_i$$

où $d_i^n \in \partial g_i(x_i^{n+1})$. (Ce sous différentiel au point x_i^{n+1} est non vide car g_i est continue).

Or :

$$\forall\, x_i \in C_i \quad \forall\, d_i \in \partial g_i(x_i) \quad \forall\, d_i^n \in \partial g_i(x_i^{n+1})$$

on a , grâce à la convexité de la fonction $\ell_i(g_i(.))$, :

$$0 = \ell_i'(g_i(x_i))(d_i, x_i - x_i^{n+1})_i \geq \ell_i'(g_i(x_i^{n+1}))(d_i^n, x_i - x_i^{n+1})_i \,.$$

Ce qui compte tenu de (A) démontre la relation (2) de (4.7.1.).

Remarque :

On peut étendre ces résultats aux procédés par approximation extérieure du chapitre II. Nous allons en voir maintenant un cas particulier.

(4.7.4.) On associe à la relaxation une méthode d'accumulation des contraintes.

Hypothèses et notations :

On suppose maintenant que C_i est défini par :

$$C_i = \{x_i \in D_i \mid g_i(v_i,x_i) \leq 0 \quad \forall\, v_i \in V_i\} \ .$$

et que :

a) D_i est une partie convexe fermée bornée de B_i

b) V_i est un ensemble compact

c) Pour tout $v_i \in V_i$ la fonctionnelle $g(v_i,.)$ est convexe et semi-continue inférieurement

c) L'ensemble C_i est défini de façon correcte c'est-à-dire que pour toute suite $\{x_i^n\}$ telle que :

$$\limsup_{n \to \infty} g_i(v_i,x_i^n) \leq 0 \quad \forall\, v_i \in V_i \ ,$$

on a :

$$\lim_{n\to\infty} d(x_i^n,C_i) = 0 \ .$$

e) Les fonctionnelles $g_i(.,x_i)$ sont équi-uniformément continues sur D_i c'est-à-dire que : $\forall i$, $\forall\, \varepsilon > 0$ il existe un voisinage U_i de 0 dans B_i tel que :

$$\forall\, x_i \in D_i \quad \forall\, v_i, v_i' \in V_i \quad \text{on ait :}$$

$$(v_i - v_i') \in U_i \Rightarrow |g_i(v_i,x_i) - g_i(v_i',x_i)| < \varepsilon \ .$$

Algorithme :

On décrit le passage de x^n à x^{n+1} .

On connait au début de l'itération des parties V_i^n de V_i auxquelles correspondent les ensembles :

$$K_i^n = \{x_i \in D_i \mid g_i(v_i, x_i) \leq 0 \quad \forall\, v_i \in V_i^n\}$$

qui sont des approximations extérieures de C_i .

On détermine successivement pour $i = 1, 2, \ldots m$: x_i^{n+1} tel que :

$$x_i^{n+1} \in K_i^n \ , \ \varphi_i^n(x_i^{n+1}) = \min(\varphi_i^n(x_i) \mid x_i \in K_i^n) \ .$$

Puis :

si $x_i^{n+1} \in C_i$ on fait $V_i^{n+1} = V_i^n$,

sinon , on détermine v_i^{n+1} tel que :

$$v_i^{n+1} \in V_i \ , \ g_i(v_i^{n+1}, x_i^{n+1}) = \max\,(g_i(v_i, x_i^{n+1}) \mid v_i \in V_i)$$

et on choisit V_i^{n+1} tel que :

$$V_i \supseteq V_i^{n+1} \supseteq V_i^n \cup \{v_i^{n+1}\} \ .$$

Montrons maintenant que ce procédé vérifie les relations (1) et (2) du (4.7.1.).

Puisque : $K_i^n \supseteq C_i$ on a par construction de x_i^{n+1} :

$$(\varphi_i^{n'}(x_i^{n+1}),\ z_i^n - x_i^{n+1})_i \geq 0 \qquad \forall\, z_i^n \in C_i$$

ce qui est la relation (2).

Soit $S_i = \{n \mid x_i^n \notin C_i\}$.

La relation (1) est triviale si le cardinal de S_i est nul ou fini. Supposons le contraire. D'après la construction de x_i^{n+1} et l'hypothèse d) la relation (1) sera vraie si on montre que :

(A) $$\limsup_{\substack{n \to \infty \\ n \in S_i}} g_i(v_i^n, x_i^n) = 0 \ .$$

Raisonnons par l'absurde. Si (A) n'était pas vraie, il existerait alors $\varepsilon > 0$ et une sous suite $\{x_i^{n_K}\}$ telle que :

(B) $$\lim_{K\to\infty} g_i(v_i^{n_K}, x_i^{n_K}) \geq \varepsilon \quad .$$

Comme V_i est compact, il existerait des entiers K, K' et K_o tels que $n_K > n_{K'} > n_{K_o}$ et tels que :

$$|g_i(v_i^{n_K}, x_i^{n_K}) - g_i(V_i^{n_{K'}}, x_i^{n_{K'}})| < \frac{\varepsilon}{3}$$

et comme $g_i(V_i^{n_K}, x_i^{n_K}) > 0$ et $g_i(V_i^{n_{K'}}, x_i^{n_{K'}}) \leq 0$ on aurait :

$$g_i(V_i^{n_K}, x_i^{n_K}) < \frac{\varepsilon}{3}$$

ce qui contredit (B) .

Remarque :

On peut faire d'autres hypothèses sur la représentation de C_i qui assurent toujours la validité de l'algorithme. (voir chapitre II) : D_i est compact et la famille de fonctionnelles $g_i(v_i,.)$ est équi-uniformément continue.

BIBLIOGRAPHIE

[1] AUSLENDER, A.
Méthodes numériques pour la décomposition et la minimisation de fonctions non différentiables.
Numerishe mathematik 18, (1971) 213-223.

[2] AUSLENDER, A.
Introduction à la théorie de l'optimisation.
Cours polycopié.
Clermond-Ferrand. Décembre 1969.

[3] AUSLENDER, A. et MARTINET, B.
Méthodes de décomposition en théorie de l'optimisation.
Séminaire d'analyse numérique. Grenoble, décembre 1971.

[4] AUSLENDER, A. et MARTINET, B.
Méthodes de décomposition pour la minimisation d'une fonctionnelle sur un espace produit.
C.R.A.S. à paraître.

[5] CEA, J.
Optimisation. Théorie et algorithmes.
Dunod 1971.

[6] DUNFORD, N. and SCHWARTZ, J.T.
Linear operators. Part I.
Interscience publishers. (1964).

[7] FRANK, M. and WOLFE, P.
An algorithm for quadratic programming.
Nav. res. Log. quat. 3, (1956) 95-110.

[8] GASTINEL, N.
Analyse numérique linéaire.
Hermann (1966).

[9] HILDRETH, C.
A quadratic programming procedure.
Nav. res. log. quat. 4, (1957) 79-85.

[10] LESCARRET, C.
Sous différentiabilité de fonctions composées.
Faculté des Sciences de Montpellier.
Séminaire d'analyse unilatérale. (1968).

[11] LEVITIN, E.S. and POLYAK, B.T.
Constrained minimization problems.
U.S.S.R. comp. math. and Phy. 6, (1966) 1-50

[12] MARTINET, B.
Convergence de certaines méthodes de relaxation
en programmation convexe.
C.R.A.S. 265, (1967) 210-212,
et Rapport CEA (1968).

[13] MARTINET, B.
Minimisation d'une fonctionnelle dans un espace
produit par une méthode de relaxation.
RIRO à paraître.

[14] MEYER, R.
The validity of a family of optimization methods.
SIAM J. Control vol. 8 n° 1 (1970) 41-54.

[15] MIELLOU, J.C.
Thèse. Octobre 1970. Grenoble.

[16] ORTEGA, J.M. and RHEINBOLDT, W.C.
Iterative solution of non linear equation in several variables.
Academic press. New-York. London 1970.

[17] POLAK, E.
Computational methods in optimization.
A unified approach. Academic press. 1971.

[18] SCHECHTER, S.
Relaxation methods for convex problems.
SIAM J. Num. Anal. 5 (1968) 601-612.

[19] TICKONOV, A.N.
Soviet math. Dokl 4 (1963) 1035-1038.
Soviet math. Dokl (1963) 1624-1627.

[20] ZANGWILL, W.I.
Non linear programming. A unified approach.
Prentice Hall, Englewood Cliffs, New-Jersey (1969).

CHAPITRE V

RECHERCHE D'UN POINT FIXE D'UNE APPLICATION PSEUDO-CONTRACTANTE.

CAS DE L'APPLICATION PROX

(5.1) INTRODUCTION.

On considère un ensemble C convexe, fermé, borné d'un espace de Hilbert H (*) et une application U de C dans C pseudo-contractante c'est-à-dire telle que :

pour tout $x,y \in C$ on a :

$$\|Ux-Uy\|^2 \leq \|x-y\|^2 - \|(U-I)x - (U-I)y\|^2 .$$

BROWDER et PETRYSHIN [2] ont montré que l'ensemble des points fixes d'une telle application U est convexe, fermé, borné et non vide. Par extension des résultats de [2] , on démontre que toute suite $\{x^n\}$ construite à partir de $x^1 \in C$ par les itérations :

$$(1) \qquad x^{n+1} = U\, x^n$$

converge faiblement vers un point fixe de U .

Nous appliquons ces résultats à deux exemples :

a) Soit une inéquation variationnelle monotone et hémicontinue :

(2) "trouver $\bar{x} \in C$: $(T\bar{x}, x-\bar{x}) \geq 0$, $\forall x \in C$" .

On lui associe l'application U qui à l'élément $z \in C$ fait correspondra l'élément unique $y \in C$ tel que :

$$(Ty,u-y) + \langle y-z,u-y \rangle \geq 0 , \forall u \in C .$$

(*) On peut étendre ces résultats au cas où H est un espace de Banach uniformément convexe.

On montre que cette application U est pseudo-contractante et que toute itération telle que (1) permet de construire une suite $\{x^n\}$ qui converge faiblement vers une solution de (2) .
(Voir Martinet [6]).

b) Soit le problème de minimisation convexe

(3) "trouver $\bar{x} \in H$ tel que :
$f(\bar{x}) = \min (f(x) \mid x \in H)$"

où f est une fonctionnelle convexe et inf-faiblement compacte.

Comme MOREAU [6], on lui associe l'application prox: U qui à $z \in H$ fait correspondre $y \in H$ tel que :

$$f(y) + \frac{1}{2} ||y-z||^2 = \min (f(x) + \frac{1}{2} ||x-z||^2 \mid x \in H).$$

L'itération (1) constitue un procédé constructif pour trouver une solution de (3).

On montre que la propriété de convergence de la suite générée par (1) est conservée si U est une application composée de plusieurs applications pseudo-contractantes. On applique ce résultat à la recherche d'un point commun à plusieurs ensembles convexes fermés bornés par un procédé de projections successives tel que celui décrit par GUBIN-POLYAK-RAIK [3] .

On donne ensuite des extensions de ces résultats quand les itérations de (1) ne sont effectuées que de façon approchée : approximation de Ux^n , problème d'optimisation résolu par un procédé où on associe la régularisation avec l'accumulation des contraintes. Dans ce cas, la propriété de convergence ne subsiste plus que pour une sous suite.

(5.2.) GENERALITES.

(5.2.1.) Définition :

Soit T une application de C dans C, on dit que cette application est :

- contractante s'il existe $k \in]0,1[$ tel que :
  $\|Tx-Ty\| \leq k\|x-y\| \quad \forall x,y \in C$,
- faiblement contractante si $\|Tx-Ty\| \leq \|x-y\| \quad \forall x,y \in C$.

L'application pseudo-contractante est un cas particulier d'application faiblement contractante.

Nous rappelons des résultats montrés par BROWDER et PETRYSHIN [2] . (cf. aussi ORTEGA-REINBOLDT [8]).

THEOREME :

Soit une application U faiblement contractante qui applique C dans C (C est un ensemble convexe fermé borné d'un espace de Hilbert H). L'ensemble M des points fixes de T est non vide convexe et fermé.

Remarque :

Si H est un espace de Banach réflexif on peut montrer de façon analogue que l'ensemble des points fixes est fermé et convexe mais cet ensemble peut être vide.

Remarque :

Les résultats du théorème ci-dessus demeurent valables si C est un ensemble fortement compact et H un espace de Banach.

THEOREME :

Soit un nombre $\lambda \in]0,1[$.
Si l'application U de C dans C est faiblement contractante, la suite $\{x^n\}$ générée à partir d'un élément initial x^1 de C par les itérations : $x^{n+1} = \lambda x^n + (1-\lambda)Ux^n$ converge faiblement vers un point fixe de U .

(5.2.2.) Théorème fondamental.

Si U est une application pseudo-contractante de C dans C (C convexe fermé borné de l'espace de Hilbert H) alors toute suite $\{x^n\}$ construite à partir de $x^1 \in C$ par les itérations : $x^{n+1} = U\,x^n$, converge faiblement vers un point fixe de U .

On note M l'ensemble des points fixes de U qui est convexe fermé, borné et non vide. Nous décomposons la démonstration de ce théorème en deux lemmes.

LEMME a :

Toute valeur d'adhérence faible de la suite $\{x^n\}$ appartient à M .

Démonstration :

Par définition de la pseudo-contraction on a pour tout $y \in M$:

$$(1) \qquad \|x^{n+1}-y\|^2 \leq \|x^n-y\|^2 - \|x^{n+1}-x^n\|^2$$

donc :
$$\begin{cases} \|x^{n+1}-y\| \leq \|x^n-y\| & \forall\, n \\ \lim\limits_{n\to\infty} \|x^n-y\| \text{ existe} \\ \lim\limits_{n\to\infty} (\|x^{n+1}-y\| - \|x^n-y\|) = 0 \end{cases}$$

D'où on tire de (1) que :

$$(2) \qquad \lim_{n\to\infty} \|x^{n+1}-x^n\| = 0 .$$

Soit $\tilde{x}$ une valeur d'adhérence faible de la suite $\{x^n\}$ (existe toujours car la suite $\{x^n\}$ est bornée), Il existe donc une sous suite $\{x^n\}_{n\in S}$ qui converge faiblement vers $\tilde{x}$.

On a :

$$\|x^n-U\tilde{x}\|^2 = \|x^n-\tilde{x}\|^2 + \|\tilde{x}-U\tilde{x}\|^2 + 2 < x^n-\tilde{x}, \tilde{x}-U\tilde{x} > ,$$

Donc :

$$(3) \qquad \|x^n - U\tilde{x}\|^2 = \|x^n - \tilde{x}\|^2 + \|\tilde{x} - U\tilde{x}\|^2 + \varepsilon_n$$

où : $\lim_{\substack{n\to\infty \\ n\notin S}} \varepsilon_n = 0$.

Or on a (inégalité du triangle et faible contraction de U):

$$\|x^n - U\tilde{x}\| \leq \|x^n - Ux^n\| + \|Ux^n - U\tilde{x}\| \leq \|x^n - x^{n+1}\| + \|x^n - \tilde{x}\| .$$

et grace à (2) :

$$\|x^n - U\tilde{x}\| \leq \|x^n - \tilde{x}\| + \eta_n$$

où : $\lim_{n\to\infty} \eta_n = 0$.

Il vient donc grace à (3) :

$$\|\tilde{x} - U\tilde{x}\|^2 + \varepsilon_n \leq 2\,\eta_n \|x^n - \tilde{x}\| + \eta_n^2 \qquad \forall\, n \in S$$

d'où :

$$\|\tilde{x} - U\tilde{x}\| = 0$$

C.Q.F.D.

LEMME b :

La suite $\{x^n\}$ converge faiblement vers un point fixe de U .

Démonstration :

Supposons que la suite $\{x^n\}$ ait deux points d'adhérence distincts y_1 et y_2 .

Posons :

$$a_i = \lim_{n\to\infty} \|x^n - y_i\| \qquad i = 1, 2 .$$

Ces limites existent grâce au lemme précédent. On suppose que :

$$(1) \qquad a_1 \geq a_2 .$$

Considérons le demi espace fermé :

$$D_\varepsilon = \{x \in H \mid < y_1 - y_2,\ x - \tfrac{1}{2}(y_1 + y_2) > \leq \varepsilon\}$$

On choisit le nombre $\varepsilon \in]0, \frac{1}{2} \|y_1-y_2\|^2[$ de sorte que :

(2) $\qquad y_1 \notin D_\varepsilon$.

Grace à (1) il existe N tel que pour tout $n \geq N$ on ait : $x^n \in D_\varepsilon$. Donc puisque l'ensemble D_ε est faiblement fermé toute valeur d'adhérence de la suite $\{x^n\}$ appartient à D_ε , ce qui contredit (2) et prouve donc que $y_1 = y_2$. On montre qu'il en est de même si $a_1 < a_2$. Ceci prouve que la suite $\{x^n\}$ a une valeur d'adhérence faible unique, donc elle converge faiblement vers cet élément.

C.Q.F.D.

Remarque :

L'élément vers lequel la suite converge faiblement dépend en général de l'élément initial des itérations.

(5.2.3.) Remarque :

On pourrait justifier ce théorème dans le cas d'une application U de C dans C vérifiant :

$$p(\|Ux-Uy\|) \leq p(\|x-y\|) - q(\|(I-U)x - (I-U)y\|) \qquad \forall x,y \in C$$

où les fonctions p et q sont des "fonctions d'influence". (C'est-à-dire que p (ou q) est une fonction non décroissante de R^+ dans R^+ telle que : $p(0) = 0$ et $p(\tau) > 0$ si $\tau > 0$).

(5.3.) RECHERCHE D'UNE SOLUTION D'UNE INEQUATION VARIATIONNELLE PAR REGULARISATION. Voir MARTINET [5] .

(5.3.1.) Application "Proxi" associée à une inéquation variationnelle monotone et hémicontinue.

Notons $\langle , \rangle$ le produit scalaire sur H et $(,)$ la dualité entre H et H'. On considère une application T de H dans H'

- monotone dans C : i.e $(Tx-Ty,x-y) \geq 0 \quad \forall x,y \in C$
- hémicontinue dans C : i.e. pour tout $x,y \in C$, l'application qui à $t \in [0,1]$ fait correspondre $T((1-t)x+ty)$ dans H' muni de la topologie faible est continue.

On cherche une solution de l'inéquation variationnelle :

"trouver $\bar{x} \in C$ tel que $(T\bar{x},x-\bar{x}) \geq 0 \quad \forall x \in C$" .

Grace aux hypothèses faites, ce problèmes a au moins une solution (voir LIONS et STAMPACCHIA [4]).

Par extension de la notion d'application prox pour un problème d'optimisation on introduit une notion analogue.

Définition :

On appelle application "proxi" l'application U qui à tout z de C fait correspondre l'élément unique y de C tel que :

(1) $\quad (Ty,u-y) + < y-z,u-y > \geq 0 \quad \forall u \in C$.

L'existence de cet élément y est facile à vérifier. Montrons son unicité.

Soit y_1 et y_2 deux solutions quelconques de (1).
On a donc :

$$(Ty_1,y_2-y_1) + < y_1-z,y_2-y_1 > \geq 0$$

$$(Ty_2,y_1-y_2) + < y_2-z,y_1-y_2 > \geq 0$$

Soit :

$$0 \geq (Ty_1-Ty_2,y_1-y_2) + \| y_1-y_2\|^2 \geq \|y_1-y_2\|^2$$

et $y_1 = y_2$ ce qui montre l'unicité.

(5.3.2.) Proposition :

L'application proxi définie plus haut est, compte tenu des hypothèses faites, (T est monotone et hémicontinue) une application pseudo-contractante.

Démonstration :

Soit $z_1, z_2 \in C$. On pose :

$$y_i = Uz_i \quad \text{et} \quad \Delta z_i = y_i - z_i \qquad \text{pour} \quad i = 1, 2.$$

U désigne l'application proxi.
Par définition on a :

$$(Ty_1, y_2-y_1) + < \Delta z_1, y_2-y_1 > \geq 0$$
$$(Ty_2, y_1-y_2) + < \Delta z_2, y_1-y_2 > \geq 0$$

Par addition de ces inégalités, il vient :

$$(Ty_1-Ty_2, y_2-y_1) + < y_1-y_2, \Delta z_2 - \Delta z_1 > \geq 0$$

Ce qui grace à la monotonie de T entraine que :

(1) $$< y_1-y_2, \Delta z_2 - \Delta z_1 > \geq 0 .$$

Donc :

$$||z_1-z_2||^2 = ||y_1 - \Delta z_1 - y_2 + \Delta z_2||^2 \geq ||y_1-y_2||^2 + ||\Delta z_2 - \Delta z_1||^2 .$$

Soit :

$$||Uz_1-Uz_2||^2 \leq ||z_1-z_2||^2 - ||(U-I)z_1 - (U-I)z_2||^2 .$$

C.Q.F.D.

Remarque :

Si T est fortement monotone i.e. :
"Il existe $k > 0$ tel que pour tout $x,y \in C$ on a :

$$(Tx-Ty, x-y) \geq k\, ||x-y||^2 " ,$$

alors l'application proxi U correspondante est contractante ; en fait on a :

$$\|Uz_1 - Uz_2\|^2 \leq \frac{1}{1+k} \|z_1 - z_2\|^2 \qquad \forall\, z_1, z_2 \in C \ .$$

(5.3.3.)

Voici maintenant une proposition facile à montrer qui justifie l'intérêt de la notion d'application proxi.

Proposition :

Toute solution d'une inéquation variationnelle monotone et hémicontinue est un point fixe de l'application proxi correspondante et réciproquement.

On en déduit que tout algorithme de recherche d'un point fixe de U permet de déterminer une solution de l'inéquation variationnelle initiale.

Corollaire :

Toute suite $\{x^n\}$ construite à partir de $x^1 \in C$ par les itérations où on détermine x^{n+1} à partir de x^n par :

$$x^{n+1} \in C : \quad (Tx^{n+1}, x - x^{n+1}) + \langle x^{n+1} - x^n, x - x^{n+1} \rangle \geq 0 \qquad \forall x \in C \ ,$$

converge faiblement vers une solution de l'inéquation variationnelle.

Remarque :

Ce procédé est à rapprocher de la méthode habituelle de régularisation d'une inéquation variationnelle que nous rappelons rapidement.

Etant donnée une suite de nombres $\{\varepsilon_n\}$ tel que $\varepsilon_n > 0$ et $\lim_{n \to \infty} \varepsilon_n = 0$, on construit la suite $\{x^n\}$ où l'élément x^n est tel que :

$$x^n \in C : \quad (Tx^n, x - x^n) + \varepsilon_n \langle x^n, x - x^n \rangle \geq 0 \qquad \forall x \in C \ .$$

Appelons M l'ensemble des solutions de l'inéquation variationnelle initiale et $\bar{x}$ l'élément M tel que :

$$\|\bar{x}\| = \min (\|x\| \mid x \in M).$$

On peut montrer que la suite $\{x^n\}$ construite ci-dessus converge faiblement vers $\bar{x}$, élément qui ne dépend évidemment pas de x^1.

(5.4.) CAS D'UN PROBLEME D'OPTIMISATION CONVEXE. RECHERCHE D'UN POINT FIXE DE L'APPLICATION Prox.

(5.4.1.)

Considérons le problème suivant :

($\mathcal{P}$) "trouver $\bar{x} \in C$ tel que $f(\bar{x}) = \min (f(x) \mid x \in H) = m$ "

où la fonctionnelle f définie sur H est convexe et faiblement inf compacte (c'est-à-dire que pour tout $a \in \mathbb{R}$, l'ensemble $\{x \in H \mid f(x) \leq a\}$ est faiblement compact).

On pose $M = \{x \in H \mid f(x) = m\}$. Il est facile de vérifier que M est convexe, fermé, borné non vide.

Définition :

Suivant MOREAU [7] on définit l'application prox : U associée à f qui à tout élément z de H fait correspondre l'élément unique y tel que :

$$f(y) + \frac{1}{2} \|y-z\|^2 = \min (f(x) + \frac{1}{2} \|x-z\|^2 \mid x \in H).$$

(5.4.2.) Proposition :

Toute solution du problème $\mathcal{P}$ est un point fixe de U et réciproquement. (cf. MOREAU [7]).

(5.4.3.) Proposition :

L'application prox est pseudo-contractante.

Démonstration :

On note $f'(x;y)$ la dérivée directionnelle de f en x pour la direction y .

Soit $z_1, z_2 \in H$. On pose :

$$y_i = Uz_i \quad \text{et} \quad \Delta z_i = y_i - z_i \qquad i = 1, 2 .$$

On a par définition de y_1 et y_2 :

$$f'(y_1;y_2-y_1) + < \Delta z_1, y_2-y_1 > \geq 0$$

$$f'(y_2;y_1-y_2) + < \Delta z_2, y_1-y_2 > \geq 0 .$$

Par addition de ces inégalités :

$$(1) \qquad f'(y_1;y_2-y_1) + f'(y_2;y_1-y_2) + < y_1-y_2, \Delta z_2 - \Delta z_1 > \geq 0 .$$

Puisque f est convexe, il vient :

$$f'(y_1;y_2-y_1) + f'(y_2;y_1-y_2) \leq 0 .$$

D'où on tire de (1) que :

$$< y_1-y_2, \Delta z_2 - \Delta z_1 > \geq 0 .$$

On termine la démonstration comme en (5.3.2.).

C.Q.F.D.

Proposition :

Si f est fortement convexe (c'est-à-dire qu'il existe $k > 0$ tel que pour tout $x,y \in H$ et tout $\lambda \in]0,1[$ on ait :

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) - k \lambda(1-\lambda) ||x-y||^2 .),$$

alors l'application prox est contractante.

(5.4.4.) Proposition :

Toute suite $\{x^n\}$ construite à partir de $x^1 \in H$ par les itérations où x^{n+1} est tel que :

$$f(x^{n+1}) + \frac{1}{2} \|x^{n+1}-x^n\|^2 = \min \left(f(x) + \frac{1}{2} \|x-x^n\|^2 \mid x \in H\right),$$

a) converge faiblement vers une solution du problème $\mathcal{P}$.

b) est telle que : $\lim_{n\to\infty} f(x^n) = m = \min (f(x) \mid x \in H)$.

Démonstration :

On a par construction :

(1) $\quad m \leq f(x^{n+1}) \leq f(x^n)$ pour tout n.

Donc :

$$x^n \in D = \{x \in H \mid f(x) \leq f(x^1)\} .$$

Ici D est un ensemble convexe fermé borné.

Les propositions (5.4.3.), (5.2.2.) et (5.4.2.) montrent alors la partie a) de la proposition.

Grâce à (1) il vient :

(2) $\quad \lim_{n\to\infty} f(x^n) \geq m$.

Comme f est convexe on a pour tout x solution de $\mathcal{P}$ et tout n :

(3) $\quad m = f(x) \geq f(x^{n+1}) + f'(x^{n+1};x-x^{n+1})$.

Or par définition de x^{n+1}, il vient :

$$f'(x^{n+1};x-x^{n+1}) + \langle x^{n+1}-x^n, x-x^{n+1} \rangle \geq 0 .$$

Par passage à la limite, puisque les éléments $(x-x^{n+1})$ sont bornés et que : $\lim_{n\to\infty} \|x^n - x^{n+1}\| = 0$, il vient :

$$\liminf_{n\to\infty} f'(x^{n+1};x-x^{n+1}) \geq 0 \qquad \forall x \in H .$$

D'où on tire grace à (3) que :

$$\lim_{n\to\infty} \sup f(x^{n+1}) \leq m .$$

Ce qui montre grace à (2) que :

$$\lim_{n\to\infty} f(x^{n}) = m .$$

C.Q.F.D.

Remarque :

Cette proposition décrit une méthode de régularisation pour résoudre le problème $\mathcal{P}$. (Voir BELLMAN [1] et MARTINET [5]).

(5.5.) CAS D'UNE APPLICATION COMPOSEE DE PLUSIEURS APPLICATIONS PSEUDO CONTRACTANTES.

(5.5.1.)

Soit p applications pseudo-contractantes U_i (i = 1, ..., p) vérifiant :

$$U_1 : C \to H \; ; \; U_p : H \to C \; ; \; U_i : H \to H \qquad i = 2, \ldots,(p-1).$$

On définit l'application A composée : $C \to C$ par :

$$A = U_p \; U_{p-1} \ldots U_1 .$$

On vérifie facilement que l'application A est faiblement contractante. (On ne peut dire en général qu'elle est pseudo-contractante). On en déduit que l'ensemble des points fixes de A est non vide, convexe, fermé et borné.

On va montrer un résultat analogue au théorème (5.2.2.) pour cette application A :

(5.5.2.) Théorème :

La suite $\{x^n\}$ construite à partir de $x^1 \in C$ par les itérations : $x^{n+1} = A\,x^n$, converge faiblement vers un point fixe de A .

Démonstration :

On note M l'ensemble des points fixes de A.

La démonstration est analogue à celle du théorème (5.2.2.) si on montre que :

$$(1) \qquad \lim_{n\to\infty} \|x^{n+1}-x^n\| = 0 .$$

On a (faible contraction de A) :

$$\|x^{n+1}-y\| \leq \|x^n-y\| \qquad \forall\, y \in M .$$

Donc :

$$(2) \qquad \lim_{n\to\infty} (\|x^{n+1}-y\| - \|x^n-y\|) = 0 \qquad \forall\, y \in M .$$

Or :

$$\|x^{n+1}-y\|^2 = \|Ax^n-Ay\|^2 = \|U_p\ldots U_1\, x^n - U_p\ldots U_1\, y\|^2$$

$$\leq \|U_{p-1}\ldots U_1\, x^n - U_{p-1}\ldots U_1\, y\|^2 - \|(I-U_p)U_{p-1}\ldots U_1 x^n - (I-U_p)U_{p-1}\ldots U_1\, y\|^2$$

puisque l'application U_p est pseudo contractante.

On a donc par récurrence :

$$(3) \qquad \|x^{n+1}-y\|^2 \leq \|x^n-y\|^2 - \|(I-U_1)x^n-(I-U_1)y\|^2 - \|(I-U_2)U_1x^n-(I-U_2)U_1y\|^2$$

$$\ldots\ldots - \|(I-U_p)U_{p-1}\ldots U_1x^n - (I-U_p)U_{p-1}\ldots U_1y\|^2 .$$

Donc grâce à (2) il vient :

$$\begin{cases} \lim_{n\to\infty} \|(I-U_1)x^n - (I-U_1)y\| = 0 \\ \lim_{n\to\infty} \|(I-U_2)U_1x^n - (I-U_2)U_1y\| = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \lim_{n\to\infty} \|(I-U_p)U_{p-1}\ldots U_1x^n - (I-U_p)U_{p-1}\ldots U_1y\| = 0 \,. \end{cases}$$

Grâce à l'inégalité du triangle :

$$\lim_{n\to\infty} \|(I-U_1)x^n - (I-U_1)y + (I-U_2)U_1x^n - (I-U_2)U_1y + \ldots.$$

$$\ldots. + (I-U_p)U_{p-1}\ldots U_1x^n - (I-U_p)U_{p-1}\ldots U_1y\| = 0$$

Soit :

$$\lim_{n\to\infty} \|x^n - y - Ax^n - Ay\| = 0$$

ce qui montre (1).

C.Q.F.D.

Remarque :

Ce théorème suggère que les résultats précédents sont valables dans un cadre plus général que celui des applications pseudo-contractantes (voir remarque (5.2.3.)).

(5.5.3.) Exemple : La méthode des projections successives pour trouver un élément commun à plusieurs convexes.

Soit p ensembles D_i de H convexes, fermés, bornés, non vides et un ensemble C de H convexe fermé, borné contenant l'ensemble : $\bigcup_{i=1,\ldots,p} D_i$.

On définit :

a) pour tout i = 1, ..., p , l'application U_i qui à $x \in C$ fait correspondre la projection de x sur D_i

b) l'application $A = U_p U_{p-1} \ldots U_1$

c) l'ensemble $B = \bigcap_{i=1,\ldots,p} D_i$.

Le théorème (5.5.2.) montre que les itérations :

(i) $x^{n+1} = Ax^n$

génèrent, à partir de $x^1 \in C$, une suite $\{x^n\}$ qui converge faiblement vers un point fixe de A .

Or l'itération (i) n'est autre que la méthode des projections successives (cf. GUBIN-POLYAK-RAIK [3]).
Nous allons préciser les propriétés de ce procédé par quelques propositions.

(5.5.3.1.) Proposition :

L'ensemble M des points fixes de A est un ensemble convexe, fermé, borné, non vide.

Démonstration :

évidente car A est faiblement contractante et applique C dans C .

(5.5.3.2.) Proposition :

Si l'ensemble B est non vide alors $M = B$.

Démonstration :

Il est évident qu'on a : $M \supseteq B$.

On montre comme dans le théorème du 5.5.2. que si l'élément z est un point fixe de A , on a pour tout $y \in B$:

$$0 \leq -\|(I-U_1)z - (I-U_1)y\|^2 - \|(I-U_2)U_1 z - (I-U_2)U_1 y\|^2 \ldots$$
$$\ldots - \|(I-U_p)U_{p-1}\ldots U_1 z - (I-U_p)U_{p-1}\ldots U_1 y\|^2 .$$

On en tire que :

$$z = U_j z \qquad \forall j = 1, \ldots, p .$$

Par conséquent, on a : $M \subseteq B$.

C.Q.F.D.

(5.5.3.3.) Proposition :

Soit $\{x^n\}$ une suite générée à partir de $x^1 \in C$ par les itérations : $x^{n+1} = Ax^n$. On a l'alternative suivante :

α) Si B est non vide , la suite $\{x^n\}$ converge faiblement vers un élément de B . On a ainsi justifié le procédé de projections successives.

β) Si B est vide, la suite $\{x^n\}$ converge faiblement vers un point fixe z de A qui correspond à un cycle. (Il existe j tel que $z \neq U_j z$).

On pourra se reporter à GUBIN-POLYAK-RAIK [3] pour une étude détaillée de ces procédés.

(5.6.) RECHERCHE DE POINTS FIXES D'UNE APPLICATION PSEUDO CONTRACTANTE PAR DES ITERATIONS APPROCHEES.

(5.6.1.) Résultats préliminaires.

(5.6.1.1.)

Soit trois suites de nombres réels non négatifs et bornés supérieurement : $\{a_n\}$, $\{b_n\}$, $\{\eta_n\}$ telles que :

$$(1) \quad \begin{cases} \lim_{n\to\infty} \eta_n = 0 \\ a_{n+1} \leq a_n - b_n + \eta_n \qquad \forall n \end{cases}$$

Nous allons montrer des résultats simples sur le comportement asymptotique des suites $\{a_n\}$ et $\{b_n\}$.

(5.6.1.2.) Proposition :

Si en plus des propriétés du (5.6.1.1.) on suppose que la série $\sum_n \eta_n$ est convergente, alors la série $\sum_n b_n$ est convergente et la suite $\{a_n\}$ possède une limite.

Démonstration :

Grâce à (1) il existe $d \in \mathbb{R}$ tel que :

$$\sum_{n=1}^{N} b_n \quad d \leq \sum_{n=1}^{N} \eta_n \qquad \forall N .$$

Donc la série $\sum_n b_n$ est convergente car la série $\sum_n \eta_n$ l'est. Par conséquent la série $\sum_n (\eta_n - b_n)$ est convergente. Supposons maintenant que la suite $\{a_n\}$ ait deux valeurs d'adhérence distinctes. On aboutit facilement à une contradiction puisque :

$$a_{n+1} \leq a_n + (\eta_n - b_n)$$

et que la série $\sum_n (\eta_n - b_n)$ est convergente.

C.Q.F.D.

(5.6.1.3.) Proposition :

Si, en plus des propriétés du (5.6.1.1.), la série $\sum_n \eta_n$ est divergente, les suites $\{a_n\}$ et $\{b_n\}$ n'ont pas en général de limites et il existe une sous suite infinie de $\{b_n\}$ qui tend vers zéro.

Démonstration :

Nous nous contentons de montrer la dernière propriété signalée ci-dessus. Soit $S = \{n \mid a_{n+1} > a_n\}$. Si l'ensemble S est vide ou fini, à partir d'un certain rang on a :

$$a_{n+1} \leq a_n$$

donc $\lim_{n\to\infty} a_n$ existe et grâce à (1) : $\lim_{n\to\infty} b_n = 0$.

Si l'ensemble S est infini. On a pour tout $n \in S$

$$0 < (a_{n+1} - a_n) + b_n \leq \eta_n .$$

Donc : $\lim_{\substack{n\to\infty \\ n\in S}} (a_{n+1}-a_n) = \lim_{\substack{n\to\infty \\ n\in S}} b_n = 0$

C.Q.F.D.

(5.6.2.)

Etant donnée une application U pseudo contractante de C dans C, on considère la suite $\{x^n\}$ générée par les itérations :

$$x^{n+1} = U x^n + \varepsilon^n$$

où $\{\varepsilon^n\}$ est une suite d'éléments de H tels que :

$$\lim_{n\to\infty} ||\varepsilon^n|| = 0 .$$

On a le résultat suivant :

(5.6.3.) Proposition :

Il existe une sous suite infinie $\{x^n\}_{n \in S}$ de la suite $\{x^n\}$ générée par le procédé (5.6.2.) telle que : toute valeur d'adhérence faible de la sous suite $\{x^n\}_{n \in S}$ soit un point fixe de U.

Démonstration :

Puisque U est pseudo contractante on a pour tout $y \in M$:

$$||Ux^n-y||^2 \leq ||x^n-y||^2 - ||Ux^n-x^n||^2 .$$

Donc puisque la suite $\{x^n\}$ est bornée on a par définition de x^{n+1} :

$$||x^{n+1}-y||^2 \leq ||x^n-y||^2 - ||x^{n+1}-x^n||^2 + \eta_n$$

où $\lim_{n\to\infty} \eta_n = 0$.

Donc grâce à la proposition (5.6.1.3.) (poser $a_n = ||x^n-y||^2$ et $b_n = ||x^{n+1}-x^n||^2$), on voit qu'il existe une sous suite infinie $\{x^n\}_{n \in S}$ telle que :

$$\lim_{\substack{n\to\infty \\ n\in S}} ||Ux^n-x^n|| = \lim_{\substack{n\to\infty \\ n\in S}} ||x^{n+1}-x^n|| = 0 .$$

On termine facilement la démonstration comme au lemme a du théorème (5.2.2.).

C.Q.F.D.

(5.6.4.) Remarque :

Si la série $\sum_n \|\varepsilon_n\|$ est convergente (hypothèse assez forte) on peut montrer alors que la suite $\{x^n\}$ générée par le procédé du (5.6.2.) converge faiblement vers un point fixe de U .

(5.6.5.) Un procédé mixte d'optimisation

On va montrer qu'on peut résoudre un problème d'optimisation convexe en associant à la régularisation définie en (5.4.) les méthodes extérieures du chapitre II.

Soit le problème :

"chercher $\bar{x} \in C$ tel que $f(\bar{x}) = \min (f(x) | x \in C)$"

où l'ensemble C est défini par :

$$C = \{x \in X |\ g(v,x) \leq 0 \qquad \forall\, v \in V\} .$$

où :

L'ensemble C est non vide.

X est une partie compacte convexe de H .

La fonctionnelle f est convexe, continue et définie sur X .

$g(v,.)$ est une famille de fonctionnelles convexes et uniformément équicontinues sur X c'est-à-dire :

$\forall\, \varepsilon > 0 \quad \exists \eta > 0 \qquad \forall\, v \in V \qquad \forall\, x,y \in X :$

$\|x-y\| \leq \eta$ on a :

$|g(v,x) - g(v,y)| \leq \varepsilon.$

On considère le procédé itératif suivant :

A l'itération n on connait un sous ensemble Vn de V et un élément x^{n-1} de X . On pose :

$$K_n = \{x \in X \mid g(v,x) \leq 0 \qquad \forall\, v \in V_n\}$$

On détermine alors x^n (problème principal) tel que :

$$x^n \in K_n \; ; \; f(x^n) + \frac{1}{2} \|x^n - x^{n-1}\|^2 = \min\left(f(x) + \frac{1}{2} \|x - x^{n-1}\|^2 \,\middle|\, x \in K_n\right)$$

Puis comme pour l'algorithme du (1.3), on choisit un ensemble V_{n+1} tel que :

$$\begin{cases} V_{n+1} = V_n & \text{si} \quad x^n \in C \; . \\ V_{n+1} = V_n \cup \{v^n\} & \text{dans le cas contraire,} \end{cases}$$

Ici l'élément v^n est choisi par un sous-problème convenable (par exemple $v^n \in V : g(v^n, x^n) \geq \sup\left(g(v, x^n) \,\middle|\, v \in V\right) - \eta_n$) .
On passe ensuite à l'itération suivante.

(5.6.6.) Théorème :

La suite $\{x^n\}$ construite par le procédé ci-dessus a, compte tenu des hypothèses du (5.6.5.), les propriétés suivantes :

α) Toute valeur d'adhérence de cette suite appartient à C .

β) $\lim\limits_{n\to\infty} f(x^n) = m = \min\left(f(x) \,\middle|\, x \in C\right)$.

Nous allons décomposer la démonstration en plusieurs lemmes.

LEMME a :

Toute valeur d'adhérence de la suite $\{x^n\}$ appartient à C et il existe une suite de réels $\{\varepsilon_n\}$: $\lim\limits_{n\to\infty} \varepsilon_n = 0$ tels que :

$$f(x^n) + \frac{1}{2} \|x^n - x^{n-1}\|^2 \leq f(x^{n-1}) + \varepsilon_n \; .$$

Démonstration :

On montre comme dans le théorème (1.5.1.) que :

(1) $\lim\limits_{n\to\infty} d(x^n, C) = 0$. ($d(x^n, c)$ désigne la distance de x^n à C).

Donc toute valeur d'adhérence de $\{x^n\}$ appartient à C . De plus

$\lim\limits_{n\to\infty} d(x^{n-1}, K_n) = 0$.

Soit z^n la projection de l'élément x^{n-1} sur K_n .
On a par définition de x^n :

$$f(x^n) + \frac{1}{2} ||x^n - x^{n-1}||^2 \leq f(z^n) + \frac{1}{2} ||z^n - x^{n-1}||^2$$

$$f(x^n) + \frac{1}{2} ||x^n - x^{n-1}||^2 \leq f(x^{n-1}) + (f(z^n) - f(x^{n-1})) + \frac{1}{2} ||z^n - x^{n-1}||^2 .$$

D'où on déduit le lemme en remarquant que la fonction f est continue sur le compact X et que :

$$\lim_{n \to \infty} ||z^n - x^{n-1}|| = 0 .$$

C.Q.F.D.

LEMME b :

Soit l'ensemble $S = \{n \mid f(x^n) > f(x^{n-1})\}$.

Si le cardinal de S est infini, on a : $\lim_{\substack{n \to \infty \\ n \in S}} f(x^n) = m$.

Démonstration :

Grâce au lemme précédent, on a :

$$(1) \qquad \lim_{\substack{n \to \infty \\ n \in S}} (f(x^n) - f(x^{n-1})) = \lim_{\substack{n \to \infty \\ n \in S}} ||x^n - x^{n-1}|| = 0 .$$

Considérons un élément $\bar{x} \in M = \{x \in C \mid f(x) = m\}$.
On a puisque f est convexe :

$$(2) \qquad m = f(\bar{x}) \geq f(x^n) + f'(x^n ; \bar{x} - x^n)$$

($f'(x;y)$ désigne ici la dérivée directionnelle de f en x dans la direction y), et comme $\bar{x} \in K_n$:

$$f'(x^n ; \bar{x} - x^n) + < x^n - x^{n-1} , \bar{x} - x^n > \geq 0 .$$

Puisque la suite $\{x^n\}$ est bornée, on en déduit grâce à (1) que :

$$\lim_{n \to \infty} \inf f'(x^n ; \bar{x} - x^n) \geq 0 ,$$

et donc il vient de (2)

$$(3) \qquad m \geq \lim_{\substack{n \to \infty \\ n \in S}} \sup f(x^n) .$$

Or toute valeur d'adhérence faible de la suite $\{x^n\}$ appartient à C, donc :

$$m \leq \lim_{n\to\infty} \inf f(x^n) .$$

D'où on tire de (3) que :

$$\lim_{\substack{n\to\infty \\ n\in S}} f(x^n) = m.$$

C.Q.F.D.

LEMME c :

$$\operatorname{Lim}_{n\to\infty} f(x^n) = m .$$

Démonstration :

On distingue trois cas :

α) S est vide ou fini : A partir d'un certain rang N on a pour tout $n \geq N$:

$$f(x^n) \leq f(x^{n+1})$$

Donc $\lim_{n\to\infty} f(x^n)$ existe est cette limite est supérieure ou égale à m.

On termine comme dans la proposition (5.4.2.).

β) Le complémentaire de S est vide ou fini. Conséquence triviale du lemme précédent.

γ) La suite S et son complémentaire ont des cardinaux infinis.

A tout $n \notin S$ assez grand on fait correspondre :

$$a(n) \in S \quad : \quad a(n) = \max (\ell \mid \ell < n, \ \ell \in S)$$

$$b(n) \in S \quad : \quad b(n) = \min (\ell \mid \ell > n, \ \ell \in S).$$

On a par définition :

$$f(x^n) \leq f(x^{a(n)}) + \varepsilon_{a(n)+1}$$
$$f(x^n) \geq f(x^{b(n)})$$

Ce qui montre le lemme par passage à la limite.

C.Q.F.D.

Remarque :

Il ne semble pas qu'on puisse montrer pour cet algorithme que la suite $\{x^n\}$ soit faiblement convergente.

BIBLIOGRAPHIE

[1] BELLMANN, R.E. ; KALABA, R.E. ; LOCKETT, J.
Numerical inversion of Laplace transform.
ELVESIER, New-York (1966) parges 143-144.

[2] BROWDER, F.E. and PETRYSHIN, W.V.
Construction of fixed points of non linear mappings in Hilbert space.
Journal of math. anal. and appl. 20, 197-228 (1967).

[3] GUBIN, L.G. ; POLYAK, B.T. et RAIK E.V.
The method of projections for finding the common point of convex sets.
USSR comp. math. and math. phys. n° 6, 1-24 (1967).

[4] LIONS, J.L. et STAMPACCHIA, G.
Variational inequalities.
C.P.A.M. vol XX , 493-519 (1967).

[5] MARTINET, B.
Régularisation d'inéquations variationnelles par approximations successives.
R.I.R.O. 154-159 (1970).

[6] MARTINET, B.
Détermination approchée d'un point fixe d'une application pseudo-contractante.
C.R.A.S. t 274 (1972) 163-165.

[7] MOREAU, J.J.
Proximité et dualité dans les espaces hilbertiens.
Bull soc. math. France 93, 273-299 (1965).

[8] ORTEGA, J.M. and REINBOLDT, W.C.
Iterative solutions of non linear equations in several variables.
Academic press. (1970) 572 pages.