



HAL
open science

Méthode des sécantes pour la résolution des systèmes d'équation algébriques non linéaires : localisation des solutions d'un système d'équation algébrique

Samira Rachidi

► **To cite this version:**

Samira Rachidi. Méthode des sécantes pour la résolution des systèmes d'équation algébriques non linéaires : localisation des solutions d'un système d'équation algébrique. Modélisation et simulation. Institut National Polytechnique de Grenoble - INPG; Université Joseph-Fourier - Grenoble I, 1976. Français. NNT: . tel-00287193

HAL Id: tel-00287193

<https://theses.hal.science/tel-00287193>

Submitted on 11 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.

THESE

présentée à

UNIVERSITE SCIENTIFIQUE ET MEDICALE DE GRENOBLE

INSTITUT NATIONAL POLYTECHNIQUE DE GRENOBLE

POUR OBTENIR LE GRADE DE

Docteur de 3ème cycle

spécialité : Analyse Numérique

Samira RACHIDI

**METHODE DES SECANTES POUR LA RESOLUTION
DES SYSTEMES D'EQUATIONS ALGEBRIQUES
NON LINEAIRES.**

**LOCALISATION DES SOLUTIONS
D'UN SYSTEME D'EQUATIONS ALGEBRIQUES.**

Thèse soutenue le 30 juin 1976 devant la Commission d'Examen

Président : N. GASTINEL

Examineurs : F. ROBERT

M. DUC JACQUET

UNIVERSITE SCIENTIFIQUE ET MEDICALE DE GRENOBLE

Monsieur Michel SOUTIF : Président

Monsieur Gabriel CAU : Vice-Président

MEMBRES DU CORPS ENSEIGNANTS DE L'U.S.M.G.

PROFESSEURS TITULAIRES

MM.	ANGLES D'AURIAC Paul	Mécanique des Fluides
	ARNAUD Paul	Chimie
	AUBERT Guy	Physique
	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
	BEAUDOING André	Clinique de Pédiatrie et Puériculture
	BERNARD Alain	Mathématiques Pures
Mme	BERTRANDIAS Françoise	Mathématiques Pures
MM.	BERTRANDIAS Jean-Paul	Mathématiques Pures
	BEZES Henri	Pathologie Chirurgicale
	BLAMBERT Maurice	Mathématiques Pures
	BOLLIET Louis	Informatique (I.U.T. B)
	BONNET Georges	Electrotechnique
	BONNET Jean-Louis	Clinique Ophtalmologique
	BONNET-EYMARD Joseph	Clinique Gastro-entérologique
Mme	BONNIER Marie-Jeanne	Chimie Générale
MM.	BOUCHERLE André	Chimie et Toxicologie
	BOUCHEZ Robert	Physique nucléaire
	BOUSSARD Jean-Claude	Mathématiques Appliquées
	BRAVARD Yves	Géographie
	CABANEL Guy	Clinique Rhumatologique et Hydrologique
	CALAS François	Anatomie
	CARLIER Georges	Biologie Végétale
	CARRAZ Gilbert	Biologie Animale et Pharmacodynamie
	CAU Gabriel	Médecine Légale et Toxicologie
	CAUQUIS Georges	Chimie Organique
	CHABAUTY Claude	Mathématiques Pures
	CHARACHON Robert	Clinique Oto-Rhino-Laryngologique
	CHATEAU Robert	Clinique de Neurologie
	CHIBON Pierre	Biologie Animale
	COEUR André	Pharmacie Chimique et Chimie Analytique
	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
	DELORMAS Pierre	Pneumo-Physiologie
	DEPORTES Charles	Chimie Minérale
	DESRE Pierre	Métallurgie

MM.	DESSAUX Georges	Physiologie Animale
	DODU Jacques	Mécanique Appliquée (I.U.T. A)
	DOLIQUE Jean-Michel	Physique des Plasmas
	DREYFUS Bernard	Thermodynamique
	DUCROS Pierre	Cristallographie
	DUGOIS Pierre	Clinique de Dermatologie et Syphiligraphie
	GAGNAIRE Didier	Chimie Physique
	GALLISSOT François	Mathématiques Pures
	GALVANI Octave	Mathématiques Pures
	GASTINEL Noël	Analyse Numérique
	GAVEND Michel	Pharmacologie
	GEINDRE Michel	Electroradiologie
	GERBER Robert	Mathématiques Pures
	GERMAIN Jean-Pierre	Mécanique
	GIRAUD Pierre	Géologie
	JANIN Bernard	Géographie
	KAHANE André	Physique Générale
	KLEIN Joseph	Mathématiques Pures
	KOSZUL Jean-Louis	Mathématiques Pures
	KRAVTCHENKO Julien	Mécanique
	KUNTZMANN Jean	Mathématiques Appliquées
	LACAZE Albert	Thermodynamique
	LACHARME Jean	Biologie Végétale
Mme	LAJZEROWICZ Janine	Physique
MM.	LAJZEROWICZ Joseph	Physique
	LATREILLE René	Chirurgie Générale
	LATURAZE Jean	Biochimie Pharmaceutique
	LAURENT Pierre-Jean	Mathématiques Appliquées
	LEDRU Jean	Clinique Médicale B
	LLIBOUTRY Louis	Géophysique
	LOISEAUX Pierre	Sciences Nucléaires
	LONGEQUEUE Jean-Pierre	Physique Nucléaires
	LOUP Jean	Géographie
Melle	LUTZ Elisabeth	Mathématiques Pures
	MALGRANGE Bernard	Mathématiques Pures
	BOUTET DE MONVEL Louis	Mathématiques Pures
	MALINAS Yves	Clinique Obstétricale
	MARTIN-NOEL Pierre	Seméiologie médicale
	MAZARE Yves	Clinique Médicale A
	MICHEL Robert	Minéralogie et Pétrographie
	MICOUD Max	Clinique Maladies Infectieuses
	MOURIQUAND Claude	Histologie
	MOUSSA André	Chimie Nucléaire
	MULLER Jean-Michel	Thérapeutique (Néphrologie)
	NEEL Louis	Physique du solide
	OZENDA Paul	Botanique
	PAYAN Jean-Jacques	Mathématiques Pures
	PEBAY-PEYROULA Jean-Claude	Physique
	RASSAT André	Chimie Systématique
	RENARD Michel	Thermodynamique
	RINALDI Renaud	Physique
	DE ROUGEMONT Jacques	Neuro-Chirurgie
	SEIGNEURIN Raymond	Microbiologie et Hygiène
	SENGEL Philippe	Zoologie
	SIBILLE Robert	Construction Mécanique (I.U.T. A)
	SOUTIF Michel	Physique Générale
	TANCHE Maurice	Physiologie

MM.	TRAYNARD Philippe	Chimie Générale
	VAILLANT François	Zoologie
	VALENTIN Jacques	Physique Nucléaire
	VAUQUOIS Bernard	Calcul Electronique
Mme	VERAIN Alice	Pharmacie Galénique
MM.	VERAIN André	Physique
	VEYRET Paul	Géographie
	VIGNAIS Pierre	Biochimie médicale
	YOCCOZ Jean	Physique Nucléaire Théorique

PROFESSEURS ASSOCIES

MM.	CLARK Gilbert	Spectrométrie Physique
	CRABBE Pierre	CERMO
	ENGLMAN Robert	Spectrométrie Physique
	HOLTZBERG Frédéric	Basses Températures
	ROST Ernest	Sciences Nucléaires

PROFESSEURS SANS CHAIRE

Melle	AGNIUS-DELDORD Claudine	Physique Pharmaceutique
	ALARY Josette	Chimie Analytique
MM.	AMBROISE-THOMAS Pierre	Parasitologie
	BELORIZKY Elie	Physique
	BENZAKEN Claude	Mathématiques Appliquées
	BIAREZ Jean-Pierre	Mécanique
	BILLET Jean	Géographie
	BOUCHET Yves	Anatomie
	BRUGEL Lucien	Energétique (I.U.T. A)
	BUISSON René	Physique (I.U.T. A)
	CONTE René	Physique (I.U.T. A)
	DEPASSEL Roger	Mécanique des Fluides
	GAUTHIER Yves	Sciences Biologiques
	GAUTRON René	Chimie
	GIDON Paul	Géologie et Minéralogie
	GLENAT René	Chimie organique
	GROULADE Joseph	Biochimie Médicale
	HACQUES Gérard	Calcul Numérique
	HOLLARD Daniel	Hématologie
	HUGONOT Robert	Hygiène et Médecine Préventive
	IDELMAN Simon	Physiologie Animale
	JOLY Jean-René	Mathématiques Pures
	JULLIEN Pierre	Mathématiques Appliquées
Mme	KAHANE Josette	Physique
MM.	KUHN Gérard	Physique (I.U.T. A)
	LE ROY Philippe	Mécanique (I.U.T. A)
	LUU DUC Cuong	Chimie Organique
	MAYNARD Roger	Physique du Solide
	PELMONT Jean	Biochimie
	PERRIAUX Jean-Jacques	Géologie et Minéralogie
	PSISTER Jean-Claude	Physique du Solide
Melle	PIERY Yvette	Physiologie Animale
MM.	RAYNAUD Hervé	M.I.A.G.
	REBECQ Jacques	Biologie (CUS)
	REVOL Michel	Urologie
	REYMOND Jean-Charles	Chirurgie Générale
	RICHARD Lucien	Biologie Végétale
Mme	RINAUDO Marguerite	Chimie Macromoléculaire
MM.	ROBERT André	Chimie Papetière

MM.	SARRAZIN Roger	Anatomie et Chirurgie
	SARROT-REYNAUD Jean	Géologie
	SIROT Louis	Chirurgie Générale
Mme	SOUTIF Jeanne	Physique Générale
MM.	STREGLITZ Paul	Anesthésiologie
	VIALON Pierre	Géologie
	VAN CUTSEM Bernard	Mathématiques Appliquées

MAITRES DE CONFERENCES ET MAITRES DE CONFERENCES AGREGES

MM.	AMBLARD Pierre	Dermatologie
	ARMAND Gilbert	Géographie
	ARMAND Yves	Chimie (I.U.T. A)
	BACHELOT Yvan	Endocrinologie
	BARGE Michel	Neuro chirurgie
	BARJOLLE Michel	MIAG
	BEGUIN Claude	Chimie organique
Mme	BERIEL Hélène	Pharmacodynamie
MM.	BOST Michel	Pédiatrie
	BOUCHARLAT Jacques	Psychiatrie adultes
Mme	BOUCHE Liane	Mathématiques (C.U.S.)
MM.	BRODEAU François	Mathématiques (I.U.T. B)
	BUTTEL Jean	Orthopédie
	CHAMBAZ Edmond	Biochimie médicale
	CHAMPETIER Jean	Anatomie et Organogénèse
	CHARDON Michel	Géographie
	CHERADAME Hervé	Chimie Papetière
	CHIAVERINA Jean	Biologie Appliquée (EFP)
	COHEN-ADDAD Jean-Pierre	Spectrométrie Physique
	COLOMB Maurice	Biochimie Médicale
	CONTAMIN Charles	Chirurgie thoracique et cardio-vasculaire
	CORDONNIER Daniel	Néphrologie
	COULOMB Max	Radiologie
	CROUZET Guy	Radiologie
	CYROT Michel	Physique du solide
	DELOBEL Claude	M.I.A.G.
	DENIS Bernard	Cardiologie
	DOUCE Roland	Physiologie Végétale
	DUSSAUD René	Mathématiques (C.U.S.)
Mme	ETERRADOSSI Jacqueline	Physiologie
MM.	FAURE Jacques	Médecine Légale
	FAURE Gilbert	Urologie
	FONTAINE Jean-Marc	Mathématiques Pures
	GAUTIER Robert	Chirurgie Générale
	GENSAC Pierre	Botanique
	GIDON Maurice	Géologie
	GROS Yves	Physique (I.U.T. A)
	GUITTON Jacques	Chimie
	HICTER Pierre	Chimie
	IVANES Marcel	Electricité
	JALBERT Pierre	Histologie
	KOLODIE Lucien	Hématologie
	KRAKOWIAK Sacha	Mathématiques Appliquées
	LE NOC Pierre	Bactériologie-virologie
	LEROY Philippe	I.U.T. A
	MACHE Régis	Physiologie Végétale
	MAGNIN Robert	Hygiène et Médecine Préventive
	MALLION Jean-Michel	Médecine du Travail
	MARECHAL Jean	Mécanique (I.U.T. A)
	MARTIN-BOUYER Michel	Chimie (C.U.S.)

M.	MICHOULIER Jean	Physique (I.U.T. A)
Mme	MINIER Colette	Physique (I.U.T. A)
MM.	NEGRE Robert	Mécanique (I.U.T. A)
	NEMOZ Alain	Thermodynamique
	NOUGARET Marcel	Automatique (I.U.T.A)
	PARAMELLE Bernard	Pneumologie
	PECCOUD François	Analyse (I.U.T. B)
	PEFFEN René	Métallurgie (I.U.T. A)
	PERRET Jean	Neurologie
	PERRIER Guy	Géophysique - Glaciologie
	PHELIP Xavier	Rhumatologie
	RACHAIL Michel	Médecine Interne
	RACINET Claude	Gynécologie et Obstétrique
	RAMBAUD André	Hygiène et Hydrologie
	RAMBAUD Pierre	Pédiatrie
Mme	RENAUDET Jacqueline	Bactériologie
MM.	ROBERT Jean-Bernard	Chimie-Physique
	ROMIER Guy	Mathématiques (I.U.T. B)
	SHOM Jean-Claude	Chimie générale
	STOEBNER Pierre	Anatomie pathologique
	VROUSOS Constantin	Radiologie

MAITRES DE CONFERENCES ASSOCIES

M.	COLE Antony	Sciences Nucléaires
----	-------------	---------------------

CHARGE DE FONCTIONS DE MAITRE DE CONFERENCES

M.	JUNIEN-LAVILLAVROY Paul	O.R.L.
----	-------------------------	--------

Fait à SAINT MARTIN D'HERES,
DECEMBRE 1975.

Président : M. NEEL Louis
 Vice-Présidents : M. BENOIT Jean
 M. BONNETAIN Lucien

PROFESSEURS TITULAIRES

MM.	BENOIT Jean	Radioélectricité
	BESSON Jean	Electrochimie
	BLOCH Daniel	Physique du solide
	BONNETAIN Lucien	Chimie Minérale
	BONNIER Etienne	Electrochimie et Electrometallurgie
	BRISSONNEAU Pierre	Physique du solide
	BUYLE-BODIN Maurice	Electronique
	COUMES André	Radioélectricité
	FELICI Noël	Electrostatique
	LESPINARD Georges	Mécanique
	MOREAU René	Mécanique
	PARIAUD Jean-Charles	Chimie-Physique
	PAUTHENET René	Physique du solide
	PERRET René	Servomécanismes
	POLOUJADOFF Michel	Electrotechnique
	SILBERT Robert	Mécanique des Fluides

PROFESSEURS ASSOCIES

MM.	RUPPERSBERG Albert, Henner	Chimie
	ROUXEL Roland	Automatique

PROFESSEURS SANS CHAIRE

MM.	BLIMAN Samuel	Electronique
	BOUVARD Maurice	Génie Mécanique
	COHEN Joseph	Electrotechnique
	DURAND Francis	Métallurgie
	FOULARD Claude	Automatique
	LACOURME Jean-Louis	Géophysique
	LANCIA Roland	Electronique
	VEILLON Gérard	Informatique Fondamentale & Appliquée
	ZADWORNÝ François	Electronique

MAITRES DE CONFERENCES

MM.	ANCEAU François	Mathématiques Appliquées
	BOUDOURIS Georges	Radioélectricité
	CHARTIER Germain	Electronique
	GUYOT Pierre	Chimie Minérale
	IVANES Marcel	Electrotechnique
	JOUBERT Jean-Claude	Physique du solide
	MORET Roger	Electrotechnique Nucléaire
	PIERRARD Jean-Marie	Hydraulique
	ROBERT François	Analyse Numérique
	SABONNADIÈRE Jean-Claude	Informatique Fondamentale & Appliquée
Mme	SAUCIER Gabrièle	Informatique Fondamentale & Appliquée

MAITRE DE CONFERENCES ASSOCIE

M. LANDAU Ioan Automatique

CHERCHEURS DU C.N.R.S. (Directeurs et Maîtres de Recherche)

MM. FRUCHART Robert Directeur de Recherche
 ANSARA Ibrahim Maître de Recherche
 CARRE René Maître de Recherche
 DRIOLE Jean Maître de Recherche
 MATHIEU Jean-Claude Maître de Recherche
 MUNIER Jacques Maître de Recherche

à mes parents

Je tiens à exprimer ma gratitude et mes remerciements :

à Monsieur GASTINEL, Professeur à l'Université Scientifique et Médicale de Grenoble, sous la direction duquel ce travail a été effectué ; ses conseils judicieux, son aide bienveillante et les longues heures qu'il a bien voulu me consacrer m'ont permis de mener à bien cette étude

à Monsieur ROBERT, Professeur à l'E.N.S.I.M.A.G., qui a accepté de faire partie du jury et pour l'aide qu'il m'a accordée dans ce travail

à Monsieur DUC JACQUET, Maître de Conférence à l'Université des Sciences Sociales de Grenoble, qui a bien voulu faire partie du Jury

à tous mes camarades de l'équipe d'analyse numérique, en Particulier Messieurs COSNARD et GALLI, qui à diverses occasions m'ont apporté leur aide

à Madame BLANC et à l'équipe du service de reprographie pour leur gentillesse et leur diligence, lors de la réalisation matérielle de cet ouvrage.

TABLE DES MATIERES

PREMIERE PARTIE : Méthode des sécantes pour la résolution des systèmes d'équations algébriques non linéaires.	
Introduction	1
Chapitre I : Formulation de la méthode des sécantes	4
Chapitre II : Etude de la convergence de la méthode des sécantes	10
I - Vitesse et ordre de convergence d'un processus itératif	10
II - Convergence locale de la méthode des sécantes	21
III - Convergence semi-locale de la méthode des sécantes..	33
Chapitre III : Etude de l'efficacité de la méthode des sécantes.	
Applications numériques	55
I - Comparaison de la méthode des sécantes avec d'autres méthodes	58
II - Calcul des itérés de la méthode des sécantes	61
III - Applications numériques	64
Bibliographie de la première partie	74

DEUXIEME PARTIE : Localisation des solutions d'un système
d'équations algébriques.

Introduction	1
Chapitre I : Cas d'un système linéaire	4
I - Cas où la matrice A est symétrique définie positive	4
II - Cas où la matrice A est symétrique régulière quelconque	7
III - Fonction de localisation de la solution d'un système linéaire	7
IV - Applications	23
Chapitre II : Cas d'un système non linéaire	37
Bibliographie	52

P R E M I E R E P A R T I E

M E T H O D E D E S S E C A N T E S P O U R L A R E S O L U T I O N D E S S Y S T E M E S
D ' E Q U A T I O N S A L G E B R I Q U E S N O N L I N E A I R E S

I N T R O D U C T I O N

Dans cette partie, nous allons développer en détail la méthode des sécantes pour la résolution d'un système d'équations algébriques non linéaires :

$$F(X) = 0$$

où F est une fonction d'un domaine D de \mathbb{R}^n dans \mathbb{R}^n .

Cette méthode n'est autre qu'une généralisation dans \mathbb{R}^n de la méthode de la sécante à une dimension qui consiste à trouver une approximation de la solution x^* de l'équation :

$$f(x) = 0 \quad \text{avec} \quad f : D \xrightarrow{\subset \mathbb{R}} \mathbb{R}$$

A partir d'une approximation x_0 de x^* , on obtient une suite d'itérés $\{x_k\}$ où le nouvel itéré x_{k+1} est la solution de l'équation linéaire :

$$L(x) = \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}} (x - x_k) + f(x_k) = 0$$

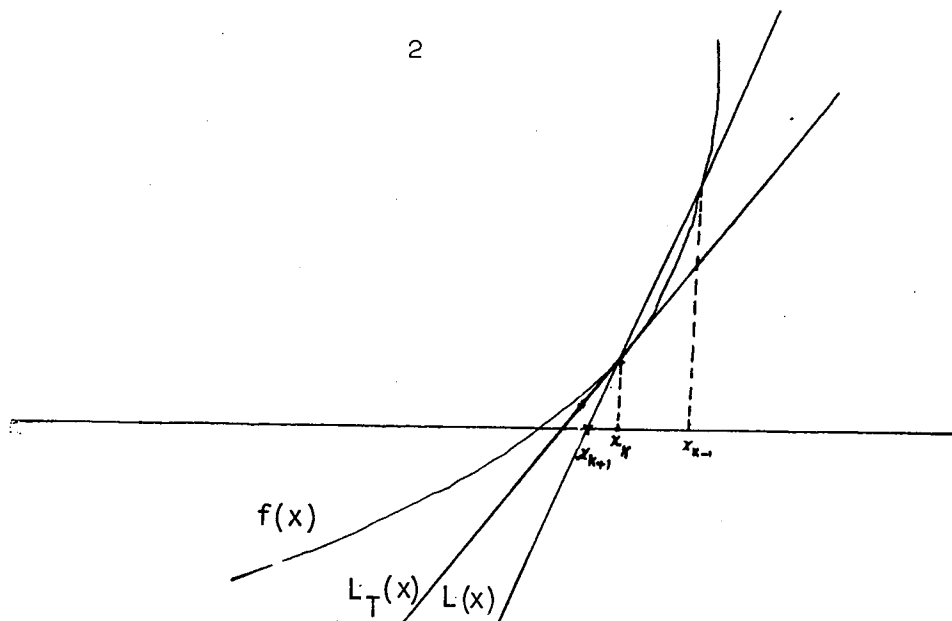
Un point important est que L peut être considéré de deux manières :

1. ou bien comme une approximation de la tangente au point x_k

$$L_T(x) = f'(x_k)(x - x_k) + f(x_k)$$

2. ou bien comme une interpolation de f entre les points x_k et x_{k-1} et ainsi x_{k+1} s'écrit :

$$x_{k+1} = x_k - \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} f(x_k)$$



La généralisation au cas de dimension n conduit à plusieurs méthodes qui dépendent du point de vue considéré :

1. ou bien on approche la matrice des dérivées partielles (Jacobien) évaluées en x_k :

$$F'(x_k) = \begin{bmatrix} \frac{\partial F_i}{\partial x_j} \end{bmatrix} \quad \begin{array}{l} i = 1, \dots, n \\ j = 1, \dots, n \end{array}$$

par la matrice $J(x_k, h_k)$ des quotients des différences.

Le nouvel itéré est alors obtenu par résolution du système d'équations linéaires $F'(x_k)(x - x_k) = -F(x_k)$.

2. ou bien on remplace chacune des surfaces f_i ($i = 1, \dots, n$) dans \mathbb{R}^{n+1} par un hyperplan, lequel interpole f_i en $n+1$ points donnés $x_{k,j}$ ($j = 0, \dots, n$) dans un voisinage de x_k . Il faut donc trouver un vecteur a_i et un scalaire α_i tels que l'application affine $L_i(x) = \alpha_i + x^T a_i$ satisfasse aux conditions :

$$L_i(x_{k,j}) = f_i(x_{k,j}) \quad j = 0, \dots, n$$

Le nouvel itéré est alors obtenu à l'intersection dans \mathbb{R}^{n+1} de ces n hyperplans avec l'hyperplan $x_{n+1} = 0$. x_{k+1} est donc la solution du système linéaire

$$L_i(x) = 0 \quad i = 1, \dots, n.$$

On remarque que le choix des points $x_{k,j}$ ($j = 0, \dots, n$) conduit à différentes méthodes.

C'est la deuxième interprétation que nous allons étudier et plus particulièrement le cas où :

$$x_{k,0} = x_k, \dots, x_{k,j} = x_{k-j}, \dots, x_{k,n} = x_{k-n}$$

Si H_k et Γ_k sont les matrices suivantes de $M_{n,n}(\mathbb{R})$:

$$H_k = (x_k - x_{k-1}, \dots, x_{k-n+1} - x_{k-n})$$

$$\Gamma_k = (F(x_k) - F(x_{k-1}), \dots, F(x_{k-n+1}) - F(x_{k-n}))$$

alors

$$L(x) = F(x_k) + \Gamma_k H_k^{-1} (x_{k+1} - x_k)$$

et ainsi la nouvelle approximation est donnée par :

$$x_{k+1} = x_k - H_k \Gamma_k^{-1} F(x_k).$$

On remarque que si F est différentiable $\Gamma_k H_k^{-1}$ peut être considéré comme une approximation du jacobien F' de F . Ainsi la méthode de la sécante peut être considérée comme une discrétisation de la méthode de Newton.

Dans le premier chapitre, on trouve la formulation de Wolfe [10] puis la formulation de Newton [3] améliorant ainsi la méthode des sécantes présentée par Wolfe.

Le deuxième chapitre est consacré aux théorèmes de convergence locale ([3] et [6]), aux théorèmes de convergence semi-locale et à l'examen de l'ordre de convergence.

Le troisième et le dernier chapitre mentionne le coût et l'efficacité de cette méthode et présente quelques résultats numériques.

C H A P I T R E I

FORMULATION DE LA METHODE DES SECANTES

Définition 1 :

Les $n+1$ points x_0, \dots, x_n de \mathbb{R}^n sont en "position générale" si les vecteurs $x_0 - x_j$ ($j = 1, \dots, n$) sont linéairement indépendants

Théorème 1 :

Si x_0, \dots, x_n sont $n + 1$ points de \mathbb{R}^n , alors les conditions suivantes sont équivalentes :

- a) x_0, \dots, x_n sont en position générale.
- b) $\forall j = 0, \dots, n$ les vecteurs $x_j - x_i$ ($i = 0, \dots, n$) et $i \neq j$ sont linéairement indépendants.
- c) La matrice $(e, X^T) \in M_{n+1, n+1}(\mathbb{R})$ où $e^T = (1, \dots, 1)$ et $X = (x_0, \dots, x_n)$, est non singulière.
- d) $\forall y \in \mathbb{R}^n$, il existe des scalaires $\alpha_0, \dots, \alpha_n$ avec $\sum_0^n \alpha_i = 1$ tels que :

$$y = \sum_0^n \alpha_i x_i$$

Démonstration :

Soit A la matrice suivante de $M_{n+1, n+1}(\mathbb{R})$:

$$A = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ x_j & x_0 - x_j & \dots & x_{j-1} - x_j & x_{j+1} - x_j & \dots & x_n - x_j \end{pmatrix}$$

alors

$$A = \begin{pmatrix} 1 & \cdots & \cdots & \cdots & 1 \\ x_j x_0 & \cdots & x_{j-1} x_{j+1} & \cdots & x_n \end{pmatrix} \begin{pmatrix} 1 & -1 & \cdots & \cdots & -1 \\ & \ddots & \ddots & \circ & 0 \\ & & \ddots & \ddots & \ddots \\ & \circ & & \ddots & \ddots \\ & & & & \cdots & 1 \end{pmatrix}$$

d'où on déduit :

$$\begin{aligned} & \det(x_0 - x_j, \dots, x_{j-1} - x_j, x_{j+1} - x_j, \dots, x_n - x_j) \\ &= \det \begin{pmatrix} 1 & \cdots & \cdots & \cdots & 1 \\ x_j x_0 & \cdots & x_{j-1} x_{j+1} & \cdots & x_n \end{pmatrix} = (-1)^j \det(e, X^T) \end{aligned}$$

ceci $\forall j = 0, \dots, n$

ce qui montre l'équivalence de a), b) et c).

De plus d) est vraie si et seulement si le système linéaire

$$\begin{pmatrix} e^T \\ X \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} 1 \\ Y \end{pmatrix} \quad (1)$$

admet une solution $\forall Y \in \mathbb{R}^n$, ce qui montre que c) implique d).

Inversement, en résolvant (1) successivement pour $Y = 0, e_1, \dots, e_n$ on voit que $\begin{pmatrix} e^T \\ X \end{pmatrix}$ est non singulière ce qui conclut le théorème.

Remarque :

Dire que x_0, \dots, x_n sont en position générale peut être interprété géométriquement par : ces points ne doivent pas être dans un sous-espace affine de dimension plus petit que n .

Ainsi pour $n = 2$, x_0, x_1, x_2 sont en position générale si et seulement si ces trois points ne sont pas colinéaires.

Théorème 2 :

Si x_0, \dots, x_n et y_0, \dots, y_n sont des points de \mathbb{R}^n , alors il existe une fonction affine unique :

$$L(x) = a + Ax \text{ où } a \in \mathbb{R}^n \text{ et } A \in M_{n,n}(\mathbb{R})$$

telle que

$$L(x_j) = y_j \quad j = 0, \dots, n.$$

si et seulement si x_0, \dots, x_n sont en position générale.

De plus A est non singulière si et seulement si y_0, \dots, y_n sont en position générale.

Démonstration :

La condition $L(x_j) = y_j \quad j = 0, \dots, n$ peut s'écrire sous forme matricielle :

$$(e, X^t) \begin{pmatrix} a^T \\ A^T \end{pmatrix} = (y_0, \dots, y_n)^T$$

où $e^T = (1, \dots, 1)$ et $X = (x_0, \dots, x_n)$

Ainsi, on constate que la première partie est une conséquence du théorème 1.

De plus $L(x_j) = y_j \quad \forall j = 0, \dots, n$ implique que :

$$A(x_j - x_0) = y_j - y_0 \quad \forall j = 1, \dots, n$$

et puisque $x_j - x_0$ ($j = 1, \dots, n$) sont linéairement indépendants, alors A est non singulière si et seulement si les vecteurs $y_j - y_0 \quad j = 1, \dots, n$ sont linéairement indépendants, car en posant :

$$X' = (x_0, x_1 - x_0, \dots, x_j - x_0, \dots, x_n - x_0)$$

$$Y' = (y_0, y_1 - y_0, \dots, y_j - y_0, \dots, y_n - y_0)$$

on a :

$$AX' = Y'$$

Or $\det X' \neq 0$ donc $\det A \neq 0$ si et seulement si $\det Y' \neq 0$, c'est-à-dire les y_i sont en position générale.

Définition 2 :

Soit $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Si les points de \mathbb{R}^n x_0, \dots, x_n et $F(x_0), \dots, F(x_n)$ sont en position générale alors le point

$$x_s = -A^{-1} a$$

où a et A satisfont :

$$a + Ax_j = F(x_j) \quad \forall j = 0, \dots, n$$

est appelé base de l'approximation des "sécantes" relativement à x_0, \dots, x_n .

Remarques :

1. Le théorème 2 assure l'existence de x_s .
2. Dans le cas de dimension 1, ces conditions sont réduites à $x_0 \neq x_1$ et $f(x_0) \neq f(x_1)$ lesquelles assurent que la sécante passant par x_0 et x_1 coupe l'axe des x .

Formulation de Wolfe [10].

Si les points x_0, \dots, x_n et $F(x_0), \dots, F(x_n)$ sont en position générale, alors la base de l'approximation des sécantes satisfait :

$$x_s = x_z = \sum_{j=0}^n z_j x_j$$

où $Z = (z_0, \dots, z_n)^T$ est la solution unique du système linéaire de $n+1$ équations à $n+1$ inconnues :

$$\begin{pmatrix} 1 & \dots & 1 \\ F(x_0) & \dots & F(x_n) \end{pmatrix} Z = (1, 0, \dots, 0)^T \quad (2)$$

Démonstration :

Comme les $F(x_j)$ $j = 0, \dots, n$ sont en position générale le théorème 1 b) assure que (2) a une solution unique qui satisfait :

$$\sum_{i=0}^n z_i = 1 \quad \text{et} \quad \sum_{i=0}^n z_i F(x_i) = 0$$

Or $F(x_i) = a + Ax_i$ donc :

$$0 = \sum_0^n z_i F(x_i) = \sum_0^n z_i (a + Ax_i) = a + A \sum_{i=0}^n z_i x_i$$

ceci entraîne

$$\sum_{i=0}^n z_i x_i = -A^{-1}a$$

car A est inversible d'après le théorème 2, et puisque x_s est la solution unique de $Ax + a = 0$ alors (1) est vraie.

Définition 3 :

Si F est une fonction de $D \subset \mathbb{R}^n$ dans \mathbb{R}^n , alors dans la suite la fonction :

$$J : E \xrightarrow{\subset \mathbb{R}^n \times L(\mathbb{R}^n)} L(\mathbb{R}^n)$$

est définie par :

$$J(x, H) = (F(x + He_n) - F(x + He_{n-1}), \dots, F(x + He_1) - F(x))H^{-1}$$

où (e_1, \dots, e_n) est la base canonique de \mathbb{R}^n et

$$E = \{(x, H) \mid x + He_i \in D \quad i = 1, \dots, n \text{ et } H \text{ non singulière}\}$$

Formulation de Newton :

On suppose pour $k \geq n$ que x_{k-n}, \dots, x_k et $F(x_{k-n}), \dots, F(x_k)$ sont en position générale et que :

$$H_k = (x_k - x_{k-1}, \dots, x_{k-j+1} - x_{k-j}, \dots, x_{k-n+1} - x_{k-n})$$

alors $J(x_k, H_k)$, donnée par la définition 3, est non singulière et la base de l'approximation des sécantes s'écrit :

$$x_s = x_{k+1} = x_k - J(x_k, H_k^{-1}) F(x_k)$$

Démonstration :

On a $F(x_{k-j+1}) = a + Ax_{k-j+1}$ et $F(x_{k-j}) = a + Ax_{k-j} \quad \forall j = 1, \dots, n$;
d'où

$$A(x_{k-j+1} - x_{k-j}) = F(x_{k-j+1}) - F(x_{k-j}) \quad \forall j = 1, \dots, n ;$$

ainsi on peut écrire :

$$AH_k = (F(x_k) - F(x_{k-1}), \dots, F(x_{k-n+1}) - F(x_{k-n}))$$

comme x_k, \dots, x_{k-n} sont en position générale, alors H est non singulière
d'où :

$$A = J(x_k, H_k)$$

les $F(x_{k-j}) \quad j = 0, \dots, n$ sont en position générale ; alors A est non
singulière et on peut écrire :

$$x_{k+1} = -A^{-1} a$$

$$F(x_k) = a + Ax_k \implies a = -Ax_k + F(x_k)$$

d'où

$$x_{k+1} = x_k - A^{-1} F(x_k)$$

ou

$$x_{k+1} = x_k - J(x_k, H_k)^{-1} F(x_k)$$

Si on pose :

$$\Gamma_k = (F(x_k) - F(x_{k-1}), \dots, F(x_{k-n+1}) - F(x_{k-n}))$$

alors :

$$J(x_k, H_k) = \Gamma_k H_k^{-1}$$

et

$$x_{k+1} = x_k - H_k \Gamma_k^{-1} F(x_k)$$

CHAPITRE II

ETUDE DE LA CONVERGENCE DE LA METHODE DES SECANTES

INTRODUCTION

Dans cette introduction nous insistons sur le fait que les problèmes les plus importants dans l'étude d'un processus itératif (pour trouver la racine x^* d'une fonction arbitraire $F(x)$ définie dans un sous ensemble de \mathbb{R}^n sont :

- a. Quand les itérés convergent-ils ?
- b. Quelle est alors la vitesse de convergence ?
- c. Quel est le critère d'efficacité ?

Plusieurs auteurs ont étudié ces problèmes dans le cas à n dimensions aussi bien que dans le cas à une dimension. Nous remarquons que l'étude de la convergence locale conduit à celle de l'ordre de convergence dont la valeur précise présente un grand intérêt pour l'étude de la vitesse de convergence et de l'efficacité d'une méthode.

C'est pourquoi avant de parler de la convergence et de l'ordre de convergence de la méthode des sécantes, nous allons faire quelques rappels sur la vitesse et l'ordre de convergence d'un processus itératif.

I - VITESSE ET ORDRE DE CONVERGENCE D'UN PROCESSUS ITERATIF.

Notons tout d'abord que l'étude la plus satisfaisante jusqu'à maintenant de l'ordre de convergence, dans le cas à n dimensions est due à Ortega et Rheinboldt [3]. A partir de deux définitions de l'ordre de convergence d'une méthode itérative, ils font une étude très intéressante et donnent des conditions suffisantes pour que certaines méthodes importantes aient

une convergence quadratique ou super linéaire.

D'autres auteurs ont donné d'autres définitions de l'ordre de convergence :
Ainsi Ostrowski [4] et Traub [7] définissent l'ordre de convergence d'une suite itérative par un nombre p , s'il existe, tel que :

$$\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^p} \rightarrow C_p$$

où la constante C_p de l'erreur asymptotique doit être $C_p \neq 0$ et $C_p \neq \infty$ (Ortega et Reinboldt montrent que cette définition est plus restreinte que leur Q -ordre).

D'autre part, Ostrowski [4] donne une autre définition de l'ordre de convergence d'une suite :

(r étant un entier ≥ 1)

$$\alpha_r = \limsup_{k \rightarrow \infty} (\|x_{k+r} - x^*\| / \|x_k - x^*\|)$$

et s'il existe un entier $r \geq 1$ tel que $\alpha_r < 1$ alors on dit que la convergence de $\{x_k\}$ vers x^* est "linéairement faible". (La convergence "linéairement faible" d'une suite est toujours une convergence R -linéaire au sens d'Ortega et Reinboldt mais l'inverse n'est pas toujours vrai).

Wall [9] et Tornheim [6] ont introduit la quantité :

$$p = \lim_{k \rightarrow \infty} (\log \|x_{k+1} - x^*\| / \log \|x_k - x^*\|)$$

comme l'ordre de convergence d'une suite $\{x_k\}$, pourvu que cette suite existe (il est facile alors de voir que p est égal à R -ordre de $\{x_k\}$ [3]).

Brent [1] et [2] définit aussi l'ordre de convergence d'une suite par :

$$\rho = \liminf_{k \rightarrow \infty} (-\log \|x_k - x^*\|)^{1/k}$$

et si $\rho > 1$, alors ρ n'est autre que le R-ordre de $\{x_k\}$ au sens d'Ortega et Rheinboldt.

Notations :

- Dans la suite φ désigne un processus itératif qui s'écrit sous la forme :

$$\begin{aligned} \varphi : x_{k+1} &= G(x_k, \dots, x_{k-m+1}) \\ k &= m-1, m, \dots \end{aligned}$$

$$\text{où } G : D^m \xrightarrow{c(R^n)^m} \mathbb{R}^n$$

- et x^* un de ses points limites.

- On note par $C(\varphi, x^*)$ l'ensemble des suites engendrées par φ et ayant même point limite x^* .

I.1 "Q-facteurs" de convergence [3].

Définition 1 :

Soit $\{x_k\} \in \mathbb{R}^n$ une suite quelconque ayant pour limite x^* alors les quantités suivantes définies $\forall p \in [1, \infty[$:

$$Q_p\{x_k\} = \begin{cases} 0 & \text{si } x_k = x^* \forall k \text{ sauf pour un nombre fini} \\ \limsup_{k \rightarrow \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^p} & \text{si } x_k \neq x^*, \text{ sauf pour un nombre fini} \\ +\infty & \text{autrement.} \end{cases}$$

sont les Q-facteurs de $\{x_k\}$, relativement à une norme $\|\cdot\|$ de \mathbb{R}^n .

Lorsqu'on considère un processus itératif et non pas une suite, il est préférable de considérer, comme indicateur de mesure de la vitesse de convergence, la vitesse de convergence la plus mauvaise d'une suite quelconque de $C(\varphi, x^*)$; et c'est pour cela que nous donnons la définition suivante.

Définition 2 :

On appelle Q-facteurs de φ en x^* relativement à une norme, dans laquelle les $Q_p\{x_k\}$ sont calculés, la quantité suivante :

$$Q_p(\varphi, x^*) = \sup\{Q_p\{x_k\} / \{x_k\} \in C(\varphi, x^*)\}$$

$$1 \leq p \leq \infty$$

Propriété 1 :

Soit $Q_p(\varphi, x^*)$, $p \in [1, \infty]$, les Q-facteurs de φ en x^* calculés dans une norme fixée dans \mathbb{R}^n , alors une et une seule des conditions suivantes est vraie :

- a. $Q_p(\varphi, x^*) = 0 \quad \forall p \in [1, \infty[$
- b. $Q_p(\varphi, x^*) = \infty \quad \forall p \in [1, \infty[$
- c. $\exists p_0 \in [1, \infty]$ tel que :

et	$Q_p(\varphi, x^*) = 0$	$\forall p \in [1, p_0[$
	$Q_p(\varphi, x^*) = \infty$	$\forall p \in [p_0, \infty[$

Démonstration :

(voir [3])

Remarque :

Le motif principal d'introduction des Q-facteurs d'un processus itératif, sans considérer uniquement les Q-facteurs d'une suite, est d'avoir une manière précise de comparaison de la vitesse de convergence de deux processus itératifs différents.

Définition 3 :

Soient φ_1 et φ_2 deux processus itératifs ayant même limite x^* et soient $Q_p(\varphi_1, x^*)$, $Q_p(\varphi_2, x^*)$ les Q-facteurs correspondants calculés avec la même norme. On dit alors que :

φ_1 est "Q-plus rapide" que φ_2 en x^* s'il existe $p \in [1, \infty[$ tel que $Q_p(\varphi_1, x^*) < Q_p(\varphi_2, x^*)$.

Remarques :

1. D'après la propriété 1, la notion de "Q-plus rapide" est bien définie et il n'est pas possible de trouver p et $p' \in [1, \infty[$ tel que :

$$Q_p(\varphi_1, x^*) < Q_p(\varphi_2, x^*)$$

et

$$Q_{p'}(\varphi_1, x^*) > Q_{p'}(\varphi_2, x^*)$$

Ceci tient au fait que, si on suppose $p' > p$ et $Q_p(\varphi_2, x^*) > Q_p(\varphi_1, x^*)$, alors $Q_p(\varphi_2, x^*) > 0$ et ainsi $Q_{p'}(\varphi_2, x^*) = +\infty$

2. La propriété 1 entraîne que la relation "Q-plus rapide" est transitive.
3. La notion "Q-plus rapide" dépend de la norme choisie dans \mathbb{R}^n . Il est possible que, dans une norme, un processus φ_2 , soit Q-plus rapide qu'un autre φ_1 et que, dans une autre norme, φ_1 soit "Q-plus rapide" que φ_2 .

Pour cela, nous allons introduire les notions suivantes indépendantes de la norme.

Définition 4 :

Soit $Q_p(\varphi, x^*)$ les Q-facteurs de φ en x^* calculés dans une norme de \mathbb{R}^n , alors :

$$Q_Q(\varphi, x^*) = \begin{cases} +\infty & \text{si } Q_p(\varphi, x^*) = 0 \quad \forall p \in [1, \infty[\\ \inf p \in [1, p[; Q_p(\varphi, x^*) = +\infty & \text{autrement} \end{cases}$$

est le Q-ordre de φ en x^* .

Propriété 2 :

Les trois relations suivantes sont indépendantes de la norme :

1. $Q_p(\varphi, x^*) = 0$
2. $0 < Q_p(\varphi, x^*) < \infty$
3. $Q_p(\varphi, x^*) = \infty$

Ainsi le Q-ordre est indépendant de la norme.

(Pour la démonstration, voir [3]).

Propriété 3 :

Soient φ_1, φ_2 deux processus itératifs ayant même point limite x^* . Si

$$O_Q(\varphi_1, x^*) > O_Q(\varphi_2, x^*)$$

alors φ_1 est Q-plus rapide que φ_2 en x^* (indépendamment de la norme).

Propriété 4 :

Soit φ un processus itératif ayant pour limite x^* :

si $Q_p(\varphi, x^*) < +\infty$ pour un $p \in [1, \infty[$ alors $O_Q(\varphi, x^*) \geq p$

si $Q_q(\varphi, x^*) > 0$ pour un $q \in [1, \infty[$ alors $O_Q(\varphi, x^*) \leq q$

et donc :

si $0 < Q_p(\varphi, x^*) < \infty$ pour $p \in [1, \infty[$ alors $O_Q(\varphi, x^*) = p$

Interprétation des résultats :

a. La comparaison de deux processus ayant même limite x^* consiste à comparer les Q-ordre de φ_1 et φ_2 :

1. s'ils sont différents, alors l'un est Q plus rapide que l'autre quelle que soit la norme.

2. s'ils sont égaux, alors on compare les Q-facteurs. Cependant, dans ce cas, le Q-plus rapide dépend de la norme.

b. Chaque fois qu'on a $Q_p(\varphi_1, x^*) < \infty$ alors $\exists k_0 \in \mathbb{N}^*$ tel que :

$$\|x_{k+1} - x^*\| \leq \gamma \|x_k - x^*\| \quad \forall k \geq k_0 \quad (\text{en choisissant } \gamma = Q_p + \varepsilon)$$

ceci $\forall \varepsilon$ et \forall la suite $\{x_k\}$ qui converge vers x^* .

c. Les processus itératifs avec un Q-ordre = 1, 2 ou 3 jouent un grand rôle dans la théorie. Notons que :

1. Si $Q_1(\varphi, x^*) = 0$ nous disons que le processus a une convergence Q-super linéaire en x^* .

(Notons que d'après la propriété 2, la notion de Q-super linéaire est indépendante de la norme).

2. Si $0 < Q_1(\varphi, x^*) < 1$ dans une norme, la convergence est appelée Q-linéaire (qui dépend de la norme).
3. Si $Q_1(\varphi, x^*) \geq 1$ dans une norme, alors φ est Q-sublinéaire dans cette norme.
4. si $Q_2(\varphi, x^*) = 0$, alors la convergence de φ est Q-superquadratique
5. Si $0 < Q_2(\varphi, x^*) < +\infty$ alors φ est Q-quadratique.
6. Si $Q_2(\varphi, x^*) = +\infty$ alors φ est Q-subquadratique.

De même pour la convergence cubique d'un processus itératif.

I.2 R-facteurs ou "Root-convergence factors" [3].

Comme dans l'étude de la convergence des séries, Ortega et Rheinboldt introduisent une définition basée sur la moyenne géométrique du vecteur erreur $\|x_k - x^*\|$.

Définition 5 :

Si $\{x_k\}$ est une suite quelconque ayant pour limite x^* , alors les quantités :

$$R_p\{x_k\} = \begin{cases} \bullet \limsup_{k \rightarrow \infty} \|x_k - x^*\|^{1/k} & \text{si } p = 1 \\ \bullet \limsup_{k \rightarrow \infty} \|x_k - x^*\|^{1/p^k} & \text{si } p > 1 \end{cases}$$

sont les R-facteurs de convergence de $\{x_k\}$ en x .

Définition 6 :

Si φ est un processus itératif ayant pour limite x^* , alors

$$R_p(\varphi, x^*) = \sup\{R_p\{x_k\} ; \{x_k\} \in \mathcal{C}(\varphi, x^*)\} \quad 1 \leq p < \infty$$

sont les R-facteurs de φ en x^* .

Propriété 5 :

Si $\{x_k\}$ est une suite qui converge vers x^* , alors $R_p\{x_k\}$ est indépendant de la norme $\forall p \in [1, \infty[$ et par conséquent $R_p(\varphi, x^*)$ l'est aussi.

Propriété 6 :

Si φ est un processus itératif ayant x^* comme point limite, alors une et une seule des conditions suivantes est vraie :

- a. $R_p(\varphi, x^*) = 0 \quad \forall p \in [1, \infty[$
- b. $R_p(\varphi, x^*) = 1 \quad \forall p \in [1, \infty[$
- c. $\exists p_0 \in [1, \infty[$ tel que $R_p(\varphi, x^*) = 0 \quad \forall p \in [1, p_0[$
et $R_p(\varphi, x^*) = 1 \quad \forall p \in]p_0, \infty[$

Définition 7 :

Si φ_1 et φ_2 sont deux processus itératifs ayant même point limite x^* ; s'il existe un $p \in [1, \infty[$ tel que :

$$R_p(\varphi_1, x^*) < R_p(\varphi_2, x^*)$$

alors φ_1 est R-plus rapide que φ_2 en x^* .

Définition 8 :

Soit φ un processus itératif ayant pour limite x^* , alors la quantité

$$O_R(\varphi, x^*) = \begin{cases} +\infty & \text{si } R_p(\varphi, x^*) = 0 \quad \forall p \in [1, \infty[\\ \inf\{p \in [1, \infty[; R_p(\varphi, x^*) = 1\} & \text{autrement.} \end{cases}$$

est appelée le R-ordre de φ en x^* .

Propriété 7 :

Soient φ_1 et φ_2 deux processus itératifs ayant même limite x^* .
Si

$$O_R(\varphi_1, x^*) > O_R(\varphi_2, x^*)$$

alors φ_1 est R-plus rapide que φ_2 en x^* .

Propriété 8 :

Soit φ un processus itératif ayant pour limite x^* ;

Si $R_p(\varphi, x^*) < 1$ pour un $p \in [1, \infty[$ alors $O_R(\varphi, x^*) \geq p$

si $R_q(\varphi, x^*) > 0$ pour un $q \in [1, \infty[$ alors $O_R(\varphi, x^*) \leq q$

donc si $0 < R_p(\varphi, x^*) < 1$, $O_R(\varphi, x^*) = p$.

Si $0 < R_1(\varphi, x^*) < 1$, alors la convergence de φ en x^* est R-linéaire

si $R_1 = 0$ ou $R_1 = 1$, alors la convergence de φ en x^* est respectivement R-superlinéaire ou R-sublinéaire.

De même, si $0 < R_2(\varphi, x^*) < 1$, alors la convergence de φ en x^* est R-quadratique.

I.3 Relations entre les Q- et les R- facteurs de convergence [3].

Propriété 9 :

Si $\{x_k\}$ est une suite de \mathbb{R}^n qui converge vers x^* telle que :

$$\left\{ \begin{array}{l} 0 < Q_p\{x_k\} < +\infty \\ \text{et} \\ 0 < R_p\{x_k\} < 1 \end{array} \right. \quad \text{pour un } p > 1.$$

Alors en choisissant une norme convenable, on a :

$$R_p\{x_k\} < Q_p\{x_k\} \text{ ou bien } R_p\{x_k\} > Q_p\{x_k\}$$

Cependant dans le cas où $p = 1$, on a toujours

$$R_1\{x_k\} \leq Q_1\{x_k\} \quad \forall \text{ la norme}$$

par conséquent, si φ est un processus itératif ayant pour limite x^* , on a :

$$R_1(\varphi, x^*) \leq Q_1(\varphi, x^*)$$

Propriété 10 :

Soit φ un processus itératif ayant pour limite x^* . S'il existe un $p \in [1, \infty[$ et une constante c_1 tels que $\forall \{x_k\} \in c(\varphi, x^*)$, on a :

$$\|x_{k+1} - x^*\| \leq c_1 \|x_k - x^*\|^p \quad \forall k \geq k_0 = k_0(\{x_k\}).$$

alors :

$$O_R(\varphi, x^*) \geq O_Q(\varphi, x^*) \geq p.$$

D'autre part, s'il existe une constante c_2 et une suite $\{x_k\} \in c(\varphi, x^*)$ telles que :

$$\|x_{k+1} - x^*\| \geq c_2 \|x_k - x^*\|^p > 0 \quad \forall k \geq k_0 = k_0(\{x_k\})$$

alors :

$$O_Q(\varphi, x^*) \leq O_R(\varphi, x^*) \leq p.$$

Ainsi, si les deux conditions sont vraies, on a :

$$O_Q(\varphi, x^*) = O_R(\varphi, x^*) = p.$$

Lemme 1 :

Quel que soit l'entier $m > 1$, le polynôme $p_m(t) = t^m - t^{m-1} - 1$ admet une seule racine r_m positive. De plus, $\tau_m \in]1, 2[$ et $\forall m \in \mathbb{N}$ on a :

$$\tau_m > \tau_{m-1} \quad \text{et} \quad \lim_{m \rightarrow \infty} \tau_m = 1$$

Proposition 1 :

Soient φ un processus itératif et $c(\varphi, x^*)$ l'ensemble des suites engendrées par φ et convergentes vers x^* et soient $\gamma_0, \dots, \gamma_n$ $n+1$ constantes non négatives. Si $\forall \{x_k\} \in c(\varphi, x^*) \exists k_0 \geq m$ tel que :

$$\|x_{k+1} - x^*\| \leq \|x_k - x^*\| \sum_{j=0}^m \gamma_j \|x_{k-j} - x^*\| \quad \forall k \geq k_0$$

alors :

$$O_R(\varphi, x^*) \geq \tau.$$

où τ est l'unique racine positive de $t^{m+1} - t^m - 1 = 0$;
si de plus il existe $\beta > 0$ et une suite $\{x_k\} \in c(\varphi, x^*)$ tels que pour $k_0 \geq m$, on a :

$$\|x_{k+1} - x^*\| \geq \beta \|x_k - x^*\| \quad \|x_{k-m} - x^*\| > 0 \quad \forall k \geq k_0$$

alors

$$O_R(\varphi, x^*) = \tau.$$

Proposition 2 [8].

Soit φ un processus itératif ayant x^* comme point limite et soient $\gamma_0, \dots, \gamma_m$ des constantes non négatives. Si \forall la suite $\{x_k\} \in c(\varphi, x^*)$ il existe un k_0 dépendant de $\{x_k\}$ et tel que :

$$\|x_{k+1} - x^*\| \leq \gamma_0 \|x_k - x^*\|^2 + \|x_k - x^*\| \sum_{j=1}^m \gamma_j \|x_{k-j} - x^*\|^2$$

pour tout $k \geq k_0$.

alors

$$O_R(\varphi, x^*) \quad \text{où } \tau \text{ est la racine positive de } t^{m+1} - t^m - 2 = 0.$$

et si de plus il existe $\gamma > 0$, $k_0 > m$ et une suite $\{x_k\} \in c(\varphi, x^*)$ tels que :

$$\|x_{k+1} - x^*\| \geq \gamma \|x_{k-m} - x^*\| \quad \|x_k - x^*\| > 0 \quad \forall k \geq k_0$$

alors :

$$O_Q(\varphi, x^*) \leq O_R(\varphi, x^*) = \tau.$$

II - CONVERGENCE LOCALE DE LA METHODE DES SECANTES.

Dans ce paragraphe, nous trouvons les conditions suffisantes pour lesquelles l'opérateur $J(x, H)$ est une approximation du jacobien $F'(x)$, si F est différentiable, ce qui va nous aider à établir le théorème de convergence locale d'Ortuga et Rheinboldt puis, au paragraphe III de ce chapitre, à établir le théorème de convergence semi-locale, et par suite à obtenir une précision de l'ordre de convergence de la suite d'itérés $\{x_k\}$ obtenue par la méthode des sécantes.

Aussi dans ce paragraphe nous trouvons le théorème de convergence locale de Tornheim, avec quelques remarques concernant les conditions suffisantes qu'il donnent et celles que nous donnons au théorème de convergence semi-locale, pour obtenir comme ordre de convergence, pour la méthode des sécantes, la racine positive τ de l'équation

$$t^{n+1} - t^n - 1 = 0$$

Définition 9 :

Soit $F : D \xrightarrow{\subset \mathbb{R}^n} \mathbb{R}^n$ une fonction G -différentiable dans $D_0 \subset D$

et soit $J : D_0 \times D_h \xrightarrow{\subset D \times L(\mathbb{R}^n)} L(\mathbb{R}^n)$

Si $0 \in L(\mathbb{R}^n)$ est un point limite de D_h et si :

$$\lim_{\substack{H \rightarrow 0 \\ H \in D_h}} J(x, H) = F'(x), \text{ uniformément } \forall x \in D_0$$

alors on dit que J est une approximation consistante de F' dans D_0 et s'il existe des constantes $c > 0$ et $r > 0$ telles que :

$$\|F'(x) - J(x, H)\| \leq c \|H\| \quad \forall x \in D_0$$

$$\text{et } \forall H \in D_h \cap S(0, r)$$

alors on dit que J est une approximation fortement consistante.

Lemme 2 :

Soit $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ une fonction continûment différentiable dans un convexe $D \subset \mathbb{R}^n$. Si F' satisfait :

$$\| F'(x) - F'(y) \| \leq \alpha \| x - y \|^p \quad \forall x, y \in D$$

avec α et p des constantes non négatives

alors $\forall x, y \in D$, on a :

$$\| F(y) - F(x) - F'(x)(y - x) \| \leq \frac{\alpha}{p+1} \| y - x \|^{p+1}$$

(pour la démonstration, voir [3]).

Proposition 3 [3] :

Supposons que F est G -différentiable dans un voisinage $S_0 \subset D$ de $x^* \in D$ pour lequel $F(x^*) = 0$, et que F' est continue, non singulière en x .

Si $J : D \times D_h \xrightarrow{\quad} L(\mathbb{R}^n)$ est une approximation consistante de F' dans S_0 , alors il existe δ et $r > 0$ tels que :

$$G(x, H) = x - J(x, H)^{-1} F(x)$$

est définie $\forall x \in S(x^*, \delta)$ et $\forall H \in D_{h'} = D_h \cap S(0, r)$ et satisfait :

$$\| x^* - G(x, H) \| \leq \omega(x, H) \| x - x^* \| \quad (I)$$

$\forall x \in S(x^*, \delta)$ et $\forall H \in D_{h'}$,

où $\omega(x, H) \rightarrow 0$ quand $x \rightarrow x^*$ et $H \rightarrow 0$, $H \in D_{h'}$.

De plus si J est une approximation fortement consistante de F' dans S_0 et si :

$$\| F'(x) - F'(x^*) \| \leq \gamma \| x - x^* \| \quad \forall x \in S_0$$

alors il existe des constantes α_1 et α_2 telles que :

$$\| x^* - G(x, H) \| \leq \alpha_1 \| x - x^* \|^2 + \alpha_2 \| H \| \| x - x^* \| \quad (II)$$

$\forall x \in S$ et $\forall H \in D_{h'}$.

Démonstration :

Posons $\beta = \|F'(x^*)^{-1}\|$ et soit $\varepsilon \in (0, 1/2\beta)$. Puisque J est une approximation consistante de F' dans S_0 , il existe un $r > 0$ tel que D_h , soit non vide et :

$$\|F'(x) - J(x, H)\| \leq \frac{1}{2} \varepsilon \quad \forall x \in S_0 \text{ et } H \in D_h,$$

De plus, d'après la continuité de F' en x^* , il existe un $\delta > 0$ tel que $S = S(x^*, \delta) \subset S_0$ et :

$$\|F'(x) - F'(x^*)\| \leq \frac{1}{2} \varepsilon \quad \forall x \in S$$

donc

$$\|F'(x^*) - J(x, H)\| \leq \varepsilon \quad \forall x \in S, H \in D_h,$$

or, $\|F'(x)^{-1}\| \leq \beta$ et $\varepsilon\beta \leq \frac{1}{2} < 1$, donc $J(x, H)^{-1}$ existe et satisfait :

$$\|J(x, H)^{-1}\| \leq \gamma = \beta/(1 - \varepsilon\beta) \quad \forall x \in S, H \in D_h,$$

d'où G est bien défini dans $S \times D_h$, et :

$$\begin{aligned} \|G(x, H) - x^*\| &= \|J(x, H)^{-1}[J(x, H)(x - x^*) - F(x)]\| \\ &\leq \eta [\|J(x, H) - F'(x)\| + \|F'(x) - F'(x^*)\|] \\ &\quad \times \|x - x^*\| + \eta \|F'(x) - F'(x^*) - F'(x^*)(x - x^*)\| \end{aligned}$$

ce qui montre I avec

$$\omega(x, H) = \eta [\|J(x, H) - F'(x)\| + \|F'(x) - F'(x^*)\| + q(x)] \quad (\text{III})$$

où

$$q(x) = \frac{\|F(x) - F(x^*) - F'(x)(x - x^*)\|}{\|x - x^*\|} \quad \text{pour } x \neq x^* \text{ et } q(x^*) = 0.$$

D'après la continuité de F' en x^* , $q(x) \rightarrow 0$ et

$$F'(x) - F'(x^*) \rightarrow 0 \quad \text{quand } x \rightarrow x^*$$

et d'après la définition 9, on a :

$$J(x, H) - F'(x) \rightarrow 0 \quad \text{quand } x \rightarrow x^* \text{ et } H \rightarrow 0.$$

Donc $\omega(x, H) \rightarrow 0$ quand $x \rightarrow x^*$ et $H \rightarrow 0$, $H \in D_h$.

Supposons maintenant que $\|F'(x) - F'(x^*)\| \leq \gamma \|x - x^*\| \quad \forall x \in S_0$;
d'après le lemme 2, on a :

$$\|q(x)\| \leq \gamma \|x - x^*\| \quad \forall x \in S$$

et II découle directement de (I) et (III) avec $\alpha_1 = 2\gamma\eta$ et $\alpha_2 = \eta c$, où c est la constante de la définition 9 de l'approximation fortement consistante.

Définition 10 :

σ étant une constante positive, on note par $K(\sigma)$ l'ensemble suivant :

$$K(\sigma) = \{H = (h^1, \dots, h^n) \in L(\mathbb{R}^n) / h^j \neq 0 \quad j = 1, \dots, n \\ \text{et tel que : } |\det(\frac{h^1}{\|h^1\|}, \dots, \frac{h^n}{\|h^n\|})| \geq \sigma\}$$

Q étant une famille de matrices de $L(\mathbb{R}^n)$, alors on dit que Q est uniformément non singulière si $Q \subset K(\sigma)$ pour un $\sigma > 0$.

Proposition 4 [3] :

Soit $Q \subset L(\mathbb{R}^n)$ une collection de matrices non singulières, alors Q est uniformément non singulière si et seulement s'il existe une constante $\alpha > 0$, telle que :

$$\|(\frac{h^1}{\|h^1\|}, \dots, \frac{h^n}{\|h^n\|})^{-1}\| \leq \alpha \quad (I) \\ \forall H = (h^1, \dots, h^n) \in Q.$$

Démonstration :

Supposons que $Q \subset K(\sigma)$ pour un $\sigma > 0$ et soit :

$$K_1 = \{A = (a_1, \dots, a_n) \in K(\sigma) / \|a_j\| = 1, j = 1, \dots, n\}$$

Il est clair que K_1 est borné. De plus, si $\{A_k\} \subset K_1$ est une suite convergente avec $A = \lim_{k \rightarrow \infty} A_k \in L(\mathbb{R}^n)$, alors $\|a_j\| = 1 (j = 1, \dots, n)$.

Et puisque le déterminant est une fonction continue de ses éléments, il s'ensuit que $|\det A| \geq \sigma$, donc $A \in K_1$ et K_1 est fermé.

D'où K_1 est un compact.

L'application $f : K_1 \rightarrow \mathbb{R}$ telle que $f(A) = \|A^{-1}\|$ est continue sur K_1 , donc il existe une constante α telle que $\|A^{-1}\| \leq \alpha$ pour toute matrice $A \in K_1$.

Mais, pour toute matrice $H \in Q$, la matrice :

$$\tilde{H} = \left(\frac{\|h^1\|}{\|h^1\|}, \dots, \frac{\|h^n\|}{\|h^n\|} \right) \text{ est aussi dans } K_1.$$

donc :

$$\|\tilde{H}^{-1}\| \leq \alpha \quad \forall H \in Q.$$

Inversement, supposons que (I) soit satisfaite ; l'ensemble

$$S = \{A \in L(\mathbb{R}^n) / \|A\| \leq \alpha\}$$

est compact, et d'après la continuité du déterminant, il existe une constante $\tau > 0$, telle que $|\det A| \leq \tau$ pour tout $A \in S$; mais $\tilde{H}^{-1} \in S$, pour tout $H \in Q$; et ainsi :

$$|\det \tilde{H}| = \frac{1}{|\det H^{-1}|} \geq \frac{1}{\tau} \quad \forall H \in Q.$$

Donc $Q \subset K(1/\tau)$, d'où Q est uniformément non singulière.

Proposition 5 [3] :

Soit $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ une fonction continûment différentiable dans un ouvert D et soit $Q \subset L(\mathbb{R}^n)$, une famille de matrices uniformément non singulière telle que la matrice nulle ait un point limite de Q . Alors, pour tout sous-ensemble compact D_0 de D , il existe $r > 0$, tel que l'application $J(x, H)$ soit bien définie dans $D_0 \times D_h$ avec $D_h = \{H \in Q / \|H\| \leq r\}$ et c'est une approximation consistante de F' dans D_0 .

De plus si

$$\|F'(x) - F'(y)\| \leq \gamma \|x - y\| \quad \forall x, y \in D \quad (I)$$

alors J est une approximation fortement consistante.

Démonstration :

Puisque D_0 est compact et que D est un ouvert, il existe un $\delta > 0$, tel que le compact $D_1 = \{x / \|x - y\| \leq \delta \text{ pour un } y \in D_0\} \subset D$.

F' est alors uniformément continue dans D_1 et, pour un $\varepsilon > 0$ donné, nous pouvons choisir $r \in (0, \delta)$ tel que :

$$\delta(h) = F(x+h) - F(x) - F'(x)h$$

satisfasse $\|\delta(h)\| \leq \varepsilon \|h\|$ chaque fois que $x \in D_0$ et $\|h\| \leq r$.

Soit maintenant β une constante telle que :

$$\|A\| \leq \max_i \|a_i\| \quad \forall A = (A_1, \dots, a_n) \in L(\mathbb{R}^n)$$

β dépend uniquement de la norme choisie (pour la L_1 -norme $\beta = 1$).

Soit α une constante telle que :

$$\left\| \left(\frac{h^1}{\|h^1\|}, \dots, \frac{h^n}{\|h^n\|} \right)^{-1} \right\| \leq \alpha$$

$$\forall H \in Q \text{ et } H = (h^1, \dots, h^n)$$

Alors pour tout $x \in D_0$ et $H \in Q$, tels que $\|h^i\| \leq r \quad i = 1, \dots, n$, nous obtenons :

$$\begin{aligned} \|J(x, H) - F'(x)\| &= \|\delta[h^1], \dots, \delta[h^n]H^{-1}\| \\ &= \left\| \left[\frac{\delta(h^1)}{\|h^1\|}, \dots, \frac{\delta(h^n)}{\|h^n\|} \right] \left(\frac{h^1}{\|h^1\|}, \dots, \frac{h^n}{\|h^n\|} \right)^{-1} \right\| \\ &\leq \alpha \left\| \left(\frac{\delta(h^1)}{\|h^1\|}, \dots, \frac{\delta(h^n)}{\|h^n\|} \right) \right\| \leq \alpha\beta\varepsilon \end{aligned}$$

Ainsi J est une approximation consistante de F' dans D_0 .

Si de plus (I) est vraie, alors on peut démontrer d'après le lemme 2, que :

$$\|\delta(h)\| \leq \frac{1}{2} \gamma \|h\|^2$$

ainsi que :

$$\|J(x, H) - F'(x)\| \leq \frac{1}{2} \alpha\beta\gamma \max_i \|h^i\|$$

Théorème 1 [3] :

Soit $F : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction continûment différentiable dans un ouvert D de \mathbb{R}^n ; supposons qu'il existe un $x^* \in D$ pour lequel $F(x^*) = 0$ et $F'(x^*)$ est non singulière. Supposons de plus que $\sigma > 0$ est choisie de telle sorte que $K(\sigma)$ soit non vide.

Alors il existe une constante $r_1 > 0$ telle que, pour toute suite $\{H_k\} \subset K(\sigma)$ avec $\|H_k\| \leq r_1$ $k = 0, 1, \dots$, la suite d'itérés de la méthode des sécantes est bien définie et $\lim_{k \rightarrow \infty} x_k = x^*$ pourvu que $\|x_0 - x^*\|$ soit assez petit.

De plus si $\lim_{k \rightarrow \infty} H_k = 0$, alors :

$$R_1\{x_k\} = Q_1\{x_k\} = 0$$

Démonstration :

Notons d'abord que $K(\sigma)$ est invariant par changement d'échelle, c'est-à-dire que si $H \in K(\sigma)$, alors $\alpha H \in K(\sigma)$ pour tout $\alpha \neq 0$, ainsi la matrice nulle est un point limite de $K(\sigma)$.

Choisissons maintenant un $\delta > 0$ tel que $S = \bar{S}(x^*, \delta) \subset D$, alors la proposition précédente montre alors qu'il existe un $r > 0$, tel que J est bien défini dans $S \times D_h$ avec :

$$D_h = \{H \in K(\sigma) / \|H\| \leq r\}$$

et c'est une approximation consistante de F' dans S .

D'après la proposition 3, on a :

$$\|x_{k+1} - x^*\| \leq \omega(x_k, H) \|x_k - x^*\| \quad \forall x \in S, H \in D_h,$$

avec $\omega(x_k, H) \rightarrow 0$ quand $x \rightarrow x^*$ et $H \rightarrow 0$.

Ainsi il existe $S' \subset S$ et $\alpha \in (0, 1)$ tels que :

$$\omega(x_k, H) \leq \alpha \quad \forall x_k \in S' \text{ et } \forall H \in D_h,$$

d'où :

$$\begin{aligned} \|x_{k+1} - x^*\| &\leq \alpha \|x_k - x^*\| \\ &\leq \alpha^{k+1} \|x_0 - x^*\| \end{aligned}$$

par suite, la suite $\{x_k\}$ reste dans S' et converge vers x^* et d'après la définition de $Q_1\{x_k\}$, on a :

$$R_1\{x_k\} \leq Q_1\{x_k\} \leq \limsup_{k \rightarrow \infty} \omega(x_k, H_k) = 0$$

Théorème 2 [3] : Convergence locale d'Ortega et Rheinboldt.

Supposons que F satisfasse les conditions de la proposition précédente alors il existe des constantes $\sigma > 0$ et $r > 0$, telles que si :

1°) $\|x_0 - x^*\|$ est suffisamment petit.

2°) La suite $\{x_k\}$ des itérés de la méthode des sécantes est bien définie.

3°) La suite des matrices H_k correspondantes à x_k reste dans l'ensemble D_h où :

$$D_h = \{H \in K(\sigma) / \|H\| \leq r\}.$$

Alors $\lim_{k \rightarrow \infty} x_k = x^*$ et $R_1\{x_k\} = Q_1\{x_k\} = 0$.

Si de plus F' est lipchitienne, c'est-à-dire il existe une constante γ telle que :

$$\|F'(x) - F'(y)\| \leq \gamma \|x - y\| \quad \forall x, y \in D$$

alors la suite $\{x_k\}$ satisfait :

$$Q_R\{x_k\} \geq \tau$$

où τ est la racine positive de $t^{n+1} - t^n - 1 = 0$

Démonstration :

La convergence et la convergence superlinéaire se déduisent immédiatement de la proposition précédente.

Supposons maintenant que F' est lipchitienne, alors il existe une constante β telle que :

$$\|H_k\| \leq \beta \sum_{j=1}^n \|x_k - x_{k-j}\| \leq n\beta \|x_k - x^*\| + \beta \sum_{j=1}^n \|x_{k-j} - x^*\|$$

Or d'après la proposition 3 :

$$\begin{aligned} \|x_{k+1} - x^*\| &\leq \alpha_1 \|x_k - x^*\|^2 + \alpha_2 \|x_k - x^*\| \|H_k\| \quad \forall k \geq k_0 \\ &\leq (\alpha_1 + n\beta\alpha_2) \|x_k - x^*\|^2 + \alpha_2\beta \|x_k - x^*\| \left(\sum_{j=1}^n \|x_{k-j} - x^*\| \right) \end{aligned}$$

et ainsi l'ordre de convergence se déduit immédiatement de la proposition 1.

Théorème 3 [6] : De convergence locale de Tornheim :

Soit $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ et supposons que :

1. x^* est une solution pour le système $F(x) = 0$
2. Les composantes f_i ($i = 1, \dots, n$) de F ont des dérivées secondes continues.
3. $\{x_k\}$ est la suite d'itérés de la méthode des sécantes suivant la formulation de Wolfe.
4. Le jacobien $J(x) = \det(\partial f_i / \partial x_j) \quad \begin{matrix} i = 1, \dots, n \\ j = 1, \dots, n \end{matrix}$ est non nul en $x = x^*$.
5. $\|x_k - x^*\| < s \|x_{k-1} - x^*\| \quad k = 1, \dots, n$ avec $s < 1$
6. V_k est définie par :

$$V_k = \det \begin{pmatrix} 1 & \dots & \dots & \dots & 1 \\ x_{k,1} & & & & x_{k-n,1} \\ \vdots & & & & \vdots \\ x_{k,n} & & & & x_{k-n,n} \end{pmatrix} = \det \begin{pmatrix} 1 & \dots & & & 1 \\ & & & & \\ & & & & \\ x_k & \dots & & & \\ & & & & x_{k-n} \end{pmatrix}$$

et $|V_k| \geq c \|x_k - x_{k-1}\| \|x_k - x_{k-2}\| \dots \|x_k - x_{k-n}\|$
avec $k = n, n+1, \dots$ et c est une constante positive.

- Alors pour x_0 assez proche de x^* , la suite $\{x_k\}$ converge vers x^* et l'ordre de convergence ρ , s'il existe, est au moins τ , où τ est la racine positive de

$$t^{n+1} - t^n - 1 = 0$$

(Démonstration voir [6]).

Remarques :

Notons tout d'abord que le théorème 2 d'Ortega et Rheinboldt et le théorème 3 de Tornheim sont de type "local". Ils supposent l'existence d'une solution x^* pour le système $F(x) = 0$ et que les approximations initiales sont suffisamment proche de x^* .

Nous rappelons que cette étude de convergence locale est importante pour la caractérisation de cette méthode et la comparaison de celle-ci avec les autres processus itératifs pour la résolution d'un système non linéaire.

Comme on l'a déjà vu, la comparaison sera faite à partir de l'ordre de convergence. Les théorèmes 2 et 3 donnent le même ordre de convergence pour la méthode des sécantes, mais les conditions ne sont pas toutes les mêmes et on peut faire les remarques suivantes :

1. Le théorème 2 utilise la formulation de Newton, le théorème 3 utilise la formulation de Wolfe. Nous rapelons que nous avons déjà établi (chapitre I) l'équivalence de ces deux formulations.
2. Les deux théorèmes supposent que F est continûment différentiable et que $F'(x^*)$ est non singulière, c'est-à-dire que les hyperplans tangents aux f_i ($i = 1, \dots, n$) ont un point d'intersection unique.
3. Il est clair que la condition 5 du théorème 3 peut être satisfaite par une numérotation convenable des points de départ x_0, \dots, x_n , à la seule condition que les $\|x_i - x^*\|$ $i = 0, \dots, n$ ne soient pas égales.
4. Le théorème 2 suppose que F' est Lipchitienne.

5. Le théorème 3 suppose que F est deux fois continûment différentiable.
6. Le théorème 2 suppose que la suite d'itérés $\{x_k\}$ obtenu par la méthode des sécantes est bien définie et qu'il existe un $\sigma > 0$ et un $r > 0$ tels que la suite $\{H_k\}$ correspondante à $\{x_k\}$ est incluse dans l'ensemble D_h où

$$D_h = \{H \in K(\sigma) / \|H\| \leq r\}.$$

Notons que le théorème 1 établi par Ortega et Rheinboldt montre l'existence d'un δ et d'un $r > 0$ tels que :

$$G(x, H) = x - J(x, H)^{-1}F(x) \text{ est bien défini}$$

est bien définie $\forall x \in S(x^*, \delta)$ et $\forall H \in D_h$.

Donc, il suffit de vérifier que la suite $\{H_k\}$ obtenue par l'itération de la méthode des sécantes est incluse dans D_h , c'est-à-dire qu'il existe un $r > 0$ et un $\sigma > 0$ tels que :

$$\|H_k\| \leq r \text{ et } H_k \in K(\sigma).$$

Or on a vu, au théorème 1, qu'il existe $S' \subset S$ telle que :

$$\|G(x_k, H_k) - x^*\| \leq \omega(x_k, H_k) \|x_k - x^*\|$$

$$\text{avec } \omega(x_k, H_k) \leq \alpha < 1.$$

Ainsi $x_{k+1} = G(x_k, H_k) \in S'$. Si δ' est le rayon de S' et si on pose $r = 2\delta'$, on a :

$$\|H_{k+1}\| = \max_{i=1, \dots, n} \|h^i\| \leq 2\delta' = r$$

Donc, pour montrer que $H_k \in D_h$, il suffit de montrer que $H_k \in K(\sigma)$. Or $H_k \in K(\sigma)$, avec $H_k = (h^1, \dots, h^n)$ si :

$$\left| \det \left(\frac{h^1}{\|h^1\|}, \dots, \frac{h^n}{\|h^n\|} \right) \right| \geq \delta$$

ou si on a :

$$|\det(h^1, \dots, h^n)| \cdot |\det(\text{diag}(1/\|h^1\|, \dots, 1/\|h^n\|))| \geq \sigma$$

ou

$$|\det(h^1, \dots, h^n)| \cdot \frac{1}{\|h^1\|} \cdot \dots \cdot \frac{1}{\|h^n\|} \geq \sigma$$

donc il suffit que :

$$|\det H_k| \geq \sigma \|h^1\| \cdot \|h^2\| \cdot \dots \cdot \|h^n\|$$

qui peut être satisfaite si H_k est non singulière ou x_k, \dots, x_{k-n} en position générale.

7. Le théorème 3 suppose que $|V_k| \geq c \|x_k - x_{k-1}\| \dots \|x_k - x_{k-n}\|$ où c est une constante positive.

D'après le théorème 1, chapitre I, on a :

$$V_k = \det \begin{pmatrix} 1 & \dots & 1 \\ x_k & \dots & x_{k-n} \end{pmatrix} = \det(x_k - x_{k-1}, \dots, x_k - x_{k-n})$$

Si on pose $H_k = (x_k - x_{k-1}, \dots, x_k - x_{k-n}) = (h^1, \dots, h^n)$.

Donc, il faut qu'on ait à chaque pas de l'itération :

$$|V_k| = |\det H_k| \geq c \|h^1\|, \dots, \|h^n\|$$

qui peut être aussi satisfaite si x_k, \dots, x_{k-n} sont en position générale $\forall k = n, n+1, \dots$

Il est clair que la condition " H_k est non singulière" ou x_k, \dots, x_{k-n} sont en position générale à chaque pas de l'itération, est la plus difficile à tester.

Cependant, dans le théorème de convergence semi-locale que nous allons donner au chapitre suivant, nous démontrons l'existence d'une racine pour $F(x) = 0$ dans une boule A et la convergence de la méthode des sécantes à partir de $n+1$ points x_0, \dots, x_n de \mathbb{R}^n convenablement choisis dans A . Il est clair que les conditions que nous proposons dans ce théorème vont être satisfaites si on est suffisamment proche de x^* racine de $F(x) = 0$, (sauf la condition F' est lipchitienne). Et nous démontrons alors qu'à chaque pas de l'itération, H_k est non singulière, c'est-à-dire x_k, \dots, x_{k-n} sont en position générale.

III - CONVERGENCE SEMI-LOCALE DE LA METHODE DES SECANTES.

Notations :

On désigne par :

$$1. \quad \|x\| = \varphi_1(x) = \sum_{i=1}^n |x_i| \text{ pour } x \in \mathbb{R}^n$$

$$2. \quad \|B\| = S_1(B)$$

$$= \max_{x \in \mathbb{R}^n} \frac{\varphi_1(Bx)}{\varphi_1(x)} = \max_j \sum_{i=1}^n |b_{ij}| = \max_j \varphi_1(B.j)$$

où B est une matrice de $M_{n,n}(\mathbb{R})$ telle que :

$$B = (b_{ij}) \quad \begin{array}{l} i = 1, \dots, n \\ j = 1, \dots, n \end{array}$$

B.j = j^{ième} colonne de B.

$$3. \quad M_k(x_k, \dots, x_{k-n}) = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ x_k & \dots & x_{k-n} \end{pmatrix}$$

$$4. \quad F : D \xrightarrow{\subset \mathbb{R}^n} \mathbb{R}^n \text{ une fonction non linéaire.}$$

Dans la suite, nous allons étudier la convergence semi-locale de la méthode des sécantes pour résoudre le système non linéaire $F(x) = 0$ où à chaque pas le nouvel itéré est donné par :

$$x_{k+1} = x_k - H_k \Gamma_k^{-1} F(x_k)$$

avec

$$H_k = (x_k - x_{k-1}, \dots, x_{k-n+1} - x_{k-n})$$

$$\Gamma_k = (F(x_k) - F(x_{k-1}), \dots, F(x_{k-n+1}) - F(x_{k-n}))$$

Théorème 4 :

Soient $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ et x_0, \dots, x_n n+1 points de \mathbb{R}^n .

On suppose qu'il existe $\eta > 0, \beta > 0, 0 < \alpha < 1, a_n > 0$ tels que si $A = \{x \in \mathbb{R}^n ; \|x - x_n\| \leq \beta\}$:

$$1^\circ) \quad \|H_n \Gamma_n^{-1}\| = (1 - \alpha)\eta < \eta$$

2°) $\forall y_0, y_1, \dots, y_n$ qui sont en position générale et appartenant à A et $\forall z_0, \dots, z_n$ en position générale et appartenant à A ,
on a :

$$(I) \quad \begin{aligned} & \left\| [F(y_n) - F(y_{n-1}), \dots, F(y_1) - F(y_0)] (y_n - y_{n-1}, \dots, y_1 - y_0)^{-1} \right. \\ & \quad \left. - [F(z_n) - F(z_{n-1}), \dots, F(z_1) - F(z_0)] (z_n - z_{n-1}, \dots, z_1 - z_0)^{-1} \right\| \\ & \leq \frac{\alpha}{2^n} \end{aligned}$$

$$3^\circ) \quad \|F(x_n)\| \leq \frac{\beta}{2^{n_n}}$$

$$4^\circ) \quad \|x_{j+1} - x_j\| \leq \frac{\beta}{2^j} \quad \forall j = 0, \dots, n-1$$

$$5^\circ) \quad \|M_n^{-1}\| \leq a_n$$

$$6^\circ) \quad c_n = a_n \cdot \beta < \frac{1}{2}$$

Alors $\forall k = n+1, \dots$, on a :

$$1. \quad \|H_{k+1}^{-1}\| < n \tag{1}$$

$$2. \quad \|x_k - x_{k-1}\| \leq n \|F(x_{n-1})\| < \frac{\beta}{2^{k-1}} \tag{2}$$

$$3. \quad \|x_k - x_n\| < \beta \tag{3}$$

$$4. \quad a_k = \frac{a_{k-1}}{1 - c_{n-1}}, \quad c_k = \frac{a_k}{2^{k-n}} = \frac{c_{k-1}}{2(1 - c_{k-1})} < \frac{1}{2} \tag{4}$$

$$5. \quad M_k \text{ est inversible et } \|M_k^{-1}\| \leq a_k \tag{5}$$

d'où H_k est inversible.

$$6. \quad \|F(x_k)\| \leq \frac{\alpha}{2^n} \|x_k - x_{k-1}\| \tag{6}$$

et la suite $\{x_k\}$ des itérés de la méthode des sécantes converge vers $x^* \in A$ racine de $F(x) = 0$.

Démonstration :

Nous allons démontrer par récurrence les conditions (1), ..., (6).

Pour $j = n + 1$:

Nous avons par hypothèse :

$$* \quad ||| H_n \Gamma_n^{-1} ||| < \eta$$

$$* \quad ||| x_{n+1} - x_n || = ||| H_n \Gamma_n^{-1} F(x_n) || \leq ||| H_n \Gamma_n^{-1} ||| ||F(x_n)|| \\ \leq \eta \frac{\beta}{2^n \eta} = \frac{\beta}{2^n} < \beta$$

* en posant :

$$\Delta_n = M_n(x_n, \dots, x_0) - M_{n+1}(x_{n+1}, \dots, x_1) \\ = \begin{pmatrix} 0 & \dots & 0 \\ x_n - x_{n+1} & \dots & x_0 - x_1 \end{pmatrix}$$

on a :

$$||| \Delta_n ||| \leq \max (||x_0 - x_1|| , \dots , ||x_n - x_{n+1}||) \\ \leq \max (\beta , \frac{\beta}{2} , \dots , \frac{\beta}{2^n})$$

$$||\Delta_n|| \leq \beta$$

d'où

$$||| M_n^{-1} \Delta_n ||| \leq ||| M_n^{-1} ||| ||| \Delta_n ||| \leq a_n \beta = c_n < \frac{1}{2}$$

donc $(I - M_n^{-1} \Delta_n)$ est inversible et comme on a :

$$M_{n+1} = M_n (I - M_n^{-1} \Delta_n)$$

alors M_{n+1} est inversible.

et par suite :

$$\begin{aligned} \|\| M_{n+1}^{-1} \|\| &\leq \|\| M_n^{-1} \|\| \|\| (I - M_n^{-1} \Delta_n)^{-1} \|\| \\ &\leq a_n \cdot (1 + c_n + c_n^2 + \dots) \\ &\leq a_n \frac{1}{1 - c_n} \end{aligned}$$

Posons :

$$a_{n+1} = \frac{a_n}{1 - c_n} \quad \text{et} \quad c_{n+1} = \frac{a_n}{1 - c_n} \cdot \frac{\beta}{2} = \frac{a_{n+1} \beta}{2}$$

on a bien :

$$c_{n+1} < \frac{1}{2}$$

Comme M_{n+1} est inversible d'après le théorème 1 du chapitre I, on a H_{n+1} est inversible.

$$* \|\| F(x_{n+1}) \|\| = \|\| F(x_{n+1}) - F(x_n) - \Gamma_n H_n^{-1} (x_{n+1} - x_n) \|\|$$

Or, on peut écrire que :

$$F(x_{n+1}) - F(x_n) = \Gamma_{n+1} H_{n+1}^{-1} (x_{n+1} - x_n)$$

en effet :

$$H_{n+1}^{-1} (x_{n+1} - x_n) = (x_{n+1} - x_n, \dots, x_2 - x_1)^{-1} (x_{n+1} - x_n) = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

par suite

$$\Gamma_{n+1} H_{n+1}^{-1} (x_{n+1} - x_n) = \Gamma_{n+1} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = F(x_{n+1}) - F(x_n)$$

ainsi

$$\begin{aligned} \|\| F(x_{n+1}) \|\| &= \|\| \Gamma_{n+1} H_{n+1}^{-1} (x_{n+1} - x_n) - \Gamma_n H_n^{-1} (x_{n+1} - x_n) \|\| \\ &\leq \|\| \Gamma_{n+1} H_{n+1}^{-1} - \Gamma_n H_n^{-1} \|\| \|\| x_{n+1} - x_n \|\| \end{aligned}$$

Or, x_{n+1}, \dots, x_1 et x_n, \dots, x_0 sont en position générale et appartiennent à A, donc :

$$\|F(x_{n+1})\| \leq \frac{\alpha}{2\eta} \|x_{n+1} - x_n\|$$

Supposons maintenant que (1) (2), ..., (6) soient vraies pour $j = n+1, \dots, k$; nous allons démontrer qu'elles sont vraies pour $j = K+1$:

En effet : (H_k étant inversible d'après (5)) :

$$\begin{aligned} 1^* H_n \Gamma_n^{-1} \Gamma_k H_k^{-1} &= I - H_n \Gamma_n^{-1} (\Gamma_n H_n^{-1} - \Gamma_k H_k^{-1}) \\ &= I - E \end{aligned}$$

$$\text{Or } \|E\| \leq \|H_n \Gamma_n^{-1}\| \|\Gamma_n H_n^{-1} - \Gamma_k H_k^{-1}\|$$

$$\leq \eta \frac{\alpha}{2\eta} = \frac{\alpha}{2} < \frac{1}{2}$$

car x_0, \dots, x_n et x_{k-n}, \dots, x_k sont en position générale et appartiennent à A.

Donc $I - E$ est inversible et :

$$H_k \Gamma_k^{-1} = H_n \Gamma_n^{-1} (I + E + \dots + E^n + \dots)$$

et

$$\|H_k \Gamma_k^{-1}\| \leq \|H_n \Gamma_n^{-1}\| (1 + \|E\| + \dots + \|E^n\| + \dots)$$

$$\|H_k \Gamma_k^{-1}\| \leq (1 - \alpha) \frac{\eta}{1 - \frac{\alpha}{2}} = \frac{1 - \alpha}{1 - \frac{\alpha}{2}} \eta$$

$$2^* \|x_{k+1} - x_k\| = \|H_k \Gamma_k^{-1} F(x_k)\| \leq \|H_k \Gamma_k^{-1}\| \|F(x_k)\|$$

$$\leq \frac{1 - \alpha}{1 - \frac{\alpha}{2}} \eta \|F(x_k)\| .$$

$$\text{Comme } \|F(x_k)\| \leq \| \Gamma_k H_k^{-1} - \Gamma_{k-1} H_{k-1}^{-1} \| \|x_k - x_{k-1}\|$$

$$\leq \frac{\alpha}{2\eta} \|x_k - x_{k-1}\| \quad (\text{d'après I})$$

$$\|x_{k+1} - x_n\| \leq \frac{1-\alpha}{1-\frac{\alpha}{2}} \eta \frac{\alpha}{2\eta} \|x_k - x_{k-1}\|$$

$$\|x_{k+1} - x_k\| \leq \frac{1-\alpha}{1-\frac{\alpha}{2}} \frac{\alpha}{2} \|x_k - x_{k-1}\| \quad (\text{II})$$

$$\leq \frac{\alpha}{2} \|x_k - x_{k-1}\| \leq \left(\frac{\alpha}{2}\right)^{k-n} \|x_{n+1} - x_n\|$$

$$\leq \left(\frac{\alpha}{2}\right)^{k-n} \frac{\beta}{2^n}$$

d'où :

$$\|x_{k+1} - x_k\| \leq \frac{\beta}{2^k}$$

$$3^* \|x_{k+1} - x_n\| \leq \|x_{k+1} - x_k\| + \|x_k - x_{k-1}\| + \dots + \|x_{n+1} - x_n\|$$

$$\leq \sum_{j=n}^k \|x_{j+1} - x_j\| \leq \sum_{j=n}^k \left(\frac{\alpha}{2}\right)^{j-n} \|x_{n+1} - x_n\|$$

$$\leq \sum_{j=0}^{k-n} \left(\frac{\alpha}{2}\right)^j \|x_{n+1} - x_n\| \leq \frac{1}{1-\frac{\alpha}{2}} \frac{\beta}{2^n}$$

$$\|x_{k+1} - x_n\| \leq \frac{\beta}{2^{n-1}} < \beta$$

4* En posant :

$$a_{k+1} = \frac{a_k}{1-c_k} \quad \text{et} \quad c_{k+1} = a_{k+1} \frac{\beta}{2^{k-n+1}}$$

on a :

$$c_{k+1} = \frac{a_k}{1-c_k} \frac{\beta}{2^{k-n+1}} = \frac{c_k}{2(1-c_k)} < \frac{1}{2}$$

$$\text{car } c_k < \frac{1}{2}$$

5* Comme (1), ..., (6) sont vraies $\forall j = n+1, \dots, k$,
alors on a :

$$\|M_k^{-1}\| \leq a_k \quad \text{et} \quad c_k = a_k \frac{\beta}{2^{k-n}} < \frac{1}{2}$$

Nous allons démontrer que M_{k+1}^{-1} existe et $\|M_{k+1}^{-1}\| \leq a_{k+1}$

En effet :

$$\Delta_k = M_k - M_{k+1} = \begin{pmatrix} 0 & & & 0 \\ & & & \\ & & & \\ x_k - x_{k+1}, & \dots, & x_{k-n} - x_{k-n+1} \end{pmatrix}$$

et

$$\begin{aligned} \|\Delta_k\| &\leq \max(\|x_k - x_{k+1}\|, \dots, \|x_{k-n} - x_{k-n+1}\|) \\ &\leq \max\left(\frac{\beta}{2^k}, \dots, \frac{\beta}{2^{k-n}}\right) \end{aligned}$$

d'où

$$\|\Delta_k\| \leq \frac{\beta}{2^{k-n}} \quad \text{et} \quad \|M_k^{-1} \Delta_k\| \leq a_k \frac{\beta}{2^{k-n}} = c_k < \frac{1}{2}$$

et ainsi $I - M_k^{-1} \Delta_k$ est inversible et par suite :

$$M_{k+1} = M_k (I - M_k^{-1} \Delta_k)$$

est inversible et

$$M_{k+1}^{-1} = M_k^{-1} (I - M_k^{-1} \Delta_k)^{-1} = (I + M_k^{-1} \Delta_k + \dots) M_k^{-1}$$

donc :

$$\|M_{k+1}^{-1}\| \leq a_k \frac{1}{1 - c_k} = a_{k+1}$$

et d'après le théorème 1 du chapitre I, H_{k+1}^{-1} existe.

$$6* \quad \|F(x_{k+1})\| = \|F(x_{k+1}) - F(x_k) - \Gamma_k H_k^{-1}(x_{k+1} - x_k)\|$$

Or on peut aussi démontrer que :

$$F(x_{k+1}) - F(x_k) = \Gamma_{k+1} H_{k+1}^{-1} (x_{k+1} - x_k)$$

donc

$$\begin{aligned} \|F(x_{k+1})\| &= \|\Gamma_{k+1} H_{k+1}^{-1} (x_{k+1} - x_k) - \Gamma_k H_k^{-1} (x_{k+1} - x_k)\| \\ &\leq \|\Gamma_{k+1} H_{k+1}^{-1} - \Gamma_k H_k^{-1}\| \|x_{k+1} - x_k\| \end{aligned}$$

d'où :

$$\|F(x_{k+1})\| \leq \frac{\alpha}{2^n} \|x_{k+1} - x_k\|$$

car $x_{k+1}, \dots, x_{k-n+1}$ et x_k, \dots, x_{k-n} sont en position générale et appartiennent à A.

Donc $\forall j = n+1, \dots$ les relations (1), (2), ..., (6) sont vraies, alors on peut écrire :

$$\begin{aligned} \|x_{k+p} - x_k\| &\leq \|x_{k+p} - x_{k+p-1}\| + \dots + \|x_{k+1} - x_k\| \\ &\leq \sum_{j=0}^{p-1} \|x_{j+k+1} - x_{j+k}\| \\ &\leq \sum_{j=0}^{p-1} \left(\frac{\alpha}{2}\right)^{j-n+k} \|x_{n+1} - x_n\| = \left(\frac{\alpha}{2}\right)^{k-n} \sum_{j=0}^{p-1} \left(\frac{\alpha}{2}\right)^j \|x_{n+1} - x_n\| \\ (7) \quad &\leq \left(\frac{\alpha}{2}\right)^{k-n} \frac{1}{1 - \frac{\alpha}{2}} \frac{\beta}{2^n} < \frac{\alpha^{k-n}}{2^k} \cdot \beta \end{aligned}$$

Cela $\forall k > n$, donc la suite $\{x_k\}$ des itérés de la méthode des sécantes est une suite de Cauchy et on a :

$$x^* = \lim_{k \rightarrow \infty} x_k$$

et de la relation (6), on tire que à la limite

$$F(x^*) = 0$$

et ainsi le théorème est démontré.

Théorème 5

Si $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ et x_0, \dots, x_n vérifiant les conditions du théorème précédent, alors :

$$R_1\{x_k\} < 1$$

et par suite la convergence est R-linéaire.

Aussi on a :

$$Q_1\{x_k\} < 1$$

et la convergence est Q-linéaire.

Démonstration :

D'après la définition 5. On a :

$$R_1\{x_k\} = \limsup_{k \rightarrow \infty} \|x_k - x^*\|^{1/k}$$

or, d'après la relation (7) du théorème précédent, on a :

$$\|x_{k+p} - x_k\| \leq \frac{\alpha^{k-n}}{2^k} \beta \text{ ceci } \forall p.$$

et lorsque $p \rightarrow \infty$, on a toujours :

$$\|x_k - x^*\| \leq \frac{\alpha^{k-n}}{2^k} \beta$$

d'où :

$$R_1\{x_k\} = \limsup_{k \rightarrow \infty} \|x_k - x^*\|^{1/k} \leq \limsup_{k \rightarrow \infty} \left(\frac{\alpha^{k-n}}{2^k} \beta \right)^{1/k}$$

$$\leq \limsup_{k \rightarrow \infty} \left(\left(\frac{\alpha}{2} \right)^k \cdot \frac{\beta}{\alpha^n} \right)^{1/k}$$

$$\leq \frac{\alpha}{2} \limsup_{k \rightarrow \infty} \left(\frac{\beta}{\alpha^n} \right)^{1/k}$$

$$\leq \frac{\alpha}{2} < 1$$

et par suite la convergence est R-linéaire.

Il reste à démontrer que la convergence est aussi Q-linéaire, c'est-à-dire que :

$$Q_1\{x_k\} = \limsup_{k \rightarrow \infty} \frac{\|x_{k+1} - x^k\|}{\|x_k - x^*\|} < 1$$

En effet, d'après la relation (II) du théorème précédent, on a :

$\forall k \geq n$

$$\|x_{k+1} - x_k\| \leq \frac{1-\alpha}{1-\frac{\alpha}{2}} \frac{\alpha}{2} \|x_k - x_{k-1}\|$$

$$\leq \gamma \|x_k - x_{k-1}\|$$

$$\text{avec } \gamma = \frac{1 - \alpha}{1 - \frac{\alpha}{2}} \frac{\alpha}{2} = \frac{(1 - \alpha)\alpha}{2 - \alpha}$$

d'où :

$$\begin{aligned} \|x_{k+p} - x_{k+1}\| &\leq \|x_{k+p} - x_{k+p-1}\| + \dots + \|x_{k+2} - x_{k+1}\| \\ &\leq \sum_{j=0}^{p-1} \gamma^j \|x_{k+2} - x_{k+1}\| \end{aligned}$$

et lorsque $p \rightarrow \infty$, on a :

$$x_{k+p} \rightarrow x^*$$

et

$$\begin{aligned} \|x_{k+1} - x^*\| &\leq \frac{1}{1 - \gamma} \|x_{k+2} - x_{k+1}\| \\ &\leq \frac{1}{1 - \gamma} \gamma \|x_{k+1} - x_k\| \\ &\leq \frac{\gamma}{1 - \gamma} [\|x_{k+1} - x^*\| + \|x_k - x^*\|] \end{aligned}$$

$$\left(1 - \frac{\gamma}{1 - \gamma}\right) \|x_{k+1} - x^*\| \leq \frac{\gamma}{1 - \gamma} \|x_k - x^*\|$$

$$\text{or } \gamma \leq \frac{\alpha}{2} \leq \frac{1}{2}$$

$$\text{d'où } \frac{\gamma}{1 - \gamma} < 1$$

$$\text{et par suite : } 1 - \frac{\gamma}{1 - \gamma} > 0$$

d'où

$$\begin{aligned} \|x_{k+1} - x^*\| &\leq \frac{\frac{\gamma}{1 - \gamma}}{1 - \frac{\gamma}{1 - \gamma}} \|x_k - x^*\| \\ &\leq \frac{\gamma}{1 - 2\gamma} \|x_k - x^*\| \end{aligned}$$

Nous allons démontrer maintenant que $\frac{\gamma}{1-2\gamma} < 1$, c'est-à-dire que $3\gamma < 1$; comme $\gamma = \frac{(1-\alpha)\alpha}{2-\alpha}$, il faut démontrer que :

$$3\gamma = \frac{3(1-\alpha)\alpha}{2-\alpha} < 1$$

c'est-à-dire :

$$\begin{aligned} 3\alpha - 3\alpha^2 &< 2 - \alpha \\ \text{ou} \\ -3\alpha^2 + 4\alpha - 2 &< 0 \end{aligned}$$

Or il est facile à vérifier que cette équation n'admet pas de racine, dor elle a le même signe que le coefficient de α^2 .

$$\text{Donc } \frac{\gamma}{1-2\gamma} < 1$$

Or :

$$\begin{aligned} Q_1\{x_k\} &= \limsup_{k \rightarrow \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} \\ &\leq \limsup_{k \rightarrow \infty} \frac{\gamma}{1-2\gamma} = \frac{\gamma}{1-2\gamma} < 1 \end{aligned}$$

donc la convergence est Q-linéaire.

Théorème 6 :

Soit F une fonction de \mathbb{R}^n dans \mathbb{R}^n et x_0, \dots, x_n $n+1$ points de \mathbb{R}^n vérifiant les conditions du théorème 4, alors la méthode des sécantes engendre une suite $\{x_k\}$ qui converge vers x^* racine de $F(x)$. Si de plus F est différentiable, si F' est continue en x^* et $F'(x^*)$ est non singulière, alors :

$$Q_1\{x_k\} = 0$$

et par suite la convergence est Q-superlinéaire.

Démonstration :

D'après le théorème 4, la suite $\{x_k\}$ converge vers x^* et $F(x^*) = 0$.

Si H_k est la matrice obtenue au $k^{\text{ième}}$ pas, on pose :

$$H_k = (h_1, \dots, h_n)$$

et

$$\tilde{H}_k = \left(\frac{h_1}{\|h_1\|}, \dots, \frac{h_n}{\|h_n\|} \right)$$

alors $\tilde{H}_k = H_k D_k^{-1}$ où $D_k = \text{diag} (\|h_1\|, \dots, \|h_n\|)$.

Si

$$P = \begin{pmatrix} 1 & 0 & \dots & 0 \\ x_k & H_k & & \end{pmatrix} \in M_{n+1, n+1}(\mathbb{R})$$

alors :

$$P = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_k & x_{k-1} & \dots & x_{k-n} \end{pmatrix} \begin{pmatrix} 1 & -1 & \circ & \\ \vdots & \vdots & \vdots & -1 \\ \circ & \vdots & \vdots & 1 \end{pmatrix}$$

et

$$P^{-1} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ \xi_1 & & & \\ \vdots & & & \\ \xi_n & & H_k^{-1} & \end{pmatrix} = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \vdots & \vdots \\ \circ & \vdots & \circ \\ \vdots & \vdots & \vdots \\ \circ & \vdots & 1 \end{pmatrix} \cdot M_k^{-1} \quad (1)$$

d'après le théorème 4, on a $\forall k$:

$$\|M_k^{-1}\| \leq a_k$$

$$\text{avec } a_k = \frac{a_{k-1}}{1 - c_{k-1}}$$

d'après (1), on a :

$$\|H_k^{-1}\| \leq (n+1) \|M_k^{-1}\| \leq (n+1)a_k$$

or

$$\|\tilde{H}_k^{-1}\| \leq \|D_k\| \|H_k^{-1}\|$$

$$\text{mais } \|D_k\| = \max(\|x_k - x_{k-1}\|, \dots, \|x_{k-n+1} - x_{k-n}\|)$$

$$\leq \frac{\beta}{2^{k-n}}$$

ainsi

$$\| \tilde{H}_k^{-1} \| \leq (n+1) \frac{\beta}{2^{k-n}} a_k \leq (n+1) c_k < \frac{n+1}{2} = \text{cte.}$$

D'après la proposition 4, la suite $\{H_k\}$ est uniformément non singulière. Et d'après la proposition 5, $J(x, H)$ est une approximation consistante de F' .

D'après la proposition 3, comme F est différentiable, comme $F'(x^*)$ est non singulière et comme H_k est une approximation consistante de F' au voisinage de x^* ; on a alors :

$$\|x_{k+1} - x^*\| \leq \omega(x_k, H_k) \|x_k - x^*\|$$

où $\omega(x_k, H_k) \rightarrow 0$ quand $x_k \rightarrow x^*$ et $H_k \rightarrow 0$.

Par suite :

$$\limsup_{k \rightarrow \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = \lim_{k \rightarrow \infty} \omega(x_k, H_k) = 0$$

donc $Q_1\{x_k\} = 0$.

Et la convergence est Q-superlinéaire. De la proposition 9, on déduit que la convergence est aussi R-superlinéaire.

Proposition 6 :

$$\text{Si } H_k = (x_k - x_{k-1}, \dots, x_{k-n+1} - x_{k-n})$$

et

$$\tilde{H}_k = (x_{k-1} - x_k, x_{k-2} - x_k, \dots, x_{k-n} - x_k)$$

alors on a :

$$J(x_k, H_k) = J(x_k, \tilde{H}_k).$$

Démonstration :

$$J(x_k, \tilde{H}_k) = \tilde{\Gamma}_k \tilde{H}_k$$

$$\text{avec } \tilde{\Gamma}_k = (F(x_{k-1}) - F(x_k), \dots, F(x_{k-n}) - F(x_k))$$

$$J(x_k, \tilde{H}_k) = \tilde{\Gamma}_k \tilde{H}_k = \tilde{\Gamma}_k P P^{-1} \tilde{H}_k^{-1} = (\tilde{\Gamma}_k P) (\tilde{H}_k P)^{-1}$$

$$\text{avec } P = - \begin{pmatrix} 1 & -1 & \dots & 0 \\ & \ddots & \ddots & -1 \\ & & \ddots & \ddots \\ 0 & & & 1 \end{pmatrix}$$

$$\text{Or } \tilde{\Gamma}_k P = (F(x_k) - F(x_{k-1}), \dots, F(x_{k-n+1}) - F(x_{k-n}))$$

$$\text{d'où } \tilde{\Gamma}_k P = \Gamma_k$$

$$\text{de même } \tilde{H}_k P = H_k$$

$$\text{donc } J(x_k, \tilde{H}_k) = \tilde{\Gamma}_k \tilde{H}_k^{-1} = \Gamma_k H_k^{-1} = J(x_k, H_k)$$

Proposition 7 :

Supposons que F soit continûment différentiable de \mathbb{R}^n dans \mathbb{R}^n et que F' satisfasse :

$$\|F'(x) - F'(y)\| \leq \varrho \|x - y\| \quad \forall x, y \in \mathbb{R}^n$$

s'il existe $a_k > 0$ et $\beta > 0$ et s'il existe $x_{k-n}, \dots, x_k \in \mathbb{R}^n$ tels que :

$$1. \quad \|M_k^{-1}\| = \left\| \begin{pmatrix} 1 & \dots & 1 \\ x_k & \dots & x_{k-n} \end{pmatrix}^{-1} \right\| \leq a_k$$

$$2. \quad \|x_{j+1} - x_j\| \leq \frac{\beta}{2^j} \quad \forall j = k-n, \dots, k-1$$

alors :

$$\|J(x_k, H_k) - F'(x_k)\| \leq \varrho \frac{\beta}{2^{k-n}} a_k \frac{\beta}{2^{k-n-1}}$$

Démonstration :

Comme $\tilde{H}_k = (x_{k-1} - x_k, \dots, x_{k-n} - x_k) \in M_{n,n}(\mathbb{R})$

Posons $\tilde{H}_k = (h_1, \dots, h_n)$

et posons $\delta(h_i) = F(x_k + h_i) - F(x_k) - F'(x_k)h_i \quad \forall i = 1, \dots, n$

alors :

$$\begin{aligned} \|\|J(x_k, \tilde{H}_k) - F'(x_k)\|\| &= \|\|(\delta(h_1), \dots, \delta(h_n))\tilde{H}_k^{-1}\|\| \\ &= \|\| \left(\frac{\delta(h_1)}{\|h_1\|}, \dots, \frac{\delta(h_n)}{\|h_n\|} \right) \left(\frac{h_1}{\|h_1\|}, \dots, \frac{h_n}{\|h_n\|} \right)^{-1} \|\| \end{aligned}$$

Posons $P = \left(\frac{h_1}{\|h_1\|}, \dots, \frac{h_n}{\|h_n\|} \right)$

$$D = \text{diag} (\|h_1\|, \dots, \|h_n\|)$$

$$Q = \left(\frac{\delta(h_1)}{\|h_1\|}, \dots, \frac{\delta(h_n)}{\|h_n\|} \right)$$

alors $\|\|J(x_k, \tilde{H}_k) - F'(x_k)\|\| = \|\|Q P^{-1}\|\| \quad \|\|Q\|\| \cdot \|\|P^{-1}\|\|$

or $\|\|P^{-1}\|\| = \|\|D \tilde{H}_k^{-1}\|\| \leq \|\|D\|\| \|\|\tilde{H}_k^{-1}\|\|$

et $\|\|D\|\| = \max_i \|h_i\| = \max_i \|x_{k-i} - x_k\| \quad (i = 1, \dots, n)$

Or $\|\|x_{k-i} - x_k\|\| = \|\|x_{k-i} - x_{k-i+1}\|\| + \dots + \|\|x_{k-1} - x_k\|\|$

$$\leq \frac{\beta}{2^{k-i}} + \frac{\beta}{2^{k-i+1}} \dots + \frac{\beta}{2^{k-1}}$$

$$\leq \frac{\beta}{2^{k-i}} \left(1 + \frac{1}{2} + \dots + \frac{1}{2^{k-i-1}} \right)$$

$$\leq \frac{\beta}{2^{k-i}} \frac{1}{1 - \frac{1}{2}} = \frac{\beta}{2^{k-i-1}}$$

D'où

$$\| \| D \| \| = \max_i \| h_i \| = \max_i \frac{\beta}{2^{k-i-1}} = \frac{\beta}{2^{k-n-1}}$$

On peut démontrer aussi que :

$$\| \| H_k^{-1} \| \| \leq a_k$$

En effet :

$$\begin{aligned} \begin{pmatrix} 1 & 0 & \dots & 0 \\ x_k & \tilde{H}_k & & \end{pmatrix} &= \begin{pmatrix} 1 & 0 & \dots & 0 \\ x_k & x_{k-1} - x_k & \dots & x_{k-n} - x_k \end{pmatrix} \\ &= \begin{pmatrix} 1 & \dots & 1 \\ x_k & x_{k-1} & \dots & x_{k-n} \end{pmatrix} \begin{pmatrix} 1 & -1 & \dots & 1 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 \end{pmatrix} \end{aligned}$$

Donc :

$$\begin{aligned} \begin{pmatrix} 1 & 0 & \dots & 0 \\ x_k & \tilde{H}_k & & \end{pmatrix}^{-1} &= \begin{pmatrix} 1 & 0 & \dots & 0 \\ \xi_1 & \tilde{H}_k^{-1} & & \\ \vdots & & & \\ \xi_n & & & \end{pmatrix} = \begin{pmatrix} 1 & -1 & \dots & -1 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & \dots & 1 \\ x_k & \dots & x_{k-n} \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 1 & 1 & \dots & 1 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 \end{pmatrix} M_k^{-1} \end{aligned}$$

d'où $\| \| \tilde{H}_k^{-1} \| \| \leq \| \| M_k^{-1} \| \| \leq a_k$

d'autre part, on a $\forall i = 1, \dots, n :$

$$\| \delta(h_i) \| = \| \delta(x_{k-i} - x_k) \| = \| F(x_{k-i}) - F(x_k) - F'(x_k)(x_{k-i} - x_k) \|$$

et comme F' est de Lipchitz, d'après le lemme 2, on a :

$$\begin{aligned} \|\delta(h_i)\| &\leq \frac{1}{2} \ell \|x_{k-i} - x_k\|^2 \\ &\leq \frac{1}{2} \ell \|h_i\|^2 \end{aligned}$$

$$\begin{aligned} \text{Or } \|Q\| &= \max_i \frac{\|\delta(h_i)\|}{\|h_i\|} \leq \max_i \frac{1}{2} \ell \|h_i\| \\ &\leq \frac{1}{2} \ell \frac{\beta}{2^{k-n-1}} \end{aligned}$$

$$\text{donc } \|Q\| \leq \ell \frac{\beta}{2^{k-n}}$$

et par suite :

$$\begin{aligned} \|J(x_k, \tilde{H}_k) - F'(x_k)\| &\leq \|Q\| \|P^{-1}\| \\ &\leq \|Q\| \|D\| \|\tilde{H}_k^{-1}\| \\ \|J(x_k, \tilde{H}_k) - F'(x_k)\| &\leq \ell \frac{\beta}{2^{k-n}} a_k \frac{\beta}{2^{k-n-1}} \end{aligned}$$

Théorème 7 :

Soit $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ une fonction continûment différentiable. On suppose qu'il existe $n+1$ points de \mathbb{R}^n x_0, \dots, x_n en position générale et des constantes :

$$\eta > 0, \beta > 0, 0 < \alpha < 1, a_n > 0 \text{ et } \ell > 0$$

tels que si $A = \{x \in \mathbb{R}^n / \|x - x_n\| \leq \beta\}$, on a :

1. F' est de lipchitz, c'est-à-dire :

$$\|F'(x) - F'(y)\| \leq \ell \|x - y\| \quad \forall x, y \in A$$

2. $\|H_n \Gamma_n^{-1}\| = (1 - \alpha)\eta < \eta$

3. $\|x_{j+1} - x_j\| < \frac{\beta}{2^j} \quad \forall j = 0, \dots, n-1$

$$4. \quad \|M_n^{-1}\| = \left\| \begin{pmatrix} 1 & \cdots & 1 \\ x_n & \cdots & x_0 \end{pmatrix} \right\| \leq a_n$$

$$5. \quad c_n = a_n \beta < \frac{1}{2}$$

$$6. \quad 2\beta \leq \frac{\beta}{2\eta}$$

alors les relations (1), ..., (6) du théorème 4 sont vraies $\forall k \geq n+1$, et par suite, la suite $\{x_k\}$ des itérés de la méthode des sécantes converge vers x^* racine de $F(x) = 0$.

Démonstration :

En remarquant qu'on a les mêmes conditions que le théorème 4, sauf (I), la démonstration sera faite aussi par récurrence et au pas $j = n+1$; on obtient de la même façon :

$$1 - \|\|H_n \Gamma_n^{-1}\|\| < \eta$$

$$2 - \|x_{n+1} - x_n\| < \frac{\beta}{2^n}$$

$$3 - \|\|M_{n+1}^{-1}\|\| \leq \frac{a_n}{1 - c_n} = a_{n+1}$$

et aussi H_{n+1}^{-1} existe

$$4 - C_{n+1} = a_{n+1} \beta \leq \frac{1}{2}$$

Il reste à démontrer que $\|F(x_{n+1})\| \leq \frac{\alpha}{2\eta} \|x_{n+1} - x_n\|$

En effet :

$$\begin{aligned} \|F(x_{n+1})\| &= \|\Gamma_{n+1} H_{n+1}^{-1}(x_{n+1} - x_n) - \Gamma_n H_n^{-1}(x_{n+1} - x_n)\| \\ &\leq \|\|\Gamma_{n+1} H_{n+1}^{-1} - \Gamma_n H_n^{-1}\|\| \|x_{n+1} - x_n\| \end{aligned}$$

or

$$\begin{aligned} \|\|\Gamma_{n+1} H_{n+1}^{-1} - \Gamma_n H_n^{-1}\|\| &\leq \|\|\Gamma_{n+1} H_{n+1}^{-1} - F'(x_{n+1})\|\| + \|\|\Gamma_n H_n^{-1} - F'(x_n)\|\| \\ &\quad + \|\|F'(x_{n+1}) - F'(x_n)\|\| \end{aligned}$$

et d'après la proposition 6, on a :

$$J(x_{n+1}, H_{n+1}) = J(x_{n+1}, \tilde{H}_{n+1}) = \Gamma_{n+1} H_{n+1}^{-1}$$

et aussi :

$$J(x_n, H_n) = J(x_n, \tilde{H}_n) = \Gamma_n H_n^{-1}$$

alors on peut écrire :

$$\| \| J(x_{n+1}, H_{n+1}) - F'(x_{n+1}) \| \| = \| \| J(x_{n+1}, \tilde{H}_{n+1}) - F'(x_{n+1}) \| \|$$

et d'après la proposition 7, comme $\| \| M_{n+1}^{-1} \| \| \leq a_{n+1}$, on a :

$$\begin{aligned} \| \| J(x_{n+1}, H_{n+1}) - F'(x_{n+1}) \| \| &\leq \frac{\ell\beta}{2^n} a_{n+1} \frac{\beta}{2^0} \\ &\leq \frac{\ell\beta}{2} c_{n+1} < \frac{\ell\beta}{4} \end{aligned}$$

De même on a :

$$\begin{aligned} \| \| J(x_n, H_n) - F'(x_n) \| \| &\leq \frac{\ell\beta}{2^0} a_n \frac{\beta}{2^{-1}} \\ &\leq 2\ell\beta c_n \leq \ell\beta \end{aligned}$$

F' étant lipchititienne, alors :

$$\| \| F'(x_{n+1}) - F'(x_n) \| \| \leq \ell \| x_{n+1} - x_n \| \leq \frac{\ell\beta}{2^n}$$

d'où

$$\begin{aligned} \| \| J(x_{n+1}, H_{n+1}) - J(x_n, H_n) \| \| &\leq \frac{\ell\beta}{4} + \ell\beta + \frac{\ell\beta}{2^n} \\ &\leq \frac{3}{2} \ell\beta \text{ car } n \geq 2 \\ &\leq \frac{\alpha}{2\eta} \quad (\text{car par hypothèse } 2\ell\beta \leq \frac{\alpha}{2\eta}) \end{aligned}$$

donc $\| \| F(x_{n+1}) \| \| \leq \frac{\alpha}{2\eta} \| x_{n+1} - x_n \|$

Supposons maintenant que (1), ..., (6) du théorème 4 soient vraies pour $j = n+1, \dots, k$; nous allons démontrer qu'elles sont aussi vraies pour $j = k+1$.

$$\begin{aligned} H_n \Gamma_n^{-1} \cdot \Gamma_k H_k^{-1} &= I - H_n \Gamma_n^{-1} (\Gamma_n H_n^{-1} - \Gamma_k H_k^{-1}) \\ &= I - E \end{aligned}$$

$$\begin{aligned} \text{Or : } \quad \|\Gamma_n H_n^{-1} - \Gamma_k H_k^{-1}\| &= \|J(x_n, H_n) - J(x_k, H_k)\| \\ &\leq \|J(x_n, H_n) - F'(x_n)\| + \|J(x_k, H_k) - F'(x_k)\| + \|F'(x_n) - F'(x_k)\| \\ &\leq \ell\beta + \frac{\ell\beta a_k}{2^{k-n}} \frac{\beta}{2^{k-n-1}} + \ell \|x_n - x_k\| \\ &\leq \ell\beta + \frac{\ell\beta c_k}{2^{k-n-1}} + \frac{\ell\beta}{2^{n-1}} \leq \ell\beta + \frac{\ell\beta}{2^{k-n}} + \frac{\ell\beta}{2^{n-1}} \\ &\leq 2\ell\beta \leq \frac{\alpha}{2\eta} \end{aligned}$$

$$\text{d'où } \|E\| \leq \frac{\alpha}{2\eta} = \frac{\alpha}{2} < \frac{1}{2}$$

et avec la même démonstration (dans le théorème 4), on trouve :

- 1°) $\|H_k \Gamma_k^{-1}\| < \eta$
- 2°) $\|x_{k+1} - x_k\| \leq \frac{\beta}{2^k}$
- 3°) $\|x_{k+1} - x_n\| \leq \frac{\beta}{2^{n-1}}$
- 4°) $a_{k+1} = \frac{a_k}{1 - c_k}$ et $c_{k+1} = \frac{a_{k+1}\beta}{2^{k-n+1}} < \frac{1}{2}$
- 5°) $\|M_{k+1}^{-1}\| \leq a_{k+1}$ et H_{k+1}^{-1} existe

et, en utilisant les propositions 6 et 7, on trouve de la même manière que précédemment :

$$\begin{aligned} \left\| J(x_{k+1}, H_{k+1}) - J(x_k, H_k) \right\| &\leq \frac{\ell\beta}{2^{k-n+1}} a_{k+1} \frac{\beta}{2^{k-n}} + \frac{\ell\beta}{2^k} + \frac{\ell\beta}{2^{k-n}} a_k \frac{\beta}{2^{k-n-1}} \\ &\leq 2\ell\beta \leq \frac{\alpha}{2\eta} \end{aligned}$$

d'où $\|F(x_{k+1})\| \leq \frac{\alpha}{2\eta} \|x_{k+1} - x_k\|$

Ainsi (1), ..., (6) du théorème 4 sont vraies $j = n+1, \dots$ et alors la suite $\{x_k\} \rightarrow x^*$ quand $k \rightarrow \infty$ et on a :

$$F(x^*) = 0$$

Ordre de convergence :

Si F est une fonction de \mathbb{R}^n dans \mathbb{R}^n et si x_0, \dots, x_n sont $n+1$ points de \mathbb{R}^n vérifiant les conditions du théorème précédent, alors la suite des itérés de la méthode des sécantes converge vers x^* racine de $F(x)$ et l'ordre de convergence ρ de cette suite est :

$$\rho \geq \tau$$

où τ est la racine positive de $t^{n+1} - t^n - 1 = 0$

Démonstration :

D'après le théorème précédent $\{x_k\} \rightarrow x^*$ et $F(x^*) = 0$, F' étant lipchilitienne, d'après la proposition 5, J est une approximation consistante forte de F' et, d'après la proposition 4, on a :

$$\|x^* - x_{k+1}\| \leq \alpha_1 \|x_k - x^*\|^2 + \alpha_2 \|H_k\| \|x_k - x^*\|$$

or $\|H_k\| = \max(\|x_k - x_{k-1}\|, \dots, \|x_{k-n+1} - x_{k-n}\|)$

$$= \|x_{k-n+1} - x_{k-1}\| \leq \|x_{k-n+1} - x^*\| + \|x_{k-n} - x^*\|$$

d'où :

$$\|x^* - x_{n+1}\| \leq \alpha_1 \|x_k - x^*\|^2 + \alpha_2 \|x_{k-n+1} - x^*\| \|x_{k-n} - x^*\| \\ \cdot \|x_k - x^*\|$$

et d'après la proposition 1 :

$$O_R\{x_k\} \geq \tau \text{ où } \tau \text{ est la racine positive de } t^{n+1} - t^n - 1 = 0$$

CHAPITRE III

ETUDE DE L'EFFICACITE DE LA METHODE DES SECANTES

APPLICATIONS NUMERIQUES

INTRODUCTION

Il est important dans l'étude des processus itératifs pour la résolution des systèmes d'équations (non linéaires), d'avoir une méthode précise pour classer ces algorithmes suivant leur "efficacité".

Plusieurs auteurs ont étudié l'efficacité des méthodes itératives pour trouver les racines des fonctions à une variable (Traub [7], Feldestein et Firestone [14], Ostrowski [4]). Mis à part l'étude faite par Shmanskii [17] et celle faite par Brent [1] et [2], nous ne connaissons pas d'autre étude de l'efficacité des processus itératifs pour la résolution des systèmes d'équations non linéaires.

La définition de l'efficacité donnée par Shamanski, puis par Brent n'est autre qu'une généralisation de l'index d'efficacité d'Ostrowski [4] :

$$E_0 = \lim_{i \rightarrow \infty} W_i \sqrt{\frac{\log \|x_{i+1} - x^*\|}{\log \|x_i - x^*\|}}$$

où W_i est le "travail" exigé pour passer de x_i à x_{i+1} , dont l'unité représente une évaluation de F ou de l'une de ses dérivées partielles, ($F(x)$ étant la fonction dont on cherche les racines).

Or $\rho = \lim_{i \rightarrow \infty} \frac{\log \|x_{i+1} - x^*\|}{\log \|x_i - x^*\|}$, s'il existe, est l'ordre de convergence

de la méthode en question. Si $W = \lim_{i \rightarrow \infty} W_i$, alors on a :

$$E_0 = \frac{1}{\rho}$$

Mais l'ordre de convergence n'est pas toujours le même quel que soit la fonction dont on cherche les racines ; Broyden [11] donne la définition suivante de l'efficacité d'un processus itératif pour la résolution d'un système particulier $F(x) = 0$:

$$R = \frac{1}{m} \log \frac{N_i}{N_f}$$

où m est le nombre total d'évaluations de F , N_i et N_f sont respectivement la norme euclidienne de F au point initial x_0 et au point final x_v .

Dans la suite, nous allons examiner la définition de Brent de l'efficacité d'une méthode itérative, puis, comme application, nous donnons des tables comparant la méthode des sécantes avec celle de Newton discrète, Brown, Brown modifié et celles de Brent. Enfin quelques exemples numériques et applications de la définition de Broyden sur ces exemples pour comparer la méthode des sécantes à d'autres méthodes.

Définition de l'efficacité de Brent [1] :

Si φ est un processus itératif pour la résolution d'un système d'équations non linéaires $F(x) = 0$, $\{x_k\}$ une suite d'itérés engendrée par φ et ayant pour limite x^* solution de $F(x) = 0$, W_{i+1} le "travail" exigé pour calculer le nouvel itéré x_{i+1} de x_i , alors l'efficacité de φ est donnée par :

$$E = \lim_{i \rightarrow \infty} \frac{1}{W_{i+1}} \log \left(\frac{\log \|x_{i+1} - x^*\|}{\log \|x_i - x^*\|} \right)$$

Remarques :

1 - $\lim_{i \rightarrow \infty} \frac{\log \|x_{i+1} - x^*\|}{\log \|x_i - x^*\|}$ si elle existe n'est autre que l'ordre de

convergence ρ de φ au sens de Brent, qui n'est autre que le R-ordre d'Ortega et Rheinboldt.

2 - Si $W = \lim_{i \rightarrow \infty} \frac{1}{W_i}$ et si ρ existe, alors l'efficacité de Brent s'écrit :

$$E = \frac{\log \rho}{W} = \log \rho^{\frac{1}{W}}$$

qui n'est autre que le logarithme de l'ordre d'efficacité E_0 d'Ostrowski.

3 - W_i peut être mesuré dans n'importe quelle unité uniforme. Dans la suite nous utilisons comme unité l'évaluation de F ; ainsi une évaluation d'une composante $f_i(x)$ de F est considérée comme $\frac{1}{n}$ unité du travail.

4 - E est indépendante de la norme choisie.

5 - Si deux processus φ et φ' ont respectivement comme efficacité E et E' et exigent respectivement $W(\epsilon)$ et $W'(\epsilon)$ unités de travail pour trouver x_i tel que $\|x_i - x^*\| < \epsilon$, alors :

$$\lim_{\epsilon \rightarrow 0} \frac{W(\epsilon)}{W'(\epsilon)} = \frac{E'}{E}$$

Ainsi φ exige E'/E fois plus de travail que φ' pour obtenir un x_i tel que $\|x_i - x^*\|$ soit plus petit qu'une constante assez petite.

De l'examen des résultats obtenus par Winograd et Wolfe [18] pour $n = 1$, Brent déduit la conjecture suivante qui est fortement plausible.

Conjecture [2] :

Il n'y a pas de méthodes localement convergentes, basées entièrement sur l'évaluation des fonctions et ayant une efficacité plus grande que $E_{\infty, n}$, avec :

$$E_{\infty, n} = \log \tau$$

où τ est l'unique racine positive de l'équation

$$\sum_{j=0}^{\infty} t^{-\frac{(n+j)!}{n! j!}} = 1$$

I - COMPARAISON DE LA METHODE DES SECANTES AVEC D'AUTRES METHODES

Dans cette partie nous comparons pour différentes valeurs de n l'efficacité de la méthode des sécantes et celle de la méthode de Newton discrète, Brown, Brown modifié et la méthode optimale S_k parmi la classe des méthodes de Brent, avec $E_{\infty, n}$.

Si n est l'ordre du système, alors on désigne dans la suite par :

- * $E(S) = \log \rho_s$ L'efficacité de la méthode des sécantes ρ_s étant son ordre de convergence ; on a déjà trouvé (chapitre II) que ρ_s est l'unique racine positive τ de l'équation

$$t^{n+1} - t^n - 1 = 0$$

D'après le chapitre I, pour calculer le nouvel itéré x_{k+1} , la méthode des sécantes exige une évaluation par pas, sauf au premier pas ; donc $W = 1$.

- * $E(N) = \frac{\log 2}{n+1}$ L'efficacité de la méthode de Newton discrète.

* $E(B) = \frac{2 \log 2}{n+3}$ l'efficacité de la méthode de Brown,

* $E_S(n)$ l'efficacité de la méthode S_k optimale de Brent,

* $E_T(n)$ l'efficacité de la méthode T_k optimale de Brent ou Brown modifiée.

TABLE 1

n	ρ_S	$E(S)$	$E_{\infty,n}$	$E(S)/E_{\infty,n}$	$E_T(n)/E_{\infty,n}$
2	1.465571	0.38228	0.43826	0.87242	0.717140
3	1.380278	0.32228	0.34143	0.94401	0.81205
4	1.324718	0.28117	0.28801	0.97633	0.87603
5	1.285198	0.25087	0.25320	0.99085	0.91251
6	1.285423	0.22747	0.22821	0.99675	0.941569
7	1.232055	0.20868	0.20875	0.99966	0.962935
8	1.213149	0.193219	0.193219	1.0000	0.98054
9	1.197492	0.180229	0.180229	1.0000	0.99224
10	1.184276	0.16912	0.16912	1.0000	1.00976
20	1.114465	0.10844	0.10844	1.0000	1.09576
50	1.058439	0.05687	0.05687	1.0000	1.20448
100	1.034430	0.03375	0.03375	1.0000	1.28106
1000	1.005261189	0.00524731	0.00524739	1.0000	1.474012

Table 2

n	$E(N)/E_{\infty,n}$	$E(B)/E_{\infty,n}$	$E_S(n)/E_{\infty,n}$
2	0.52719	0.63263	0.6817
3	0.50753	0.675018	0.7048
4	0.48133	0.688309	0.7161
5	0.45625	0.68438	0.7227
6	0.433903	0.67688	0.7243
7	0.41505	0.66409	0.7253
8	0.39859	0.65369	0.7264
9	0.38459	0.64226	0.72787
10	0.37259	0.62968	0.7285
20	0.30438	0.55701	0.7417
50	0.23898	0.45885	0.7672
100	0.20334	0.398553	0.7872
1000	0.1319617	0.263923	0.8432355

Remarques :

1 - D'après la table 1, nous remarquons que pour $n \geq 8$

$$E(S) \approx E_{\infty,n}$$

2 - Pour n assez grand, on a :

$$E_{\infty,n} \approx \frac{\log n}{n} \quad E(S)/E_{\infty,n} \approx 1$$

$$E(T_k)/E_{\infty,n} \approx 2 \quad E(N)/E_{\infty,n} \approx \frac{\log 2}{\log n}$$

$$E(B)/E_{\infty,n} \approx \frac{2 \log 2}{\log n} \quad E_S(n)/E_{\infty,n} \approx 1$$

3 - A part la méthode T_k ou Brown modifiée, la méthode des sécantes est la plus efficace parmi ces méthodes et surtout par rapport à la méthode de Newton, ce que montre la table 2.

4 - D'après la table 1, nous déduisons que la méthode T_k a une efficacité plus grande que $E_{\infty, n}$, d'où une efficacité plus grande que celle de la méthode des sécantes pour n assez grand.

Mais d'autre part, la méthode T_k exige $O(n^3)$ opérations par pas [13] et la méthode des sécantes, comme on va le voir après, exige $O(n^2)$ opérations par pas, sauf au premier pas. Or pour n assez grand l'ordre d'opérations n'est pas négligeable.

Ainsi la méthode T_k est plus efficace, au sens de Brent, que la méthode des sécantes parmi cette classe des méthodes itératives pour la résolution des systèmes d'équations non linéaires, mais elle exige beaucoup plus d'opérations.

Maintenant existe-t-il des méthodes plus efficaces au sens de Brent que la méthode des sécantes et exigeant moins d'opérations ; c'est un problème ouvert.

II - CALCUL DES ITERES DE LA METHODE DES SECANTES

On a déjà vu (chapitre I) que l'itéré de la méthode des sécantes obtenu à partir de $n+1$ points donnés n'est autre que l'intersection dans \mathbb{R}^{n+1} de n hyperplans qui interpolent respectivement les n composants f_i ($i = 1, \dots, n$) de F en ces $n+1$ points.

Nous avons ainsi n systèmes à résoudre :

$$[j] \begin{cases} \alpha_1^j x_{k-n}^1 + \alpha_2^j x_{k-n}^2 + \dots + \alpha_n^j x_{k-n}^n + \alpha_{n+1}^j f_j(x_{k-n}) = 1 \\ \dots \\ \alpha_1^j x_i^1 + \alpha_2^j x_i^2 + \dots + \alpha_n^j x_i^n + \alpha_{n+1}^j f_j(x_i) = 1 \\ \dots \\ \alpha_1^j x_k^1 + \alpha_2^j x_k^2 + \dots + \alpha_n^j x_k^n + \alpha_{n+1}^j f_j(x_k) = 1 \end{cases}$$

où $k-n \leq i \leq k$ et $x_i = (x_i^1, x_i^2, \dots, x_i^n) \quad \forall j = 1, \dots, n$

alors

$$\Gamma_{n+1}^{-1} = \Gamma_n'^{-1} + \mu \Gamma_n'^{-1} U e^T \Gamma_n'^{-1}$$

$$\text{où } \mu = \frac{-1}{1 + e^T \Gamma_n'^{-1} U}$$

En faisant cette permutation pour calculer

$$x_{n+2} = x_{n+1} - H_{n+1} \Gamma_{n+1}^{-1} F(x_{n+1})$$

il faut

additions	multiplications	divisions
$n(4n-1)$	$n(4n+2)$	1

Donc la méthode des sécantes exige $O(n^2)$ opérations par pas sauf au premier

III - APPLICATIONS NUMERIQUES

Dans l'ordre de comparer la performance des diverses méthodes utilisées pour résoudre un système particulier d'équations algébriques non linéaires, nous allons prendre tant que possible les mêmes conditions initiales.

La performance de ces méthodes va être résumée pour chaque problème "test" dans une table, laquelle donne la norme Euclidienne initiale et finale N_i et N_f de F , le nombre total m d'évaluations de fonctions exigé pour réduire la norme de F à N_f , ce nombre d'évaluations comprend aussi le nombre d'évaluations exigé pour l'approximation de Jacobien dans chaque cas, et donne aussi l'efficacité de Broyden R , définit par :

$$R = \frac{1}{m} \log \frac{N_i}{N_f}$$

(comme on l'a vu dans l'introduction de ce chapitre).

Exemple 1

Fonction de Rosenberg [16]

Cette fonction est souvent utilisée dans les problèmes de minimisation, c'est une fonction à 2 variables avec une composante linéaire :

$$\begin{aligned} f_1(x_1, x_2) &= 1 - x_1 \\ f_2(x_1, x_2) &= 10(x_2 - x_1^2) \end{aligned}$$

Le point (1, 1) est la racine de cette fonction.

Les valeurs initiales utilisées par les méthodes de Newton, Broyden I, Broyden II [11], Powell [16], Brown [13], Brown modifié [1] et Brent [1] sont :

$$\left. \begin{aligned} x_0^1 &= -1,2 \\ x_0^2 &= 1,0 \end{aligned} \right\} N_0 = 4.919$$

Les valeurs initiales utilisées par la méthode des sécantes sont :

$$\left. \begin{aligned} x_0^1 &= -1,2 \\ x_0^2 &= 1,0 \end{aligned} \right\} N_0 = 4.919$$

$$\left. \begin{aligned} x_1^1 &= 0,5 \\ x_1^2 &= 1,1 \end{aligned} \right\} N_1 = 8,53$$

$$\left. \begin{aligned} x_2^1 &= 0,5 \\ x_2^2 &= 0,9 \end{aligned} \right\} N_2 = 6,532$$

Table 3 - Résultats de l'exemple

Méthode	m	N_f		R
Newton	12	7.451	10^{-8}	1.500
Broyden I	59	4.729	10^{-10}	0.3909
Broyden II	39	2.547	10^{-10}	0.673
Powell	70	1.000	10^{-10}	0.154
Brown	8	0.000		-
Brown modifié	8	0.000		-
Brent	8	0.000		-
Sécantes	6	0.000		-

Exemple 2 [11] :

$$f_1 = -(3 + \alpha x_1)x_1 + 2x_2 - 1$$

⋮

$$f_i = x_{i-1} - (3 + \alpha x_i)x_i + 2x_{i+1} - 1, \quad i = 2, \dots, n-1$$

⋮

$$f_n = x_n - (3 + \alpha x_n)x_n - 1$$

* Les valeurs initiales utilisées par les méthodes de Newton, Broyden I et II et Brown sont :

$$x_0^i = -1.0 \quad \forall i = 1, \dots, n$$

pour $n = 5$ $N_0 = 1.91$

$n = 10$ $N_0 = 2.121$

$n = 10$ $N_0 = 2.646$

Les valeurs initiales utilisées par la méthode des sécantes, c'est-à-dire les valeurs de x_j^i $i = 1, \dots, n$ et $j = 0, \dots, n$ sont comprises entre -0.4 et -1.2.

Pour $n = 5$

$$\begin{array}{lll} N_0 = 2.49 & N_1 = 2.56 & N_2 = 2.01 \\ N_3 = 1.83 & N_4 = 1.74 & N_5 = 1.91 \end{array}$$

Pour $n = 10$

$$\begin{array}{lll} N_0 = 3.257 & N_1 = 2.922 & N_2 = 2.499 \\ N_3 = 2.88 & N_4 = 2.971 & N_5 = 2.13 \\ N_6 = 3.27 & N_7 = 3.209 & N_8 = 3.149 \\ N_9 = 3.576 & N_{10} = 2.121 & \end{array}$$

Pour $n = 20$

$$\begin{array}{llll} N_0 = 3.98 & N_1 = 3.47 & N_2 = 3.838 & N_3 = 4.108 \\ N_4 = 4.37 & N_5 = 4.56 & N_6 = 3.08 & N_7 = 5.3 \\ N_8 = 3.73 & N_9 = 4.66 & N_{10} = 2.819 & N_{11} = 5.11 \\ N_{12} = 5.01 & N_{13} = 4.7 & N_{14} = 3.522 & N_{15} = 3.53 \\ N_{16} = 5.6 & N_{17} = 3.9 & N_{18} = 3.01 & N_{19} = 5.09 \\ & & & N_{20} = 2.646 \end{array}$$

TABLE 3 - RESULTAT DE L'EXEMPLE 1

méthode	$n = 5$	$\alpha = -0, 1$		R
	m	N_f		
Newton	24	1.931	10^{-7}	0.6711
Broyden I	11	1.991	10^{-7}	1.5524
Broyden II	19	4.557	10^{-8}	0.9237
Brown	12	1.837	10^{-7}	1.3464
Sécantes	11	2.849	10^{-9}	1.8475

Table 4 - Résultats de l'exemple 2

	n = 10	$\alpha = -0,5$		
Méthode	m	N_f		R
Newton	44	2.356	10^{-6}	0.3344
Broyden I	18	1.597	10^{-7}	0.9112
Broyden II	34	2.602	10^{-7}	0.46805
Brown	26	9.064	10^{-8}	0.65262
Sécantes	20	7.870	10^{-9}	0.9786

Table 5 - Résultats de l'exemple 2

	n = 20	$\alpha = -0,5$		
Méthodes	m	N_f		R
Broyden I	29	3.611	10^{-7}	0.545
Broyden II	64	4.185	10^{-6}	0.24468
Brown	46	2.077	10^{-7}	0.3556
Sécantes	41	7.886	10^{-8}	0.4225
Newton	105	2.098	10^{-7}	0.1557

Exemple 3 [15] :

$$f_1(x_1, x_2) = -13 + x_1 + ((-x_2 + 5)x_2 - 2)x_2$$

$$f_2(x_1, x_2) = -29 + x_1 + ((x_2 + 1)x_2 - 14)x_2$$

valeurs initiales pour la méthode de Newton, Broyden I et II et Brent sont :

$$x_0^1 = 15. \quad x_0^2 = -2$$

Valeurs initiales pour la méthode des sécantes :

$$\left. \begin{array}{l} x_0^1 = 14. \\ x_0^2 = -25 \end{array} \right\} N_0 = 34.205$$

$$\left. \begin{array}{l} x_1^1 = 15 \\ x_2^1 = -3 \end{array} \right\} N_1 = 27.857$$

$$\left. \begin{array}{l} x_2^1 = 15 \\ x_2^2 = -2 \end{array} \right\} N_2 = 35.4478$$

Table 6 - Résultats de l'exemple 3

Méthode	m	N_f	R
Newton	87	2.188 10^{-6}	0.1908
Broyden I	diverge	-	-
Broyden II	diverge	-	-
Brown	25	1.192 10^{-7}	0.7804
Sécantes	13	2.232 10^{-6}	1.275

Exemple 4 :

Fonction singulière de Powell [12] :

$$\begin{aligned} f_1(x_1, x_2, x_3, x_4) &= x_1 + 10x_2 & f_2(x_1, x_2, x_3, x_4) &= \sqrt{5}(x_3 - x_4) \\ f_3(x_1, x_2, x_3, x_4) &= (x_2 - 2x_3)^2 & f_4(x_1, x_2, x_3, x_4) &= \sqrt{10}(x_1 - x_4)^2 \end{aligned}$$

Cette fonction admet $(0, 0, 0, 0)$ comme solution et elle a une propriété intéressante, c'est que le Jacobien est singulier en $(0, 0, 0, 0)$.

Le point de départ pour la méthode de Brown, Brown modifié, Brent, Powell est :

$$x_0 = (3, -1, 0, 1) \quad N_0 = 14.66287$$

Les points de départ de la méthode des sécantes sont :

$$x_0 = (3.148, -1.245, 0.157, 1.78) \quad N_0 = 12.03284$$

$$x_1 = (3, -1.8, 0.78, 1.1) \quad N_1 = 16.144408$$

$$x_2 = (3.25, -1.987, 0, 1.753) \quad N_2 = 18.904991$$

$$x_3 = (3.875, -1.896, 0.578, 1.986) \quad N_3 = 21.24999$$

$$x_4 = (3, -1, 0, 1) \quad N_4 = 14.66287$$

Table 7 - Résultat de l'exemple 4

Méthode	m	N_f	R
Powell	26	4.0 10^{-6}	0.52416
Brown	51	0.0	-
Brown modifié	49	1.2 10^{-6}	0.30269
Brent	45	1.9 10^{-6}	0.31939
Sécantes	11	2.27 10^{-5}	0.91448

Exemple 5

Système d'équations algébriques associé à un circuit électronique.

$$\left\{ \begin{array}{l} f_1 = vp_1 (1/R_5 + 1/R_7 + 1/R_9) + 1/R_7 \cdot x_3 + 1/R_9 \cdot x_4 + 1/R_7 \cdot x_6 ; \\ f_2 = -x_1 \left(\frac{1}{R_1} + \frac{1}{R_2} \right) + x_2 \cdot \frac{1}{R_2} - x_5 \cdot \frac{1}{R_1} + \frac{1}{R_1} E_1 ; \\ f_3 = x_1 \cdot \frac{1}{R_2} - x_2 \left(\frac{1}{R_2} + \frac{1}{R_3} \right) + \frac{1}{R_3} \cdot x_4 ; \\ f_4 = \frac{1}{R_7} \cdot vp_1 - x_3 \left(\frac{1}{R_4} + \frac{1}{R_7} + \frac{1}{R_8} \right) + x_4 \left(\frac{1}{R_4} + \frac{1}{R_8} \right) - x_6 \left(\frac{1}{R_7} + \frac{1}{R_8} \right) + J_4 \\ f_5 = vp_1 \cdot \frac{1}{R_9} + x_2 \cdot \frac{1}{R_3} + x_3 \left(\frac{1}{R_4} + \frac{1}{R_8} \right) - x_4 \left(\frac{1}{R_3} + \frac{1}{R_4} + \frac{1}{R_8} + \frac{1}{R_9} \right) + x_6 \cdot \frac{1}{R_8} + J_2 - J_3 \\ f_6 = -x_1 \cdot \frac{1}{R_1} - x_5 \cdot \frac{1}{R_1} + E_1 \cdot \frac{1}{R_1} - J_1 ; \\ f_7 = vp_1 \cdot \frac{1}{R_7} - x_3 \left(\frac{1}{R_7} + \frac{1}{R_8} \right) + x_4 \cdot \frac{1}{R_8} - x_6 \left(\frac{1}{R_6} + \frac{1}{R_7} + \frac{1}{R_8} \right) + J_3 + J_4 \end{array} \right.$$

$$\begin{array}{ll} J_1 = 10^{-7} (e^{(34-x_5)} - 1) & J_2 = 10^{-7} [(e^{(34 \cdot x_4)} - 1) - (e^{45(30-x_4)} - 1)] \\ J_3 = 1,28 * 10^{-10} (e^{(24,55 \cdot x_6)} - 1) & J_4 = 0.105 * J_3 \end{array}$$

$$\frac{1}{R_1} = 10^3 \quad \frac{1}{R_2} = 0.04 \quad \frac{1}{R_3} = 0.2 \quad \frac{1}{R_4} = 0.5 \cdot 10^{-3}$$

$$\frac{1}{R_5} = 0.025 \quad \frac{1}{R_6} = 0.67 * 10^{-5} \quad \frac{1}{R_9} = 10^{-15}$$

$$R_7 = \text{Max} \{60, P_2 * P_3 / (1 + (J_3 + 10^{-30}) / 1.97E-6) * 0.5\}$$

$$R_8 = \text{Max} \{1, (1 - P_2) * P_3\}$$

$$P_1 = x_4 - vp_1 - 10$$

$$P_2 = 0.695 - 87 * 10^{-5} * P_1$$

$$P_3 = 6.75 * 10^3 + 51 * P_1$$

On résoud ce système dans 3 cas : cas où $E1 = 0,1$, cas où $E1 = 0,01$ et cas où $E1 = 0,001$. Dans les 3 cas les 7 composants de 8 vecteurs de départ ont des valeurs variant entre 0 et 1.

1 - cas où $E1 = 0.1$

Racine de $F(x) = 0$:

$$\begin{array}{ll} x_1^* = -9.631038 \cdot 10^{-6} & x_5^* = -0.001512 \\ x_2^* = -0.001516 & x_6^* = 0.10151 \\ x_3^* = -0.001513 & x_7^* = 0.000436 \\ x_4^* = 0.001507 & \end{array}$$

$$\text{norme de } F(x^*) = \|F(x^*)\| = 2.06605 \cdot 10^{-6}$$

La convergence est obtenue après 4 itérations, le nombre d'évaluations de F est 12.

2 - cas où $E1 = 0.01$

Racine de $F(x) = 0$:

$$\begin{array}{ll} x_1^* = -2.2 \cdot 10^{-5} & x_5^* = 0.00314 \\ x_2^* = -0.00316 & x_6^* = 0.013 \\ x_3^* = -0.00314 & x_7^* = 0.0009 \\ x_4^* = -0.003132 & \end{array}$$

$$\|F(x^*)\| = 1.851408 \cdot 10^{-11}$$

converge après 5 itérations ; le nombre d'évaluations est 13.

3 - cas où $E1 = 0.001$

Racine de $F(x) = 0$:

$$\begin{array}{ll} x_1^* = -1.4311 \cdot 10^{-5} & x_5^* = -0.0022308 \\ x_2^* = -0.002253 & x_6^* = 0.003253 \end{array}$$

.73.

$$x_3^* = -0.002230$$

$$x_7^* = 0.0006450$$

$$x_4^* = 0.002239$$

$$\|F(x^*)\| = 2.86662 * 10^{-11}$$

converge après 5 itérations.

B I B L I O G R A P H I E

- [1] BRENT (1972)
The computational complexity of iterative methods for systems of non linear equations, proc. symposium on the complexity of computation.
- [2] BRENT (1973)
Some efficient algorithms for solving systems of non linear equations.
SIAM J. Numer. Anal. Vol 10 N° 2, April 1973, pp. 327-344
- [3] ORTEGA et RHEINBOLDT (1970)
Iterative solution of non linear equations in several variables.
Academic Press, New York.
- [4] OSTROWSKI (1960)
Solution of equations and systems of equations.
Academic Press, New York
- [5] STEWART and GRAGG (1974)
A stable variant of the secant method for solving non linear equations.
- [6] TORNHEIM (1964)
Convergence of multipoint iterative methods.
J. Assoc. Comput. Mach 11, pp. 210-220.

- [7] TRAUB (1964)
Iterative methods for the solution of equations.
Prentice - Hall, Englewood Cliffs.
- [8] VOIGT (1971)
Orders of convergence for iterative processes
S.I.A.M. J. Num. Anal., 8, pp. 222-243
- [9] WALL (1956)
The order of an iterative formular.
Math. Tables aids comput. 10, pp. 167-168.
- [10] WOLFE (1959)
The second method for simultaneous non linear equations.
Comm. A.C.M. 2, pp. 12-13
- [11] BROYDEN (1965)
A class of methods ofr solving non linear simultaneous equations.
Math. Comp. 19, pp. 577-593
- [12] COSNARD (1975)
A comparaisn of four methods for solving systems of non linear
equations.
T.R. 75-248. Department of Computer sc., Cornell University
I thaca, NY 14853.
- [13] BROWN (1966)
A quadratically convergent method for solving simultaneous
non linear equations.
Ph. D. thesis Purdue University. Department of Computer sc.
- [14] FELDSTEEN and FIRESTONE (1969)
A study of Oshowski efficiency for composite iteration algorithms.
Proc. 24th. National Conference of the A.C.M., pp. 145-155.

[15] FREUDENSTEIN and ROTH

Numerical solution of systems of non linear equations.
J. Assoc. Comput. Mach. 10, pp. 550-556.

[16] POWELL (1965)

A method for minimizing a sum of squares of non linear functions
without calculating derivatives computer journal vol. 7, pp. 303-
307.

[17] SHAMANSKII (1967)

A modification of Newton's method
Ukrain, Math. Zh. 19, pp. 133-138.

[18] WINOGRAD and WOLFE (1971)

Optimal iterative processes.
Rep. RC. 3511, IBM Research, Center, Yorktown, Heights, N.Y.

DEUXIEME PARTIE

LOCALISATION DES SOLUTIONS D'UN SYSTEME D'EQUATIONS ALGEBRIQUES

I N T R O D U C T I O N

Soit $F(x)$ une fonction définie et continue de \mathbb{R}^n dans \mathbb{R}^n et soit Ω une racine de F . On considère dans le premier chapitre que F est de la forme :

$$F(x) = Ax - b$$

où A est une matrice non singulière de $M_{n,n}(\mathbb{R})$ et $b \in \mathbb{R}^n$.

Dans le deuxième chapitre, F sera de la forme :

$$F(x) = \begin{pmatrix} x^T B_1 x \\ \vdots \\ x^T B_i x \\ \vdots \\ x^T B_n x \end{pmatrix} + Ax - b$$

où les B_i $i = 1, \dots, n \in M_{n,n}(\mathbb{R})$ sont symétriques semi-définies positives, A symétrique définie positive et $b \in \mathbb{R}^n$.

On s'intéresse à la localisation de Ω dans une boule de centre 0 et de rayon ρ , en donnant dans le premier cas des conditions nécessaires et suffisantes et dans le deuxième cas des conditions suffisantes pour que Ω soit à l'intérieur de $B(0, \rho)$.

Théorème d'existence des solutions d'un système dans une boule :

Soit $F : x \rightarrow F(x)$ une application définie et continue dans $B(0, \rho)$ et à valeurs dans \mathbb{R}^n , avec :

$$B(0, \rho) = \{x \in \mathbb{R}^n / \|x\| \leq \rho\}$$

$$\|x\| = \sqrt{\sum x_i^2}$$

Théorème 1 : ([1] p. 141 ex 10)

Si pour aucun point de ∂B (c'est-à-dire tel que $\|x\| = \rho$), $F(x) = \alpha x$ avec $\alpha > 0$, alors il existe dans $B(0, \rho)$ au moins un x^* tel que $F(x^*) = 0$ ($\in \mathbb{R}^n$).

Démonstration :

C'est une conséquence du théorème de Brouwer [1] :

"toute application continue d'un convexe fermé et borné (de \mathbb{R}^n) dans lui-même a un point fixe".

En effet, soit $x \in B(0, \rho)$ et $y = x + f(x)$

1) si $\|y\| \leq \rho$ ($y \in B$), on pose $\varphi(x) = x + F(x)$

2) Si $\|y\| \leq \rho$ on pose $\varphi(x) = \lambda(x + F(x))$ avec $\lambda > 0$

et tel que $\|\varphi(x)\| = \rho$ c'est-à-dire $\lambda = \frac{\rho}{\|x + F(x)\|}$

On considère l'application :

$$\varphi : x \rightarrow \varphi(x) = \rho \frac{x + F(x)}{\|x + F(x)\|} \psi \left(\frac{\|x + F(x)\|}{\rho} \right)$$

avec $\psi(u) = u$ où $0 \leq u \leq 1$ et $\psi(u) = 1$ si $u > 1$.

Il est évident que cette application est continue (le cas $\|x + F(x)\| = 0$ est dans la 1ère définition où $\varphi(x) = x + F(x)$) de $B(0, \rho)$ dans elle-même donc il existe un point fixe tel que $\varphi(x^*) = x^*$.

Si $\|x^* + F(x^*)\| \leq \rho$ alors $x^* = x^* + F(x^*) \implies F(x^*) = 0$

Si $\|x^* + F(x^*)\| > \rho$ alors $\frac{\rho}{\|x^* + F(x^*)\|} (x^* + F(x^*)) = x^*$

d'où :

1) $x^* \in \partial B$

2) $\frac{\rho}{\|x^* + F(x^*)\|} \cdot F(x^*) = \left(1 - \frac{\rho}{\|x^* + F(x^*)\|}\right) x^*$

c'est-à-dire $F(x^*) = \alpha x^*$ avec $\alpha > 0$, ce qui est impossible.

Donc cela prouve que le point fixe de φ est nécessairement tel que $F(x^*) = 0$.

CHAPITRE I

CAS D'UN SYSTEME LINEAIRE

A étant une matrice non singulière de $M_{n,n}(\mathbb{R})$ et $b \in \mathbb{R}^n$. Dans la suite, on désigne par Ω la solution du système $Ax - b = 0$.

I - CAS OU LA MATRICE A EST SYMETRIQUE DEFINIE POSITIVE

Dans la suite, nous considérons le produit scalaire :

$$\phi(x) = - \langle x, Ax - b \rangle$$

il est évident que : $|\phi(x)| \leq \|x\| \|Ax - b\| \quad \forall x \in \mathbb{R}^n$

Si $\|x\| = \rho$, on a donc :

$$\phi(x) = - \langle x, Ax - b \rangle \leq \|x\| \|Ax - b\| ; \forall x \in \partial B(0, \rho).$$

Théorème 2 :

Soit A une matrice symétrique définie positive et $\in M_{n,n}(\mathbb{R})$,
b un vecteur quelconque de \mathbb{R}^n . Si Ω est la solution de

$$Ax - b = 0$$

alors pour que $\rho > \|\Omega\|$, il faut et il suffit qu'il existe $\varepsilon_\rho > 0$, tel que :

$$- \langle x, Ax - b \rangle \leq (1 - \varepsilon_\rho) \|x\| \|Ax - b\| \\ \forall x \text{ tel que } \|x\| = \rho.$$

Démonstration :

1°) La condition est nécessaire :

Supposons que $\rho > \|\Omega\|$ et étudions (si $\|x\| = \rho$) la fonction :

$$x \xrightarrow{\in \partial B(0, \rho)} \frac{-\langle x, Ax - b \rangle}{\|x\| \|Ax - b\|} = \psi(x)$$

$\|Ax - b\| \neq 0$ pour $x \in \partial B(0, \rho)$ donc :

ψ continue sur le compact $B(0, \rho)$, donc est bornée et ψ atteint ses bornes ainsi $\exists \alpha, x_0$ tels que :

$$\alpha = \frac{-\langle x_0, Ax_0 - b \rangle}{\|x_0\| \|Ax_0 - b\|} = \max_{\|x\| = \rho} \frac{-\langle x, Ax - b \rangle}{\|x\| \|Ax - b\|} \leq 1$$

α est nécessairement < 1 , sinon

$$-\langle x_0, Ax_0 - b \rangle = \|x_0\| \|Ax_0 - b\| \quad (I)$$

On est dans le cas où, pour deux vecteurs non nuls, l'égalité (I) dans l'inégalité de Schwartz est atteinte, donc il existe une constante positive $\rho > 0$ telle que :

$$-px_0 = Ax_0 - b \quad (II)$$

A étant symétrique définie positive, on pose

$$A = Q^T D Q \quad \text{où } D \text{ est diagonale à éléments positifs.}$$

On voit alors que (II) implique :

$$-px_0 = Q^T D Q x_0 - b = Q^T D Q (x_0 - \Omega)$$

ou

$$Q^T D Q x_0 + px_0 = Q^T D Q \Omega$$

$$D Q x_0 + p Q x_0 = D Q \Omega \quad (III)$$

$$Q x_0 + p D^{-1} x_0 = Q \Omega \quad (IV)$$

Q étant unitaire, alors $\|Qx_0\| = \|x_0\|$
 et $\|Q\Omega\| = \|\Omega\|$

Si on pose $x'_0 = Qx_0$ et $\Omega' = Q\Omega$, on voit que (IV) s'écrit :

$$x'_0 + pD^{-1}x'_0 = \Omega' \quad (V)$$

en posant $x'_0 = (\xi'_i)$, $\Omega' = (\omega'_i)$ $i = 1, \dots, n$
 $\lambda_1, \dots, \lambda_n$ les valeurs propres de A (toutes positives), alors
 d'après (V) on a :

$$\omega'_i = \left(1 + \frac{p}{\lambda_i}\right) \xi'_i \implies |\xi'_i| < |\omega'_i| \quad \text{car } p > 0$$

donc $\rho = \|x'_0\| = \|x_0\| < \|\Omega'\| = \|\Omega\|$ ce qui est impossible.

Ainsi, si $\rho > \|\Omega\|$, il existe $\varepsilon_\rho > 0$ tel que $\alpha = 1 - \varepsilon_\rho$;
 c'est-à-dire $-\langle x, Ax - b \rangle \leq (1 - \varepsilon_\rho) \|x\| \|Ax - b\|$.

2°) La condition est suffisante :

Supposons qu'il existe $\varepsilon_\rho > 0$ tel que :

$$-\langle x, Ax - b \rangle \leq (1 - \varepsilon_\rho) \|x\| \|Ax - b\| \quad (VI)$$

pour $\|x\| = \rho$

alors $\rho > \|\Omega\|$; en effet :

La fonction $F(x) = -(Ax - b)$ est continue et définie dans tout \mathbb{R}^n . Supposons
 que sur la frontière de $B(0, \rho)$, il existe $\alpha > 0$ et x_0 tel que $\|x_0\| = \rho$
 et $\alpha x_0 = -(Ax - b) = F(x_0)$, puisque on a (VI) alors :

$$+\alpha \|x_0\|^2 \leq (1 - \varepsilon_\rho) \|x_0\|^2 \cdot \alpha$$

ce qui implique $1 \leq 1 - \varepsilon_\rho$ avec $\varepsilon_\rho > 0$, ce qui est absurde ; donc il
 existe (théorème 1) une solution de $-(Ax - b) = 0$ dans $B(0, \rho)$.

II - CAS OU A EST UNE MATRICE REGULIERE QUELCONQUE

Soit $Ax - b = 0$ (I),

un système linéaire où A est régulière quelconque.

On sait que (I) a même solution que le système :

$$A^T Ax - A^T b = 0$$

qui, lui, a une matrice $A^T A$ du 1er membre symétrique définie positive. Donc on a, en remarquant que :

$$- \langle x, A^T Ax - A^T b \rangle = - \langle Ax, Ax - b \rangle$$

Théorème 3 :

Si Ω est la solution du système régulier $Ax - b = 0$, pour que $\rho > \|\Omega\|$ il faut et il suffit qu'il existe $\varepsilon_\rho > 0$, tel que :

$$- \langle Ax, Ax - b \rangle \leq (1 - \varepsilon_\rho) \|x\| \|A^T Ax - A^T b\|$$

$$\text{si } \|x\| = \rho$$

III - FONCTION DE LOCALISATION DE LA SOLUTION D'UN SYSTEME LINEAIRE.

Soit $\rho > 0$ et $Ax - b = 0$ un système linéaire régulier. On pose pour $\rho \neq \|\Omega\|$:

$$\lambda(A, B)(\rho) = \max_{\|x\|=\rho} \frac{- \langle x, Ax - b \rangle}{\|x\| \|Ax - b\|}$$

pour A symétrique définie positive.

On posera aussi pour A régulière quelconque :

$$\Lambda(A, b) = \lambda(AA^T, A^T b)$$

Définition :

La fonction :

$$\rho \longrightarrow \lambda(A, b)(\rho) \quad (\text{respectivement } \Lambda(A, b)(\rho))$$

définie quelque soit $\rho > 0$, est dite fonction de localisation simple (respectivement totale).

Définition :

La fonction :

$$\rho \longrightarrow \lambda(A, b)(\rho) \text{ (respectivement } \Lambda(A, b)(\rho))$$

définie quelque soit $\rho > 0$, est dite fonction de localisation simple (respectivement totale).

(On peut poser $\lambda(A, b)(0) = 1$ (respectivement $\Lambda(A, b)(0) = 1$) et $\lambda(A, b)(\|\Omega\|) = 1$ pour étendre la définition à \mathbb{R}^+).

Etude de $\lambda(\rho) = \lambda(A, b)(\rho)$: (Cas où A est symétrique définie positive).

A étant symétrique définie positive, on pose dans la suite :

$$A = Q^T D Q$$

où Q est unitaire et D diagonale à éléments positifs.

$$y = Qx \implies \|y\| = \|x\|$$

$$\omega = Q\Omega \implies \|\omega\| = \|\Omega\|$$

$\lambda_1, \dots, \lambda_n$ les valeurs propres de A (toutes positives).

3.1 Interprétation géométrique de $\lambda(\rho)$:

A étant symétrique définie positive, alors on a l'interprétation géométrique suivante de la fonction de localisation $\lambda(\rho)$:

$$\begin{aligned} \lambda(\rho) &= \max_{\|x\| = \rho} - \frac{\langle x, Ax - b \rangle}{\|x\| \|Ax - b\|} \\ &= \max_{\|x\| = \rho} - \frac{\langle x, Ax - b \rangle}{\|x\| \|Ax - b\|} \\ &= \max_{\|x\| = \rho} (-\cos(x, Ax - b)) \end{aligned}$$

$$\text{Or } Ax - b = \text{grad}\left(\frac{1}{2} x^T Ax - x^T b\right)$$

$$\text{et } \frac{1}{2} x^T Ax - x^T b = \frac{1}{2} x^T Ax - x^T A\Omega = (x - \Omega)^T A(x - \Omega) - \Omega^T A\Omega$$

$$\frac{1}{2} x^T Ax - x^T b = \frac{1}{2} \|A^{1/2}(x - \Omega)\|^2 - \|A^{1/2}\Omega\|^2$$

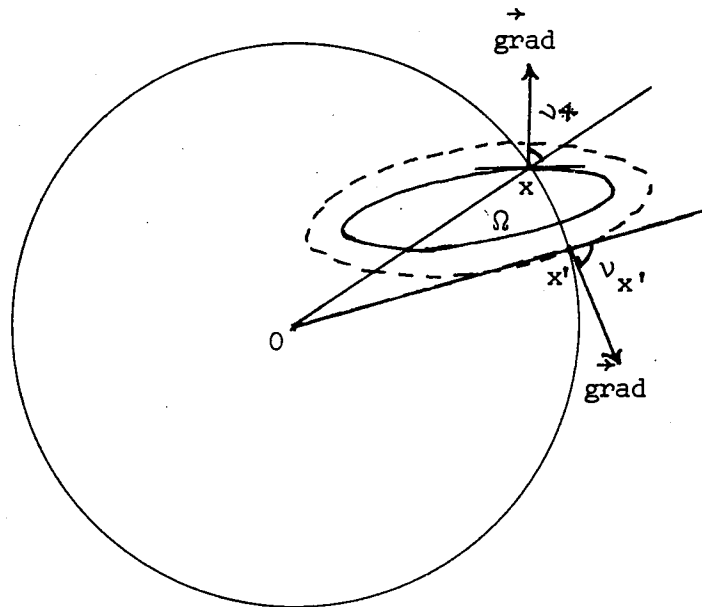
d'où

$$\frac{1}{2} x^T Ax - x^T b = \text{cte}$$

est une famille d'ellipsoïde homothétique et de centre Ω .
Soit (E) cette famille.

Ainsi :

$$\lambda(\rho) = \max_{\|x\|=\rho} [-\cos(x, \text{grad}(\frac{1}{2} x^T Ax - x^T b))]$$



Autrement dit, en chaque point $x \in \partial B(0, \rho)$, il passe un ellipsoïde $\epsilon \in (E)$.
Si ν_x est l'angle que fait le vecteur x avec le gradient en x de cet ellipsoïde, alors :

$$\lambda(\rho) = \max_{\|x\|=\rho} (-\cos \nu_x)$$

III.2 Propriété 1 :

Si ρ_1, ρ_2 sont tels que :

$$1) \rho_1 > \rho_2 > \|\Omega\|$$

$$2) \lambda(\rho_i) \geq 0 \quad i = 1, 2$$

alors :

$$\lambda(\rho_1) \leq \lambda(\rho_2)$$

où

$$\lambda(\rho_i) = \max_{\|x\| = \rho} \frac{-\langle x, Ax - b \rangle}{\|x\| \|Ax - b\|} \quad i = 1, 2$$

Démonstration :

A étant symétrique définie positive, posons :

$$A = Q^T D Q, \quad y = Qx, \quad \omega = Q\Omega$$

avec Q unitaire et D diagonale à éléments positifs.

Alors on a :

$$\lambda(\rho) = \max_{\|x\| = \rho} \frac{-\langle x, Ax - b \rangle}{\|x\| \|Ax - b\|} = \max_{\|y\| = \rho} \frac{-\langle y, D(y - \omega) \rangle}{\|y\| \|D(y - \omega)\|}$$

$$\begin{aligned} \text{Or } -\langle y, D(y - \omega) \rangle &= -\left(y - \frac{\omega}{2}\right)^T D \left(y - \frac{\omega}{2}\right) + \frac{1}{4} \omega^T D \omega \\ &= -\|D^{1/2} \left(y - \frac{\omega}{2}\right)\|^2 + \|D^{1/2} \frac{\omega}{2}\|^2 \end{aligned}$$

d'où :

$$\lambda(\rho) = \max_{\|y\| = \rho} \frac{-\|D^{1/2} \left(y - \frac{\omega}{2}\right)\|^2 + \|D^{1/2} \frac{\omega}{2}\|^2}{\|y\| \|D(y - \omega)\|}$$

Soit u le point de $\partial B(0, \rho_1)$ en lequel $\lambda(\rho_1)$ est atteinte, on a alors :

$$\lambda(\rho_1) = \frac{-\|D^{1/2} \left(u - \frac{\omega}{2}\right)\|^2 + \|D^{1/2} \frac{\omega}{2}\|^2}{\|u\| \|D(u - \omega)\|}$$

Nous allons démontrer qu'il existe un $z \in \partial B(0, \rho_2)$ tel que :

$$-\frac{\|D^{1/2}(u - \frac{\omega}{2})\|^2 + \|D^{1/2} \frac{\omega}{2}\|^2}{\|u\| \|D(u - \omega)\|} \leq \frac{-\|D^{1/2}(z - \frac{\omega}{2})\|^2 + \|D^{1/2} \frac{\omega}{2}\|^2}{\|z\| \|D(z - \omega)\|} \quad (1)$$

En effet :

les numérateurs de deux membres de (1) étant positifs (car $\lambda(\rho_1) \geq 0$), alors, pour satisfaire (1), il suffit que :

$$\begin{cases} 1^\circ) & -\|D^{1/2}(u - \frac{\omega}{2})\|^2 + \|D^{1/2} \frac{\omega}{2}\|^2 \leq -\|D^{1/2}(z - \frac{\omega}{2})\|^2 + \|D^{1/2} \frac{\omega}{2}\|^2 \\ 2^\circ) & \|u\| \|D(u - \omega)\| \geq \|z\| \|D(z - \omega)\| \end{cases}$$

ou

$$\begin{cases} 1^\circ) & \|D^{1/2}(u - \frac{\omega}{2})\|^2 \geq \|D^{1/2}(z - \frac{\omega}{2})\|^2 \\ 2^\circ) & \|D(u - \omega)\| \geq \|D(z - \omega)\| \quad (\text{car } \|u\| = \rho_1 > \rho_2 = \|z\|) \end{cases}$$

Il suffit donc :

$$\begin{cases} 1^\circ) & \left\{ \sum_1^n \lambda_i (u_i - \frac{\omega_i}{2})^2 \geq \sum_1^n \lambda_i (z_i - \frac{\omega_i}{2})^2 \right. \\ 2^\circ) & \left. \sum_1^n \lambda_i^2 (u_i - \omega_i)^2 \geq \sum_1^n \lambda_i^2 (z_i - \omega_i)^2 \right. \end{cases}$$

comme les $\lambda_i > 0$ ($i = 1, \dots, n$), alors il suffit de choisir un $z \in \partial B(0, \rho_2)$ tel que :

$$\begin{cases} 1^\circ) & |u_i - \frac{\omega_i}{2}| \geq |z_i - \frac{\omega_i}{2}| \\ 2^\circ) & |u_i - \omega_i| \geq |z_i - \omega_i| \end{cases} \quad (2)$$

On peut toujours construire un $z \in \partial B(0, \rho_2)$ et vérifiant (2) ; en effet :

Soit $I_r = \{i \in \{1, \dots, n\} / |\omega_i| \geq |v_i|\}$

Comme $\|\omega\| < \|u\| = \rho_1$, si r est le cardinal de I_r , alors on a nécessairement $r < n$.

Si on prend un z tel que :

$$1^\circ) \quad z_i = u_i \quad \forall i \in I_r$$

$$2^\circ) \quad z_i \in [\omega_i, u_i] \text{ si } i \notin I_r \text{ et } \omega_i y_i \leq 0$$

$$3^\circ) \quad z_i \in [-\omega_i, u_i] \text{ si } i \notin I_r \text{ et } \omega_i y_i \leq 0$$

et tel que :

$$\|z\|^2 = \sum z_i^2 = \rho_2^2 \quad (\text{qui est toujours possible car } \rho_2 < \rho_1).$$

On peut vérifier facilement que, dans le cas où $i \in I_r$ ou le cas $i \notin I_r$, z_i vérifie (2); donc (1) est vérifiée, c'est-à-dire :

$$\lambda(\rho_1) \leq \frac{-\langle z, D(z - \omega) \rangle}{\|z\| \|D(z - \omega)\|}$$

et comme $z \in \partial B(0, \rho_2)$, on a :

$$\lambda(\rho_2) = \max_{\|y\| = \rho_2} \frac{-\langle y, D(y - \omega) \rangle}{\|y\| \|D(y - \omega)\|} \geq \frac{-\langle z, D(z - \omega) \rangle}{\|z\| \|D(z - \omega)\|}$$

et par suite :

$$\lambda(\rho_1) \leq \lambda(\rho_2)$$

et $\lambda(\rho)$ est décroissante pour tout ρ tel que $\lambda(\rho) \geq 0$.

III.3 Propriété 2

Il existe un ρ_0 unique pour lequel $\lambda(\rho) = 0$.

Démonstration :

Supposons qu'il existe x_0 et x_1 tels que :

$$1) \quad \|x_0\| = \rho_0 \quad \|x_1\| = \rho_1$$

avec $\rho_1 > \rho_0$

$$2) \quad \max_{\|x\| = \rho_0} \frac{-\langle x, Ax - b \rangle}{\|x\| \|Ax - b\|} = \frac{-\langle x_0, Ax - b \rangle}{\|x_0\| \|Ax_0 - b\|} = 0$$

$$\max_{\|x\| = \rho_1} \frac{-\langle x, Ax - b \rangle}{\|x\| \|Ax - b\|} = \frac{-\langle x_1, Ax_1 - b \rangle}{\|x_1\| \|Ax_1 - b\|} = 0$$

donc $\langle x_1, Ax_1 - b \rangle = 0$ et $\langle x_0, Ax_0 - b \rangle = 0$

et $\forall z \in \partial B(0, \rho_0)$, on a :

$$\frac{-\langle z, Az - b \rangle}{\|z\| \|Az - b\|} \leq 0 \quad (1)$$

Posons $y = Qx_1 \Rightarrow \|y\| = \|x_1\| = \rho_1 > \rho_0$

$$\text{Or } -\langle x_1, Ax_1 - b \rangle = -\langle y, D(y - \omega) \rangle = -\|D^{1/2}(y - \frac{\omega}{2})\|^2 + \|D^{1/2} \frac{\omega}{2}\|^2 = 0$$

$$\text{ou } -\sum \lambda_i (y_i - \frac{\omega_i}{2})^2 + \sum \lambda_i (\frac{\omega_i}{2})^2 = 0$$

Si $\|y\| = \rho_1 > \rho_0$, on peut donc construire un point $z_0 \in \partial B(0, \rho_0)$, tel que :

$$-\sum \lambda_i (z_i^0 - \frac{\omega_i}{2})^2 + \sum \lambda_i (\frac{\omega_i}{2})^2 > 0 \quad (\text{les } \lambda_i > 0)$$

c'est-à-dire tel que :

$$\left| z_i^0 - \frac{\omega_i}{2} \right| < \left| y_i - \frac{\omega_i}{2} \right|$$

qui est toujours possible car $\rho_0 < \rho_1$ (la construction de z_0 ressemble à la construction de z dans la propriété 1).

Donc :

$$-\frac{\langle z_0, D(z_0 - \omega) \rangle}{\|z_0\| \|D(z_0 - \omega)\|} > 0$$

ce qui est en contradiction avec (1).

Donc, on a nécessairement $\rho_0 = \rho_1$ et par suite la racine ρ_0 de $\lambda(\rho)$ est unique.

III.4 Propriété 3 : Limite à l'infini

D'après les propriétés 1) et 2), on a $\forall \rho > \rho_0$ ($\lambda(\rho_0) = 0$)

$$\max_{\|x\| = \rho} \frac{-\langle x, Ax - b \rangle}{\|x\| \|Ax - b\|} < 0$$

donc :

$$\frac{\langle x, Ax - b \rangle}{\|x\| \|Ax - b\|} = \frac{\langle y, D(y - \omega) \rangle}{\|x\| \|D(y - \omega)\|} > 0$$

avec $y = Qx$ et $\omega = Q\Omega$.

Ceci pour tout x tel que $\|x\| = \rho > \rho_0$.

Posons :

$$z(y) = \frac{y^T D(y - \omega)}{\|y\| \|D(y - \omega)\|} = \frac{\frac{y^T Dy}{\|y\| \|Dy\|} - \frac{y^T d}{\|y\| \|Dy\|}}{\frac{D(y - \omega)}{\|Dy\|}}$$

qui est > 0 pour un y tel que $\|y\|$ est assez grand ($\|y\| > \rho_0$)

On sait que $\forall y$ tel que $\|y\| = \rho$, on a :

$$1^\circ) \quad \|Dy\| - \|D\| \leq \|D(y - \omega)\| \leq \|Dy\| + \|D\omega\|$$

ou

$$1 - \frac{\|D\omega\|}{\|Dy\|} \leq D(y - \omega) \leq 1 + \frac{\|D\omega\|}{\|Dy\|} \quad (1)$$

2°) Il existe y_1 et $y_0 \in B(0, \rho)$ tels que :

$$Z(y_1) = \min_{\|y\| = \rho} Z(y)$$

d'où

$$Z(y_1) \leq Z(y_0) \quad (2)$$

et

$$\frac{y_0^T D y_0}{\|y_0\| \|Dy_0\|} = \min_{\|y\| = \rho} \frac{y^T D y}{\|y\| \|Dy\|}$$

d'où

$$\frac{y_0^T D y_0}{\|y_0\| \|Dy_0\|} \leq \frac{y_1^T D y_1}{\|y_1\| \|Dy_1\|} \quad (3)$$

Ainsi :

$$\frac{\frac{y_0^T D y_0}{\|y_0\| \|Dy_0\|} - \frac{y_1^T D \omega}{\|y_1\| \|Dy_1\|}}{\frac{D(y_1 - \omega)}{\|Dy_1\|}} \leq \frac{\frac{y_1^T D y_1}{\|y_1\| \|Dy_1\|} - \frac{y_1^T D \omega}{\|y_1\| \|Dy_1\|}}{\frac{D(y_1 - \omega)}{\|Dy_1\|}}$$

$$= Z(y_1) \leq Z(y_0)$$

$$Z(y_0) = \frac{\frac{y_0^T D y_0}{\|y_0\| \|Dy_0\|} - \frac{y_0^T D \omega}{\|y_0\| \|Dy_0\|}}{\frac{\|D(y_0 - \omega)\|}{\|Dy_0\|}}$$

(Ceci d'après (2) et (3)).

Et d'après (1), on a :

$$\frac{\frac{y_0^T D y_0}{\|y_0\| \|Dy_0\|} - \frac{y_1^T D \omega}{\|y_1\| \|Dy_1\|}}{1 + \frac{\|D\omega\|}{\|Dy_1\|}} \leq Z(y_1) \leq \frac{\frac{y_0^T D y_0}{\|y_0\| \|Dy_0\|} - \frac{y_0^T D \omega}{\|y_0\| \|Dy_0\|}}{1 - \frac{\|D\omega\|}{\|Dy_0\|}} \quad (4)$$

Or $\forall \varepsilon > 0$ assez petit, il existe un k assez grand tel que pour tout y_1, y_0 ; $\|y_1\| = \|y_0\| > k$.

On a :

$$a) \frac{y_1^T D \omega}{\|y_1\| \|Dy_1\|} \leq \varepsilon \quad \text{et} \quad \frac{y_0^T D \omega}{\|y_0\| \|Dy_0\|} \leq \varepsilon$$

$$b) \frac{\|D\omega\|}{\|Dy_1\|} \leq \varepsilon \quad \text{et} \quad \frac{\|D\omega\|}{\|Dy_0\|} \leq \varepsilon$$

d'où (4) devient pour $\|y_1\| = \|y_0\| > k$

$$\frac{\frac{y_0^T D y_0}{\|y_0\| \|Dy_0\|}}{1 + \varepsilon} \leq Z(y_1) \leq \frac{\frac{y_0^T D y_0}{\|y_0\| \|Dy_0\|} + \varepsilon}{1 - \varepsilon}$$

et lorsque $\rho \longrightarrow \infty$, $\varepsilon \longrightarrow 0$;

$$\text{et limite } Z(y_1) = \limite \frac{y_0^T D y_0}{\|y_0\| \|Dy_0\|}$$

mais

$$\forall \rho > 0 \quad \min_{\|y\|=\rho} \frac{y^T D y}{\|y\| \|Dy\|} = \frac{y_0^T D y_0}{\|y_0\| \|Dy_0\|} = \text{cte}$$

et si on pose $D^{1/2}y = x$, on aura :

$$\min \frac{y^T D y}{\|y\| \|Dy\|} = \frac{x^T x}{\|D^{-1/2}x\| \|D^{1/2}x\|} = \min \frac{\|x\|^2}{\|D^{-1/2}x\| \|D^{1/2}x\|} = \frac{2\sqrt{\lambda_M \lambda_m}}{\lambda_m + \lambda_M}$$

C'est la constante de Kantorovich [2].

λ_M et λ_m sont respectivement la plus grande et la plus petite valeur propre de D.

Donc :

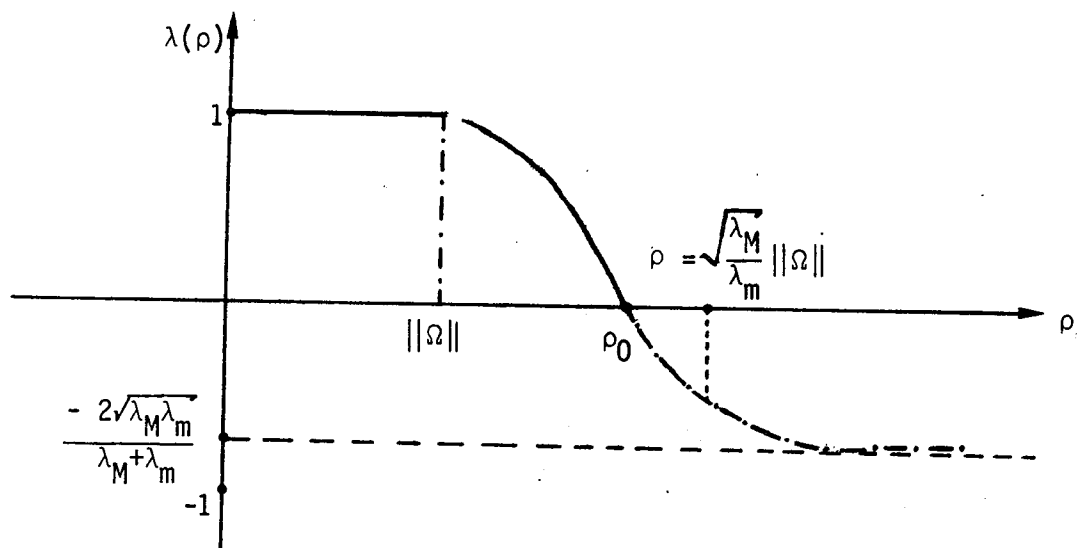
$$\limite_{\rho \rightarrow \infty} Z(y_1) = \lim_{\rho \rightarrow \infty} \min_{\|y\|=\rho} \frac{\langle y, D(y - \omega) \rangle}{\|y\| \|D(y - \omega)\|} = \frac{2\sqrt{\lambda_M \lambda_m}}{\lambda_m + \lambda_M}$$

et ainsi :

$$\lim_{\rho \rightarrow \infty} \max_{\|y\|=\rho} - \frac{\langle y, D(y - \omega) \rangle}{\|y\| \|D(y - \omega)\|} = - \frac{2\sqrt{\lambda_M \lambda_m}}{\lambda_m + \lambda_M}$$

$\lambda(\rho)$ étant continue, alors à partir des propriétés 1), 2), et 3) et le théorème 2, on va essayer de tracer graphiquement cette courbe, en faisant la remarque qu'il semblerait que $\lambda(\rho) \rightarrow 1$ quand $\rho \rightarrow \|\Omega\|$.

Graphe de $\lambda(\rho)$:



Théorème 4 :

Une condition suffisante pour que $\lambda(\rho)$ soit négatif (c'est-à-dire il existe une racine à l'intérieur de $B(0, \rho)$) est que :

$$\rho \geq \sqrt{\frac{\lambda_M}{\lambda_m}} \|\Omega\|$$

où λ_M et λ_m sont la plus grande et la plus petite valeur propre de A .

Démonstration :

Soit $Ax - b = 0$ un système linéaire, A symétrique définie positive.

Posons $A = Q^T D Q$, $y = Qx$ et $\omega = Q\Omega$.

$$\begin{aligned} - \langle x, Ax - \Omega \rangle &= - \langle x, Q^T D Q (x - \Omega) \rangle = - \langle Qx, DQx - D\omega \rangle \\ &= - \langle y, D(y - \omega) \rangle = - y^T D y + y^T D \omega \end{aligned}$$

$$\begin{aligned}
&= - \left[y - \frac{\omega}{2} \right]^T D \left[y - \frac{\omega}{2} \right] + \frac{1}{4} \omega^T D \omega \\
&= - \left\| D^{1/2} \left(y - \frac{\omega}{2} \right) \right\|^2 + \left\| D^{1/2} \frac{\omega}{2} \right\|^2
\end{aligned}$$

Dès que $\left\| D^{1/2} \left(y - \frac{\omega}{2} \right) \right\|^2 \geq \left\| D^{1/2} \frac{\omega}{2} \right\|^2$ (1)

le produit scalaire reste toujours négatif ou nul.

Or $\left\| D^{1/2} X \right\| = \frac{\|X\|}{\|D^{-1/2}\|}$; donc (1) est nécessairement réalisé si :

$$\left\| y - \frac{\omega}{2} \right\| \geq \|D^{-1/2}\| \left\| D^{1/2} \frac{\omega}{2} \right\|$$

Si de plus on remarque que :

$$\left\| D^{1/2} y - D^{1/2} \frac{\omega}{2} \right\| \geq \left\| D^{1/2} y \right\| - \left\| D^{1/2} \frac{\omega}{2} \right\|$$

On voit que l'on réalise (1), si :

$$\left\| D^{1/2} y \right\| \geq \left\| D^{1/2} \frac{\omega}{2} \right\|$$

Il suffit de prendre

$$\|y\| \geq \|D^{-1/2}\| \left\| D^{1/2} \frac{\omega}{2} \right\|$$

c'est-à-dire

$$\rho \geq \|D^{-1/2}\| \left\| Q^T D^{1/2} Q \right\|$$

mais comme

$$\left\| D^{1/2} \frac{\omega}{2} \right\| \leq \left\| D^{1/2} \right\| \left\| \frac{\omega}{2} \right\|$$

il suffit alors de prendre :

$$\|y\| \geq \|D^{-1/2}\| \left\| D^{1/2} \right\| \left\| \frac{\omega}{2} \right\|$$

Or

$$\begin{aligned}\|D^{1/2}\| &= \sqrt{\lambda_M} \\ \|D^{-1/2}\| &= \frac{1}{\sqrt{\lambda_m}} \\ \|\omega\| &= \|\Omega\|\end{aligned}$$

Donc, pour avoir $\lambda(\rho) \leq 0$, il suffit de prendre :

$$\rho \geq \sqrt{\frac{\lambda_M}{\lambda_m}} \|\Omega\|$$

III.5 Recherche de ρ_0 tel que $\lambda(\rho_0) = 0$

si

$$\lambda(\rho_0) = \max_{\|x\| = \rho_0} \frac{-\langle x, Ax - b \rangle}{\|x\| \|Ax - b\|} = 0$$

alors on a

$$-\langle x_0, Ax_0 - b \rangle = 0$$

où x_0 est le point en lequel $\lambda(\rho_0)$ est atteinte.

Ainsi la recherche de ρ_0 pour lequel $\lambda(\rho_0) = 0$ revient à trouver :

$$(P) \begin{cases} \max \|x\| \\ \text{si } \langle x, Ax - b \rangle = 0 \end{cases}$$

(on a déjà trouvé que ρ_0 est unique - propriété 2).

Posons $A = Q^T D Q$, $y = Qx$, Q

avec Q unitaire et D diagonale dont les éléments ne sont en outre que les valeurs propres de A .

D'où,

$$\langle x, Ax - b \rangle = y^T D y - y^T D \omega = 0 \quad (1)$$

et le problème (P) devient

$$(P') \quad \max \phi(y) = y^T y + \gamma(y^T D y - y^T D \omega)$$

Pour que le problème (P') ait une solution, il faut que ses dérivées partielles soient nulles.

C'est-à-dire :

$$\frac{\partial \phi}{\partial y_i} = 2y_i + 2\gamma\lambda_i y_i - \gamma\omega_i = 0$$

d'où

$$y_i = \frac{\gamma\lambda_i \omega_i}{2(1+\gamma\lambda_i)} \quad \forall i = 1, \dots, n$$

En remplaçant dans (1), on obtient :

$$\sum_1^n \lambda_i \frac{\gamma^2 \lambda_i^2 \omega_i^2}{4(1+\gamma\lambda_i)^2} - \sum_1^n \lambda_i \frac{\gamma}{2} \frac{\lambda_i \omega_i^2}{(1+\gamma\lambda_i)} = 0$$

$$\sum_1^n \left(\frac{\gamma}{2} \frac{\lambda_i^3 \omega_i^2}{(1+\gamma\lambda_i)^2} - \frac{\lambda_i^2 \omega_i^2}{1+\gamma\lambda_i} \right) = 0$$

$$\sum_1^n \frac{\gamma\lambda_i^3 \omega_i^2 - 2\lambda_i^2 \omega_i^2 - 2\gamma\lambda_i^3 \omega_i^2}{2(1+\gamma\lambda_i)^2} = 0$$

$$\sum_1^n \frac{\gamma\lambda_i^3 \omega_i^2 + 2\gamma\lambda_i^2 \omega_i^2}{2(1+\gamma\lambda_i)^2} = 0$$

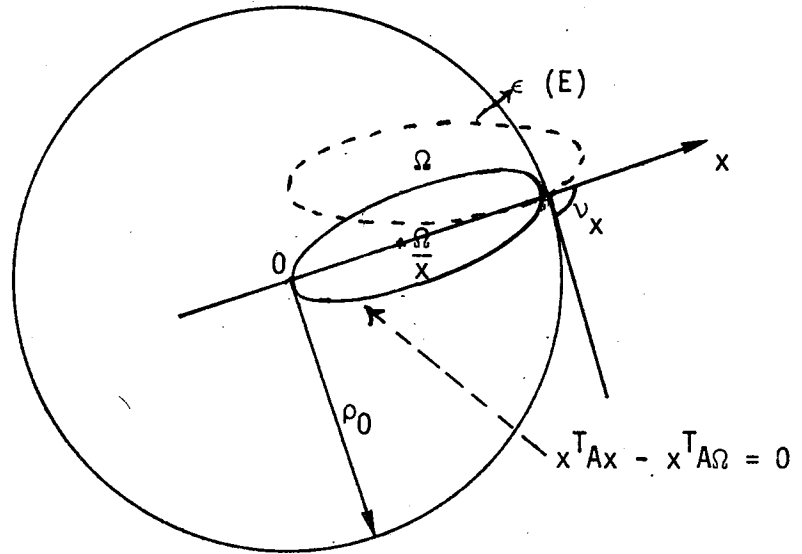
d'où

$$\sum_1^n \frac{\gamma\lambda_i^2 \omega_i^2}{(1+\gamma\lambda_i)^2} = -\frac{\gamma}{2} \sum_1^n \frac{\gamma\lambda_i^3 \omega_i^2}{(1+\gamma\lambda_i)^2}$$

et γ sera la plus grande racine de cette équation.

Remarques :

- 1 - Géométriquement le lieu des points x tels que $x^T A x - x^T A \omega = 0$ est l'ellipsoïde de centre $\frac{\Omega}{2}$ et passant par 0 et Ω , d'où la recherche de la solution du problème (P) revient à trouver le rayon ρ_0 de la plus grande boule tangente à l'ellipsoïde $x^T A x - x^T A \omega = 0$



- 2 - Le théorème 4 donne un ρ_1 pour lequel $\lambda(\rho_1) \leq 0$ et par suite on a $\Omega \in B(0, \rho_1)$. Il est clair, d'après la propriété 2, que $B(0, \rho_0) \subset B(0, \rho_1)$ mais la recherche de ρ_0 ou d'un ρ tel que $\lambda(\rho) > 0$ n'est pas très simple. Pour cela nous donnons dans la suite une application pour la localisation des systèmes linéaires particuliers, en utilisant le théorème 4.

IV - APPLICATIONS

IV - Application à une localisation de la solution d'un système "somme"

Soient

$$\left. \begin{array}{l} Ax - b = 0 \\ A'x - b' = 0 \end{array} \right\} \quad (I)$$

deux systèmes linéaires avec A et A' réguliers.

$$(A + A')x - (b + b') = 0 \quad (II)$$

est le système somme de deux systèmes (I).

Supposons que A + A' est régulière et posons Ω , Ω' et Σ les solutions des systèmes (1) et (2).

Théorème 5 :

S'il existe un $\rho_1 > 0$ tel que $\lambda(\rho_1) \leq 0$ et $\rho_2 > 0$ tel que $\lambda'(\rho_2) \leq 0$, on a alors :

$$\|\Sigma\| \leq \max(\rho_1, \rho_2)$$

Démonstration :

On a par définition

$$\begin{array}{ll} - \langle x, Ax - b \rangle \leq \lambda(A, b)(\rho) \|Ax - b\| \rho & \text{si } \|x\| = \rho \\ - \langle x, A'x - b' \rangle \leq \lambda'(A', b')(\rho) \|A'x - b'\| \rho & \text{si } \|x\| = \rho \end{array}$$

Donc en ajoutant

$$\begin{array}{l} - \langle x, (A + A')x - (b + b') \rangle \leq \rho [\lambda(\rho) \|Ax - b\| + \lambda'(\rho) \|A'x - b'\|] \quad (1) \\ \text{si } \|x\| = \rho \end{array}$$

Or si $\rho \geq \rho_1 \implies \lambda(\rho) \leq 0$

De même

si $\rho \geq \rho_2 \implies \lambda'(\rho) \leq 0$

Si $\rho_0 = \max(\rho_1, \rho_2)$, alors pour tout $\rho \geq \rho_0$ on a :

$$\lambda(\rho) \leq 0 \quad \text{et} \quad \lambda'(\rho) \leq 0$$

d'où, dans ces conditions, le deuxième membre de (1) est négatif et on a :

$$- \langle x, A'x - b' \rangle \leq 0 \quad \forall x \text{ tel que } \|x\| = \rho \geq \rho_0$$

donc

$$\|\Sigma\| \leq \rho_0 = \max(\rho_1, \rho_2).$$

Corrolaire 1 :

Si A et A' , les matrices de deux systèmes (I), sont symétriques définies positives, alors le système "somme" (II) admet une solution et on a :

$$\|\Sigma\| \leq \max\left(\sqrt{\frac{\lambda_M}{\lambda_m}} \|\Omega\|, \sqrt{\frac{\lambda'_M}{\lambda'_m}} \|\Omega'\|\right)$$

avec λ_M et λ_m (resp. λ'_M et λ'_m) sont respectivement la plus grande et la plus petite valeur propre de A (resp. de A')

Démonstration :

(II) admet une solution car $A+A'$ est aussi symétrique définie positive.

Le reste est une conséquence du théorème précédent. En effet :

$$\lambda\left(\sqrt{\frac{\lambda_M}{\lambda_m}} \|\Omega\|\right) \leq 0 \quad \text{et} \quad \lambda'\left(\sqrt{\frac{\lambda'_M}{\lambda'_m}} \|\Omega'\|\right) \leq 0$$

d'où

$$\|\Sigma\| \leq \max\left(\sqrt{\frac{\lambda_M}{\lambda_m}} \|\Omega\|, \sqrt{\frac{\lambda'_M}{\lambda'_m}} \|\Omega'\|\right)$$

$$B_k = \begin{pmatrix} \alpha_1 & & & & \\ & \ddots & & & \\ & & \beta_k & & \\ & & & \ddots & \\ & & & & \beta_n \\ & & & & & \alpha_n \end{pmatrix}$$

telle que :

$$1^\circ) \alpha_i \geq 0 \quad \forall i = 1, \dots, n$$

$$2^\circ) \alpha_i \geq |\beta_i| \quad \forall i = k, \dots, n$$

$$3^\circ) \alpha_i \geq |\beta_{k+i-1}| \quad \forall i = 1, \dots, n-k+1$$

$$\text{et si } \alpha_M^k = \max_i \alpha_i$$

alors $\forall \omega \in \mathbb{R}^n$

$$- \langle x, B_k(x - \omega) \rangle \leq \frac{1}{2} \alpha_M^k \|\omega\|^2$$

Démonstration :

La matrice B_k s'écrit :

$$B_k = M_k + C$$

avec

$$M_k = \begin{pmatrix} & & & & & \\ & & & & & \\ & & |\beta_k| & & & \\ & & & \ddots & & \\ \beta_k & & & & |\beta_n| & \\ & & \alpha_{n-k+2} & & & \\ & & & \ddots & & \\ & & & & \beta_n & \\ & & & & & \alpha_n \end{pmatrix} \quad \text{et} \quad C = \begin{pmatrix} \alpha_1 - |\beta_k| & & & & \\ & \alpha_{n-k+1} - |\beta_n| & & & \\ & & 0 & & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix}$$

et d'après le lemme 1, on a :

$$B_k = Q_k^t D_k + C \quad (1)$$

et en posant

$$y = Qx$$

on obtient

$$- \langle x, B_k(x - \omega) \rangle \leq -\frac{1}{2} y^T D_k y + y^T D_k \omega - \frac{1}{2} x^T c x + x^T C \omega$$

Or D_k et c sont des matrices diagonales à éléments positifs, d'où :

$$\begin{aligned} - \langle x, B_k(x - \omega) \rangle &\leq -\frac{1}{2} \|D_k^{1/2}(y - \omega)\|^2 + \frac{1}{2} \|D_k^{1/2}\omega\|^2 - \frac{1}{2} \|c^{1/2}(x - \omega)\|^2 + \frac{1}{2} \|c^{1/2}\omega\|^2 \\ &\leq \frac{1}{2} \|D_k^{1/2}\omega\|^2 + \frac{1}{2} \|c^{1/2}\omega\|^2 = \frac{1}{2} \|D^{1/2}\omega\|^2 \end{aligned}$$

avec

$$D = D_k + C = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}$$

d'où

$$- \langle x, B_k(x - \omega) \rangle \leq \frac{1}{2} \alpha_M^K \|\omega\|^2$$

Proposition 2

Si B est une matrice de $M_{n,n}(\mathbb{R})$ telle que

$$B = \sum_{k \in I} B_k \quad \text{où} \quad I \subset \{1, 2, \dots, n\}$$

et si $\forall k \in I, B_k$ vérifie les conditions de la proposition 1,

alors pour tout $\omega \in \mathbb{R}^n$ on a

$$- \langle x, B(x - \omega) \rangle \leq \frac{1}{2} \left(\sum_{k \in I} \alpha_M^k \right) \|\omega\|^2$$

Démonstration :

$$\begin{aligned}
 - \langle x, B(x - \omega) \rangle &= -x^T B x + x^T B \omega \\
 &= -x^T \left(\sum_{k \in I} B_k \right) (x - \omega) \\
 &= \sum_{k \in I} -x^T B_k (x - \omega) \\
 &= \sum_{k \in I} -\langle x, B_k (x - \omega) \rangle
 \end{aligned}$$

et d'après la proposition 1, on a $\forall k \in I$

$$-\langle x, B_k (x - \omega) \rangle \leq \frac{1}{2} \alpha_M^k \|\omega\|^2$$

d'où

$$-\langle x, B(x - \omega) \rangle \leq \frac{1}{2} \left(\sum_{k \in I} \alpha_M^k \right) \|\omega\|^2$$

Proposition 3 :

Si A est une matrice de $M_{n,n}(\mathbb{R})$ triangulaire inférieure telle que

$$A = \sum_{k \in I} B_k + D$$

avec

1°) $I \subset \{1, \dots, n\}$

2°) B_k vérifie les conditions de la proposition 1 ($\forall k \in I$)

3°) La matrice D est diagonale

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

avec $\lambda_i > 0 \quad \forall i = 1, \dots, n$

Et, si b est un vecteur quelconque de \mathbb{R}^n et ω la solution du système

$$Ax - b = 0$$

$$\text{alors pour tout } \rho \geq \frac{1}{2} \frac{\sqrt{\lambda_M} + \sqrt{\lambda_m + 2\alpha_M}}{\sqrt{\lambda_m}} \|\omega\|$$

où λ_M et λ_m sont respectivement la plus grande et la plus petite valeur propre de D et $\alpha_M = \sum_{k \in I} \alpha_M^k$

On a :

$$- \langle x, A(x - \omega) \rangle \leq 0$$

$$\forall x / \|x\| = \rho$$

c'est-à-dire $\lambda(\rho) \leq 0$.

Démonstration :

La matrice A est triangulaire à éléments diagonaux positifs, donc $Ax - b = 0$ admet toujours une solution $\omega \in \mathbb{R}^n$

$$- \langle x, A(x - \omega) \rangle = -x^T (\sum B_k)(x - \omega) - x^T D(x - \omega)$$

D'après la proposition 2, on a :

$$\begin{aligned} - \langle x, A(x - \omega) \rangle &\leq \frac{1}{2} \sum_{k \in I} \alpha_M^k \|\Omega\|^2 - x^T D(x - \omega) \\ &\leq \frac{1}{2} \alpha_M \|\omega\|^2 - \|D^{1/2}(x - \frac{\omega}{2})\|^2 + \|D^{1/2} \frac{\omega}{2}\|^2 \end{aligned}$$

car D est diagonale à éléments > 0 .

$$- \langle x, A(x - \omega) \rangle \leq - \|D^{1/2}(x - \frac{\omega}{2})\|^2 + \|H^{1/2} \frac{\omega}{2}\|^2 \quad (1)$$

avec

$$H = \begin{pmatrix} \lambda_1 + 2\alpha_M & & & \\ & \circ & & \\ & \vdots & & \\ \circ & & \ddots & \\ & & & \lambda_n + 2\alpha_M \end{pmatrix}$$

Alors, pour que $- \langle x, A(x - \omega) \rangle \leq 0$, il suffit d'avoir :

$$\|D^{1/2}(x - \frac{\omega}{2})\| \geq \|H^{1/2} \frac{\omega}{2}\|$$

$$\text{Or } \|D^{1/2}(x - \frac{\omega}{2})\| \geq \|D^{1/2}x\| - \|D^{1/2}\frac{\omega}{2}\| \geq \frac{\|x\|}{\|D^{-1/2}\|} - \|D^{1/2}\|\|\omega\|$$

$$\text{et } \|H^{1/2}\frac{\omega}{2}\| \leq \|H^{1/2}\|\|\frac{\omega}{2}\| = (\sqrt{\lambda_M} + 2\alpha_M)\|\frac{\omega}{2}\|$$

Ainsi, il suffit que :

$$\frac{\|x\|}{\|D^{-1/2}\|} \geq \|D^{1/2}\|\|\frac{\omega}{2}\| + (\sqrt{\lambda_M} + 2\alpha_M)\|\frac{\omega}{2}\|$$

ou bien

$$\|x\| \geq \|D^{-1/2}\|(\|D^{1/2}\| + \sqrt{\lambda_M} + 2\alpha_M)\|\frac{\omega}{2}\|$$

d'où si

$$\rho = \|x\| \geq \frac{1}{2} \frac{\sqrt{\lambda_M} + \sqrt{\lambda_M} + 2\alpha_M}{\sqrt{\lambda_m}} \|\omega\|$$

on a

$$- \langle x, A(x - \omega) \rangle \leq 0 \quad \forall x \text{ tel que } \|x\| = \rho$$

ou

$$\lambda(\rho) \leq 0$$

Remarque :

Si les B_k sont triangulaires supérieurs, on a les mêmes conditions pour avoir $\lambda(\rho) \leq 0$; d'où le théorème suivant.

Théorème 6 :

Si A est une matrice de $M_{n,n}(\mathbb{R})$ de la forme

$$A = \sum_{k \in I} B_k + D_1 + \sum_{k \in I'} B'_k + D_2$$

et telle que :

- a) $\forall k \in I$ ou $\in I'$, on a B_k et ${}^t B_k'$ triangulaires inférieures et vérifiant les conditions de la proposition 1,
- b) D_1 et D_2 sont diagonales à éléments strictement positifs.
Alors on a :

1) A est inversible

2) $\forall b \in \mathbb{R}^n$ le système $Ax - b = 0$ est le système "somme" de deux systèmes

$$\left. \begin{array}{l} A_1 x - b_1 = 0 \\ A_2 x - b_2 = 0 \end{array} \right\} (I)$$

avec $A_1 = \sum_{k \in I} B_k + D_1$, $A_2 = \sum_{k \in I'} B_k' + D_2$ et $b = b_1 + b_2$.

3) Si ω_1 et ω_2 sont les solutions de deux systèmes (I) et Σ est la solution du système "somme" $Ax - b = 0$, alors on a :

$$\|\Sigma\| \leq \frac{1}{2} \max \left(\frac{\sqrt{\lambda_M^1} + \sqrt{\lambda_m^1 + 2\alpha_M'}}{\sqrt{\lambda_m^1}} \|\omega_1\|, \frac{\sqrt{\lambda_M^2} + \sqrt{\lambda_m^2 + 2\alpha_M'}}{\sqrt{\lambda_m^2}} \|\omega_2\| \right)$$

avec λ_M^1, λ_m^1 sont la plus grande et la petite valeur propre de D_1 , λ_M^2, λ_m^2 sont celles de D_2 , $\alpha_k = \sum_{k \in I} \alpha_M^k$
 $\alpha_k' = \sum_{k \in I'} \alpha_M^k$.

Démonstration :

Si A vérifie a) et b), alors on peut vérifier facilement que A est une matrice à diagonale strictement dominante, d'où A est inversible.

Si $\lambda_1(\rho)$ est la fonction de localisation du système $A_1 x - b_1 = 0$ et $\lambda_2(\rho)$ est celle de $A_2 x - b_2 = 0$.

D'après la proposition précédente, on a :

$$\lambda_1(\rho) \leq 0 \quad \text{pour tout } \rho \geq \rho_1 = \frac{1}{2} \frac{\sqrt{\lambda_M^1} + \sqrt{\lambda_M^1 + 2\alpha_M}}{\sqrt{\lambda_m^1}} \quad \|\omega_1\|$$

$$\lambda_2(\rho) \leq 0 \quad \text{pour tout } \rho \geq \rho_2 = \frac{1}{2} \frac{\sqrt{\lambda_M^2} + \sqrt{\lambda_M^2 + 2\alpha_M'}}{\sqrt{\lambda_m^2}} \quad \|\omega_2\|$$

et d'après le théorème 5, on a :

$$\|\Sigma\| \leq \max(\rho_1, \rho_2)$$

Exemple :

Nous savons que les méthodes itératives sont pratiquement utilisables pour des systèmes linéaires à $n \geq 100$, et généralement les systèmes à n assez grand sont creux. Suivant la forme de la matrice, on décompose le système en "somme" de m systèmes dont on arrive facilement à calculer leurs fonctions de localisation.

Soit le système $Ax - b = 0$ avec $A \in M_{n,n}(\mathbb{R})$ et $b \in \mathbb{R}^n$ tels que

$$A = \begin{pmatrix} 7 & -3 & 0 & \dots & 1 \\ 1 & 5 & \dots & \circ & \vdots \\ \vdots & \vdots & \dots & \vdots & \vdots \\ \vdots & \circ & \dots & \vdots & \vdots \\ 1 & 0 & \dots & 1 & 7 \end{pmatrix}$$

Ce système admet $\Sigma = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ comme solution.

Le système $Ax - b = 0$ peut s'écrire comme la somme de 2 systèmes

$$A_1x - b_1 = 0$$

$$A_2x - b_2 = 0$$

$$\omega_1 = \begin{pmatrix} 1.2 \\ \vdots \\ 1.2 \\ 0.24 \end{pmatrix} \quad \omega_2 = \begin{pmatrix} 0.37 \\ 0.5 \\ \vdots \\ 0.5 \end{pmatrix}$$

d'où :

$$\|\omega_1\|^2 = (n-1)(1.2)^2 + (0.24)^2$$

$$\|\omega_2\|^2 = (n-1)(0.5)^2 + (0.37)^2$$

et

$$\|\Sigma\| \leq \frac{1}{2} \max(3 \|\omega_1\|, (1 + \sqrt{17}) \|\omega_2\|)$$

$$\|\Sigma\| \leq \frac{3}{2} \|\omega_1\| \quad (\text{car } n \geq 2)$$

$$\|\Sigma\| \leq \frac{3}{2} \sqrt{(n-1) 1.44 + 0.0576} = \rho_0$$

Or

$$\|\Sigma\| = \sqrt{n}$$

pour n assez grand ($n \geq 4$) on a aussi

$$\rho_0 \leq \frac{3}{2} \sqrt{n} = \frac{3}{2} \|\Sigma\|$$

Remarque :

Si $Ax = b$ (I)

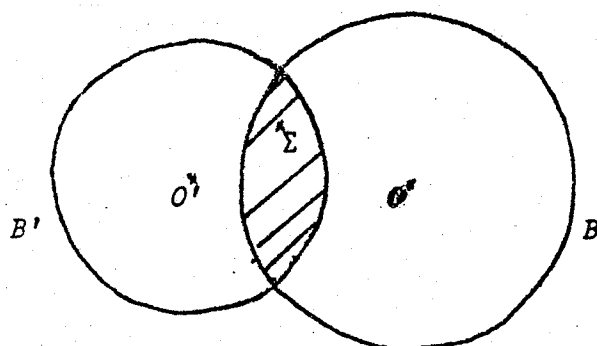
est le système somme de deux systèmes :

$$A_1 x = b_1 \quad (II)$$

$$A_1 x = b_2$$

d'après le théorème 5, nous pouvons localiser la racine Σ de (I) dans une boule B de rayon ρ .

Si on fait une translation d'origine $O' = O + h$, nous remarquons qu'on peut localiser Σ dans une autre boule B' de centre O' , et Σ appartiendra alors à l'intersection de B et B' ;



ceci en tenant compte que le système

$$(A_1 + A_2)(x + h) = b_1 + b_2 + A_1 h + A_2 h \quad (I')$$

est le système somme de deux systèmes

$$\begin{aligned} A_1(x + h) &= b_1 + A_1 h = b'_1 \\ A_2(x + h) &= b_2 + A_2 h = b'_2 \end{aligned} \quad (II')$$

et (II') n'est autre que (II), mais c'est le second membre qui a changé.

CHAPITRE II

CAS D'UN SYSTEME NON LINEAIRE

Théorème 7 :

Soit $F(x)$ une fonction non linéaire continue et définie de \mathbb{R}^n dans \mathbb{R}^n , et soit Ω une racine de $F(x)$, alors pour que $\rho > \|\Omega\|$ il suffit qu'il existe un $\varepsilon_\rho > 0$ tel que :

$$\langle x, F(x) \rangle \leq (1 - \varepsilon_\rho) \|x\|^2 \|F(x)\| \quad (1)$$

$$\forall x \mid \|x\| = \rho$$

Démonstration :

Supposons pour un $\rho > 0$ (1) est vraie, alors on a $\rho > \|\Omega\|$ en effet : $x \longrightarrow F(x)$ est une fonction continue et définie dans tout \mathbb{R}^n , supposons que sur la frontière de $\mathbb{B}(0, \rho)$ il existe $\alpha > 0$, et x_0 tels que $\|x_0\| = \rho$ et $\alpha x_0 = F(x_0)$.

D'après (1), on a :

$$x_0^T \alpha x_0 \leq (1 - \varepsilon_\rho) \|x_0\| \|\alpha x_0\|$$

$$\alpha \|x_0\|^2 \leq (1 - \varepsilon_\rho) \|x_0\|^2$$

$$1 \leq (1 - \varepsilon_\rho)$$

Or, $\varepsilon_\rho > 0$, ce qui est absurde. Donc, il existe d'après le théorème de Brouwer une solution de $F(x)$ à l'intérieur de $\mathbb{B}(0, \rho)$.

EXEMPLE D'UN SYSTEME NON LINEAIRE

Soit

$$F(x) = - \begin{pmatrix} x^T B_1 x \\ \vdots \\ x^T B_i x \\ \vdots \\ x^T B_n x \end{pmatrix} - A x + b = 0 \quad (2)$$

une fonction de \mathbb{R}^n dans \mathbb{R}^n . C'est une fonction définie et continue dans tout \mathbb{R}^n .

On suppose dans la suite que $\forall i = 1, \dots, n$ $B_i \in M_{n \times n}$ est symétrique définie non négative et A symétrique définie positive.

Posons :

1° - Ω une racine de $F(x)$ alors

$$b = \begin{pmatrix} \Omega^T B_1 \Omega \\ \vdots \\ \Omega^T B_n \Omega \end{pmatrix} + A\Omega$$

2° - λ_i ($i=1, \dots, n$) les valeurs propres de A , λ_M et λ_m étant respectivement la plus grande et la plus petite valeur propre de A

3° - β_i^j $j=1, \dots, n$ les valeurs propres de B_i ($i=1, \dots, n$)

4° - $\beta_i = \max (\beta_i^j)$ $j=1, \dots, n$

5° - $\delta = \|\Omega\|$

6° - $\beta = \max_i \beta_i$ ($i=1, \dots, n$).

On va essayer dans la suite de trouver des conditions suffisantes pour localiser Ω (racine de $F(x)$ (2)) en utilisant le théorème précédent.

Il suffit alors de trouver des conditions pour que :

$$\frac{\langle x, F(x) \rangle}{\|x\| \|F(x)\|} \leq 0 < 1 \quad \text{c'est-à-dire } \langle x, F(x) \rangle \leq 0$$

$$\forall x \in \partial B(0, \rho)$$

ou

$$x^T \left[- \begin{pmatrix} x^T B_1 x \\ \vdots \\ x^T B_n x \end{pmatrix} - Ax + \begin{pmatrix} \Omega^T B_1 \Omega \\ \vdots \\ \Omega^T B_n \Omega \end{pmatrix} + A\Omega \right] \leq 0$$

ce qui revient à

$$- x^T \begin{pmatrix} x^T B_1 x - \Omega^T B_1 \Omega \\ \vdots \\ x^T B_n x - \Omega^T B_n \Omega \end{pmatrix} - x^T A(x - \Omega) \leq 0 \quad (3)$$

CAS OU LES VALEURS PROPRES DE B_i SONT ASSEZ PETITES (par rapport à celles de A)

Théorème 8 :

$$\text{Soit } F(x) = - \begin{pmatrix} x^T B_1 x \\ \vdots \\ x^T B_n x \end{pmatrix} - Ax + b$$

où $\forall i = 1, \dots, n$ B_i est symétrique définie non négative, A est symétrique définie positive.

Posons : $\alpha_1 = \min(\sqrt{n} \beta, \sum_{i=1}^n \beta_i)$, $\alpha_2 = \min(\sqrt{n} \lambda_M, \sum_{i=1}^n \lambda_i)$

$$(I) \left\{ \begin{array}{l} \beta_0 = \frac{-2\alpha_2 \delta + \sqrt{4\alpha_2^2 \delta^2 + 4\delta^2 \lambda_m^2}}{4\delta^2} \\ \rho_0 = \frac{\lambda_m + \sqrt{\lambda_m^2 + 4\alpha_1 \delta(\alpha_1 \delta + \alpha_2)}}{2\alpha_1} \\ \rho_1 = \frac{\lambda_m - \sqrt{\lambda_m^2 + 4\alpha_1 \delta(\alpha_1 \delta + \alpha_2)}}{2\alpha_1} \end{array} \right.$$

si $\alpha_1 \leq \beta_0$

alors pour tout ρ compris entre ρ_0 et ρ_1 on a (3) est toujours négative $\forall x$ tel que $\|x\| = \rho$

Ainsi

$$\frac{\langle x, F(x) \rangle}{\|x\| \|F(x)\|} \leq 0 < 1$$

et par suite, d'après le théorème 7, il existe une racine à l'intérieur de la boule $B(0, \rho)$.

Démonstration :

Il faut démontrer que

$$-x^T \begin{pmatrix} x^T B_1 x - \Omega^T B_1 \Omega \\ \vdots \\ x^T B_n x - \Omega^T B_n \Omega \end{pmatrix} - x^T A(x - \Omega) \leq 0$$

ou

$$\gamma = -x^T \begin{pmatrix} x^T B_1 x - \Omega^T B_1 \Omega \\ \vdots \\ x^T B_n x - \Omega^T B_n \Omega \end{pmatrix} \leq x^T A(x - \Omega) = \gamma$$

Or

$$\gamma \leq \left| x^T \begin{pmatrix} x^T B_1 x - \Omega^T B_1 \Omega \\ \vdots \\ x^T B_n x - \Omega^T B_n \Omega \end{pmatrix} \right| = |x^T z|$$

$$\gamma \leq \sum_{i=1}^n |x_i| |z_i| \quad \text{avec } z_i = x^T B_i x - \Omega^T B_i \Omega$$

$\forall i=1, \dots, n$ B_i est symétrique définie non négative alors on peut poser

$$\begin{aligned} B_i &= Q_i^T D_i Q_i & i &= 1, \dots, n \\ y_i &= Q_i x & \text{avec } \|y_i\| &= \|x\| \\ w_i &= Q_i \Omega & \|w_i\| &= \|\Omega\| = \delta \end{aligned}$$

$$\begin{aligned} \text{alors } |z_i| &= \left| \|\rho_i^{1/2} y_i\|^2 - \|\rho_i^{1/2} \omega_i\|^2 \right| \\ &\leq \beta_i \|y_i\|^2 + \beta_i \|\omega_i\|^2 \quad \forall i=1, \dots, n \\ &\leq \beta_i (\rho^2 + \delta^2) \end{aligned}$$

$$\begin{aligned} \gamma &\leq \sum_{i=1}^n |x_i| \cdot \beta_i (\rho^2 + \delta^2) \\ &\leq (\rho^2 + \delta^2) \sum_{i=1}^n \beta_i |x_i| \\ \gamma &\leq \beta (\rho^2 + \delta^2) \sum |x_i| \leq \sqrt{n} \beta \rho (\rho^2 + \delta^2) \end{aligned}$$

ou

$$\gamma \leq \rho (\rho^2 + \delta^2) \sum_{i=1}^n \beta_i$$

Si $\alpha_1 = \min (\sum \beta_i, \sqrt{n} \beta)$

$$\text{on a } \boxed{\gamma \leq \alpha_1 \rho (\rho^2 + \delta^2)} \quad (4)$$

D'autre part, A est symétrique définie positive, posant :

$$A = Q^T D Q$$

$$y = Q x$$

$$w = Q \Omega$$

$$\begin{aligned} Y &= x^T A x - x^T A \Omega = y^T D y - y^T D w \\ &= \sum \lambda_i y_i^2 - \sum \lambda_i y_i \omega_i \\ &\geq \lambda_m \rho^2 - \sum \lambda_i y_i \omega_i \\ &\geq \lambda_m \rho^2 - \delta \left| \sum \lambda_i y_i \right| \\ Y &\geq \lambda_m \rho^2 - \sqrt{n} \lambda_M \delta \rho \end{aligned}$$

De même

$$Y \geq \lambda_m \rho^2 - \delta \rho \sum_{i=1}^n \lambda_i$$

$$\text{alors si } \alpha_2 = \min(\sqrt{n} \lambda_m, \sum_{i=1}^n \lambda_i)$$

on a

$$\boxed{Y \geq (\lambda_m \rho^2 - \alpha_2 \delta) \rho} \quad (5)$$

D'après (4) et (5), on a $\gamma \leq Y$ si

$$\alpha_1 \rho (\rho^2 + \delta^2) \leq (\lambda_m \rho - \alpha_2 \delta) \rho$$

ou si

$$\alpha_1 (\rho^2 + \delta^2) \leq \lambda_m \rho - \alpha_2 \delta$$

c'est-à-dire

$$\alpha_1 \rho^2 - \lambda_m \rho + \delta(\alpha_1 \delta + \alpha_2) \quad (6)$$

doit être négative.

Or, si ρ_0, ρ_1 sont les racines de (6), alors celle-ci est négative pour tout ρ tel que $\rho_0 \leq \rho \leq \rho_1$ (car $\alpha_1 \geq 0$)

où

$$\rho_1 = \frac{\lambda_m \pm \sqrt{\lambda_m^2 - 4 \alpha_1 \delta (\alpha_1 \delta + \alpha_2)}}{2 \alpha_1}$$

Pour que ρ_0 et ρ_1 existent, il faut que :

$$\lambda_m^2 - 4 \alpha_1 \delta (\alpha_1 \delta + \alpha_2) \geq 0$$

c'est-à-dire :

$$4 \delta^2 \alpha_1^2 + 4 \alpha_2 \delta \alpha_1 - \lambda_m^2 \leq 0 \quad (7)$$

Or, (7) est négative entre les racines.

$$\beta_0 = \frac{-2\alpha_2\delta + \sqrt{4\alpha_2^2\delta^2 + 4\delta^2\lambda_m^2}}{4\delta^2}$$

étant la plus grande racine de (7)

et pour $\alpha_1 = 0$ (7) est négative, donc (7) est toujours négative pour tout α_1 tel que

$$0 \leq \alpha_1 \leq \beta_0 \quad (8)$$

Ainsi pour tout $\alpha_1 \leq \beta_0$ ($\alpha_1 \geq 0$), il existe ρ_0 et ρ_1 tel que (6) est négative pour tout ρ tel que $\rho_0 \leq \rho \leq \rho_1$

Donc $\gamma \leq Y$, c'est-à-dire $\langle x, F(x) \rangle \leq 0 \quad \forall x / \|x\| = \rho$

Donc d'après le théorème 7, il existe une racine à l'intérieur de la boule $IB(0, \rho)$.

APPLICATION A UNE LOCALISATION DE SOLUTION D'UN SYSTEME "SOMME"

Soient

$$\begin{pmatrix} x^T B_1 x \\ \vdots \\ x^T B_n x \end{pmatrix} + Ax - b = 0 \quad (1)$$

$$\begin{pmatrix} x^T B'_1 x \\ \vdots \\ x^T B'_n x \end{pmatrix} + A'_x - b' = 0 \quad (2)$$

2 systèmes non linéaires où $\forall i=1, \dots, n$ $B_i, B'_i \in M_{n \times n}$ sont symétriques définies non négatives, A et $A' \in M_{n \times n}$ sont symétriques définies positives

$$\begin{pmatrix} x^T (B_1 + B'_1) x \\ \vdots \\ x^T (B_n + B'_n) x \end{pmatrix} + (A + A')x - (b + b') = 0 \quad (3)$$

est le système "somme" de 2 systèmes (1) et (2). Les matrices $(B_i + B'_i)$ ($i=1, \dots, n$) sont symétriques définies non négatives et $A+A'$ est symétrique définie positive.

On désigne par Ω, Ω' et Σ des solutions pour (1), (2) et (3).

$\alpha_1, \alpha_2, \beta_0, \rho_0, \rho_1$ sont les constantes du système (1) et qui sont donnés par les formules (I) du théorème 8.

$\alpha'_1, \alpha'_2, \beta'_0, \rho'_0, \rho'_1$ sont ceux du système (2).

Théorème 9 :

Si $\alpha_1, \alpha_2, \beta_0, \rho_0, \rho_1$ et $\alpha'_1, \alpha'_2, \beta'_0, \rho'_0, \rho'_1$ sont les constantes de deux systèmes (1) et (2) données par (I) du théorème 8, et si en plus :

$$\alpha_1 \leq \beta_0$$

$$\alpha'_1 \leq \beta'_0$$

et $\rho'' = \max(\rho_0, \rho'_0)$ est tel que :

$$\rho'' \leq \rho_1$$

$$\rho'' \leq \rho'_1$$

alors on a : $\|\Sigma\| \leq \rho''$

Démonstration :

$$\text{Posons } \lambda(\rho) = \max_{\|x\|=\rho} \frac{\langle x, F(x) \rangle}{\|x\| \|F(x)\|}$$

Alors on a :

$$- \langle x, \begin{pmatrix} x^T B_1 x \\ \vdots \\ x^T B_n x \end{pmatrix} + Ax - b \rangle \geq \lambda(\rho) \|F(x)\| \rho \quad \text{si } \|x\| = \rho$$

$$- \langle x, \begin{pmatrix} x^T B_1 x \\ \vdots \\ x^T B_n x \end{pmatrix} + A'x - b' \rangle \leq \lambda'(\rho) \|F'(x)\| \rho \quad \text{si } \|x\| = \rho$$

donc en ajoutant

$$- \langle x, \begin{pmatrix} x^T (B_1 + B_1') x \\ \vdots \\ x^T (B_n + B_n') x \end{pmatrix} + (A+A')x - (b+b') \rangle \leq \rho [\lambda(\rho) \|F(x)\| + \lambda'(\rho) \|F'(x)\|]$$

pour $\|x\| = \rho$

Or par hypothèse :

$$\rho_0'' \in [\rho_0, \rho_1] \implies \lambda(\rho_0'') \leq 0$$

$$\rho_0'' \in [\rho_0', \rho_1'] \implies \lambda(\rho_0'') \leq 0$$

(d'après le théorème 8).

donc $\langle x, F(x) + F'(x) \rangle \leq 0$ pour tout x tel que $\|x\| = \rho_0''$.

Ainsi :

$$\frac{\langle x, F(x) + F'(x) \rangle}{\|x\| \|F(x) + F'(x)\|} \leq 0$$

donc $\|\Sigma\| < \rho_0''$ et par suite $\Sigma \in \mathcal{B}(0, \rho)$

Exemple :

Soit

$$(1) \begin{pmatrix} x^T B_1 x \\ \vdots \\ x^T B_n x \end{pmatrix} + Ax - b = 0$$

où B_1 est symétrique définie non négative

A symétrique définie positive.

$$(2) \begin{pmatrix} 0 \\ \vdots \\ x^T B_1 x \\ \vdots \\ x^T B_n x \end{pmatrix} + A_1 x - b_1 = 0 \quad (i = 1, \dots, n).$$

A_i est symétrique définie positive et tel que $\sum_{i=1}^n A_i = A$ et $\sum b_i = b$.

Le système (1) est le système "somme" de n systèmes (2).

Or le système (2) est facile à résoudre (on a $n - 1$ équations linéaires et une équation polynomiale du second degré).

Ainsi, si les n systèmes (2) vérifient les conditions du théorème 9, alors le système (1) admet une racine à l'intérieur de la boule $B(0, \rho)$

où

$$\rho = \max_i (\rho_0^i) \quad \text{avec} \quad \rho \leq \min_i (\rho_1^i)$$

(où ρ_0^i, ρ_1^i correspondant au $i^{\text{ième}}$ système de (2) sont donnés par les formules (I) du théorème 8).

Théorème 10 :

$$F(x) = \begin{pmatrix} x^T B_1 x \\ \vdots \\ x^T B_n x \end{pmatrix} + Ax - b$$

où

$B_i \in M_{n,n}$ est symétrique définie non négative ($i = 1, \dots, n$)

$A \in M_{n,n}$ est symétrique définie positive

$b \in \mathbb{R}^n$

Soient β_i la plus grande valeur propre de B_i ($i = 1, \dots, n$)

$$\beta = \max (\beta_i)$$

λ_m la plus petite valeur propre de A

$$\alpha = \min (\sqrt{n}\beta, \sum_{i=1}^n \beta_i)$$

Si

$$\alpha \leq \frac{\lambda_m^2}{4 \sum |b_i|}$$

alors $\forall \rho \in [\rho_0, \rho_1]$ avec

$$\rho_0 = \frac{\lambda_m \pm \sqrt{\lambda_m^2 - 4\alpha \sum |b_i|}}{2\alpha^2}$$

$$\rho_1 = \frac{\lambda_m \pm \sqrt{\lambda_m^2 - 4\alpha \sum |b_i|}}{2\alpha^2}$$

on a :

$$\|\Omega\| \leq \rho \text{ où } \Omega \text{ est une racine de } F(x).$$

Démonstration :

On a vu dans le théorème 8 qu'il suffit de démontrer :

$$-x^T \begin{pmatrix} x^T B_1 x \\ \vdots \\ x^T B_n x \end{pmatrix} \leq x^T A x - x^T b \quad (1)$$

$$\text{or } x^T A x - x^T b \geq \lambda_m \rho^2 - \rho \sum |b_i| = (\lambda_m - \sum |b_i|) \rho$$

De même :

$$\gamma = -x^T \begin{pmatrix} x^T B_1 x \\ \vdots \\ x^T B_n x \end{pmatrix} \leq |x^T Z| = |\sum x_i z_i|$$

$$\text{où } z_i = x^T B_i x = y_i^T D_i y_i$$

$$\leq \beta_i \rho^2 \quad (B_i \text{ symétrique définie non négative})$$

$$\implies \gamma \leq \sum x_i \beta_i \rho^2$$

$$\leq \rho^2 \beta \sqrt{n} \rho = \rho^3 \sqrt{n} \beta$$

$$\text{de même } \gamma \leq \rho^3 \sum \beta_i$$

$$\text{si } \alpha = \min(\sqrt{n}, \sum \rho_i)$$

on a :

$$\gamma \leq \alpha \rho^3$$

Ainsi il suffit que :

$$\alpha \rho^3 \leq (\lambda_m \rho - \sum |b_i|) \rho$$

c'est-à-dire :

$$\alpha \rho^2 \leq (\lambda_m \rho - \sum |b_i|)$$

$$\alpha \rho - \lambda_m \rho + \sum |b_i| \leq 0 \quad (2)$$

Or les racines de (2) :

$$\rho_0 = \frac{\lambda_m \pm \sqrt{\lambda_m^2 - 4\alpha \sum |b_i|}}{2\alpha}$$

$$\rho_1 = \frac{\lambda_m - \sqrt{\lambda_m^2 - 4\alpha \sum |b_i|}}{2\alpha}$$

existent car $\alpha \leq \frac{\lambda_m^2}{4\sum |b_i|}$ par hypothèse.

α étant positive, alors (2) est négative pour tout ρ tel que $\rho_0 \leq \rho \leq \rho_1$. Ainsi (1) est vérifié $\forall x$ tel que $\|x\| = \rho$ et par suite $F(x)$ admet une racine à l'intérieur de la boule $IB(0, \rho)$.

Remarque : D'après les théorèmes 9 et 10, on remarque que si les valeurs propres des B_i ($i = 1, \dots, n$) sont assez petites, alors $F(x)$ admet une solution qu'on peut localiser.

Application :

Dans cette partie, nous prenons quelques exemples particuliers de la forme :

$$F(x) = \begin{pmatrix} x^T B_1 x \\ \vdots \\ x^T B_n x \end{pmatrix} + Ax - b = 0$$

et nous localisons les racines en utilisant le théorème 10. En prenant maintenant des points de départ appartenant à la boule déjà trouvée et qui contient au moins une racine, nous appliquons la méthode des sécantes pour résoudre le système $F(x) = 0$.

Dans la suite, nous donnons une table qui, pour chaque exemple, donne la petite valeur propre λ_m , α , ρ_0 qui est le rayon de la boule qui contient une racine, norme de x^* et norme de $F(x^*)$ où x^* est une racine de $F(x) = 0$ et donne aussi le nombre d'itérations m exigé par la méthode des sécantes pour converger vers x^* .

EXEMPLES :Exemple 1

$$\begin{aligned} y_1 &= 0,01 x_2^2 + 2 x_1 - 10 \\ y_2 &= 0,01 x_3^2 + 5 x_2 + x_3 + 5 \\ y_3 &= 0,01 x_1^2 + 0,01 x_2^2 + x_2 + 3 x_3 - 5 \end{aligned}$$

Exemple 2

$$\begin{aligned} y_1 &= 0,01 x_1^2 + 10 x_1 + 2 x_2 + 80 \\ y_2 &= 0,1 x_2^2 + 2 x_1 + 20 x_2 - 87,5 \\ y_3 &= 0,1 x_3^2 + 9 x_3 - 12,5 \end{aligned}$$

Exemple 3

$$\begin{aligned} y_1 &= 0,2 x_1^2 - 0,2 x_1 x_2 + 0,2 x_2^2 + 5 x_1 + 2 x_2 + 1 \\ y_2 &= 0,3 x_3^2 + 2 x_1 + 7 x_2 - 2 x_3 - 1 \\ y_3 &= 0,1 x_2^2 - 0,1 x_1 x_2 + 0,1 x_3^2 - 2 x_2 + 5 x_3 - 1 \end{aligned}$$

Exemple 4

$$\begin{aligned}
 y_1 &= 0,05 (x_1^2 - 2 x_1 x_2 + x_2^2) + 5 x_1 - 3 x_2 - 2 \\
 y_2 &= 0,05 (x_1^2 + 2 x_1 x_2 + x_2^2) - 3 x_1 + 8 x_2 - x_3 + 1 \\
 y_3 &= 0,05 x_2^2 + 0,05 x_3^2 + x_2 + 5 x_3 + 2
 \end{aligned}$$

Exemple 5

$$\begin{aligned}
 y_1 &= 0,05 (x_1^2 - 2 x_1 x_2 + x_2^2) + 3 x_1 + 2 x_2 - 0,6 \\
 y_2 &= 0,015 x_1^2 + 0,015 x_2^2 + 0,015 x_3^2 + 2 x_1 + 5 x_2 - 2 x_3 + 0,4 \\
 y_3 &= 0,01 x_2^2 + 0,01 x_3^2 - 2 x_2 + 3 x_3 - 1
 \end{aligned}$$

Exemple 6

$$\begin{aligned}
 y_1 &= 0,5 (x_1^2 + x_4^2 + x_5^2) + 6 x_1 - 2 x_2 - 5 \\
 y_2 &= 0,25 (x_2^2 + x_3^2) - 2 x_1 + 6 x_2 + 3,75 \\
 y_3 &= 0,75 (x_2^2 + x_5^2) + 5 x_3 + 2 x_4 + 1,25 \\
 y_4 &= 0,4 (x_3^2 + x_5^2) + 2 x_3 + 7 x_4 - 2 x_5 + 5 \\
 y_5 &= 0,6 (x_1^2 + x_3^2 + x_5^2) - 2 x_4 + 5 x_5 - 5
 \end{aligned}$$

Exemple 7

$$\begin{aligned}
 y_1 &= 0,03 (x_1^2 + x_4^2) + 3 x_1 + 2 x_2 + 1 \\
 y_2 &= 0,002 (x_2^2 + x_5^2 + x_1^2) + 2 x_1 + 5 x_2 - 2 x_3 - 1 \\
 y_3 &= 0,005 (x_2^2 + x_3^2) - 2 x_2 + 3 x_3 + 1 \\
 y_4 &= 0,001 (x_4^2 + x_9^2) + x_4 - 0,5 \\
 y_5 &= 0,003 (x_4^2 + x_6^2 + x_8^2) + 5 x_5 + 2 x_6 - 1,5 \\
 y_6 &= 0,005 (x_5^2 + x_6^2 + x_8^2) + 2 x_5 + 7 x_6 - 2 x_7 - 1 \\
 y_7 &= 0,001 (x_3^2 + x_7^2) - 2 x_6 + 5 x_7 + 1 \\
 y_8 &= 0,002 (x_7^2 + x_9^2 + x_3^2) + 2 x_8 - 0,5 \\
 y_9 &= 0,001 (x_7^2 + x_9^2) + 3 x_9 - x_{10} + 1,5 \\
 y_{10} &= 0,002 (x_5^2 + x_{10}^2) - x_9 + 3 x_{10} + 1
 \end{aligned}$$

- Table -

	λ_m	α	ρ_0	$\ x^*\ $	$\ F(x^*)\ $	m
Exemple 1	2	0,03	12	5,5778	$3,1401 \times 10^{-16}$	9
Exemple 2	9	0,1	30	10,5597	0	6
Exemple 3	3	0,6	2,3	0,57828	$3,7751 \times 10^{-16}$	7
Exemple 4	3	0,2	9,6	0,54793	$2,3263 \times 10^{-16}$	8
Exemple 5	1	0,125	4	0,3824	$5,578 \times 10^{-15}$	6
Exemple 6	3	2,5	10	divergé		
Exemple 7	1	0,15	40	1,3478	$5,5776 \times 10^{-15}$	9

Remarque :

La table 1 montre que la méthode des sécantes converge pour la plupart des exemples considérés, avec des points initiaux x_0, \dots, x_n appartenant à la boule trouvée, contenant une racine de $F(x) = 0$; d'où nous précisons que la localisation d'une racine de $F(x) = 0$, nous donne un moyen efficace de bien démarrer l'itération (pour le cas de la méthode des sécantes, ou d'autres méthodes itératives).

B I B L I O G R A P H I E

- [1] S. LEFSCHETZ (1949)

Introduction to topology.

Princeton University Press, New Jersey.

- [2] N. GASTINEL (1968 - 1969)

Les méthodes générales des "projections" pour la solution de problèmes linéaires et les constantes-rectangles d'espaces normés (Séminaire - Grenoble, N° 70).

A N N E X E

PROCEDURE ALGOLW POUR LA RECHERCHE DES RACINES D UNE FONCTION
NON LINEAIRE PAR LA METHODE DES SECANTES

CETTE PROCEDURE UTILISE UNE PROCEDURE
PROCEDURE EVALFONCT(INTEGER N,U;LONG REAL ARRAY X,F(*));
CETTE PROCEDURE D'EVALUATION DE N COMPOSANTES DE F DONNE AUSSI LA
POSSIBILITE DE RESOUDRE M SYSTEMES A LA FOIS,F EST LA FONCTION
A EVALUER , U LE NUMERO DU SYSTEME;

```
BEGIN
  PROCEDURE PRODMAT(LONG REAL ARRAY A(*,*);LONG REAL ARRAY B,C(*)
    INTEGER N);
    BEGIN
      FOR I:= 1 UNTIL N DO
        BEGIN C(I):= 0L;
          FOR J:= 1 UNTIL N DO C(I):= C(I)+A(I,J)*B(J);
        END;
      END PRODMAT;
```

COMMENT N ETANT L'ORDRE DU SYSTEME $F(I)=0(I= 1....N)$ OU $F(I,*)=0$,
LES TABLEAUX $(X(I,J))$, $(F(I,J))$ REPRESENTENT RESPECTIVEMENT LES N
COMPOSANTES DE N+2 VECTEURS $(X(.,J))$ ET N+2 VECTEURS F EVALUES
RESPECTIVEMENT EN N+2 VECTEURS $(X(.,J))$ $J=0...N+1$,
 $H(I,J)=X(I,J)-X(I,J-1)$ $I,J=1,...N$, $M(I,J)=F(I,J)-F(I,J-1)$ $I,J=1,...N$,
SI B EST L'INVERSE DE M LA FORMULE DE LA METHODE DES SECANTES
S'ECRIT : $X(.,N+1)=X(.,N)-H*B*F(.,N)$.
LE TEST D'ARRET EST :OU BIEN NORME DE F<EPS OU BIEN LE NOMBRE
D'ITERATION EST EGAL A L'ENTIER ITERMAX.
LES DONNEES SONT: LA VALEUR DE N (L'ORDRE DU SYSTEME) ET LES N
COMPOSANTES DE N+1 VECTEURS DE DEPART $X(I)$ $I=0,...N$;

```
BEGIN
  INTEGER V,U,N;
  U:= 1;
  PREMIER: READ(N);
  IF N=999 THEN GOTO DERNIER;
  WRITE("N",N);
  WRITE("U",U);
  V:=1;
  BEGIN
    LONG REAL ARRAY X,F(1::N,0::N+1);
    LONG REAL ARRAY B(1::N,1::N);
    LONG REAL ARRAY O,C(1::N);
    FOR J:=0 UNTIL N DO
      FOR I:=1 UNTIL N DO
        READON (X(I,J));
      END;
  END;
```

```

LONG REAL ARRAY M(1::N,1::N);
FOR K:= 0 UNTIL N DO
EVALFONCT(N,U,X(*,K),F(*,K));
FOR I:= 1 UNTIL N DO
FOR J:= 1 UNTIL N DO
M(I,J):= F(I,N-J+1)-F(I,N-J) ;

```

COMMENT CALCUL ,AU PREMIER PAS DE LA METHODE DES SECANTES,
DE LA MATRICE B INVERSE DE M PAR SIMPLE ELIMINATION DE GAUSS;

```

BEGIN
INTEGER I,J,L; LONG REAL R,TB; LONG REAL ARRAY M1(1::N,1::2*N);
FOR I:= 1 UNTIL N DO
FOR J:= 1 UNTIL N DO
BEGIN M1(I,J):= M(I,J); M1(I,N+J):=0L; END;
FOR I:= 1 UNTIL N DO M1(I,N+1):=1.L;
FOR L:= 1 UNTIL N-1 DO
BEGIN
NORMAL: BEGIN IF M1(L,L)=0.L THEN GOTO ECHANGE;
FOR I:= L+1 UNTIL N DO
BEGIN R:= M1(I,L)/M1(L,L);
FOR J:= L+1 UNTIL N+L DO
M1(I,J):= M1(I,J)-R*M1(L,J);
END;
END; GOTO RETOUR;
ECHANGE: BEGIN INTEGER Q;
Q:= L+1; WHILE (M1(Q,L)=0.L) AND (Q<=N) DO Q:= Q+1;
IF Q= N+1 THEN GOTO IMPOSSIBLE ;
FOR J:= L UNTIL N+Q DO
BEGIN R:= M1(L,J); M1(L,J):= M1(Q,J); M1(Q,J):= R;
END; GOTO NORMAL END;
RETOUR: END;
BEGIN INTEGER I,J,L; IF M1(N,N)=0.L THEN GOTO IMPOSSIBLE ;
FOR I:= N STEP -1 UNTIL 1 DO
FOR J:= N STEP -1 UNTIL 1 DO
BEGIN TB:= 0.L; FOR L:= N STEP -1 UNTIL I+1 DO
TB:= TB-B(L,J)*M1(I,L);
B(I,J):= (TB+M1(I,N+J))/M1(I,I);
END; END; END; GOTO A2; END;
A1:
FOR I:= 1 UNTIL N DO
BEGIN O(I):= B(N,I);

FOR J:= N STEP -1 UNTIL 2 DO
B(J,I):= B(J-1,I); B(I,I):= O(I); END;

```

COMMENT CALCUL AU KIEME PAS DE LA METHODE DES SECANTES ,DE LA
MATRICE B INVERSE DE M PAR PERTURBATION ;

```

BEGIN LONG REAL R; LONG REAL ARRAY B1(1::N,1::N);
FOR J:= 1 UNTIL N DO
O(J):= F(J,N+1)-F(J,N)-F(J,1)+F(J,0); R:= 0.L;
FOR I:= 1 UNTIL N DO
R:= R+B(1,I)*O(I); R:= -1.L/(1+R); PRODMAT(B,O,C,N);
FOR I:= 1 UNTIL N DO
FOR J:= 1 UNTIL N DO

```

```

B1(I,J):= R*C(I)*B(1,J);
FOR I:= 1 UNTIL N DO
FOR J:= 1 UNTIL N DO
B(I,J):= B(1,J)+B1(I,J);
FOR I:= 1 UNTIL N DO
FOR J:= 0 UNTIL N DO
BEGIN X(I,J):= X(I,J+1); F(I,J):= F(I,J+1); END; END;

```

COMMENT AYANT LA MATRICE B ON CALCULE A CHAQUE PAS LE
NOUVEL ITERE : $X(.,N+1) = X(.,N) - H * B * F(.,N)$;

```

A2: PRODMAT(B,F(*,N),C,N);
BEGIN LONG REAL ARRAY H(1::N,1::N);
FOR I:= 1 UNTIL N DO
FOR J:= 1 UNTIL N DO
H(I,J):= X(I,N-J+1)-X(I,N-J); PRODMAT(H,C,O,N); END;
BEGIN INTEGER I,ITERMAX; LONG REAL R,EPS;
FOR I:= 1 UNTIL N DO
BEGIN
X(I,N+1):= X(I,N)-O(I);
END;
EVALFONCT(N,U,X(*,N+1),F(*,N+1));
FOR I:= 1 UNTIL N DO
WRITE("F(XN+1)",I,"=",F(I,N+1));
R:=0.L; FOR I:= 1 UNTIL N DO
R:= R + X(I,N+1)**2; R:= SQRT(R);
WRITE ("NORME DE XN ",I,"=",R);
R:= 0.L;
FOR I:= 1 UNTIL N DO
R:= R+F(I,N+1)**2; R:= SQRT(R);
WRITE("NORME DE F(XN)",I,"=",R);
IF R < 1'4L THEN GOTO IMPOSSIBLE ;
EPS := 1'-12; ITERMAX:= 20;
IF ((R<EPS) OR (V=ITERMAX)) THEN GOTO A3;
END;
V:= V+1; GOTO A1;
A3:
WRITE(" NOMBRE DES ITERATIONS ",V) ;
FOR I:= 1 UNTIL N DO
WRITE("RACINE DE F(X)",X(I,N+1));
IMPOSSIBLE:
U:= U+1;
GOTO PREMIER ;
END;
END;DERNIER:
END.

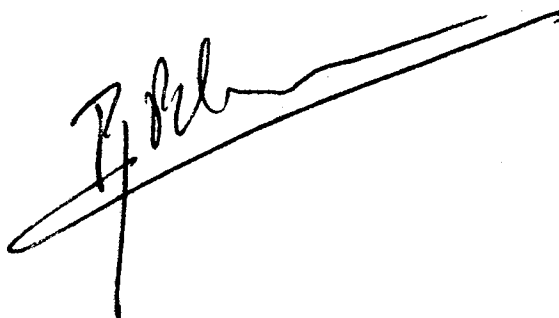
```

Dernière page d'une thèse

VU

Grenoble, le

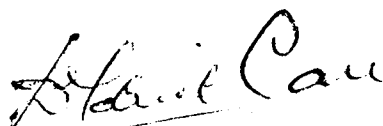
Le Président de la thèse

A handwritten signature in black ink, appearing to be 'P. Del', written over a horizontal line that extends to the right.

Vu, et permis d'imprimer,

Grenoble, le

Le Président de l'Université
Scientifique et Médicale

A handwritten signature in black ink, appearing to be 'Louis Cau', written over a horizontal line that extends to the left.