N° d'ordre : 918

THÈSE

présentée à

L'UNIVERSITÉ DES SCIENCES ET TECHNIQUES DE LILLE

pour obtenir le titre de

DOCTEUR DE 3eme CYCLE

(Mathématiques Appliquées)

par

Mohammedi EL HALLABI

SUR LA CONVERGENCE DE FORMULES DE METRIQUE VARIABLE PARAMETREES

Soutenue le 9 octobre 1981 devant la Commission d'Examen

Membres du Jury :	**Président**	**C. BREZINSKI**
	Rapporteur	**J.C. FIOROT**
	Examinateurs	**C. LEMARECHAL**
		F. MIGNOT

PROFESSEURS 1ERE CLASSE

M. BACCHUS Pierre	Mathématiques
M. BEAUFILS Jean Pierre (dét.)	Chimie
M. BIAYS Pierre	G.A.S.
M. BILLARD Jean	Physique
M. BONNOT Ernest	Biologie
M. BOUGHON Pierre	Mathématiques
M. BOURIQUET Robert	Biologie
M. CELET Paul	Sciences de la Terre
M. COEURE Gérard	Mathématiques
M. CONSTANT Eugène	I.E.E.A.
M. CORDONNIER Vincent	I.E.E.A.
M. DEBOURSE Jean Pierre	S.E.S.
M. DELATTRE Charles	Sciences de la Terre
M. DURCHON Maurice	Biologie
M. ESCAIG Bertrand	Physique
M. FAURE Robert	Mathématiques
M. FOURET René	Physique
M. GABILLARD Robert	I.E.E.A.
M. GRANELLE Jean Jacques	S.E.S.
M. GRUSON Laurent	Mathématiques
M. GUILLAUME Jean	Biologie
M. HECTOR Joseph	Mathématiques
M. HEUBEL Joseph	Chimie
M. LABLACHE COMBIER Alain	Chimie
M. LACOSTE Louis	Biologie
M. LANSRAUX Guy	Physique
M. LAVEINE Jean Pierre	Sciences de la Terre
M. LEBRUN André	C.U.E.E.P.
M. LEHMANN Daniel	Mathématiques
Mme LENOBLE Jacqueline	Physique
M. LHOMME Jean	Chimie
M. LOMBARD Jacques	S.E.S.
M. LOUCHEUX Claude	Chimie
M. LUCQUIN Michel	Chimie
M. MAILLET Pierre	S.E.S.
M. MONTREUIL Jean	Biologie
M. PAQUET Jacques	Sciences de la Terre
M. PARREAU Michel	Mathématiques
M. PROUVOST Jean	Sciences de la Terre

.../...

Professeurs 1ère classe (suite)

M. SALMER Georges	I.E.E.A.
Mme SCHWARTZ Marie Hélène	Mathématiques
M. SEGUIER Guy	I.E.E.A.
M. STANKIEWICZ François	Sciences Economiques
M TILLIEU Jacques	Physique
M. TRIDOT Gabriel	Chimie
M. VIDAL Pierre	I.E.E.A.
M. VIVIER Emile	Biologie
M. WERTHEIMER Raymond	Physique
M. ZEYTOUNIAN Radyadour	Mathématiques

P R O F E S S E U R S 2ème C L A S S E

M. AL FAKIR Sabah	Mathématiques
M. ANTOINE Philippe	Mathématiques
M. BART André	Biologie
Mme BATTIAU Yvonne	Géographie
M. BEGUIN Paul	Mathématiques
M. BELLET Jean	Physique
M. BKOUCHE Rudolphe	Mathématiques
M. BOBE Bernard	S.E.S.
M. BODART Marcel	Biologie
M. BOILLY Bénoni	Biologie
M. BONNELLE Jean Pierre	Chimie
M. BOSQ Denis	Mathématiques
M. BREZINSKI Claude	I.E.E.A.
M. BRUYELLE Pierre (Chargé d"enseignement)	Géographie
M. CAPURON Alfred	Biologie
M. CARREZ Christian	I.E.E.A.
M. CHAMLEY Hervé	E.U.D.I.L.
M. CHAPOTON Alain	C.U.E.E.P.
M. COQUERY Jean Marie	Biologie
Mme CORSIN Paule	Sciences de la Terre
M. CORTOIS Jean	Physique
M. COUTURIER Daniel	Chimie
Mle DACHARRY Monique	Géographie
M. DEBRABANT Pierre	E.U.D.I.L.
M. DEGAUQUE Pierre	I.E.E.A.
M. DELORME Pierre	Biologie
M. DEMUNTER Paul	C.U.E.E.P.
M. DE PARIS Jean-Claude	Mathématiques

.../...

M. DEVRAINNE Pierre	Chimie
M. DHAINAUT André	Biologie
M. DORMARD Serge	S.E.S.
M. DOUKHAN Jean Claude	E.U.D.I.L.
M. DUBOIS Henri	Physique
M. DUBRULLE Alain	Physique
M. DUEE Gérard	Sciences de la Terre
M. DYMENT Arthur	Mathématiques
M. FLAMME Jean Marie	E.U.D.I.L.
M. FONTAINE Hubert	Physique
M. GERVAIS Michel	S.E.S.
M. GOBLOT Rémi	Mathématiques
M. GOSSELIN Gavriel	S.E.S.
M. GOUDMAND Pierre	Chimie
M. GREVET Patrice	S.E.S.
M. GUILBAULT Pierre	Biologie
M. HANGAN Théodor	Mathématiques
M. HERMAN Maurice	Physique
M. JACOB Gérard	I.E.E.A.
M. JACOB Pierre	Mathématiques
M. JOURNEL Gérard	E.U.D.I.L.
M. KREMBEL Jean	Biologie
M. LAURENT François	I.E.E.A.
Mlle LEGRAND Denise	Mathématiques
Mlle LEGRAND Solange	Mathématiques (Calais)
Mme LEHMANN Josiane	Mathématiques
M. LEMAIRE Jean	Physique
M. LENTACKER Firmin	G.A.S.
M. LEVASSEUR Michel	I.P.A.
M. LHENAFF René	G.A.S.
M. LOCQUENEUX Robert	Physique
M. LOSFELD Joseph	I.E.E.A.
M. LOUAGE Francis	E.U.D.I.L.
M. MACKE Bruno	Physique
M. MAIZIERES Christian	I.E.E.A.
Mlle MARQUET Simone	Mathématiques
M. MESSELYN Jean	Physique
M. MIGEON Michel	E.U.D.I.L.
M. MIGNOT Fulbert	Mathématiques
M. MONTEL Marc	Physique
Mme NGUYEN VAN CHI Régine	G.A.S.
M. PARSY Fernand	Mathématiques
Mlle PAUPARDIN Colette	Biologie

Professeurs 2ème classe (suite)

M. PERROT Pierre	Chimie
M. PERTUZON Emile	Biologie
M. PONSOLLE Louis	Chimie
M. PORCHET Maurice	Biologie
M. POVY Lucien	E.U.D.I.L
M. RACZY Ladislas	I.E.E.A.
M. RICHARD Alain	Biologie
M. RIETSCH François	E.U.D.I.L.
M. ROGALSKI Marc	M.P.A.
M. ROUSSEAU Jean-Paul	Biologie
M. ROY Jean-Claude	Biologie
M. SALAMA Pierre	S.E.S.
Mme SCHWARZBACH Yvette (CCP)	M.P.A.
M. SCHAMPS Joël	Physique
M. SIMON Michel	S.E.S.
M. SLIWA Henri	Chimie
M. SOMME Jean	G.A.S.
Mle SPIK Geneviève	Biologie
M. STERBOUL François	E.U.D.I.L.
M. TAILLIEZ Roger	Institut Agricole
M. TOULOTTE Jean-Marc	I.E.E.A.
M. VANDORPE Bernard	E.U.D.I.L.
M. WALLART Francis	Chimie
M. WATERLOT Michel	Sciences de la Terre
Mme ZINN JUSTIN Nicole	M.P.A.

CHARGES DE COURS

M. TOP Géard	S.E.S.
M. ADAM Michel	S.E.S.

CHARGES DE CONFERENCES

M. DUVEAU Jacques	S.E.S.
M. HOFLACK Jacques	I.P.A.
M. LATOUCHE Serge	S.E.S.
M. MALAUSSENA DE PERNO Jean-Louis	S.E.S.
M. OPIGEZ Philippe	S.E.S.

REMERCIEMENTS

Monsieur le Professeur BREZINSKI, dont j'ai suivi avec intérêt le Séminaire d'Analyse Numérique et d'Optimisation, a bien voulu me faire l'honneur de présider le Jury. Je le prie de croire à ma profonde gratitude.

Je suis très reconnaissant envers Monsieur le Professeur MIGNOT qui a bien voulu s'intéresser à ce travail et le juger.

Je remercie très vivement Monsieur LEMARECHAL, Ingénieur de Recherche à l'INRIA, pour avoir accepté de s'intéresser à ce travail, de le juger et de m'avoir fourni le code M1QN1.

Que Monsieur FIOROT, qui m'a proposé ce sujet et dont les conseils, les suggestions et les encouragements permanents m'ont été si précieux, veuille bien trouver ici l'expression de toute ma reconnaissance.

Que tous mes amis trouvent ici l'expression de ma gratitude pour leur présence chaleureuse et permanente.

Je voudrais remercier les membres du Laboratoire de Calcul de l'Université des Sciences et Techniques de Lille I pour l'accueil qu'ils m'ont réservé et pour les facilités matérielles qu'ils m'ont accordées.

Je remercie Monsieur HAYEZ, du Centre Interuniversitaire de Traitement de l'Information qui m'a souvent aidé lors de difficultés de programmation.

J'adresse enfin mes remerciements à Madame CARON et Mademoiselle Bénédicte FIEVET qui ont tapé cette thèse avec soin et célérité, ainsi que Mr et Mme DEBOCK qui, avec les mêmes qualités, en ont tirés les épreuves.

A mon père,

A ma mère,

A toute ma famille.

TABLE DES MATIÈRES

NOTATIONS et DEFINITIONS

$\mathbb{R}^n$ espace vectoriel de dimension n construit sur le corps des réels $\mathbb{R}$,

f fonction réelle définie sur $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$,

$\nabla f(x_k) = g(x_k) = g_k$ gradient de f au point x_k,

$\nabla^2 f(x_k) = G(x_k) = G_k$ hessien de f au point x_k, symétrique,

$s_k = x_{k+1} - x_k$ différence de points successifs,

$y_k = g_{k+1} - g_k$ différence de gradients successifs,

A^t la transposée de la matrice A,

A^{-1} la matrice inverse de la matrice A

$A^{\frac{1}{2}}$ la matrice "racine carrée" de A

$||A||$ la norme euclidienne de A

$||x||$ la norme euclidienne de x dans $\mathbb{R}^n$

$<s, y>$ le produit scalaire des vecteurs s et y

INTRODUCTION

Ce travail entre dans le cadre de la minimisation d'une fonction réelle continuement différentiable sans contrainte, i.e. minimiser f(x) pour x parcourant $\mathbb{R}^n$. Ce type de problème, propre en soi, sert souvent pour résoudre des problèmes de minimisation avec contraintes. En effet, beaucoup de méthodes proposées jusqu'ici en optimisation avec contraintes conduisent à remplacer le problème donné par une suite de problèmes sans contraintes. C'est le cas des méthodes de pénalités (Carol [8], Fiacco et Mc-Cormick [18], Lootsma [35], Zangwill [49]) et des méthodes des multiplicateurs (Bertsekas [3, 4], Di Pillo et Grippo [16], Hestenes [29], Miele, Moseley, Levy et Coggins [36], Powell [42], Rockafellar [44]).

Les algorithmes proposées consistent à déterminer le successeur du point courant x_k sous la forme $x_{k+1} = x_k + \lambda_k d_k$, où d_k, vecteur de $\mathbb{R}^n$, est la direction et λ_k, un réel strictement positif, est le pas. Pour la détermination de la direction, une condition nécessaire pour obtenir la convergence est que le produit scalaire de d_k et du gradient de f au point x_k, noté g_k, soit négatif. On dit alors que d_k est une direction de descente. Pour la détermination du pas, appelée recherche linéaire, nous exigeons au moins que $f(x_k + \lambda_k d_k)$ soit strictement inférieure à $f(x_k)$. De plus, pour assurer une bonne vitesse de convergence, des conditions sur la dérivée directionnelle en x_{k+1} sont souvent demandées.

Pour la rechercher linéaire, plusieurs méthodes ont été proposées : citons la minimisation exacte dans la direction d_k, la recherche linéaire d'Armijo [2], la recherche linéaire de Goldstein [27], et la recherche linéaire de Wolfe [48]. Cette dernière sera avantageusement utilisée dans ce travail.

Pour la détermination de la direction, depuis plusieurs années, différentes méthodes ont été définies. La plus ancienne (Cauchy [9]) consiste à prendre pour direction l'opposé du gradient. Les méthodes dites de gradient conjugué définissent la direction d_k par $d_k = -g_k + \gamma_k d_{k-1}$, où γ_k est un réel strictement positif. Les diverses méthodes de gradient conjugué correspondent aux différentes manières de choisir le réel γ_k. Ces méthodes sont dues à Hestenes et Steifel [28], et depuis, elles ont fait l'objet de travaux de la part de nombreux auteurs dont, par exemple, Fletcher et Reeves [25], Polak et Ribière [41] et Mukai [38].

La méthode de Newton définit la direction par $d_k = G_k^{-1} g_k$. Cette définition est déduite du développement de Taylor d'ordre deux de f : $q_k(s) = f(x_k) + \langle s, g_k\rangle + \frac{1}{2}\langle s, G_k s\rangle$ où $s = x - x_k$. Le successeur de x_{k+1} est choisi comme zéro du gradient $\nabla q_k(s) = G_k s + g_k$, ce qui donne $x_{k+1} = x_k - G_k^{-1} g_k$. Le pas λ_k est identiquement égal à 1. Sous certaines hypothèses, (voir Kantorovich et Akilov [32]), cette méthode a une convergence locale d'ordre 2.

Les méthodes dites de métrique variable ou de quasi-Newton, dont fait partie le travail proposé ici, ont été introduites par Davidon [11]. Elles consistent à remplacer, dans la définition de la direction de Newton, l'inverse du hessien par une matrice H_k, obtenue de manière itérative. L'inverse d'une telle matrice, notée B_k, est une approximation du hessien. Le travail de Davidon a été repris par Fletcher et Powell [24] qui ont donné la formule bien connue de la DFP à savoir (H-formule) :

$$H_{k+1} = H_k - \frac{(H_k y_k)(H_k y_k)^t}{\langle H_k\, y_k, y_k\rangle} + \frac{s_k\, s_k^t}{\langle s_k, y_k\rangle}, \quad (0.1)$$

Une autre formule, appelée BFGS, abondamment utilisée, a été trouvée indépendamment par Broyden [7], Fletcher [22], Goldfarb [26] et Shanno [46]. Nous pouvons l'écrire sous la forme suivante donnée par Shanno [46] : (H-formule)

$$H_{k+1} = H_k - (1-r_k)\,\frac{(H_k y_k)(H_k y_k)^t}{\langle H_k\, y_k,\, y_k\rangle} + \frac{(s_k - r_k H_k y_k)(s_k r_k H_k y_k)^t}{\langle s_k - r_k H_k y_k,\, y_k\rangle} \tag{0.2}$$

$$\text{où } r_k = \frac{\langle s_k,\, y_k\rangle}{\langle s_k,\, y_k\rangle + \langle H_k y_k,\, y_k\rangle}$$

Ces deux matrices font parties de la classe des matrices de Broyden [6] Cette dernière famille groupe les matrices de Huang [30], qui vérifient la condition dite de quasi-Newton suivante :

$$H_{k+1}\, y_k \;= s_k$$

Comme $d_{k+1} = H_{k+1}\, g_{k+1}$ doit être une direction de descente, on exige des matrices H_{k+1} d'être symétriques strictement définies positives. Cette propriété est vérifiée par les matrices de Broyden pourvu que le produit scalaire $\langle s_k, y_k\rangle$ soit strictement positif (propriété qui dépend de la recherche linéaire utilisée) et que la matrice H_k soit symétrique strictement définie positive. Remarquons que les matrices inverses, B_{k+1}, peuvent être obtenues explicitement à l'aide de la formule de Sherman-Morrison (cf. [39]).

Pour d'autres études sur les matrices de métrique variable, et plus particulièrement des matrices de D.F.P et de BFGS, le lecteur peut se référer à Dennis [13], Dennis et Moré [14, 15], Dixon [17], Lenard [34], Powell [43], Ritter [45].

Dans tous les travaux cités, les résultats de convergence des méthodes de métriques variable sont obtenus sous des hypothèses de convexité et des conditions portant sur le hessien de la fonction économique.

Dans ce travail, des hypothèses plus faibles seront utilisées à savoir f de classe C^1 et f bornée inférieurement. Elles permettront d'obtenir la convergence.

Après avoir donné, au chapitre I, quelques remarques sur la recherche linéaire de Wolfe et la formule de BFGS, nous définissons, au chapitre II, à l'aide de paramètres réels, deux B-corrections ainsi que deux H-corrections de rang 1 d'une matrice symétrique strictement définie positive. Ces deux types de corrections, inspirées des formules de D.F.P et de BFGS, nous permettront, au chapitre III, de définir une famille de matrices de métrique variable B_k, dites de rang 2, uniformément majorées ainsi que leurs inverses. Ceci permettra dans le chapitre IV de démontrer la convergence de telles méthodes de métrique variable. Nous montrerons, sous des conditions plus fortes, que cette convergence est au moins R-linéaire (cf. [39], chapitre 9). Une caractérisation des paramètres assurant les résultats du chapitre III ainsi que la convergence des suites de matrices sont données au chapitre V. Au chapitre VI, des résultats numériques, fort encourageants, obtenus en utilisant l'une des matrices du chapitre III, sont données en comparaison avec ceux obtenus avec la formule de B.F.G.S. Nous avons utilisé le code M1QN1 de Modulopt, fourni par C. Lemaréchal. C'est une version équivalente du code VA13A de la bibliothèque de Harwell (M.J.P. Powell) conçu pour le B.F.G.S.

Ce travail a fait l'objet d'une note : MM. FIOROT et EL HALLABI [19] et d'une communication au NATO ARI on N.L. Opt. : FIOROT and EL HALLABI [20].

C H A P I T R E I

QUELQUES REMARQUES SUR LA

RECHERCHE LINEAIRE DE WOLFE

ET LES MATRICES DE

B F G S

I.1 - DEFINITION

Si x_k est un point de $\mathbb{R}^n$ et d_k un vecteur de $\mathbb{R}^n$, nous dirons que d_k est une direction de descente au point x_k si le produit scalaire des vecteurs g_k et d_k est strictement négatif.

I.2 - RELATION ENTRE LA RECHERCHE LINEAIRE ET LE PRODUIT SCALAIRE DES VECTEURS s_k, y_k

Nous rappelons l'algorithme de recherche linéaire de Wolfe [48] :
Etant donné d_k, une direction de descente au point x_k, c_1 et c_2 deux réels vérifiant :

$$0 < c_1 < c_2 < 1 \tag{I.2.1}$$

le pas λ_k sera choisi tel que :

$$f(x_k+\lambda_k d_k) \le f(x_k) + \lambda_k\, c_1 <g_k,\, d_k> \tag{I.2.2}$$

$$<g(x_k+\lambda_k d_k).d_k> \ge c_2 <g_k,\, d_k> \tag{I.2.3}$$

Dans un algorithme général de métrique variable, on utilise une recherche linéaire telle que le produit scalaire de s_k et y_k soit strictement positif. Ceci permet d'obtenir des matrices B_k, $k \in \mathbb{N}$, strictement définies positives. Ainsi Dennis et Moré [15] démontrent que si B_k est une matrice symétrique strictement définie positive et si B_{k+1} est une matrice de la famille de Broyden alors le produit scalaire $<s_k, y_k>$ est strictement positif si et seulement si la matrice B_{k+1} (symétrique) est strictement définie positive. Dans le même article, les auteurs signalent que si d_k est une direction de descente alors le produit scalaire $<s_k, y_k>$ est strictement positif si et seulement s'il existe β_k dans $]0, 1[$ tel que le produit scalaire $<s_k, g_{k+1}>$ soit supérieur à $\beta_k <s_k, g_k>$. Nous pouvons considérer d'après ce qui précède que, dans un algorithme de métrique variable utilisant une matrice de Broyden, le deuxième critère de choix de la recherche linéaire de Wolfe (I.2.3) a une certaine importance.

I.3 - DEFINITION (cf [39])

On dit que la suite $\{x_k,\ k \in \mathbb{N}\}$ converge R - linéairement vers x^* si :

$$||x_k - x^*|| \leq c\ \theta^k \qquad \text{(I.3.1)}$$

où θ est un réel de l'intervalle $]0, 1[$ et c une constante.

I.4 - DEFINITION (cf [39])

On dit que la suite $\{x_k,\ k \in \mathbb{N}\}$ converge superlinéairement vers x^* si :

$$\lim_{k \to +\infty} \frac{||x_{k+1} - x^*||}{||x_k - x^*||} = 0 \qquad \text{(I.4.1)}$$

I.5 - DEFINITION (cf [39])

On dit que la suite $\{x_k \ /\ k \in \mathbb{N}\}$ converge à l'ordre $p > 1$ vers x^* si :

$$\lim_{k \to +\infty} \frac{||x_{k+1} - x^*||}{||x_k - x^*||^p} = a \qquad \text{(I.5.1)}$$

où a est un réel non nul. Si $p = 2$ la convergence est dite quadratique.

Dans les différents algorithmes de métrique variable proposés, une des hypothèses, capitable pour la convergence de ces algorithmes, est la convexité globale de la fonction économique (voir [15], [34], [40], [43], [45]). Néanmoins, les directions de recherches linéaires sont les seules parties de $\mathbb{R}^n$ qui interviennent effectivement dans la progression de ces algorithmes.

Ceci se trouve illustré dans le travail de Dennis et Moré [15] où les auteurs ont établi une condition nécessaire et suffisante de convergence superlinéaire à savoir :

$$\lim_{k \to +\infty} ||[B_k - G(x^*)]u_k|| = 0 \qquad (I.5.2)$$

où $u_k = s_k/||s_k||$. C'est à dire que pour avoir une convergence superlinéaire, il suffit (et il faut) que B_k converge vers le hessien de f uniquement suivant la direction de recherche linéaire qui est alors, expliquent les auteurs, dans [15], asymptotiquement une direction de Newton. Dans le cas où f est quadratique strictement convexe, la matrice B_{n+1} est en général l'inverse du hessien de f. Nous pourrions donc nous attendre à ce que, seules, les propriétés de la fonction f, restreintes à la direction de recherche linéaire, aient une importance pour la convergence globale. En fait cette convergence a jusqu'ici été obtenue, comme nous l'avons déjà dit, sous l'hypothèse de convexité de f.

Les lemmes I.6 et I.8 ont été supprimés

Suite page 11

I.7 - PROPOSITION

Si A est une matrice (n, n) symétrique définie positive alors pour tout x de $\mathbb{R}^n$ nous avons :

$$||Ax||^2 \leq ||A|| <x, Ax> \qquad (I.7.1)$$

Preuve :

L'opérateur $||A||I_n - A$ est défini positif d'après la définition de la norme. Il commute avec A. Nous en déduisons, d'après ([33], 1f, 1.12) que l'opérateur produit $A(||A||I-A)$ est défini positif, i.e. pour tout x dans $\mathbb{R}^n$ nous avons $<x, A(||A||I-A)x>$ positif ou nul, qui, compte tenu de la symétrie de A, donne le résultat.

Dans [15] les auteurs démontrent que la condition nécessaire et suffisante de convergence superlinéaire (I.5.2) est réalisée pour la DFP et la BFGS sous les hypothèses suivantes : f de classe C^2, $g(x^*) = 0$, $G(x^*)$ symétrique strictement définis positif et lipschitzien dans un voisinage convexe compact V de x^*, $\sum_{k=o}^{+\infty} ||x_k - x^*|| < +\infty$. Ils obtiennent comme résultat intermédiaire que les matrices B_k ainsi que leurs inverses sont uniformément bornées. Pour la BFGS, Powell, dans [43], a montré la convergence de la série ci-dessus, en ajoutant l'hypothèse que f est uniformément strictement convexe dans V, i.e. il existe deux constantes m_1 et m_2 strictement positives telles que :

$$m_1 ||y||^2 \leq <y, G(x) y> \leq m_2||y||^2$$

pour tout x dans V et tout y dans $\mathbb{R}^n$.

Sous ces hypothèses, nous donnons le lemme I.9 suivant, qui, bien qu'il n'ait jamais été publié est connu de certains auteurs (par exemple [47]).

I.9 - LEMME

La suite matricielle $\{\Delta B_k = B_{k+1} - B_k\}_{k\in\mathbb{N}}$ (où B_k est définie par la formule de BFGS) converge vers zéro.

Preuve :

De la propriété (I.5.2), de la continuité du hessien et de la convergence de x_k vers x^* nous déduisons que

$$\lim_{k\to+\infty} ||[B_k - G(z_k)]u_k|| = 0 \tag{I.9.1}$$

où z_k est donné par (I.5.3). Ecrivons la relation (I.9.1) sous la forme :

$$\lim_{k\to+\infty} ||B_k u_k - \frac{y_k}{||s_k||}|| = 0 \tag{I.9.2}$$

Posons $u'_k = B_k u_k$ et $v'_k = y_k/||s_k||$ (I.9.3)

1) Il existe deux constantes m' et M' strictement positives, telles que $||u'_k|| \leq M'$ et $m' \leq \langle u_k, u'_k\rangle \leq M'$. Ceci résulte du fait que les matrices B_k, $k \in \mathbb{N}$, et leurs inverses sont uniformément majorées.

2) Il existe m et M deux constantes strictement positives telles que

$$||v'_k|| \leq M \quad \text{et} \quad m \leq \langle u_k, v'_k\rangle \leq M.$$

Ceci est conséquence de (I.9.3) et de :

(i) La suite $\{x_k, k \in \mathbb{N}\}$ converge vers x^*, minimum local strict de f,

(ii) G(x) est continue sur V voisinage convexe compact de x^*.

3) La suite $w_k = \langle u'_k, u_k\rangle^{\frac{1}{2}} - \langle v'_k, u_k\rangle^{\frac{1}{2}}$ converge vers zéro.

Nous avons :

$$\langle u'_k, u_k\rangle^{\frac{1}{2}} - \langle v'_k, u_k\rangle^{\frac{1}{2}} = \frac{\langle u'_k - v'_k, u_k\rangle}{\langle u'_k, u_k\rangle^{\frac{1}{2}} + \langle v'_k, u_k\rangle^{\frac{1}{2}}}$$

qui donne

$$|w_k| \leq \frac{||u'_k - v'_k||}{<u'_k, u_k>^{\frac{1}{2}} + <v'_k, u_k>^{\frac{1}{2}}} \tag{I.9.4}$$

qui d'après 1) et 2) implique que $|w_k|$ est inférieur ou égal à $\dfrac{||u'_k - v'_k||}{\sqrt{m'} + \sqrt{m}}$

qui converge vers zéro d'après (I.9.2) et (I.9.3).

4) La suite $\dfrac{u'_k}{<u'_k, u_k>^{\frac{1}{2}}} - \dfrac{v'_k}{<v'_k, u_k>^{\frac{1}{2}}}$ converge vers zéro :

On a :

$$\frac{u'_k}{\sqrt{<u'_k, u_k>}} - \frac{v'_k}{\sqrt{<v'_k, u_k>}} = \frac{\sqrt{<v'_k, u_k>}\, u'_k - \sqrt{<u'_k, u_k>}\, v'_k}{\sqrt{<u'_k, u_k> <v'_k, u_k>}}$$

$$= \frac{\sqrt{<v'_k, u_k>}\,(u'_k - v'_k) + (\sqrt{<u'_k, u_k>} - \sqrt{<v'_k, u_k>})\, v'_k}{\sqrt{<u'_k, u_k> <v'_k, u_k>}}$$

et en passant aux normes, nous obtenons :

$$\left|\left|\frac{u'_k}{\sqrt{<u'_k, u_k>}} - \frac{v'_k}{\sqrt{<v'_k, u_k>}}\right|\right| \leq \frac{\sqrt{M}\,||u'_k - v'_k|| + M|\sqrt{<u'_k, u_k>} - \sqrt{<v'_k, u_k>}}{\sqrt{mm'}}$$

et nous concluons d'après (I.9.2), (I.9.3) et 3).

5) La suite ΔB_k converge vers zéro.

Nous avons

$$\Delta B_k = \frac{y_k\, y_k^t}{<s_k, y_k>} - \frac{(B_k s_k)(B_k s_k)^t}{<B_k s_k, s_k>} \tag{I.9.5}$$

Dans le deuxième membre de (I.9.5) ci-dessus, nous divisons dans chaque terme le numérateur et le dénominateur par $||s_k||^2$, ce qui donne d'après (I.9.3) :

$$\Delta B_k = \left(\frac{v'_k}{\sqrt{<v'_k, u_k>}}\right)\left(\frac{v'_k}{\sqrt{<v'_k, u_k>}}\right)^t - \left(\frac{u'_k}{\sqrt{<u'_k, u_k>}}\right)\left(\frac{u'_k}{\sqrt{<u'_k, u_k>}}\right)^t$$

Posons $\bar{u}_k = \dfrac{u'_k}{\sqrt{\langle u'_k, u_k\rangle}}$ et $\bar{v}_k = \dfrac{v'_k}{\sqrt{\langle v'_k, u_k\rangle}}$, (I.9.6)

ce qui donne

$\Delta B_k = \bar{v}_k \bar{v}_k^t - \bar{u}_k \bar{u}_k^t$, or nous avons :

$$\bar{v}_k \bar{v}_k^t - \bar{u}_k \bar{u}_k^t = (-\bar{u}_k + \bar{v}_k)\bar{u}_k^t + \bar{v}_k(-\bar{u}_k + \bar{v}_k)^t,$$

d'où $||\Delta B_k|| \leq ||\bar{u}_k - \bar{v}_k|| \; (||\bar{u}_k|| + ||\bar{v}_k||)$ qui converge vers zéro d'après (I.9.6), 1), 2) et 4).

C H A P I T R E   II

B - CORRECTION DE RANG 1

H - CORRECTION DE RANG 1

INTRODUCTION

Dans ce chapitre, nous nous proposons de corriger une matrice symétrique strictement définie positive par une matrice anti-scalaire, i.e. de la forme uu^t où u est un vecteur de $\mathbb{R}^n$. Pour simplifier son utilisation ultérieure, cette correction sera traitée sous deux formes différentes : l'une dite B-correction, l'autre dite H-correction. Ces deux appellations sont à rapprocher des appellations de B-formule et H-formule de métrique variable. Des conditions pour que la matrice résultant d'une telle correction soit strictement définie positive seront données ainsi que des relations entre les normes de la matrice et de sa corrigée. Ces relations seront utilisées par la suite pour les formules de rang 2.

II.1 - DEFINITION

Si la direction d_k est fournie comme solution du système linéaire $B_k\, d_k = -g_k$ nous dirons qu'elle est obtenue par une B-formule.

II.2 - DEFINITION

Si la direction d_k est définie par $d_k = -H_k g_k$, nous dirons qu'elle est obtenue par une H-formule.

II.3 - LA PREMIERE B-CORRECTION

II.3.0 - Hypothèses :

Soient :

- B une matrice symétrique, strictement définie positive, d'inverse H,
- s un vecteur de $\mathbb{R}^n$, non identiquement nul,

- α un réel strictement positif,
- β un réel positif ou nul,
- ε un entier prenant les valeurs +1 ou -1.

Sous les hypothèses ci-dessus, considérons la matrice symétrique suivante :

$$\tilde{B} = \alpha B + \varepsilon\beta \frac{(Bs)(Bs)^t}{<Bs,s>} \qquad \text{(II.3.0)}$$

Rappelons la B-formule de la BFGS :

$$B^+ = B - \frac{(Bs)(Bs)^t}{<Bs,s>} + \frac{yy^t}{<s,y>}$$

La matrice $\tilde{B}$ peut-être considérée comme une perturbation d'une formule incomplète de B^+ ci-dessus.

II.3.1 - ***Proposition***

B étant strictement définie positive, dans les cas suivants :

(i) $\varepsilon = 1$, $\alpha > 0$ et $\beta \geq 0$ (II.3.1.1)

(ii) $\varepsilon = -1$, $0 \leq \beta < \alpha$ (II.3.1.2)

$\tilde{B}$ est strictement définie positive.

Preuve :

Calculons $<z, \tilde{B}z>$ pour z un vecteur quelconque de $\mathbb{R}^n$.

Nous avons $<z, \tilde{B}z> = \alpha<z, Bz> + \varepsilon\beta \frac{<z,Bs>^2}{<s,Bs>}$

(i) si $\varepsilon = 1$, pour α strictement positif et β positif ou nul $\tilde{B}$ est strictement définie positive.

(ii) si $\varepsilon = -1$, comme B_k est symétrique strictement définie positive $<z, Bz>^{\frac{1}{2}}$ est une norme de z, noté $||z||_B$ et $<z, s>_B = <z, Bs>$ est le produit scalaire de z et s relativement à cette norme.

Nous avons :

$$\langle z, \tilde{B}z\rangle = \{\alpha||z||_B^2\ ||s||_B^2 - \beta\langle s, z\rangle_B^2\} \ / ||s||_B^2$$

De l'inégalité de Cauchy-Schwartz $\langle z, s\rangle_B^2 \leq ||z||_B^2\ ||s||_B^2$, nous déduisons que si β est strictement inférieure à α, alors $\langle z, \tilde{B}z\rangle$ est strictement positif pour tout z non nul de $\mathbb{R}^n$.

II.3.2 - Proposition

Sous les hypothèses précédentes nous avons les relations suivantes entre les normes de B et de $\tilde{B}$:

(i) si $\varepsilon = 1$, $\alpha > 0$ et $\beta \geq 0$ $\quad ||\tilde{B}|| \leq (\alpha+\beta)\ ||B||$ (II.3.2.1)

(ii) si $\varepsilon = -1$, $0 \leq \beta < \alpha$ $\quad ||\tilde{B}|| \leq \alpha||B||$ (II.3.2.2)

Preuve :

Soit $B^{\frac{1}{2}}$ la racine carrée de la matrice B, elle est symétrique strictement définie positive. Comme $\langle Bs, s\rangle = ||B^{\frac{1}{2}}s||^2$, nous obtenons

$$\tilde{B} = \alpha(B^{\frac{1}{2}})^2 + \varepsilon\ \beta\ B^{\frac{1}{2}} \frac{(B^{\frac{1}{2}}s)(B^{\frac{1}{2}}s)^t}{||B^{\frac{1}{2}}s||^2} B^{\frac{1}{2}} \tag{II.3.2.3}$$

Posons
$$u = \left(\frac{\beta}{\alpha}\right)^{\frac{1}{2}} \frac{B^{\frac{1}{2}}s}{||B^{\frac{1}{2}}s||} \tag{II.3.2.4}$$

Nous obtenons $\tilde{B} = \alpha\ B^{\frac{1}{2}}\ [I + \varepsilon uu^t]B^{\frac{1}{2}}$, d'où $||\tilde{B}|| \leq \alpha\ ||B^{\frac{1}{2}}||^2\ ||I + \varepsilon uu^t||$.

Nous savons que $||B^{\frac{1}{2}}||^2 = ||(B^{\frac{1}{2}})^2|| = ||B||$.

Par suite nous avons

$$||\tilde{B}|| \leq \alpha\ ||B||\ ||I + \varepsilon uu^t|| \tag{II.3.2.5}$$

1) Si $\varepsilon = 1$, pour x de norme 1 nous avons

$$||(I + uu^t)x||^2 = ||x||^2 + \langle u, x\rangle^2 \, ||u||^2 + 2 \langle u, x\rangle^2$$

$$\leq 1 + ||u||^4 + 2||u||^2 = (1 + ||u||^2)^2$$

Ce qui donne $||(I + uu^t) \, x|| \leq 1 + ||u||^2$ pour tout x de norme 1 de $\mathbb{R}^n$;

l'égalité est atteinte pour $x = \frac{u}{||u||}$, d'où

$$||I + uu^t|| = 1 + ||u||^2 \qquad \text{(II.3.2.6)}$$

D'après (II.3.2.4) nous avons alors

$$||I + uu^t|| = 1 + \frac{\beta}{\alpha} = \frac{\alpha+\beta}{\alpha}$$, ce qui, avec (II.3.2.5) permet de

conclure.

2) Si $\varepsilon = -1$, pour x de norme 1, nous avons

$$||(I - uu^t) \, x||^2 = ||x||^2 + ||(uu^t) \, x||^2 - 2 \langle x, uu^t x\rangle$$

$$= 1 + \langle u, x\rangle^2 \, ||u||^2 - 2\langle u, x\rangle^2$$

$$= 1 + \langle u, x\rangle^2 \, (||u||^2 - 2).$$

Or d'après (II.3.2.4), $||u||^2 = \frac{\beta}{\alpha}$ qui d'après (II.3.2.2) est inférieur à 2, par suite $||(I - uu^t)x|| \leq 1$. La valeur 1 est atteinte pour tout vecteur x de norme 1, perpendiculaire à u.

Par suite

$$||I - uu^t|| = 1 \qquad \text{(II.3.2.7)}$$

Ce qui compte tenu de (II.3.2.5) permet de conclure. En fait nous pouvons prendre $\alpha = \beta (\alpha \neq 0)$ dans (ii).

II.3.3 - Formule de Sherman-Morrison (cf. 39)

Si A est une matrice (n, n) régulière, v et w deux vecteurs de $\mathbb{R}^n$, alors la matrice $A + vw^t$ est régulière si et seulement si $1 + \langle w, A^{-1} v\rangle$

est non nul, et nous avons :

$$[A + vw^t]^{-1} = A^{-1} - \frac{A^{-1}\, vw^t\, A^{-1}}{1+\langle w,\, A^{-1}v\rangle} \qquad \text{(II.3.3.1)}$$

D'après la proposition II.3.1, $\tilde{B}$ est régulière. Notons $\tilde{H}$ son inverse. En appliquant la formule (II.3.3.1) ci-dessus à la matrice $\tilde{B}$ donnée en (II.3.1) nous obtenons

$$\tilde{H} = \frac{1}{\alpha} H - \varepsilon \frac{\beta}{\alpha(\alpha+\varepsilon\beta)} \frac{ss^t}{\langle Bs,\, s\rangle} \qquad \text{(II.3.3.2)}$$

II.3.4 - Proposition

Sous les hypothèses précédentes, nous avons les relations suivantes entre les normes de H et de $\tilde{H}$ respectivement inverses de B et de $\tilde{B}$.

(i) si $\varepsilon = 1$, $\alpha > 0$ et $\beta \geq 0$ $\quad ||\tilde{H}|| \leq \frac{1}{\alpha} ||H||$ (II.3.4.1)

(ii) si $\varepsilon = -1$, $0 \leq \beta < \alpha$ $\quad ||\tilde{H}|| \leq \frac{1}{\alpha-\beta} ||H||$ (II.3.4.2)

<u>Preuve</u> :

Soit $H^{\frac{1}{2}}$ la matrice racine carrée de H. Elle est symétrique, par suite $||H^{\frac{1}{2}}||^2 = ||H||$. La matrice $\tilde{H}$ peut s'écrire sous la forme suivante

$$\tilde{H} = \frac{1}{\alpha} H^{1/2} \left(I - \frac{\beta (B^{\frac{1}{2}}s)(B^{\frac{1}{2}}s)^t}{(\alpha+\varepsilon\beta)||B^{\frac{1}{2}}s||^2} \right) H^{\frac{1}{2}}$$

qui donne

$$||\tilde{H}|| \leq \frac{1}{\alpha} ||H|| \left|\left| I - \varepsilon \frac{\beta}{\alpha+\varepsilon\beta} \frac{(B^{\frac{1}{2}}s)(B^{\frac{1}{2}}s)^t}{||B^{\frac{1}{2}}s||^2} \right|\right| \qquad \text{(II.3.4.3)}$$

Pour les deux valeurs de ε, le réel $\alpha+\varepsilon\beta$ est strictement positif.

Si nous posons

$$u = \left(\frac{\beta}{\alpha+\varepsilon\beta}\right)^{\frac{1}{2}} \frac{B^{\frac{1}{2}}s}{||B^{\frac{1}{2}}s||}$$

nous avons

$$||u||^2 = \frac{\beta}{\alpha+\varepsilon\beta} \qquad \text{(II.3.4.4)}$$

L'inégalité (II.3.4.3) s'écrit alors

$$||\tilde{H}|| \leq \frac{1}{\alpha} ||H|| \; ||I - \varepsilon uu^t|| \qquad \text{(II.3.4.5)}$$

1) si $\varepsilon = 1$, (II.3.4.5) devient $||\tilde{H}|| \leq \frac{1}{\alpha} ||H|| \; ||I - uu^t||$

et nous pouvons conclure d'après (II.3.2.7).

2) Si $\varepsilon = -1$, (II.3.4.5) devient $||\tilde{H}|| \leq \frac{1}{\alpha} ||H|| \; ||I + uu^t||$ et

d'après (II.3.2.6) nous avons alors

$||\tilde{H}|| \leq \frac{1}{\alpha} ||H|| \; (1 + ||u||^2) = \frac{1}{\alpha} ||H|| \; (1 + \frac{\beta}{\alpha-\beta})$, ce qui donne

$||\tilde{H}|| \leq \frac{1}{\alpha-\beta} ||H||$.

II.4 - LA DEUXIEME B-CORRECTION

II.4.0 - Hypothèses

Soient :

i - B une matrice (n, n) symétrique strictement définie positive, d'inverse H,

ii - s, y deux vecteurs de $\mathbb{R}^n$ tels que leur produit scalaire $<s, y>$ soit strictement positif,

iii - α un réel strictement positif

iv - β un réel positif ou nul

v - ε un entier prenant les valeurs +1 ou -1.

Sous les hypothèses ci-dessus, considérons la matrice symétrique suivante

$$\underset{\sim}{B} = \alpha B + \varepsilon\beta \frac{(Bs)\ (Bs)^t}{<Bs,s> + <s,y>} \tag{II.4.0}$$

En remplaçant dans la H-formule de la BFGS, donnée par Shanno [46], (H, y, s) par (B, s, y) nous obtenons la B-formule de la D.F.P, à savoir :

$$B^+ = B - (1-r)\frac{(Bs)(Bs)^t}{<Bs,s>} + \frac{(y-rBs)(y-rBs)^t}{<y-rBs,\ y>}$$

où $$r = \frac{<s,\ y>}{<s,\ y> + <Bs,\ s>}$$

Elle peut s'écrire encore :

$$B^+ = B - \frac{(Bs)(Bs)^t}{<Bs,\ s> + <s,\ y>} + \frac{(y-rBs)(y-rBs)^t}{<y-rBs,\ y>}.$$

Ainsi la matrice $\underset{\sim}{B}$ définie par (II.4.0) peut être considérée comme une perturbation d'une formule incomplète de la matrice B^+ ci-dessus.

Donnons quelques résultats dont les démonstrations seront déduites du paragraphe ultérieur II.5.

II.4.1 - <u>*Proposition*</u>

B étant une matrice symétrique strictement définie positive, dans les deux cas suivants

(i) $\varepsilon = 1,\ \alpha > 0\quad \beta \geq 0$ (II.4.1.1.)

(ii) $\varepsilon = -1,\ 0 \leq \beta \leq \alpha\ (\alpha \neq 0)$ (II.4.1.2)

$\underset{\sim}{B}$ est strictement définie positive.

II.4.2 - <u>*Proposition*</u>

Sous les hypothèses précédentes, nous avons les relations suivantes entre les normes de B et de $\underset{\sim}{B}$

(i) si $\varepsilon = 1$, $\alpha > 0$, $\beta \geq 0$ $\quad ||\underset{\sim}{B}|| \leq (\alpha+\beta)\, ||B||$ (II.4.2.1)

(ii) si $\varepsilon = -1$, $0 \leq \beta \leq \alpha$ $(\alpha \neq 0)$, $||\underset{\sim}{B}|| \leq \alpha\, ||B||$ (II.4.2.2)

Soit $\underset{\sim}{H}$ la matrice inverse de $\underset{\sim}{B}$. D'après la formule de Sherman-Morrison (II.3.3.1) $\underset{\sim}{H}$ est donné par

$$\underset{\sim}{H} = \frac{1}{\alpha} H - \varepsilon \frac{\beta}{\alpha} \frac{ss^t}{(\alpha+\varepsilon\beta)\, \langle Bs, s\rangle + \alpha \langle s, y\rangle} \qquad \text{(II.4.2.3)}$$

II.4.3 - Proposition

Sous les hypothèses précédentes, nous avons les relations suivantes entre les normes de H et $\underset{\sim}{H}$:

(i) si $\varepsilon = 1$, $\alpha < 0$ et $\beta \geq 0$ $\quad ||\underset{\sim}{H}|| \leq \frac{1}{\alpha} ||H||$ (II.4.3.1)

(ii) si $\varepsilon = -1$, $0 \leq \beta < \alpha$, $||\underset{\sim}{H}|| \leq \frac{1}{\alpha-\beta}\, ||H||$ (II.4.3.2)

II.5 - LA PREMIERE H-CORRECTION

II.5.0 - Hypothèses

Soient :

i - H' une matrice (n, n) symétrique, strictement définie positive, d'inverse B',

ii - y un vecteur de $\mathbb{R}^n$ non identiquement nul,

iii - δ un réel strictement positif,

iv - γ un réel positif ou nul,

v - ε' un entier prenant les valeurs +1 ou -1.

Sous les hypothèses ci-dessus, considérons la matrice symétrique suivante :

$$\hat{H}' = \delta H' + \varepsilon' \gamma \frac{(H'y)(H'y)^t}{\langle H'y, y\rangle} \qquad \text{(II.5.0)}$$

Elle peut être considérée comme une perturbation d'une formule incomplète de la H-formule de la DFP rappelée en introduction (0.1).

II.5.1 - Proposition

H' étant strictement définie positive, dans les cas suivants :

(i) $\varepsilon' = 1, \quad \delta > 0$ et $\gamma \geq 0$ (II.5.1.1)

(ii) $\varepsilon' = -1, \; 0 \leq \gamma < \delta$ (II.5.1.2)

$\hat{H}'$ est strictement définie positive.

Preuve : Elle est identique à celle de la proposition II.3.1 ; il suffit de remplacer dans cette dernière (ε, α, β, s, B) par (ε', δ, γ, y, H).

II.5.2 - Proposition

Sous les hypothèses précédentes, nous avons les relations suivantes entre les normes de H' et $\hat{H}'$:

(i) si $\varepsilon' = 1$, $\delta > 0$ et $\gamma \geq 0$ $\quad ||\hat{H}'|| \leq (\delta+\gamma)\, ||H'||$ (II.5.2.1)

(ii) si $\varepsilon' = -1$, $0 \leq \gamma < \delta$ $\quad ||\hat{H}'|| \leq \alpha\, ||H'||$ (II.5.2.2)

Preuve :

Elle est la même que pour la proposition (II.3.2).

En utilisant la formule de Sherman-Morrison (II.3.3.1) nous obtenons $\hat{B}'$ la matrice inverse de $\hat{H}'$ suivante :

$$\hat{B}' = \frac{1}{\delta} B' - \varepsilon \frac{\gamma}{\delta(\delta+\varepsilon'\gamma)} \frac{yy^t}{\langle Hy,y\rangle}$$

II.5.3 - Proposition

Sous les hypothèses précédentes, nous avons les relations suivantes entre les normes de B' et $\hat{B}'$:

(i) $\varepsilon' = 1 \quad \delta > 0$ et $\gamma \geq 0$ $\qquad ||\hat{B}'|| \leq \frac{1}{\delta} ||B'||$ (II.5.3.1)

(ii) $\varepsilon' = -1 \quad 0 \leq \gamma < \delta$ $\qquad ||\hat{B}'|| \leq \frac{1}{\delta-\gamma} ||B'||$ (II.5.3.2)

Preuve

Elle est identique à la démonstration de la proposition (II.3.4) : il suffit de remplacer dans cette dernière $(\varepsilon, \alpha, \beta, B, s)$ par $(\varepsilon', \delta, \gamma, H, y)$.

II.6 - LA DEUXIEME H-CORRECTION

II.6.0 - *Hypothèses*

Faisons les hypothèses suivantes :

i - H' est une matrice (n, n) symétrique strictement définie positive d'inverse B',

ii - s, y sont deux vecteurs de $\mathbb{R}^n$ dont le produit scalaire est strictement positif,

iii - δ est un réel strictement positif,

iv - γ est un réel positif ou nul,

v - ε' est un entier prenant les valeurs +1 ou -1.

Sous les hypothèses précédentes II.6.0, considérons la matrice symétrique suivante :

$$\bar{H}' = \delta H' + \varepsilon'\gamma \frac{(H'y)(H'y)^t}{<H'y,y> + <s,y>} \tag{II.6.0}$$

Elle peut être considérée comme étant une perturbation de la H-formule de la BFGS donnée en (0.2). En effet (0.2) peut s'écrire comme suit

$$H^+ = H - \frac{(Hy)(Hy)^t}{<Hy,y> + <s, y>} + \frac{(s-rHy)(s-rHy)^t}{(s-rHy, y)}$$

où $\quad r = \dfrac{<s, y>}{<s, y> + <Hy, y>}$

où l'indice k a été omis dans le membre de droite et l'indice k+1 a été remplacé par le signe + dans le membre de gauche.

II.6.1 - Proposition

H' étant strictement défini positive, dans les cas suivants :

(i) $\varepsilon' = 1, \quad \delta > 0$ et $\gamma \geq 0$ (II.6.1.1)

(ii) $\varepsilon' = -1, \; 0 \leq \gamma \leq \delta \; (\delta \neq 0)$ (II.6.1.2)

$\bar{H}'$ est strictement définie positive.

Preuve :

Pour z un vecteur de $\mathbb{R}^n$, calculons $<z, \bar{H}'z>$; nous avons

$$<z, \bar{H}'z> = \delta <z, H'z> + \varepsilon'\gamma \frac{<H'y, z>^2}{<H'y,y> + <s, y>} \tag{II.6.1.3}$$

Comme H' est symétrique, strictement définie positive, la fonction $z \to <z, H'z>^{\frac{1}{2}}$ est une norme, notée $||z||_{H'}$, $<z, y>_{H'}$ est le produit scalaire des vecteurs z et y relativement à cette norme. L'égalité (II.6.1.3) s'écrit donc

$$<z, \bar{H}'z> = [\delta \, ||z||^2_{H'} \, ||y||^2_{H'} + \varepsilon' \gamma <z, y>^2_{H'} + \delta \, ||z||^2_{H'} <s, y>] \, / \, D \tag{II.6.1.4}$$

où $D = <s, y> + ||y||^2_{H'}$ qui est strictement positif d'après l'hypothèse II.6.0 (ii).

1) si $\varepsilon' = 1$, il est évident que pour δ strictement positif et γ positif ou nul $\bar{H}'$ est strictement positif puisque H' l'est.

2) Si $\varepsilon' = -1$, l'égalité (II.6.1.4) s'écrit :

$$<z, \bar{H}'z> = [\delta \, ||z||^2_{H'} \, ||y||^2_{H'} - \gamma <z, y>^2_{H'} + \delta \, ||z||^2_{H'} <s, y>] \, / \, D$$

De l'hypothèse (II.6.1. ii) et de l'inégalité de Cauchy-Schwarz nous déduisons que si γ, réel positif, est inférieur ou égal à δ, réel strictement positif, alors $<z, \bar{H}'z>$ est strictement positif pour tout z non nul de $\mathbb{R}^n$.

II.6.2 - Proposition

Sous les hypothèses précédentes nous avons les relations suivantes entre les normes de H' et $\bar{H}'$:

(i) si $\varepsilon' = 1$, $\delta > 0$ et $\gamma \geq 0$ $\quad ||\bar{H}'|| \leq (\gamma+\delta)\, ||H'||$ (II.6.2.1)

(ii) si $\varepsilon' = -1$, $0 \leq \gamma \leq \delta$, $\quad ||\bar{H}'|| \leq \delta\, ||H'||$ (II.6.2.2)

Preuve :

Soit T la racine carrée de H', elle est symétrique strictement définie positive. L'égalité (II.6.0) s'écrit alors :

$$\bar{H}' = \delta\, T \left[I + \varepsilon' \frac{\gamma}{\delta} \frac{(Ty)(Ty)^t}{||Ty||^2 + <s, y>} \right] T \tag{II.6.2.3}$$

Posons $\quad v = \left(\frac{\gamma}{\alpha}\right)^{\frac{1}{2}} \frac{Ty}{D}$ (II.6.2.4)

où $\quad D = (||Ty||^2 + <s, y>)^{\frac{1}{2}}$, nous avons alors :

$$\bar{H}' = \delta\, T\, [I + \varepsilon' vv^t] T.$$

En passant aux normes et en tenant compte de $||T||^2 = ||H'||$ nous obtenons alors $\quad ||\bar{H}'|| \leq \delta\, ||H'||\, ||I + \varepsilon vv^t||$ (II.6.2.5)

1) Si $\varepsilon' = 1$, l'inégalité (II.6.3.5) devient $||\bar{H}'|| \leq \delta\, ||H'||\, ||I + vv^t||$, ce qui donne d'après (II.3.2.6) $||\bar{H}'|| \leq \delta\, ||H'||\, (1 + ||v||^2)$. Or nous avons (II.6.3.4) qui donne

$$||v||^2 = \frac{\gamma}{\delta} \frac{||Ty||^2}{||Ty||^2 + <s, y>} < \frac{\gamma}{\delta}, \text{ d'où}$$

$$||\bar{H}'|| < \delta\, (1 + \frac{\gamma}{\delta})\, ||H'|| = (\delta + \gamma)\, ||H'||$$

2) si $\varepsilon' = -1$, l'inégalité (II.6.3.5) s'écrit

$$||\bar{H}'|| \leq \delta \, ||H'|| \, ||I - vv^t||$$

qui permet de conclure d'après (II.3.2.7) car $||v||^2$ est inférieur à 2.

D'après la proposition II.6.1 $\bar{H}'$ est inversible. Son inverse $\bar{B}'$ est donné par la formule de Sherman-Morrison (II.3.3.1). Nous avons :

$$\bar{B}' = \frac{1}{\delta} B' - \varepsilon' \frac{\gamma}{\delta} \frac{yy^t}{(\delta + \varepsilon'\gamma) <H'y, y> + \delta <s, y>} \qquad (II.6.2.6)$$

Comme $(B')^{\frac{1}{2}} = T^{-1}$, nous pouvons écrire :

$$\bar{B}' = \frac{1}{\delta} T^{-1} \left[I - \varepsilon' \frac{\gamma(Ty)(Ty)^t}{(\delta + \varepsilon'\gamma) \, ||Ty||^2 + \delta <s, y>} \right] T^{-1} \qquad (II.6.2.7)$$

II.6.3 - Proposition

Sous les hypothèses précédentes, nous avons les relations suivantes entre les normes de B' et $\bar{B}'$:

i) si $\varepsilon' = 1$, $\delta > 0$ et $\gamma \geq 0$ $\quad ||\bar{B}'|| \leq \frac{1}{\delta} ||B'||$ (II.6.3.1)

ii) si $\varepsilon' = -1$, $0 \leq \gamma < \delta$ $\quad ||\bar{B}'|| \leq \frac{1}{\delta - \gamma} ||B'||$ (II.6.3.2)

Preuve :

1) Si $\varepsilon' = 1$,

posons $$v = \left(\frac{\gamma}{(\delta+\gamma) \, ||Ty||^2 + \delta <s, y>} \right)^{\frac{1}{2}} Ty$$

l'égalité (II.6.2.7) s'écrit $\bar{B}' = \frac{1}{\delta} T^{-1}[I - vv^t] \, T^{-1}$, ce qui d'après (II.3.2.7) permet de conclure.

2) si $\varepsilon' = -1$,

posons $$v = \left(\frac{\gamma}{(\delta-\gamma) \, ||Ty||^2 + \delta <s, y>} \right)^{\frac{1}{2}} Ty,$$ l'égalité (II.6.2.7) s'écrit alors $\bar{B}' = \frac{1}{\delta} T^{-1} [I + vv^t] T^{-1}$, qui d'après (II.3.2.6) donne

$$||\bar{B}'|| \leq \frac{1}{\delta} ||T^{-1}||^2 (1 + ||v||^2) \qquad \text{(II.6.3.3)}$$

or nous avons $||T^{-1}||^2 = ||B'||$ et

$$||v||^2 = \frac{\gamma ||Ty||^2}{(\delta-\gamma) ||Ty||^2 + \delta <s, y>} < \frac{\gamma}{\delta-\gamma}$$

ce qui, compte tenu de (II.6.3.3), permet de conclure.

Remarque

Les quatre corrections précédentes sont d'un même type : les matrices corrigées sont toutes de la forme $A \pm uu^t$ ou A est une matrice (n, n) et u un vecteur colonne de $\mathbb{R}^n$. Elles ont été introduites sous différentes formes pour être utilisées comme telles dans la suite.

CHAPITRE III

CONSTRUCTION DES MATRICES DE METRIQUE VARIABLE PARAMETREES : CORRECTION DE RANG DEUX

INTRODUCTION

Dans ce chapitre, partant d'une matrice (n, n) symétrique strictement définie positive A, nous nous proposons de construire une matrice A' de la forme $A' = \mu A + \varepsilon uu^t - \varepsilon' vv^t$ où μ est un réel strictement positif, ε et ε' sont deux entiers prenant indépendamment chacun les valeurs +1 ou -1, u et v sont des vecteurs colonnes de $\mathbb{R}^n$. La matrice A' sera dite déduite de A par une correction de rang 2 (si $\varepsilon = \varepsilon'$, u sera différent de v). Nous montrons que les formules de rang 2, dépendant de paramètres, son uniformément majorées, ainsi que leurs inverses pour des choix judicieux de tels paramètres.

III.1 - HYPOTHESES

Soient :

i - B une matrice (n, n) symétrique strictement définie positive, d'inverse H,

ii - s et y deux vecteurs de $\mathbb{R}^n$ dont le produit scalaire $\langle s, y\rangle$ est strictement positif,

iii - α et δ deux réels strictement positifs,

iv - β et γ deux réels positifs,

v - ε et ε' deux entiers tels que $\varepsilon = \pm\, \varepsilon' = \pm\, 1$.

III.2 - CONSTRUCTION DE B-FORMULES DE RANG DEUX

Cette construction se fait en plusieurs étapes :

1) A partir des données $(B, \alpha, \beta, s, \varepsilon)$ Nous définissons $\tilde{B}$ (respectivement $\underset{\sim}{B}$) par la B-correction donnée en (II.3.0) (respectivement (II.4.0)) dans les deux cas étudiés dans le chapitre précédent, i.e. pour

i) $\varepsilon = 1,\ \alpha > 0,\ \beta \geq 0$ (III.2.1)

ii) $\varepsilon = -1,\ 0 \leq \beta < \alpha$ (III.2.2)

2) A l'aide de la formule de Sherman-Morrison (II.3.3.1) nous calculons $\tilde{H}$ (respectivement $\underset{\sim}{H}$) matrice inverse de $\tilde{B}$ (respectivement $\underset{\sim}{B}$) (cf (II.3.3.2), (respectivement II.4.2.3)).

3) A partir de $\tilde{H}$ (respectivement $\underset{\sim}{H}$) qui est symétrique strictement définie positive et la donnée de δ, γ, s, y et ε' nous définissons par l'une des deux H-corrections données en (II.5.0), (II.6.0) deux matrices $\bar{\tilde{H}}$ et $\hat{\tilde{H}}$ (respectivement $\bar{\underset{\sim}{H}}$ et $\hat{\underset{\sim}{H}}$) qu'on notera $\bar{H}$ et $\hat{H}$, et ceci dans les deux cas :

i) $\varepsilon' = 1$, $\delta > 0$, $\gamma \geq 0$

ii) $\varepsilon' = -1$, $0 \leq \gamma < \delta$

qui implique que $\bar{H}$ et $\hat{H}$ sont régulières.

4) Enfin $\bar{B}$ et $\hat{B}$ seront définies comme les matrices inverses de $\bar{H}$ et $\hat{H}$ respectivement, inverses obtenues par la formule de Sherman-Morrison rappelée en (II.3.3.1).

Cette construction est schématisée par le diagramme suivant pour l'une des formules.

$$B$$

$$\tilde{B} = \alpha B + \varepsilon\beta \, \frac{(Bs)(Bs)^t}{\langle Bs,s\rangle}$$

$$\varepsilon = +1,\ \alpha>0,\ \beta\geq 0 \qquad\qquad \varepsilon = -1,\ 0 \leq \beta < \alpha$$

$$\tilde{H} = [\tilde{B}]^{-1}$$

$$\bar{H} = \delta\tilde{H} + \varepsilon'\gamma \, \frac{(\tilde{H}y)(\tilde{H}y)^t}{\langle \tilde{H}y,\ y\rangle + \langle s,\ y\rangle}$$

$$\varepsilon' = 1,\ \delta>0,\ \gamma\geq 0 \qquad\qquad \varepsilon' = -1,\ 0 \leq \gamma < \delta$$

$$\bar{B} = [\bar{H}]^{-1} = \frac{\alpha}{\delta}\, B + \varepsilon\, \frac{\beta}{\delta}\, \frac{(Bs)(Bs)^t}{\langle Bs,s\rangle} - \varepsilon'\, \frac{\gamma}{\delta}\, \frac{yy^t}{(\delta+\varepsilon'\delta)\langle \tilde{H}y,y\rangle+\delta\langle s,y\rangle}$$

Diagramme : Formule (III.2.3)

Nous obtenons ainsi quatre B-formules :

$$\bar{B} = \frac{1}{\delta}\left[\alpha B + \varepsilon\beta\,\frac{(Bs)(Bs)^t}{<Bs,s>}\right] - \varepsilon'\,\frac{\gamma}{\delta}\,\frac{yy^t}{(\delta+\varepsilon'\delta)\,<\tilde{H}y,y> + \delta\,<s,y>} \qquad \text{(III.2.3)}$$

$$\bar{\bar{B}} = \frac{1}{\delta}\left[\alpha B + \varepsilon\beta\,\frac{(Bs)(Bs)^t}{<Bs,s> + <s,y>}\right] - \varepsilon'\,\frac{\gamma}{\delta}\,\frac{yy^t}{(\delta+\varepsilon'\delta)<\underset{\sim}{H}y,y>+\delta<s,y>} \qquad \text{(III.2.4)}$$

$$\hat{B} = \frac{1}{\delta}\left[\alpha B + \varepsilon\beta\,\frac{(Bs)(Bs)^t}{<Bs,s>}\right] - \varepsilon'\,\frac{\gamma}{\delta}\,\frac{yy^t}{(\delta+\varepsilon'\gamma)\,<\tilde{H}y,y>} \qquad \text{(III.2.5)}$$

$$\hat{\hat{B}} = \frac{1}{\delta}\left[\alpha B + \varepsilon\beta\,\frac{(Bs)(Bs)^t}{<Bs,s> + <s,y>}\right] - \varepsilon'\,\frac{\gamma}{\delta}\,\frac{yy^t}{(\delta+\varepsilon'\gamma)<\underset{\sim}{H}y,y>} \qquad \text{(III.2.6)}$$

où $\tilde{H}$ (respectivement $\underset{\sim}{H}$) est la matrice inverse de la matrice entre crochet de la même formule.

Remarque

Formellement la formule de la BFGS est un cas particulier de la première formule $\bar{B}$ donnée en (III.2.3). En effet si dans cette dernière nous faisons $\varepsilon = \varepsilon' = -1$ et $\alpha = \beta = \delta = \gamma = 1$ on retrouve la formule de BFGS. Néanmoins les paramètres α, β, γ et δ, tous égaux, ne vérifient pas les conditions requises pour la définition de $\bar{B}$ (cf. diagramme).

III.3 - PROPOSITION

Sous les hypothèses précédentes, nous avons les relations suivantes entre les normes de B et $\bar{B}$ (respectivement B et $\hat{B}$) :

i) si : 1) $\varepsilon = 1$, $\alpha > 0$, $\beta \geq 0$

2) $\varepsilon' = 1$, $\delta > 0$, $\gamma \geq 0$

alors $\qquad ||\bar{B}|| \leq \frac{\alpha+\beta}{\delta}\,||B||$ (III.3.1)

(respectivement $||\hat{B}|| \leq \frac{\alpha+\beta}{\delta}\,||B||$) (III.3.2)

ii) si 1) $\varepsilon = 1,\ \alpha > 0,\ \beta \geq 0$

2) $\varepsilon' = -1,\ 0 \leq \gamma < \delta$

alors $\qquad ||\bar{B}|| \leq \frac{\alpha+\beta}{\delta-\gamma} ||B||$ (III.3.3)

(respectivement $||\hat{B}|| \leq \frac{\alpha+\beta}{\delta-\gamma} ||B||$) (III.3.4)

iii) si 1) $\varepsilon = -1,\ 0 \leq \beta < \alpha$

2) $\varepsilon' = 1,\ \delta > 0,\ \gamma \geq 0$

alors $\qquad ||\bar{B}|| \leq \frac{\alpha}{\delta} ||B||$ (III.3.5)

(respectivement $||\hat{B}|| \leq \frac{\alpha}{\delta} ||B||$) (III.3.6)

iv) si 1) $\varepsilon = -1,\ 0 \leq \beta < \alpha$

2) $\varepsilon' = -1,\ 0 \leq \gamma < \delta$

alors $\qquad ||\bar{B}|| \leq \frac{\alpha}{\delta-\gamma} ||B||$ (III.3.7)

(respectivement $||\hat{B}|| \leq \frac{\alpha}{\delta-\gamma} ||B||$) (III.3.8)

<u>Preuve</u>

i) Dans la proposition II.6.4 (respectivement II.5.3) nous remplaçons dans (II.6.4.1) (respectivement (II.5.3.1) $\bar{B}'$ par $\bar{B}$ (respectivement $\hat{B}'$ par $\bar{B}$) et B' par $\tilde{B}$ ou $\underset{\sim}{B}$ selon que $\bar{B}$ (respectivement $\hat{B}$) est définie par (III.2.3) (respectivement (III.2.5)) ou (III.2.4) (respectivement (III.2.6)). Nous avons alors $||\bar{B}|| \leq \frac{1}{\delta} ||\tilde{B}||$ ou $||\bar{B}|| \leq \frac{1}{\delta} ||\underset{\sim}{B}||$. Or d'après (II.3.2.1) ou (II.4.2.1) nous avons $||\tilde{B}|| \leq (\alpha+\beta)\ ||B||$ ou $||\underset{\sim}{B}|| \leq (\alpha+\beta)\ ||B||$. Ce qui permet de conclure .

ii) Pour l'inégalité (III.3.3) (respectivement (III.3.4)) nous utilisons (II.6.4.2) (respectivement (II.5.3.2)) et (II.3.2.1) ou (II.4.2.1).

iii) L'inégalité (III.3.5) (respectivement (III.3.6)) est obtenue en utilisant (II.6.4.1) (respectivement (II.5.3.1)) et (II.3.2.2) ou (II.4.2.2).

iv) De même l'inégalité (III.3.7) (respectivement III.3.8) est obtenue en utilisant cette fois (II.6.4.2) (respectivement (II.5.3.2)) et (II.3.2.2) ou (II.4.2.2).

III.4 - PROPOSITION

Sous les hypothèses précédentes nous avons les relations suivantes entre les normes de H et $\bar{H}$ (respectivement H et $\hat{H}$) :

i) Si 1) $\varepsilon = 1$, $\alpha > 0$ et $\beta \geq 0$

2) $\varepsilon' = 1$, $\delta > 0$ et $\gamma \geq 0$

alors $$||\bar{H}|| \leq \frac{\delta+\gamma}{\alpha} ||H|| \qquad \text{(III.4.1)}$$

(respectivement $$||\hat{H}|| \leq \frac{\delta+\gamma}{\alpha} ||H||) \qquad \text{(III.4.2)}$$

ii) Si 1) $\varepsilon = 1$, $\alpha > 0$ et $\beta \geq 0$

2) $\varepsilon' = -1$, $0 \leq \gamma < \delta$

alors $$||\bar{H}|| \leq \frac{\delta}{\gamma} ||H|| \qquad \text{(III.4.3)}$$

(respectivement $$||\hat{H}|| \leq \frac{\delta}{\alpha} ||H||) \qquad \text{(III.4.4)}$$

iii) Si 1) $\varepsilon = -1$, $0 \leq \beta < \alpha$

2) $\varepsilon' = 1$, $\delta > 0$ $\gamma \geq 0$

alors $$||\bar{H}|| \leq \frac{\delta+\gamma}{\alpha-\beta} ||H|| \qquad \text{(III.4.5)}$$

(respectivement $||\hat{H}|| \leq \frac{\delta+\gamma}{\alpha-\beta}\ ||H||$) (III.4.6)

iv) si 1) $\varepsilon = -1$, $0 \leq \beta < \alpha$

2) $\varepsilon' = -1$, $0 \leq \gamma < \delta$

alors $||\bar{H}|| \leq \frac{\delta}{\alpha-\beta}\ ||H||$ (III.4.7)

(respectivement $||\hat{H}|| \leq \frac{\delta}{\alpha-\beta}\ ||H||$) (III.4.8)

Preuve :

Pour i) dans la proposition II.6.3 (respectivement II.5.2) nous remplaçons dans (II.6.3.1) (respectivement (II.5.2.1)) $\bar{H}'$ par $\bar{H}$ (respectivement $\hat{H}'$ par $\hat{H}$) et H' par $\tilde{H}$ si $\bar{B}$ (respectivement $\hat{B}$) est donnée par (III.2.3) (respectivement (III.2.5)) ou H' par $\utilde{H}$ si $\bar{B}$ (respectivement $\hat{B}$) est définie par (III.2.4) (respectivement (III.2.6)).

Nous avons alors $||\bar{H}|| \leq (\delta+\gamma)\ ||\tilde{H}||$ ou $||\bar{H}|| \leq (\delta+\gamma)\ ||\utilde{H}||$ (respectivement $||\hat{H}|| \leq (\delta+\gamma)\ ||\tilde{H}||$ ou $||\hat{H}|| \leq (\delta+\gamma)\ ||\utilde{H}||$). Or d'après (II.3.4.1) nous avons $||\tilde{H}|| \leq \frac{1}{\alpha}\ ||H||$ et d'après (II.4.3.1) $||\utilde{H}|| \leq \frac{1}{\alpha}||H||$. Ce qui permet de conclure.

Pour la démonstration de ii) nous utiliserons (II.6.3.2) (respectivement (II.5.2.2)), (II.3.4.1) ou (II.4.3.1), et pour celle de iii) nous utiliserons (II.6.3.1) (respectivement (II.5.2.1)), (II.3.4.2) ou (II.4.3.2), et enfin pour celle de iv) nous utiliserons (II.6.3.2) (respectivement (II.5.2.2)), (II.3.4.2) ou (II.4.3.2).

III.5 - DEFINITION DES MATRICES DE METRIQUE VARIABLE

Les matrices B_k de métrique variable qui seront utilisées par la suite pour définir une direction de descente d_k telle que $B_k\ d_k = -g_k$

sont données ci-après. L'utilisation de ces matrices est justifiée par les propositions III.6 et III.7. La matrice B_o, d'inverse H_o, étant symétrique strictement définie positive, B_k est choisie dans une famille engendrée par l'une des quatre B-formules suivantes :

$$1)\ B_{k+1} = \frac{1}{\delta_k}\left[\alpha_k\ B_k + \varepsilon\beta_k \frac{(B_k s_k)(B_k s_k)^t}{\langle B_k s_k,\ s_k\rangle}\right]$$

$$- \varepsilon' \frac{\gamma_k}{\delta_k} \frac{y_k y_k^t}{(\delta_k+\varepsilon'\gamma_k)\ \langle H_k^* y_k, y_k\rangle + \delta_k\ \langle s_k,\ y_k\rangle} \qquad \text{(III.5.1)}$$

$$2)\ B_{k+1} = \frac{1}{\delta_k}\left[\alpha_k\ B_k + \varepsilon\beta_k \frac{(B_k s_k)(B_k s_k)^t}{\langle B_k s_k, s_k\rangle + \langle s_k, y_k\rangle}\right]$$

$$- \varepsilon' \frac{\gamma_k}{\delta_k} \frac{y_k y_k^t}{(\delta_k+\varepsilon'\gamma_k)\ \langle H_k^* y_k, y_k\rangle + \delta_k\ \langle s_k, y_k\rangle} \qquad \text{(III.5.2)}$$

$$3)\ B_{k+1} = \frac{1}{\delta_k}\left[\alpha_k\ B_k + \varepsilon\beta_k \frac{(B_k s_k)(B_k s_k)^t}{\langle B_k s_k,\ s_k\rangle}\right]$$

$$- \varepsilon' \frac{\gamma_k}{\delta_k} \frac{y_k y_k^t}{(\delta_k+\varepsilon'\gamma_k)\ \langle H_k^* y_k,\ y_k\rangle} \qquad \text{(III.5.3)}$$

$$4)\ B_{k+1} = \frac{1}{\delta_k}\left[\alpha_k\ B_k + \varepsilon\beta_k \frac{(B_k s_k)(B_k s_k)^t}{\langle B_k s_k,\ s_k\rangle + \langle s_k, y_k\rangle}\right]$$

$$- \varepsilon' \frac{\gamma_k}{\delta_k} \frac{y_k y_k^t}{(\delta_k+\varepsilon'\gamma_k)\ \langle H_k^* y_k, y_k\rangle} \qquad \text{(III.5.4)}$$

où

- H_k^* est la matrice inverse de la matrice entre **crochets de la même formule** Elle est donnée explicitement par la formule de Sherman-Morrison rappelée en (II.3.3.1).
- α_k, β_k et ε sont des paramètres définissant une B-correction, i.e. vérifiant (II.3.1.1) ou (II.3.1.2),

. δ_k, γ_k et ε' sont des paramètres définissant une H-correction i.e. vérifiant (II.6.1.1) ou (II.6.1.2).

Dans toute la suite, $\{B_k,\ k \in \mathbb{N}\}$ désignera une suite de matrices définies intérativement par l'une des quatre formules ci-dessus et $\{H_k,\ k \in \mathbb{N}\}$ la suite des inverses.

III.6 - PROPOSITION

Dans chacun des quatre cas suivants :

ℓ_1) $(\varepsilon, \varepsilon') = (1, 1)$, $0 < \alpha_k$, $0 \leq \beta_k$, $0 < \delta_k$, $0 \leq \gamma_k$ et $\dfrac{\alpha_k - \delta_k + \beta_k}{\delta_k}$ terme général d'une série absolument convergente,

ℓ_2) $(\varepsilon, \varepsilon') = (+1, -1)$, $0 < \alpha_k$, $0 \leq \beta_k$, $0 \leq \gamma_k < \delta_k$ et $\dfrac{\alpha_k - \delta_k + \beta_k + \gamma_k}{\delta_k - \gamma_k}$ terme général d'une série absolument convergente,

ℓ_3) $(\varepsilon, \varepsilon') = (-1, 1)$, $0 \leq \beta_k < \alpha_k$, $0 < \delta_k$, $0 \leq \gamma_k$ et $\dfrac{\alpha_k - \delta_k}{\delta_k}$ terme général d'une série absolument convergente,

ℓ_4) $(\varepsilon, \varepsilon') = (-1, -1)$, $0 \leq \beta_k < \alpha_k$, $0 \leq \gamma_k < \delta_k$ et $\dfrac{\delta_k - \alpha_k - \gamma_k}{\alpha_k}$ terme général d'une série absolument convergente,

la suite $\{B_k,\ k \in \mathbb{N}\}$ est uniformément majorée.

<u>Preuve</u> :

Reportons-nous à la proposition III.3, nous avons $\bar{B}$ (ou $\hat{B}$) n'est autre que B_{k+1} définie par l'une des formules de III.5. Ainsi :

Dans le cas ℓ_1), nous avons $||B_{k+1}|| \leq \dfrac{\alpha_k + \beta_k}{\delta_k} ||B_k||$, d'où

$||B_{k+1}|| \leq \left(\prod_{j=o}^{k} \dfrac{\alpha_j + \beta_j}{\delta_j} \right) ||B_o||$. Si le produit infini de terme général $\dfrac{\alpha_j + \beta_j}{\delta_j}$, $j \in \mathbb{N}$, est convergent alors la suite $\{B_k,\ k \in \mathbb{N}\}$ est uniformément majorée ; or ce produit infini converge si et seulement si la série de terme général $\frac{\alpha_j+\beta_j}{\delta_j} - 1 = \frac{\alpha_j-\delta_j+\beta_j}{\delta_j}$ est absolument convergente.

Dans le cas ℓ_2) nous avons $||B_{k+1}|| \leq \frac{\alpha_k+\beta_k}{\delta_k-\gamma_k}\, ||B_k||$, et dans le cas ℓ_3) nous avons $||B_{k+1}|| \leq \frac{\alpha_k}{\delta_k}\, ||B_k||$. La suite de la démonstration est identique à celle du cas ℓ_1). Dans le cas ℓ_4) nous avons $||B_{k+1}|| \leq \frac{\alpha_k}{\delta_k-\gamma_k}\, ||B_k||$.

Si le produit infini de terme général $\frac{\alpha_k}{\delta_k-\gamma_k}$ où, $\frac{\delta_k-\gamma_k}{\alpha_k}$, $k \in \mathbb{N}$, converge alors la suite $\{B_k,\ k \in \mathbb{N}\}$ est uniformément majorée. Or ce produit infini converge si et seulement si le produit infini de terme général $\frac{\delta_k-\gamma_k}{\alpha_k}$ converge.

III.7 - PROPOSITION

Dans chacun des quatre cas suivants :

ℓ'_1) $(\varepsilon,\varepsilon') = (1,\ 1)$, $0 < \alpha_k$, $0 \leq \beta_k$, $0 < \delta_k$, $0 \leq \gamma_k$ et $\frac{\delta_k-\alpha_k+\gamma_k}{\alpha_k}$ terme général d'une série absolument convergente,

ℓ'_2) $(\varepsilon,\varepsilon') = (1,\ -1)$, $0 < \alpha_k$, $0 \leq \beta_k$, $0 \leq \gamma_k < \delta_k$ et $\frac{\delta_k-\alpha_k}{\alpha_k}$ terme général d'une série absolument convergente,

ℓ'_3) $(\varepsilon,\varepsilon') = (-1,\ 1)$, $0 \leq \beta_k < \alpha_k$, $0 < \delta_k$, $0 \leq \gamma_k$ et $\frac{\delta_k-\alpha_k+\gamma_k+\beta_k}{\alpha_k-\beta_k}$ terme général d'une série absolument convergente,

ℓ'_4) $(\varepsilon,\varepsilon') = (-1,\ -1)$, $0 \leq \beta_k < \alpha_k$, $0 \leq \gamma_k < \delta_k$ et $\frac{\alpha_k-\delta_k-\beta_k}{\delta_k}$ terme général d'une série absolument convergente,

la suite $\{H_k,\ k \in \mathbb{N}\}$ est uniformément majorée.

Preuve :

Reportons-nous à la proposition III.4 où $\bar{H}$ et $\hat{H}$ ne sont autres que H_{k+1}, matrice inverse de B_{k+1} définie par l'une des formules de III.5 .

Dans le cas ℓ'_1) nous avons d'après (III.4.1) ou (III.4.2) $||H_{k+1}|| \leq$

$\frac{\delta_k+\gamma_k}{\alpha_k} ||H_k||$, ce qui donne $||H_{k+1}|| \leq (\prod_{j=o}^{k} \frac{\delta_j+\gamma_j}{\alpha_j}) ||H_o||$.

Dans le cas ℓ'_2) nous avons d'après (III.4.3) ou (III.4.4) $||H_{k+1}|| \leq$

$\frac{\delta_k}{\alpha_k} ||H_k||$; ce qui donne $||H_{k+1}|| \leq (\prod_{j=o}^{k} \frac{\delta_j}{\alpha_j}) ||H_o||$.

Dans le cas ℓ'_3) nous avons d'après (III.4.5) ou (III.4.6) $||H_{k+1}|| \leq$

$\frac{\delta_k+\gamma_k}{\alpha_k-\beta_k} ||H_k||$; ce qui donne $||H_{k+1}|| \leq (\prod_{j=o}^{k} \frac{\delta_j+\gamma_j}{\alpha_j-\beta_j}) ||H_o||$.

Dans le cas ℓ'_4), nous avons, d'après (III.4.7) ou (III.4.8) $||H_{k+1}|| \leq$

$\frac{\delta_k}{\alpha_k-\beta_k} ||H_k||$, ce qui donne $||H_{k+1}|| \leq (\prod_{j=o}^{k} \frac{\delta_j}{\alpha_j-\beta_j}) ||H_o||$.

Ces inégalités permettent de conclure.

CONCLUSION

Nous avons défini quatre familles de matrices symétriques paramètrées $\{B_k, k \in \mathbb{N}\}$. Les paramètres vérifient certaines conditions rendant les matrices B_k :

- strictement définies positives,
- uniformément majorées, ainsi que leurs inverses.

REMARQUE

Nous avons défini au chapitre III quatre familles de matrices de métrique variable déduites d'une matrice symétrique strictement définie positive par une correction de rang 2. Nous avons remarqué que la famille engendrée par (III.2.3) contient formellement la B-formule de BFGS.

Si dans III.5 nous remplaçons (B, s, y, H) par (H, y, s, B) nous obtenons quatre familles de matrices déduites de H, symétrique strictement définie positive, par une correction de rang 2. Ces familles sont de H-formules, dont la première contient formellement la H-formule de D.F.P. (pour les paramètres $\alpha = \delta = \beta = \gamma = 1$ et $\varepsilon = \varepsilon' = -1$). Ces H-formules sont symétriques, strictement définies positives, (ainsi que leurs inverses) pourvu que H le soit. Des conditions analogues à celles des propositions III.6 et III.7 peuvent être obtenues pour que ces matrices soient uniformément majorées ainsi que leurs inverses.

C H A P I T R E IV

CONVERGENCE DE L'ALGORITHME

INTRODUCTION

Dans ce chapitre nous donnerons un théorème de convergence globale pour une fonction bornée inférieurement et continûment différentiable. Nous démontrerons sous d'autres hypothèses plus fortes que cette convergence est au moins R-linéaire.

Considérons le programme mathématique suivant :

$$\text{Min } \{f(x) \ / \ x \in \mathbb{R}^n\}$$

Remarque

Dans toute la suite les conditions ℓ_i et ℓ'_i définie par les propositions III.6 et III.7 seront notées L_i, pour $i = 1, 2, 3$ ou 4.

IV.1 - HYPOTHESES

Soit x_o un point quelconque de $\mathbb{R}^n$. Dans toute la suite nous ferons les hypothèses suivantes :

h_1) la fonction économique f sera supposée continuement différentiable,

h_2) l'ensemble $E_o = \{x \in \mathbb{R}^n \ / \ f(x) \leq f(x_o)\}$ sera supposé compact.

IV.2 - DEFINITION DE L'ALGORITHME DE METRIQUE VARIABLE

Soient :

i) ε, ε' deux entiers prenant indépendamment les valeurs +1 ou -1,

(ii) $\{\alpha_k, k \in \mathbb{N}\}$, $\{\beta_k, k \in \mathbb{N}\}$, $\{\gamma_k, k \in \mathbb{N}\}$, $\{\delta_k, k \in \mathbb{N}\}$

quatre suites réelles vérifiant la condition L_i, $i = 1, 2, 3$ ou 4 définie ci-dessus.

iii) c_1 et c_2 deux réels tels que $0 < c_1 < c_2 < 1$, $c_1 < \frac{1}{2}$.

Initialisation

Nous nous donnons x_o un point quelconque de $\mathbb{R}^n$ et B_o une matrice (n,n) d'inverse H_o, symétrique strictement définie positive (exemple $B_o = I$).

Iteration k

1) Disposant d'un point x_k et d'une matrice B_k, symétrique strictement définie positive, si $g_k \neq 0$, nous définissons la direction d_k comme suit :

$$d_k = - H_k \, g_k$$

où H_k est l'inverse d'une B-formule définie au chapitre précédent. Pratiquement d_k est définie comme la solution du système linéaire $B_k \, d_k = -g_k$.
L'arrêt (théorique) est obtenu pour $g_k = 0$.

2) Nous cherchons dans la direction d_k, un point x_{k+1} de la forme $x_k + \lambda_k \, d_k$ où le pas λ_k est donné par la recherche linéaire de Wolfe à savoir :

i) $f(x_k + \lambda_k \, d_k) \leq f(x_k) + c_1 \, \lambda_k \, \langle g_k, d_k \rangle$ (IV.2.1)

ii) $\langle g_{k+1}, d_k \rangle \geq c_2 \, \langle g_k, d_k \rangle$ (IV.2.2)

3) Partant de s_k, y_k, ε, ε', α_k, β_k, γ_k et δ_k nous définissons B_{k+1} par l'une des quatre formules de III.5, et nous retournons en 1) avec $k = k+1$.

IV.3 - LEMME

Si $\{A_k, k \in \mathbb{N}\}$ est une suite de matrices (n, n) symétriques strictement définies positives telles que $||A_k^{-1}|| \leq M$, alors tout point d'accumulation A, s'il existe, de la suite $\{A_k,\ k \in \mathbb{N}\}$ est une matrice symétrique définie positive.

Preuve :

Il existe $N' \subset N$ tel que $A_k \xrightarrow[k \in N']{} A$

La matrice A est évidemment symétrique, définie positive. Montrons qu'elle est strictement définie positive. Soit λ_1 la plus petite valeur de A ; nous pouvons lui associer un vecteur propre unitaire z_1. Nous avons alors $\langle z_1, Az_1\rangle = \lambda_1$. Si $\lambda_1 < \frac{1}{M}$ nous posons $\eta = \frac{1}{2}\left(\frac{1}{M} - \lambda_1\right)$; il existe alors un entier k_o tel que pour tout k dans N' plus grand que k_o nous ayons $\langle z_1, A_k\, z_1\rangle \leq \lambda_1 + \eta$. Comme $\inf \{\langle z, A_k\, z\rangle \;/\; ||z|| \leq 1\}$ n'est autre que la plus petite valeur propre de A_k, noté λ_1^k et que λ_1^k est supérieure à $\frac{1}{M}$, nous obtenons une contradiction :

$$\frac{1}{M} \leq \lambda_1^k = \inf \{\langle z, A_k\, z\rangle \;/\; ||z|| \leq 1\} \leq \lambda_1 + \eta < \frac{1}{M}$$

Nous allons donner le théorème de convergence de l'algorithme.

IV.4 - THEOREME DE CONVERGENCE

Sous les hypothèses h_1 et h_2 précédentes et pour chacune des conditions L_i, i = 1, 2, 3, 4, tout point d'accumulation x^* de la suite $\{x_k,\ k \in \mathbb{N}\}$ définie par l'algorithme IV.2 est tel que $\nabla f(x^*) = 0$.

Preuve

Soit M une borne supérieure (fournie par la proposition III.7) de la suite $\{H_k,\ k \in \mathbb{N}\}$. De $d_k = -\, H_k\, g_k$ et du fait que la fonction f est continuement différentiable dans le compact E_0 contenant toute la suite

$\{x_k,\ k \in \mathbb{N}\}$ nous déduisons que la suite $\{d_k,\ k \in \mathbb{N}\}$ est uniformément bornée par un réel ρ. Notons B_ρ la boule de centre zéro et de rayon ρ.

Définissons l'application multivoque $M : E_o \times B_\rho \to \mathcal{P}(E_o)$ défini par :

$$M(x, d) = \left\{\begin{array}{l} y \in E_o \ /\ y = x + \lambda d,\ \lambda \geq 0,\ f(x + \lambda d) \leq f(x) + \lambda c_1 <g(x), d>, \\ \qquad\qquad\qquad\qquad\qquad <g(x + \lambda d).d> \geq c_2 <g(x), d> \end{array}\right\}$$

Elle est sup-continue sur $E_o \times (B_\rho - \{0\})$. En effet pour toute suite $\{(x_m, d_m),\ m \in \mathbb{N}\}$ convergeant vers (x, d), avec d non nul, et toute suite $\{y_m \in M(x_m, d_m),\ m \in \mathbb{N}\}$ convergeant vers y, nous avons bien y dans $M(x, d)$. Pour cela il suffit d'utiliser la définition de λm suivante :

$$\lambda m = \frac{<y_m - x_m, d_m>}{||d_m||^2}$$

et le fait que f est continûment différentiable.

Soit x^* un point d'accumulation de la suite $\{x_k,\ k \in \mathbb{N}\}$; un tel point existe d'après l'hypothèse h_2 (E_o compact). Montrons que $g^* = g(x^*) = 0$. De même des suites $\{d_k,\ k \in \mathbb{N}\}$ et $\{H_k,\ k \in \mathbb{N}\}$ appartenant respectivement aux compacts $B_\rho \subset \mathbb{R}^n$ et $K \subset \mathbb{R}^{n^2}$, nous pouvons extraire des sous-suites convergentes vers d^* et H^* respectivement. Nous pouvons supposer que (x_k, d_k, H_k) converge vers (x^*, d^*, H^*) pour $k \in N' \subset \mathbb{N}$. D'après le lemme IV.3 H^* est une matrice symétrique strictement définie positive. La sous-suite $\{d_k = - H_k\, g_k,\ k \in N'\}$ converge vers $d^* = - H^*\, g^*$. Si nous supposons que g^* est non nul alors d^* est non nul. Considérons alors la sous-suite $\{x_{k+1} \in M(x_k, d_k),\ k \in N'\}$. Nous pouvons en extraire une sous-suite convergente vers x^{**}. Soit $\{x_{k+1},\ k \in N'' \subset N'\}$ une telle sous-suite. Comme M est sup-continue, x^{**} appartient à $M(x^*, d^*)$, i.e. $f(x^{**}) - f(x^*) \leq \lambda^*\, c_1$ $<g^*, d^*>$ et $<g^{**}, d^*> \geq c_2 <g^*, d^*>$. De plus $<g^*, d^*>$ est strictement négatif et par suite $f(x^{**})$ est strictement inférieur à $f(x^*)$. Or nous savons en raison de la décroissance de la suite $\{f(x_k) \ / \ k \in \mathbb{N}\}$ que $f(x^*) = f(x^{**})$, d'où la contradiction, par suite $g(x^*) = 0$.

IV.5 - DEFINITION D'UN ALGORITHME DE PENTE

Si dans un algorithme général d'optimisation, les directions de recherche linéaire d_k sont telles qu'il existe ν dans $]0, 1]$ vérifiant $- \langle d_k, g_k \rangle \geq \nu \, ||g_k|| \; ||d_k||$ pour tout entier k, alors l'algorithme est dit de pente.

IV.6 - PROPOSITION

Soient $\{H_k, \ k \in \mathbb{N}\}$ une suite de matrices symétriques strictement définies positives et deux constantes M et M' telles que pour tout entier k $||H_k|| \leq M$ et $||B_k|| \leq M'$ alors il existe une constante μ strictement positive telle que si $d_k = - H_k \, g_k$ alors pour tout k nous avons :

$$- \langle d_k, g_k \rangle \geq \mu \, ||d_k|| \; ||g_k|| \qquad \text{(IV.6.1)}$$

Preuve :

Notons $\lambda_1^k \leq \lambda_2^k \leq \ldots \leq \lambda_n^k$ les valeurs propres de H_k. Nous savons que $||H_k|| = \lambda_n^k$ et $||B_k|| = \frac{1}{\lambda_1^k}$. Nous obtenons $- \langle d_k, g_k \rangle = \langle H_k \, g_k, g_k \rangle \geq \lambda_1^k \, ||g_k||^2 = \lambda_1^k \, ||g_k|| \; ||B_k \, d_k||$ et nous avons $||B_k \, d_k|| \geq (\lambda_n^k)^{-1} \, ||d_k||$; nous en déduisons que $- \langle d_k, g_k \rangle \geq \lambda_1^k (\lambda_n^k)^{-1} \, ||g_k|| \; ||d_k||$. De $\lambda_1^k \geq (M')^{-1}$ et $\lambda_n^k \leq M$, nous déduisons que $\lambda_1^k (\lambda_n^k)^{-1} \geq (MM')^{-1}$, par suite nous obtenons

$$- \langle g_k, d_k \rangle \geq \mu \, ||g_k|| \; ||d_k|| \quad \text{où } \mu = (MM')^{-1}$$

Remarquons que le réel μ dépend de la norme euclidienne.

Sous certaines hypothèses, il est démontré (cf. Théorème IV.7) que tout algorithme de pente converge au moins R-linéairement si la recherche linéaire utilisée est

i) exacte, i.e. on détermine $\lambda_k \in \operatorname{argmin} \{f(x_k + \lambda d_k) \;/\; \lambda \geq 0\}$ (IV.6.2)

ii) la recherche linéaire de Goldstein [27] :

Etant donné $\alpha \in]0, 1[$ nous cherchons λ_k positif tel que

$$\lambda_k (1-\alpha) <g_k, d_k> \leq f(x_k + \lambda_k d_k) - f(x_k) \leq \lambda_k \alpha <g_k, d_k> \qquad \text{(IV.6.3)}$$

iii) la recherche linéaire d'Armijo [2] :

Etant donné $\alpha \in]0, 1[$, $\beta \in]0, 1[$ et $\bar{\rho} > 0$, nous cherchons λ_k sous la forme $\beta^j \bar{\rho}$ où j est le plus petit entier tel que

$$f(x_k + \lambda_k d_k) - f(x_k) \leq \alpha \bar{\rho} \beta^j <g_k, d_k> \qquad \text{(IV.6.4)}$$

IV.7 - THEOREME (cf [40])

Soit $\{x_k, k \in \mathbb{N}\}$ une suite donnée par un algorithme de pente ν de recherche linéaire (IV.6.2), (IV.6.3) ou (IV.6.4). Si

i) x_k converge vers un minimum local de f, $\hat{x}$.

ii) f deux fois continuement differentiable dans un voisinage convexe V de $\hat{x}$,

iii) il existe m_1 et m_2 deux constantes réelles telles que $0 < m_1 \leq m_2$ et

$$m_1 ||y||^2 \leq <y, G(x) y> \leq m_2 ||y||^2 \qquad \text{(IV.7.1)}$$

pour tout x dans V et tout y dans $\mathbb{R}^n$, alors pour k assez grand nous avons :

$$||x_k - \hat{x}|| \quad \sqrt{\frac{2}{m_1} f(x_o) - f(\hat{x})} \quad \theta^{\frac{k}{2}}$$

et $$f(x_k) - f(\hat{x}) \leq [f(x_o) - f(\hat{x})] \, \theta^k$$

où $\theta = 1 - \nu^2 (\frac{m_1}{m_2})^2$ pour la recherche linéaire exacte,

$\theta = 1 - 4 \alpha^2 \nu^2 (\frac{m_1}{m_2})^2$ pour la recherche linéaire de Goldstein

$\theta = 1 - 2 \beta^j \bar{\rho} \alpha (\frac{m_1}{m_2})^2$ pour la recherche linéaire d'Armijo

Remarque :

Dans tous les cas θ dépend quadratiquement du nombre de conditionnement de la matrice des dérivées secondes de f.

Plus récemment Allwright [1], a amélioré la constante θ précédente.

IV.8 - THEOREMES ([1])

Sous les hypothèses du théorème précédent nous avons

$$f(x_k) - f(\hat{x}) \leq [f(x_o) - f(\hat{x})]\ \phi^k \text{ où}$$

$\phi = 1 - \nu^2 \dfrac{m_1}{m_2}$ pour la recherche linéaire exacte,

$\phi = 1 - 4\ \alpha^2\ \nu^2\ \dfrac{m_1}{m_2}$ pour la recherche linéaire de Goldstein

$\phi = 1 - 4\ \alpha(1 - \alpha)\ \beta\ \nu^2\ \dfrac{m_1}{m_2}$ pour la recherche linéaire d'Armijo.

Comme la recherche linéaire de Wolfe est largement utilisée, nous allons étendre le résultat du théorème IV.8.

IV.9 - Théorème

Soit x^* un point d'accumulation d'une suite engendrée par un **algorithme** de pente (cf IV.5) et la recherche linéaire de Wolfe. S'il existe un voisinage $V(x^*)$ tel que :

i) f soit deux fois continuement différentiable dans $V(x^*)$,

ii) il existe m_1 et m_2 deux constantes strictement positive vérifiant $m_1\ ||y||^2 \leq \langle y, G(x)\ y\rangle \leq m_2\ ||y||^2$,

iii) il existe un entier, k^*, tel que la sous-suite $\{x_k, k \geq k^*\}$ soit contenue dans $V(x^*)$, alors pour $k \geq k^*$, nous avons

$$f(x_k) - f(x^*) \leq [f(x_o) - f(x^*)]\ \theta^k$$

$$||x_k - x^*|| \leq \sqrt{\frac{2}{m_1}\ (f(x_o) - f(x^*))}\ \ \theta^{\frac{k}{2}}$$

où $$\theta = 1 - 2c_1\ (1-c_2)\ \frac{m_1}{m_2}\ \mu^2$$

Nous pouvons supposer que $k^* = 0$ pour la commodité de la démonstration.

Preuve du théorème

D'après l'hypothèse ii) nous pouvons écrire

$$f(x_{k+1}) \leq f(x_k) + \lambda_k <g_k,\ d_k> + \frac{1}{2}\ \lambda_k^2\ m_2\ ||d_k||^2.$$

Nous avons aussi :

$$<g_{k+1},\ d_k> = <g_k,\ d_k> + \lambda_k\ <d_k,\ G(\xi'_k)\ d_k> \tag{IV.9.1}$$

or nous avons, d'après le deuxième critère de choix de Wolfe, i.e.

$$<g_{k+1},\ d_k> \geq c_2\ <g_k,\ d_k> \tag{IV.9.2}$$

De (IV.9.1) et (IV.9.2) nous déduisons

$$(1 - c_2)\ <g_k,\ d_k> \geq -\ \lambda_k\ <d_k,\ G(\xi'_k)\ d_k>,$$

ce qui donne, d'après l'hypothèse ii)

$$(1 - c_2)\ <g_k,\ d_k> \geq -\ \lambda_k\ m_2\ ||d_k||^2$$

d'où $$\lambda_k \geq -\ \frac{1-c_2}{m_2}\ \frac{<g_k,\ d_k>}{||d_k||^2} \tag{IV.9.3}$$

D'après le premier critère de choix de Wolfe, nous avons $f(x_{k+1}) - f(x_k) \leq c_1\ \lambda_k <g_k,\ d_k>$, qui, d'après (IV.9.3) donne $f(x_{k+1}) - f(x_k) \leq -\ \frac{c_1(1 - c_2)}{m_2}\ \frac{<g_k,\ d_k>^2}{||d_k||^2}$ (IV.9.4)

L'hypothèse ii) permet d'écrire pour $x \in V(x^*)$, $u \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$.

$$f(x) + \lambda<g(x),u> + \frac{1}{2}\ m_1\ \lambda^2\ ||u||^2 \leq f(x+\lambda u) \leq f(x)+\lambda<g(x),u>+ \frac{1}{2}\ m_2\ \lambda^2||u||^2.$$

Suivant Allwright [1], pour x fixé dans $V(x^*)$, nous avons :

$$f(x) - \frac{||g(x)||^2}{2m_1} = \min \{f(x) + <g(x), u> + \frac{1}{2} m_1 ||u||^2 \ / \ u \in \mathbb{R}^n\}$$

$$\leq \min \{f(x + u) \ / \ u \in \mathbb{R}^n\}$$

$$f(x) - \frac{||g(x)||^2}{2m_1} \leq f(x^*).$$

En particulier dans l'inégalité précédente en posant : $x = x_k$, nous obtenons :

$$||g_k||^2 \geq 2m_1 \ [f(x_k) - f(x^*)].$$

Ce qui, avec l'inégalité (IV.9.4) donne

$$1 + \frac{f(x_{k+1}) - f(x_k)}{f(x_k) - f(x^*)} \leq 1 - 2c_1 \ (1-c_2) \ \frac{m_1}{m_2} \left[\frac{<g_k, \ d_k>}{||g_k|| \ ||d_k||}\right]^2 \qquad \text{(IV.9.5)}$$

et nous avons (IV.6.1), $- <g_k, \ d_k> \geq \mu \ ||g_k|| \ ||d_k||$ qui implique :

$$\left[\frac{<g_k, d_k>}{||d_k|| \ ||g_k||}\right]^2 \geq \mu^2 \qquad \text{(IV.9.6)}$$

De (IV.9.5) et (IV.9.6) nous déduisons

$$1 + \frac{f(x_{k+1}) - f(x_k)}{f(x_k) - f(x^*)} \leq 1 - 2 \ c_1 \ (1 - c_2) \ \frac{m_1}{m_2} \ \mu^2$$

Nous obtenons ainsi

$$f(x_{k+1}) - f(x^*) \leq (1 - 2c_1 \ (1-c_2) \ \frac{m_1}{m_2} \ \mu^2) \ [f(x_k) - f(x^*)] \qquad \text{(IV.9.7)}$$

Posons $\qquad \theta = 1 - 2c_1 \ (1 - c_2) \ \frac{m_1}{m_2} \ \mu^2, \qquad$ (IV.9.8)

l'inégalité (IV.9.7) ci-dessus s'écrit alors :

$f(x_k) - f(x^*) \leq \theta^k \ [f(x_o) - f(x^*)]$ qui est le premier résultat du théorème.

Nous avons

$f(x_k) - f(x^*) = <g^*, \ x_k - x^*> + \frac{1}{2} <x_k - x^*, \ G(\bar{\xi}_k) \ (x_k - x^*)>$,où $\bar{\xi}_k \in \ ]x_k, \ x^*[$, or $g^* = 0$, d'où $f(x_k) - f(x^*) = \frac{1}{2} <x_k-x^*, \ G(\bar{\xi}_k)(x_k-x^*)>$ et d'après l'hypothèse ii) nous avons :

$$m_1 \, ||x_k - x^*||^2 \leq 2\,[f(x_k) - f(x^*)] \leq m_2 \, ||x_k - x^*||^2 \qquad \text{(IV.9.9)}$$

Par suite nous obtenons $||x_k - x^*|| \leq \frac{2}{m_1}\,[f(x_o) - f(x^*)]\,\theta^k$ qui donne le deuxième résultat du théorème :

$$||x_k - x^*|| \leq \sqrt{\frac{2}{m_1}\,[f(x_o) - f(x^*)]}\;\theta^{\frac{k}{2}} \qquad \text{(IV.9.10)}$$

IV.10 - BORNES DU PAS λ_k

Nous allons donné quelques propriétés des itérés d_k, g_k et s_k. Dans toute la suite nous supposerons vérifiées les hypothèses du théorème IV.9. Ces hypothèses nous ont permis, avec l'utilisation de la recherche linéaire de Wolfe, de trouver une borne inférieure pour le pas λ_k (cf. IV.9.3)

$$\lambda_k \geq -\frac{1-c_2}{m_2}\,\frac{<g_k,\, d_k>}{||d_k||^2}$$

Nous obtenons, compte tenu de $B_k\, d_k = -g_k$,

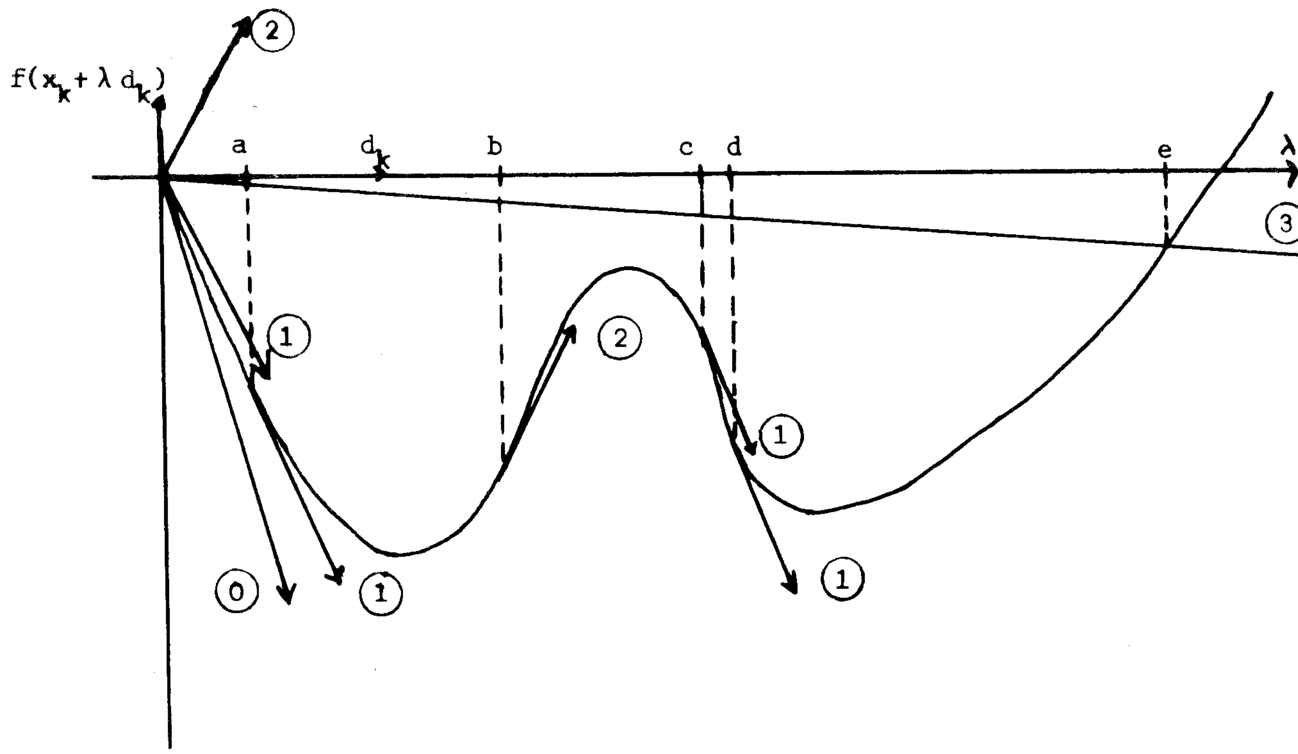
$$\lambda_k \geq \frac{1-c_2}{m_2}\,\frac{<B_k d_k, d_k>}{||d_k||^2} \geq \frac{1-c_2}{m_2}\,\frac{1}{M} \qquad \text{(IV.10.1)}$$

où M est un majorant uniforme de $||H_k||$, $k \in \mathbb{N}$.

Avec le critère de Wolfe nous ne pouvons rien dire quant à une éventuelle borne supérieure uniforme de λ_k. Pour avoir une telle borne, nous renforçons le critère de Wolfe par le suivant :

$$|<g_{k+1},\, d_k>| \leq c_2\,|<g_k,\, d_k>| \qquad \text{(IV.10.2)}$$

Ce critère cité dans [23] est schématisé par la figure suivante où nous avons tenu compte du premier critère de Wolfe :



(0) Demi-droite de pente $<g_k, d_k>$

(1) Demi-droite de pente $c_2 <g_k, d_k>$

(2) Demi-droite de pente $-c_2 <g_k, d_k>$

(3) Demi-droite de pente $c_1 <g_k, d_k>$

Avec une recherche linéaire de Wolfe le domaine successeur est :

$$[a, c] \cup [d, e]$$

Ici il est égal à $[a, b] \cup [d, e]$.

Cette deuxième recherche linéaire nous permettra de trouver une borne supérieure de la suite λ_k.

L'inégalité (IV.10.2) est équivalente à :

$$c_2 <g_k, d_k> \leq <g_{k+1}, d_k>$$

et $\quad <g_{k+1}, d_k> \leq -c_2 <g_k, d_k>$

La première inégalité a été utilisée pour établir (IV.10.1). La deuxième inégalité et la relation :

$$<g_{k+1}, d_k> = <g_k, d_k> + \lambda_k <d_k, G(z_k)\, d_k>$$

où $z_k \in]x_k, x_{k+1}[$ entrainent :

$$\lambda_k <d_k, G(z_k)\, d_k> \leq -\ (1 + c_2) <g_k, d_k>$$

et compte tenu de (IV.7.1) nous obtenons

$$\lambda_k\ m_1\ ||d_k||^2 \leq -\ (1 + c_2) <g_k, d_k> \text{ et}$$

par conséquent

$$\lambda_k \leq -\ \frac{1+c_2}{m_1}\ \frac{<g_k, d_k>}{||d_k||^2}$$

En remplaçant g_k par $- B_k d_k$, il vient :

$$\lambda_k\ \ \frac{1+c_2}{m_1}\ \frac{<B_k d_k, d_k>}{||d_k||^2}$$

De $||B_k|| \leq M'$, nous déduisons :

$$\lambda_k \leq \frac{1+c_2}{m_1}\ M'$$

CONCLUSION

i) Compte tenu des hypothèses du théorème IV.9, à l'itération k, λ_k admet donc les bornes suivantes :

$$-\ \frac{1-c_2}{m_2}\ \frac{<g_k, d_k>}{||d_k||^2} \leq \lambda_k \leq -\ \frac{1+c_2}{m_1}\ \frac{<g_k, d_k>}{||d_k||^2}$$

ii) Compte tenu des hypothèses du théorème IV.9 et du fait que les matrices B_k et H_k sont uniformément majorées, nous obtenons les bornes uniformes pour le pas λ_k (k assez grand) suivantes :

$$\boxed{(1-c_2)\ \frac{1}{Mm_2} \leq \lambda_k \leq (1+c_2)\ \frac{M'}{m_1}}$$

où m_1 et m_2 sont respectivement minorant et majorant des valeurs propres du hessien de f dans un voisinage de stricte convexité de x^*, M^{-1} et M' sont respectivement minorant et majorant uniformes des valeurs propres des matrices B_k, $k \in \mathbb{N}$.

Dans [15] les auteurs ont établi, sous certaines hypothèses, une condition nécessaire et suffisante de convergence superlinéaire pour une suite $\{x_k, k \in \mathbb{N}\}$ engendrée par $x_{k+1} = x_k - \lambda_k B_k^{-1} g_k$ (où B_k est une matrice régulière) et convergeant vers x^*. Cette condition est la suivante :

$$\lim_{k \to +\infty} \left\| [\lambda_k^{-1} B_k - G(x^*)] \frac{s_k}{\|s_k\|} \right\| = 0 \qquad \text{(IV.10.4)}$$

(cf. théorème 3.1 page 51 de [15]).

Une conséquence de ce résultat est que la suite des directions d_k, converge vers zéro : Il suffit de voir la démonstration du théorème 6.4 page 65 de [15] où nous remplaçons (6.10) par (IV.10.4) ci-dessus.

Nous allons voir dans la suite que la série de terme général $\|d_k\|$, d_k étant définie dans l'algorithme IV.2. est convergente.

IV.11 - Corollaire

Sous les hypothèses du théorème IV.9, les séries de termes généraux $\|x_k - x^*\|^{\beta}$ et $\|d_k\|^{\beta'}$ où β et β' sont deux réels strictement positifs quelconques, sont convergentes.

Preuve :

D'après (IV.9.10) nous avons $||x_k - x||^{\beta} \leq c(\theta^{\frac{\beta}{2}k})$, et comme θ appartient à l'intervalle]0, 1[nous avons $\sum_{k=1}^{\infty} ||x_k - x^*||^{\beta} < +\infty$, pour tout β strictement positif.

La direction d_k est définie par $d_k = - H_k g_k$, qui s'écrit $d_k = - H_k(g_k - g^*)$ ou encore $d_k = - H_k[G(\xi_k).(x_k - x^*)]$ où ξ_k appartient à $]x_k, x^*[$. En passant aux normes, nous en déduisons que pour tout réel strictement positif β', $||d_k||^{\beta'} \leq c' \, ||x_k - x^*||^{\beta'}$, d'où le résultat.

CONSEQUENCE

Sous les hypothèses du théorème IV.9, la suite d_k converge vers zéro.

C H A P I T R E V

CARACTERISATION DES PARAMETRES

INTRODUCTION

Dans les chapitres précédents nous avons établi des conditions suffisantes pour que les matrices B_k et leurs inverses H_k soient :

a) strictement définies positives pourvu que B_o le soit,

b) uniformément majorées.

Ces conditions sont rappelées par le tableau V.I de la page suivante.

Nous allons donner pour chaque couple $(\varepsilon, \varepsilon')$ un théorème de caractérisation des suites vérifiant les conditions suffisantes donnant les résultats a) et b) ci-dessus.

Dans toute la suite, il sera supposé que B_o est une matrice symétrique strictement définie positive.

V.1 - TABLEAU

$(\varepsilon, \varepsilon')$	1	1 bis	2	2 bis	3	4
(1, 1)	$0 \leq \beta_k$	$0 < \alpha_k$	$0 \leq \gamma_k$	$0 < \delta_k$	$\sum_{k=o}^{+\infty} \frac{\|\alpha_k-\delta_k+\beta_k\|}{\delta_k} < +\infty$	$\sum_{k=o}^{+\infty} \frac{\|\delta_k-\alpha_k+\gamma_k\|}{\alpha_k} < +\infty$
(1, -1)	$0 \leq \beta_k$	$0 < \alpha_k$	$0 \leq \gamma_k$	$\gamma_k < \delta_k$	$\sum_{k=o}^{+\infty} \frac{\|\alpha_k-\delta_k+\gamma_k+\beta_k\|}{\delta_k-\gamma_k} < +\infty$	$\sum_{k=o}^{+\infty} \frac{\|\delta_k-\alpha_k\|}{\alpha_k} < +\infty$
(-1, 1)	$0 \leq \beta_k$	$\beta_k < \alpha_k$	$0 \leq \gamma_k$	$0 < \delta_k$	$\sum_{k=o}^{+\infty} \frac{\|\alpha_k-\delta_k\|}{\delta_k} < +\infty$	$\sum_{k=o}^{+\infty} \frac{\|\delta_k-\alpha_k+\gamma_k+\beta_k\|}{\alpha_k-\beta_k} < \infty$
(-1, -1)	$0 \leq \beta_k$	$\beta_k < \alpha_k$	$0 \leq \gamma_k$	$\gamma_k < \delta_k$	$\sum_{k=o}^{+\infty} \frac{\|\delta_k-\alpha_k-\gamma_k\|}{\alpha_k} < +\infty$	$\sum_{k=o}^{+\infty} \frac{\|\alpha_k-\beta_k-\delta_k\|}{\delta_k} < +\infty$

V.2 - Théorème

Soit $(\varepsilon, \varepsilon') = (-1, -1)$

Nous nous donnons 2 suites (a_k) et (b_k) vérifiant $(a_k+1)(b_k+1) \leq 1$ et telles que :

c_1) $b_k > -1$,

c_2) $a_k > -1$,

c_3) a_k et le terme d'une série absolument convergente,

c_4) b_k est le terme d'une série absolument convergente.

Nous déterminons les suites (α_k) et (δ_k), $k \in \mathbb{N}$ telles que :

c_5) $\delta_k > 0$,

c_6) $\alpha_k \leq \frac{1}{1+b_k} \delta_k$,

c_7) $\alpha_k \geq (1+a_k)\ \delta_k$,

et nous définissons les suites (β_k) et (γ_k), $k \in \mathbb{N}$, par :

d_1) $\beta_k = \alpha_k - \delta_k\ (1+a_k)$,

d_2) $\gamma_k = \delta_k - \alpha_k\ (1+b_k)$.

Alors les conditions c_j, $j = 1, 2, \ldots, 7$ et les définitions d_1 et d_2 caractérisent les suites qui vérifient les conditions 1, 1 bis, 2, 2 bis, 3 et 4 du tableau V.1, i.e. les conditions suffisantes établies précédemment qui rendent H_k et B_k :

a) strictement définies positives

b) uniformément bornées.

Preuve du théorème :

La propriété $(1+a_k)(1+b_k) \leq 1$ entraine que l'ensemble des couples (α_k, δ_k) vérifiant c_5, c_6 et c_7 est non vide, ces couples sont contenus dans le cône engendré par les deux demi-droites de pentes positives $(1+a_k)$ et $(1+b_k)$

et inclus dans l'orthan positif.

a) Les conditions c_j, $j = 1, 2, \ldots, 7$ et les définitions d_1 et d_2 entrainent les conditions 1, 1 bis, 2, 2 bis, 3 et 4 du tableau V.1. Nous avons : d_1 et c_7 entrainent 1

d_1, c_5, c_2 entrainent 1 bis

d_2, c_6 et c_1 entrainent 2

d_2, c_5, c_7, c_2 et c_1 entrainent 2 bis

d_2, c_5, c_7, c_4 et c_2 entrainent 3

d_1, c_5 et c_3 entrainent 4.

b) Réciproquement les conditions 1, 1 bis, 2, 2 bis, 3 et 4 et les définitions d_1 et d_2 entrainent les conditions c_j, $j = 1, 2, \ldots, 7$. Nous avons : 1, 1 bis, 2 bis et d_2 entrainent c_1

1 bis, 2, 2 bis et d_1 entrainent c_2

2, 2 bis, 4 et d_1 entrainent c_3

1, 1 bis, 3 et d_2 entrainent c_4

2 et 2 bis entrainent c_5

1, 1 bis, 2, 2 bis et d_2 entrainent c_6

1 et d_1 entrainent c_7

V.3 - Théorème

Pour $(\varepsilon, \varepsilon') = (+1, +1)$, nous nous donnons deux suites (a_k) et (b_k), $k \in \mathbb{N}$, telles que : $(a_k+1)(a_k+1) \geq 1$ pour tout entier positif k,

et : c_1) $b_k > -1$

c_2) a_k est le terme d'une série absolument convergente

c_3) b_k est le terme d'une série absolument convergente.

Nous déterminons les suites (α_k) et (δ_k), $k \in \mathbb{N}$, telles que :

$c_4)$ $\delta_k > 0$,

$c_5)$ $\alpha_k \leq (1+a_k)\ \delta_k$,

$c_6)$ $\alpha_k \geq \frac{1}{1+b_k}\ \delta_k$,

et nous définissons les suites (β_k) et (γ_k), $k \in \mathbb{N}$, par :

$d_1)$ $\beta_k = -\ \alpha_k + (1+a_k)\ \delta_k$,

$d_2)$ $\gamma_k = -\ \delta_k + (1+b_k)\ \alpha_k$.

Alors les conditions c_j, $j = 1, 2, \ldots, 6$ et les définitions d_1 et d_2 caractérisent les suites qui vérifient les conditions 1, 1 bis, 2, 2 bis, 3 et 4 du tableau V.1, i.e. les conditions suffisantes établies précédemment qui rendent H_k et B_k :

a) strictement définies positives

b) uniformément majorées.

Preuve :

Les conditions c_1, c_4 et c_6 impliquent que α_k est strictement positif d'où 1 bis, et alors avec la condition c_5, $1+a_k$ est strictement positif. Pour k fixé la propriété $(1+a_k)(1+b_k) \geq 1$ entraine que l'ensemble des couples (α_k, δ_k) vérifiant c_4, c_5 et c_6 est non vide, ces couples sont contenus dans le cône engendré par les deux demi-droites de pentes $(1+a_k)$ et $(1+b_k)^{-1}$, positives, et inclus dans l'orthan positif.

a) Les conditions c_j, $j = 1, 2, \ldots, 6$ et les définitions d_1 et d_2 entrainent les conditions 1, 1 bis, 2, 2 bis, 3 et 4 du tableau V.1 :

En effet nous avons :

d_1 et c_5 entrainent 1

c_4, c_6 et c_1 entrainent 1 bis

d_2, c_1 et c_6 entrainent 2

c_4 entraine 2 bis

d_1, c_4 et c_2 entrainent 3

d_2, c_4, c_6, c_3 et c_1 entrainent 4

b) Réciproquement, les conditions 1, 1 bis, 2, 2 bis, 3 et 4 et les définitions d_1 et d_2 entrainent les conditions c_j, $j = 1, 2, \ldots, 6$

1 bis, 2, 2 bis et d_2 entrainent c_1 et c_6

d_1, 2 bis et 3 entrainent c_2

d_2, 1 bis et 4 entrainent c_3

2 bis entraine c_4

d_1 et 1 entrainent c_5

V.4 - THEOREME

Pour $(\varepsilon, \varepsilon') = (1, -1)$, nous nous donnons deux suites (a_k) et (b_k), $k \in \mathbb{N}$, telles que : $a_k + b_k + a_k\, b_k \geq 0$ pour tout entier positif k et :

c_1) $b_k > -1$

c_2) a_k est le terme général d'une série absolument convergente,

c_3) b_k est le terme général d'une série absolument convergente.

Nous déterminons les suites (α_k) et (β_k), $k \in \mathbb{N}$, telles que :

c_4) $\alpha_k > 0$

c_5) $\beta_k \geq 0$

c_6) $\beta_k \leq (a_k + b_k + a_k\, b_k)\, \alpha_k$

et nous définissons les suites (δ_k) et (γ_k), $k \in \mathbb{N}$ par :

d_1) $\delta_k = (1+a_k)\, \alpha_k$,

d_2) $\gamma_k = \delta_k - \dfrac{\alpha_k+\beta_k}{1+b_k}$

Alors les conditions c_j, $j = 1, 2, \ldots, 6$ et les définitions d_1 et d_2 caractérisent les suites qui vérifient les conditions 1, 1 bis, 2, 2 bis, 3 et 4 du tableau V.1, i.e. les conditions suffisantes établies précédemment qui rendent H_k et B_k :

a) strictement définies positives

b) uniformément majorées.

Preuve :

Pour k fixé, la propriété $(a_k + b_k + a_k\, b_k) \geq 0$ entraine que l'ensemble des couples (α_k, β_k) vérifiant c_4, c_5 et c_6 est non vide, c'est le cône engendré par les demi-droites, de pentes $a_k + b_k + a_k\, b_k$ et zéro, inclus dans l'orthan positif.

a) Les conditions c_j, $j = 1, 2, \ldots, 6$, et les définitions d_1 et d_2 entrainent les conditions 1, 1 bis, 2, 2 bis, 3 et 4 du tableau V.1 :

c_5 entraine 1

c_4 entraine 1 bis

d_1, d_2, c_1, c_6 entrainent 2

d_2, c_1, c_4 et c_5 entrainent 2 bis

d_2, c_1, c_4, c_5 et c_3 entrainent 3

d_1, c_2 et c_4 entrainent 4

b) Réciproquement, les conditions 1, 1 bis, 2, 2 bis, 3 et 4 et les définitions d_1 et d_2 entrainent les conditions c_j, $j = 1, 2, \ldots, 6$:

1, 1 bis, 2 bis et d_2 entrainent c_1

1 bis, 4 et d_1 entrainent c_2

1, 1 bis, 2 bis, 3 et d_2 entrainent c_3

1 bis entraine c_4

1 entraine c_5

1, 1 bis, 2, 2 bis, d_1 et d_2 entrainent c_6

V.5 - THEOREME

Pour $(\varepsilon, \varepsilon') = (-1, 1)$ nous nous donnons deux suites (a_k), (b_k), $k \in \mathbb{N}$, telles que $(a_k + b_k + a_k b_k) \geq 0$ pour tout entier positif k et

c_1) $b_k > -1$

c_2) a_k est le terme général d'une série absolument convergente,

c_3) b_k est le terme général d'une série absolument convergente.

Nous déterminons les suites (δ_k) et (γ_k), $k \in \mathbb{N}$, telles que :

c_4) $\delta_k > 0$

c_5) $\gamma_k \geq 0$

c_6) $\gamma_k \leq (a_k + b_k + a_k b_k)\, \delta_k$

et nous définissons les suites (α_k) et (β_k), $k \in \mathbb{N}$, par :

d_1) $\alpha_k = (1 + a_k)\, \delta_k$

d_2) $\beta_k = \alpha_k - \dfrac{\delta_k+\gamma_k}{1+b_k}$

Alors les conditions c_j, $j = 1, 2, \ldots, 6$ et les définitions d_1 et d_2 caractérisent les suites qui vérifient les conditions 1, 1 bis, 2, 2 bis, 3 et 4 du tableau V.1, i.e. les conditions suffisantes établies précédemment qui rendent H_k et B_k :

a) strictement définies positives,

b) uniformément majorées.

Preuve :

Pour k fixé, la propriété $(a_k + b_k + a_k b_k) \geq 0$ entraine que l'ensemble des couples (δ_k, γ_k) vérifiant c_4, c_5 et c_6 n'est pas vide, ces couples sont contenus dans le cône engendré par les demi-droites de pentes, positives, $a_k + b_k + a_k b_k$ et zéro, inclus dans l'orthan positif.

La preuve se déduit de la remarque suivante : si dans le tableau V.1 nous remplaçons β, γ, δ, α respectivement par γ, β, α, δ dans les conditions 1, 1 bis, 2, 2 bis, 3 et 4 du cas $(\varepsilon, \varepsilon') = (1, -1)$, nous obtenons les conditions 2, 2 bis, 1, 1 bis, 4 et 3 du cas $(\varepsilon, \varepsilon') = (-1, 1)$.

Nous avons caractérisé les suites réalisant les conditions suffisantes qui rendent les matrices B_k, $k \in \mathbb{N}$, bien conditionnées. Nous allons établir avec une condition supplémentaire que la suite $\{B_k, k \in \mathbb{N}\}$ est convergente. Pour cela nous allons utiliser le lemme suivant (élaboré à la suite d'une discussion avec A. Draux).

V.6 - LEMME

Soit (r_k), $k \in \mathbb{N})$, une suite dans $\mathbb{R}^m$ (ou dans un Banach) définie par $r_{k+1} = \nu_k r_k + t_k$ où ν_k est un réel et t_k un élément $\mathbb{R}^m$ vérifiant :

(i) $1 - \nu_k$, $k \in \mathbb{N}$, est le terme général d'une série absolument convergente,

(ii) t_k, $k \in \mathbb{N}$, est le terme général d'une famille absolument sommable,

(iii) la suite (r_k), $k \in \mathbb{N}$, est uniformément majorée, i.e. il existe un réel ω strictement positif tel que

$$||r_k|| \leq \omega \text{ pour tout } k \text{ dans } \mathbb{N}$$

Alors la suite (r_k), $k \in \mathbb{N}$, est convergente.

Preuve :

On a l'égalité $\sum_{i=1}^{k} (\nu_{i-1} r_{i-1} + t_{i-1})$, elle s'écrit encore :

$$r_k = - \sum_{i=1}^{k-1} r_i + \sum_{i=1}^{k} (\nu_{i-1} \, r_{i-1} + t_{i-1})$$

ou

$$r_k = - \sum_{i=1}^{k-1} r_i + \sum_{i=2}^{k} \nu_{i-1} \, r_{i-1} + \sum_{i=o}^{k-1} t_i + \nu_o \, r_o$$

c'est-à-dire :

$$r_k = \nu_o \, r_o + \sum_{i=1}^{k-1} (\nu_i - 1) \, r_i + \sum_{i=o}^{k-1} t_i \qquad (V.6.1)$$

Les hypothèses (i) à (iii) permettent de conclure.

Les matrices B_{k+1} définies au chapitre III par (III.5.1), (III.5.2), (III.5.3) et (III.5.4) peuvent s'écrire sous la forme générale suivante :

$$B_{k+1} = \alpha'_k \, B_k + \beta'_k \, u_k \, u_k^t + \gamma'_k \, v_k \, v_k^t \qquad (V.7.0)$$

V.7 - PROPOSITION

Si les suites $\{\alpha_k, k \in \mathbb{N}\}$,$\{\beta_k, k \in \mathbb{N}\}$, $\{\gamma_k, k \in \mathbb{N}\}$ et $\{\delta_k, k \in \mathbb{N}\}$ vérifient les conditions du tableau V.1 et si de plus $\sum_{k=o}^{\infty} |\alpha_k - \delta_k| < +\infty$ et $\underline{\delta} \leq \delta_k \leq \bar{\delta}$ pour tout entier k, alors les suites de matrices $\{B_k, k \in \mathbb{N}\}$ définies par (III.5.1), (III.5.2), (III.5.3) et (III.5.4) sont convergentes.

Preuve :

Elle se fait en trois étapes. Les deux premières seront faites pour l'une des suites de matrices définies en III.5, par exemple, la suite définie par (III.5.1). Les résultats restent vrai pour les autres.

Compte tenu de (V.7.0) et de (III.5.1) nous avons :

$$u_k\, u_k^t = \frac{(B_k s_k)(B_k s_k)^t}{(B_k s_k, s_k)} \;,\; \varepsilon\, \frac{\beta_k}{\delta_k} = \beta'_k \;,\; -\varepsilon'\, \frac{\gamma_k}{\delta_k} = \gamma'_k \text{ et}$$

$$v_k\, v_k^t = \frac{y_k y_k^t}{(\delta_k + \varepsilon' \gamma_k)\, \langle H_k^* y_k, y_k\rangle + \delta_k\, \langle s_k, y_k\rangle}$$

a) La suite $\{u_k\, u_k^t,\ k \in \mathbb{N}\}$ est uniformément majorée.

Comme dans $\mathbb{R}^{n^2}$ toutes les normes sont équivalentes, nous montrons que cette suite est uniformément majorée au sens de la norme de Frobenius, i.e. $||A||_F^2 = \mathrm{tr}(AA^T)$.

Nous avons $||u_k u_k^t||_F = ||u_k||^2 = \dfrac{||B_k s_k||^2}{\langle B_k s_k, s_k\rangle}$

Comme B_k est symétrique définie positive, nous avons d'après la proposition I.7

$$\frac{||B_k s_k||^2}{\langle B_k s_k, s_k\rangle} \le ||B_k|| \le M' \tag{V.7.1}$$

d'où le résultat.

b) La suite $\{v_k\, v_k^t,\ k \in \mathbb{N}\}$ est uniformément majorée, nous avons :

$$||v_k v_k^t||_F = \frac{||y_k||^2}{(\delta_k + \varepsilon' \gamma_k)\, \langle H_k^* y_k,\ y_k\rangle + \delta_k\, \langle s_k, y_k\rangle} \;,$$

Comme $\delta_k\, \langle s_k, y_k\rangle$ est positif, nous avons :

$$||v_k v_k^t||_F = \frac{||y_k||^2}{(\delta_k + \varepsilon' \gamma_k)\, \langle H_k^* y_k,\ y_k\rangle} \tag{V.7.2}$$

D'après les théorèmes précédents V.2, V.3, V.4, V.5, la condition $\underline{\delta} \le \delta_k \le \bar{\delta}$ entraine qu'il existe deux constantes strictement positives telles que $\underline{\alpha} \le \alpha_k \le \bar{\alpha}$. Nous déduisons, alors, des conditions données dans les colonnes 3 et 4 du tableau V.1 que β_k et γ_k sont termes généraux de séries convergentes. Par suite (V.7.2) entraine

$$||v_k v_k^t||_F \leq \frac{||y_k||^2}{\frac{1}{2}\,\underline{\delta}\,\langle H_k^* y_k, y_k\rangle}$$

or H_k^*, qui vaut soit $\tilde{H}_k$ ou $\underset{\sim}{H}_k$ définies au chapitre II, est définie positive. Comme β_k est terme général d'une série convergente et que α_k est uniformément bornée (voir plus haut), $\tilde{B}_k$ et $\underset{\sim}{B}_k$ sont uniformément majorées (cf. les prop. II.3.2, II.4.2 et III.6). Par suite la plus petite valeur propre de H_k^* est uniformément minorée et strictement positive. D'où

$$||v_k v_k^t||_F \leq \frac{2}{\underline{\delta}}\, C \text{ où } C \text{ est une constante.}$$

Remarque :

La démonstration pour les autres formules se fait de manière identique car nous obtenons toujours l'inégalité (V.7.2)

c) Les suites de matrices $\{B_k,\ k \in \mathbb{N}\}$ définies en III.2 sont convergentes. Posons $D_k = \beta'_k\, u_k u_k^t + \gamma'_k\, v_k v_k^t$. (V.7.0) s'écrit alors $B_{k+1} = \alpha'_k\, B_k + D_k$. Comme nous avons $\sum |\alpha_n - \delta_n| < +\infty$, $\delta_k \geq \underline{\delta}$ et $\alpha'_k = \frac{\alpha_k}{\delta_k}$ l'hypothèse (i) du lemme V.6 est satisfaite. Les résultats a) et b) précédents et le fait que γ_k et β_k sont les termes de séries convergentes impliquent la condition ii) est satisfaite. Enfin le choix global des paramètres réalise l'hypothèse (iii) du même lemme. Donc il existe une matrice B^* telle que :

$$\lim_{k\to+\infty} B_k = B^*.$$

BIBLIOGRAPHIE

[1] ALLWRIGHT, J.C., *"Global convergence rates for descent algorithms : Improved results for strictly connex problems"*, J.O.T.A., Vol. 26, n° 3, pp. 367-372, 1978.

[2] ARMIJOL, L., *"Minimization of functions having continuous partial derivatives"*, Pacific J. Maths, Vol. 16, pp. 1-3, 1968.

[3] BERTSEKAS, D.P., *"Combined primal-dual for penalty methods for constrained minimization"*, S.I.A.M. J. Control, Vol. 13, pp. 521-544, 1975.

[4] BERTSEKAS, D.P., *"Multiplier methods : A survey"*, Automatica, Vol. 26, pp. 133-145, 1976.

[5] BOLAND, W.R. and KOWALIK, J.S., *"Extended congujate-gradient methods with restarts"*, J.O.T.A., Vol. 28, n° 1, pp. 1-9, 1979.

[6] BROYDEN, C.G., *"Quasi-Newton methods and their application to function minimization"*, Maths. Comput., Vol. 21, pp. 368-381, 1967.

[7] BROYDEN, C.G., *"The convergence of a class of double rank minimization algorithme"*, Parties I et II, J.I.M.A., Vol. 6, pp. 76-90 et 222-231, 1970.

[8] CAROLL, C.W., *"The created response surface technique for optimizing nonlinear restrained systems"*, Ops. Res., Vol. 9, pp. 269-276, 1961.

[9] CAUCHY, A. *"Méthode générale pour la résolution des systèmes d'équations simultannées"*, Comp. Rend. Acad. Sci., pp. 536-538, 1847.

[10] CORNWELL, L.W., KOCMAN, M.G., PROSSER, M.F., "*Computational experience with Davidon's Least-square algorithm*", J.O.T.A., Vol. 31, n° 1, pp. 27-39, 1980.

[11] DAVIDON, W.C., "*Variable metric method for minimization*", AEC Res. and Develop. Rep. ANL-5990 (Rev.), 1959.

[12] DENNIS, J.E. J.R., "*On some methods based on Broyden's secant approximation to the hessian*", in : Numerical Method for Nonlinear Optimization, F.A. Lootsma ed., pp. 19-34, Academic Press, London (1972).

[13] DENNIS, J.E. J.R., "*On the convergence of Broyden's method for nonlinear systems of equations*", Maths. Comput., Vol. 25, pp. 559-567, 1971.

[14] DENNIS, J.E. J.R., and MORE, J.J., "*A characterization of superlinear convergence and its applications to quasi-Newton methods*", Maths. Comput., Vol. 28, pp. 549-560, 1974.

[15] DENNIS, J.E. J.R., and MORE, J.J., "*Quasi-Newton methods, motivation and theory*", S.I.A.M. Review, Vol. 19, n° 1, pp. 46-89, 1977.

[16] DI PILLO, G., and GRIPPO, L., "*A new class of augmented lagrangians in nonlinear programming*", S.I.A.M. J. Control and Optimization, Vol. 17, n° 5, pp. 618-628, 1972.

[17] DIXON, L.C.W., "*Variable metric algorithms : Necessary and sufficient conditions for identical behavior on nonquadratic functions*", J.O.T.A., Vol. 10, pp. 34-40, 1972.

[18] FIACCO, A.V., and Mc CORMICK, G.P., "*Nonlinear Programming : Sequential unconstrained minimization techniques*", John Wiley, New-York, 1968.

[19] FIOROT, J.C., et EL HALLABI, M. *"Sur la convergence de méthodes utilisant des formules paramétrées de métrique variable"*, Publication n° 53, Laboratoire de Calcul, Université des Sciences et Techniques de Lille I, 1981.

[20] FIOROT, J.C., and EL HALLABI, M. *"On the convergence of methods using parametric variable metric formulae"*, NATO ARI on Nonlinear Optimization, Cambridge, England, July 13th - 24th, 1981.

[21] FIOROT, J.C., et EL HALLABI, M. *"Quelques essais numériques sur une famille de métrique variable paramétrée. Comparaison avec la B.F.G.S."*, Publication n° 54, Laboratoire de Calcul, Université des Sciences et Techniques de Lille I, 1981.

[22] FLETCHER, R., *""A new approach to variable metric algorithms"*, Comp. J., Vol. 13, pp. 317-322, 1970.

[23] FLETCHER, R., *"Unconstrained Optimization"*, Vol. 1, Joh, Wiley Sons, Ltd, 1980.

[24] FLETCHER, R., and POWELL, M.J.D., *"A rapidly convergent descent method for minimization"*, Comp. J., Vol. 6, pp. 163-168, 1963.

[25] FLETCHER, R., and REEVES, C.M., *"Function minimization by conjugate gradient"*, Comp. J., Vol. 7, pp. 149-154, 1964.

[26] GOLDFARB, D., *"A family of variable metric method derived by variational means"*, Maths. Comput., Vol. 24, pp. 23-26, 1970.

[27] GOLDSTEIN, AA., and PRICE, J.F., *"An effective algorithms for minimization"*, Num. Math., Vol. 10, pp. 184-189, 1967.

[28] HESTNES, M.R., and STEIFEL, E. *"Methods of conjugate gradients for solving linear systems"*, J. Res. Nat. Bur-Stand., Vol. 49, pp. 409-436, 1952.

[29] HESTENES, M.R., *"Multipliers and gradients methods"*, Comuting methods in optimization problems, Vol. 2, pp. 143-163, Zadeh, L.A., Neustadt, L.W., Balakrishnan, A.V. eds., A.P., New-York, 1969.

[30] HUANG, H.Y., *"Unified approach to quadratically convergent algorithms for function minimization"*, J.O.T.A., Vol. 5, n° 6, pp. 405-423, 1970.

[31] HUANG, H.Y., and LEVY A.V., *"Numerical experiments on quadratically convergent algorithms for function minimization"*, J.O.T.A., Vol. 6, n° 3, pp. 269-282, 1970.

[32] KANTOROVICH, L., and AKILOV, G., *"Functional analysis in normed spaces"*, Fizmatgiz, Moscow, trad. par Brown D. et Roberston A., Pergamon Press, Oxford, 1964.

[33] KORGANOF, A., et PAVEL, M., *"Eléments de théorie des matrices caraées et rectangulaires en analyse numérique"*, Eds. Dunod, Paris, 1967.

[34] LENARD, M., *"Practical convergence conditions for the Davidon-Fletcher-Powell method"*, Math. Prog., Vol. 9, pp. 69-86, 1975.

[35] LOOTSMA, F.A., *"Boundary properties of penalty functions for constrained minimization"*, Thèse à l'Université d'Eindhoven, The Netherlands (1970).

[36] MIELE, A., MOSELEY, A., LEVY, A.V., and COGGINS, G.M., *"On the method of multipler for mathematical programming problems"*, J.O.T.A., Vol. 10, pp. 1-33, 1972.

[37] MORE, J.J., GARBOW, B.S., HILLSTROM, K.E., *"Testing unconstrained optimization software"*, ACM Trans. Maths. Soft., Vol. 7, n° 1, pp. 17-41, 1981.

[38] MUKAI, H., *"Readly implementable conjugate gradient methods"*, Math. Prog. Vol. 17, pp. 298-319, 1979.

[39] ORTEGA, J.M., and REINBOLDT, W.C., *"Iterative solution of non-linear equations in several variables"*, A.P., New-York, 1970.

[40] POLAK, E., *"Computational methods in Optimization : A Unified Approach"*, A.P., New-York, 1971.

[41] POLAK, E., et RIBIERE, G., *"Note sur la convergence de méthodes de directions conjuguées"*, Rev. Fran. Inform. Rech. Oper., Vol. 16, pp. 35-43, 1969.

[42] POWELL, M.J.P., *"A method for nonlinear constraints in minimization problems"*, Optimization, R. Fletcher, Eds, A.P., New-York, pp. 283-298, 1969.

[43] POWELL, M.J.D., *"Some global convergence properties of a variable metric algorithm for minimization without exact lines searches"*, Non linear Programming, S.I.A.M. AMS Proceeding, Vol. 9, American Mathematical Society, Providence, R.I., 1976.

[44] ROCKAFELLAR, R.T., *"The multiplier method of Hestenes and Powell applied to convexe programming"*, J.O.T.A., Vol. 12, pp 555-562, 1973.

[45] RITTER, K., *"On the rate of superlinear convergence of a class of variable metric methods"*, Num. Math., Vol. 35, pp. 293-313, 1980.

[46] SHANNO, D.F., *"Conditionning of quasi-Newton methods for function minimization"*, Maths. Comp., Vol. 24, pp. 647-656, 1970.

[47] STENHAUG, T., Communication orale, 1981.

[48] WOLFE, P., *"Convergence conditions for ascent methods"*, S.I.A.M. Review, Vol. 11, pp. 226-235, 1969.

[49] ZANGWILL, W.I., *"Non linear programming via penalty function"*, management Sc., Vol. 13, pp. 344-358, 1967.

C H A P I T R E VI

ESSAIS NUMERIQUES

Les fonctions testées sont quelques-unes de celles couramment utilisées dans la littérature. Un petit nombre d'expériences numériques ont été reproduites ici, dans FIOROT et EL HALLABI [21] on trouvera un échantillon plus important de tels essais. Ils ont été effectués sur CII 80. Nous avons utilisé le code M1QN1 de la bibliothèque MODULOPT fourni par C. Lemaréchal (INRIA), il est équivalent au code VA13A écrit par M.J.D. Powell pour la bibliothèque de Harwell. Ce code a été conçu pour la BFGS.

La formule (III.5.1) a été essentiellement utilisées pour les valeurs suivantes :

$\varepsilon = \varepsilon' = -1$, $\alpha_k = \delta_k = 1$, $\beta_k = \gamma_k = \eta^k$ où η est un réel de $]0, 1[$.

Les tests d'arrêt sont :

i) $||g_k||^2 \leq 10^{-25}$

ii) ou $f(x_k + \lambda_k\ \phi_k) > f(x_k)$ pour $\lambda_k\ ||d_k|| \leq 10^{-16}$.

Dans les tableaux ci-après, les symboles IT, NF sont respectivement les nombres d'itérations et de calcul de la fonction (et du gradient). La comparaison des résultats est systématiquement faite avec la formule de la BFGS. Le temps de calcul est donné en dixièmes de seconde.

Fonction test de : Rosenbrock [5]

$$f(x) = 100(x_2-x_1^2)^2 + (1-x_1)^2 + 90(x_4-x_3^2)^2$$

$$+ (1-x_3)^2 + 10.1[(x_2-1)^2 + (x_4-1)^2] + 19.8(x_2-1)(x_4-1)$$

Point optimal théorique $\overset{*}{x}(i) = 1, \quad i = 1, 2, 3, 4.$

Valeur optimale théorique de $f = 0.$

Point de départ $x_0(1) = x_0(3) = -3.$

$x_0(2) = x_0(4) = -1.$

Valeur initiale de $f = 0.2\ 10^5.$

Formule $(\mathrm{III}.5.1)$ avec

$\varepsilon = \varepsilon' = -1$

$\alpha_k = \delta_k = 1$

$\beta_k = \gamma_k = 1/k^{1.25}$

B. F. G.S

IT	NF	VALEUR DE F (B. F. G.S)	NF	VALEUR DE F (Formule (III.5.1) avec …)
0	1	.191920000000000D+05	1	.191920000000000D+05
1	2	.118721571221521D+05	2	.118721571221521D+05
2	4	.948169353118564D+04	4	.948169353118564D+04
3	5	.551271595802321D+04	7	.944453046414249D+04
4	6	.981172468770879D+03	10	.744689607452567D+04
5	7	.157987735297254D+03	11	.417667905256003D+04
6	8	.181212012678833D+02	12	.487032633939048D+03
7	9	.111245742033431D+02	15	.106486330464413D+02
8	10	.106153003833355D+02	20	.836806526143234D+01
9	11	.100212277374460D+02	21	.507483190110897D+01
10	12	.972361986310809D+01	23	.412112334574267D+01
11	13	.958481100183761D+01	27	.411636973195800D+01
12	14	.943119763207672D+01	31	.411636508687956D+01
13	15	.917862135205089D+01	33	.411636374753117D+01
14	16	.877611608175349D+01	37	.411414952450069D+01
15	17	.823572440243176D+01	39	.410312968239621D+01
16	18	.795213966143415D+01	41	.402547924025291D+01
17	19	.784466742076066D+01	44	.402405836880548D+01
18	20	.784092973672404D+01	46	.402378954330568D+01
19	21	.784084418081088D+01	49	.402378777381616D+01
20	22	.784082929956590D+01	51	.402378732643564D+01
21	23	.784081005888527D+01	55	.402321016945229D+01
22	25	.784070995933814D+01	57	.401915638501314D+01
23	30	.783091371917631D+01	59	.399531257718957D+01
24	31	.781602769664108D+01	61	.376792851058896D+01
25	32	.779132910158152D+01	62	.340412307781651D+01
26	35	.769128513205125D+01	63	.304891069170424D+01
27	36	.763281056939005D+01	65	.293130710545409D+01
28	39	.750739467734747D+01	66	.274665171136403D+01
29	40	.744827948732897D+01	67	.270089322661219D+01
30	42	.708169058872875D+01	68	.262957817018794D+01
31	43	.693576701887294D+01	69	.253311454666366D+01
32	46	.678091363283215D+01	70	.245211458422010D+01
33	47	.653556315027896D+01	72	.234745376743408D+01
34	48	.616639117078953D+01	73	.218715620600044D+01
35	49	.587573810133962D+01	74	.196841062179819D+01
36	53	.506788217079266D+01	75	.169466153872453D+01
37	54	.482437914655729D+01	77	.158528796523897D+01

B.F.G.S

Formule (III.5.1) avec

$\varepsilon = \varepsilon' = -1$, $\alpha_k = \delta_k = 1$

$\beta_k = \gamma_k = 1/k^{1.25}$

IT	NF	VALEUR DE F
38	55	.449321110226852D+01
39	56	.415801380493768D+01
40	57	.385690240978354D+01
41	58	.354490644206732D+01
42	59	.329105101964988D+01
43	60	.299902436763142D+01
44	62	.285584971590158D+01
45	63	.263147174677473D+01
46	64	.234815183837570D+01
47	66	.217950442287283D+01
48	67	.206127548334131D+01
49	68	.191185612241971D+01
50	69	.167969096158350D+01
51	71	.158867972739197D+01
52	72	.143608561421916D+01
53	73	.120297122365755D+01
54	74	.960143882614493D+00
55	76	.830218662065349D+00
56	77	.701478631366175D+00
57	78	.509806075491892D+00
58	79	.418224115077422D+00
59	80	.310076889190983D+00
60	82	.266151972736550D+00
61	84	.704975179361515D-01
62	85	.392245968657491D-01
63	86	.166732649648412D-01
64	87	.669701830950759D-02
65	88	.212804965361407D-02
66	89	.744832879753420D-03
67	90	.267942319247620D-04
68	91	.316247324980623D-05
69	92	.840987220105416D-06
70	93	.654240843722495D-07
71	94	.191646623352128D-08
72	95	.136264952630658D-10
73	96	.132493292679025D-12
74	97	.254246829910434D-14
75	98	.431646198540373D-19
76	99	.787421443918310D-23
77	100	.114924518427551D-25
78	101	.873170414466509D-29

IT	NF	VALEUR DE F
38	79	.149201912003521D+01
39	81	.116439649367656D+01
40	83	.960774076642629D+00
41	84	.929128747630622D+00
42	86	.621317508561012D+00
43	87	.440501725634249D+00
44	88	.384959677094873D+00
45	89	.304728218247728D+00
46	90	.254330868623621D+00
47	92	.220745237928106D+00
48	94	.610697772513574D-01
49	96	.393586043101170D-01
50	97	.163688313178173D-01
51	99	.479660660162418D-02
52	101	.311404693891764D-02
53	102	.727783656121216D-03
54	104	.318408421573274D-04
55	107	.231965477358915D-05
56	109	.173508128711815D-06
57	111	.103868048948370D-07
58	113	.658143965629716D-08
59	114	.308214813645510D-08
60	115	.448890282779743D-09
61	117	.347454240601106D-10
62	119	.138945133978910D-11
63	122	.204164496073723D-13
64	125	.166409033592961D-14
65	127	.404922733365312D-16
66	130	.282720843612030D-18
67	133	.504049201792093D-19
68	135	.613319139248641D-26
69	141	.308217816431165D-27
70	144	.895064386074302D-29

Formule (III.5.1) avec

B.F.G.S

$\varepsilon = \varepsilon' = -1$, $\alpha_k = \delta_k = 1$

$\beta_k = \gamma_k = 1/k^{1,25}$

NORME GRADIENT CARRE

.74646397030036OD-26	.10148141863920 5D-25

TEMPS CALCUL

4	4

POINT OPTIMAL

X(1)= .1000000000000000D+01	X(1)= .1000000000000000D+01
X(2)= .1000000000000000D+01	X(2)= .9999999999999999D+00
X(3)= .9999999999999999D+00	X(3)= .1000000000000000D+01
X(4)= .9999999999999998D+00	X(4)= .1000000000000000D+01

Formule (III.5.1) avec

$\varepsilon = \varepsilon' = -1$

$\alpha_k = \delta_k = 1$

$\beta_k = \gamma_k = (0.999)^k$

B. F. G. S

IT	NF	VALEUR DE F	IT	NF	VALEUR DE F
0	1	.191920000000000D+05	0	1	.191920000000000D+05
1	2	.118721571221521D+05	1	2	.118721571221521D+05
2	4	.948169353118564D+04	2	3	.293270746258427D+04
3	5	.551271595802321D+04	3	5	.196925364366256D+03
4	6	.981172468770879D+03	4	6	.114692596010112D+03
5	7	.157987735297254D+03	5	7	.513602104813088D+02
6	8	.181212012678833D+02	6	8	.302925463038664D+02
7	9	.111245742033431D+02	7	10	.278468760843142D+02
8	10	.106153003833355D+02	8	12	.240818116620959D+02
9	11	.100212277374460D+02	9	15	.145259647430014D+02
10	12	.972361986310809D+01	10	16	.711709319415641D+01
11	13	.958481100183761D+01	11	18	.364962198988523D+01
12	14	.943119763207672D+01	12	20	.302939566025583D+01
13	15	.917862135205089D+01	13	22	.301975936116656D+01
14	16	.877611608175349D+01	14	23	.301020182151647D+01
15	17	.823572440248176D+01	15	25	.297658288408944D+01
16	18	.795213966143415D+01	16	27	.277542430863097D+01
17	19	.784466742076066D+01	17	28	.255413534655550D+01
18	20	.784092973672404D+01	18	29	.228428122515563D+01
19	21	.784084418081088D+01	19	31	.214062358034319D+01
20	22	.784082929956590D+01	20	33	.209354442019123D+01
21	23	.784081005888527D+01	21	34	.204173401955146D+01
22	25	.784070995933814D+01	22	36	.193851087135177D+01
23	30	.783091371917631D+01	23	38	.147426112175755D+01
24	31	.781602769664108D+01	24	39	.103501608352966D+01
25	32	.779132910158152D+01	25	41	.891504609634490D+00
26	35	.769128513205125D+01	26	42	.678143530214940D+00
27	36	.763281056939005D+01	27	44	.635256227868527D+00
28	39	.750739467734747D+01	28	46	.381202186130225D+00
29	40	.744827948732897D+01	29	48	.240672885404000D+00
30	42	.708169058872875D+01	30	49	.121632126156454D+00
31	43	.693576701887294D+01	31	51	.631919435147899D-01
32	46	.678091363283215D+01	32	53	.445000581315894D-01
33	47	.653556315027896D+01	33	54	.226323418118906D-01
34	48	.616639117078953D+01	34	56	.169684443433649D-01
35	49	.587573810133962D+01	35	57	.886091389696182D-02
36	53	.506788217079266D+01	36	58	.204437163373319D-02
37	54	.482437914655729D+01	37	60	.720229065502085D-04

B.F.G.S

Formule (III.5.1) avec

$\varepsilon = \varepsilon' = -1$, $\alpha_k = \delta_k = 1$

$\beta_k = \gamma_k = (0.999)^k$

IT	NF	VALEUR DE F
38	55	.449321110226852D+01
39	56	.415801380493768D+01
40	57	.385690240978354D+01
41	58	.354490644206732D+01
42	59	.329105101964988D+01
43	60	.299902436763142D+01
44	62	.285584971590158D+01
45	63	.263147174677473D+01
46	64	.234815183837570D+01
47	66	.217950442287283D+01
48	67	.206127548334131D+01
49	68	.191185612241971D+01
50	69	.167969096158350D+01
51	71	.158867972739197D+01
52	72	.143608561421916D+01
53	73	.120297122365755D+01
54	74	.960143882614493D+00
55	76	.830218662065349D+00
56	77	.701478631366175D+00
57	78	.509806075491892D+00
58	79	.418224115077422D+00
59	80	.310076889190983D+00
60	82	.266151972736550D+00
61	84	.704975179361515D-01
62	85	.392245968657491D-01
63	86	.166732649648412D-01
64	87	.669701830950759D-02
65	88	.212804965361407D-02
66	89	.744832879753420D-03
67	90	.267942319247620D-04
68	91	.316247324980623D-05
69	92	.840987220105416D-06
70	93	.654240843722495D-07
71	94	.191646623352128D-08
72	95	.136264952630658D-10
73	96	.132493292679025D-12
74	97	.254246829910434D-14
75	98	.431646198540373D-19
76	99	.787421443918310D-23
77	100	.114924518427551D-25
78	101	.873170414466509D-29

IT	NF	VALEUR DE F
38	62	.437917919761314D-04
39	63	.599190802609435D-05
40	65	.970941759420180D-06
41	67	.820676481727754D-08
42	69	.320129175919194D-08
43	71	.577403982275680D-09
44	73	.630594550877653D-14
45	76	.407833072558062D-15
46	78	.195067959759122D-16
47	81	.981547318195159D-21
48	84	.398730140977380D-24
49	87	.196753003118600D-30

B.F.G.S

Formule (III.5.1) avec

$\varepsilon = \varepsilon' = -1$, $\alpha_k = \delta_k = 1$

$\beta_k = \gamma_k = (0.999)^k$

NORME GRADIENT CARRE

.74646397030036OD-26 .35051833686667OD-28

TEMPS CALCUL

4 2

POINT OPTIMAL

X(1)= .100000000000000D+01	X(1)= .100000000000000D+01
X(2)= .100000000000000D+01	X(2)= .100000000000000D+01
X(3)= .999999999999999D+00	X(3)= .100000000000000D+01
X(4)= .999999999999998D+00	X(4)= .100000000000000D+01

Fonction test de : MIELE [31]

$$f(x) = [\exp(x_1) - x_2]^4 + 100(x_2-x_3)^6 + tg^4(x_3-x_4) + x_1^8 + (x_4-1)^2.$$

Point optimal théorique $\overset{*}{x}(1) = 0.$

$\overset{*}{x}(i) = 1 \quad i = 2, 3, 4.$

Valeur optimale théorique de f = 0.

Point de départ $\overset{o}{x}(1) = 5.$

$\overset{o}{x}(i) = 10 \quad i = 2, 3, 4.$

Valeur initiale de $f = 0.367\ 10^9.$

Pour cette fonction, avec le même point de départ, la formule (III.5.1) utilisée avec différents jeux de paramètres donne différents points stationnaires dont le minimum absolu, seul point fourni par la BFGS.

B. F. G. S

Formule (III.5.1) avec

$\varepsilon = \varepsilon' = 1$

$\alpha_k = \delta_k = 1$

$\beta_k = \gamma_k = (0.9)^k$

IT	NF	VALEUR DE F
0	1	.367427433351379D+09
1	2	.411409726240057D+08
2	4	.111844664954110D+08
3	5	.645975235728765D+07
4	6	.281022397094109D+07
5	7	.135702308502179D+07
6	8	.622701126432615D+06
7	9	.290776549027957D+06
8	10	.138463725762094D+06
9	11	.727334535240319D+05
10	12	.470688221183263D+05
11	13	.389443973834940D+05
12	14	.367723913807655D+05
13	15	.356530228575160D+05
14	16	.338203016949014D+05
15	18	.252875483933167D+05
16	20	.746087215344569D+04
17	21	.521442414779559D+04
18	23	.264020782962858D+04
19	24	.178322511919431D+04
20	25	.164479826286243D+04
21	26	.162177833265397D+04
22	27	.161882630226520D+04
23	28	.161489617351539D+04
24	30	.159672550246553D+04
25	33	.939986082877449D+03
26	34	.796621561670043D+03
27	35	.746797413882691D+03
28	36	.704657463920658D+03
29	37	.642095076112820D+03
30	39	.585987620404651D+03
31	41	.585061680700559D+03
32	42	.584174203608074D+03
33	43	.583935078793080D+03
34	45	.583433978406307D+03
35	48	.547793140043782D+03
36	51	.522709070804069D+03
37	53	.522554424180662D+03

IT	NF	VALEUR DE F
0	1	.367427433351379D+09
1	2	.411409726240057D+08
2	4	.416798093779681D+05
3	14	.768956944052443D+03
4	20	.453930506220659D+03
5	22	.375028291236582D+03
6	26	.374942778163706D+03
7	30	.374942777295166D+03
8	36	.374883743987487D+03
9	40	.325706336164598D+03
10	41	.310065089130642D+03
11	43	.304823787614742D+03
12	45	.302462428150771D+03
13	48	.302374466003475D+03
14	51	.302374156960953D+03
15	52	.302373963129955D+03
16	54	.302373931795567D+03
17	57	.302373931008973D+03
18	65	.295770309597118D+03
19	67	.216688611218805D+03
20	69	.203003144811065D+03
21	70	.188742713330605D+03
22	72	.177341607494914D+03
23	74	.173857950208054D+03
24	76	.173773415503847D+03
25	77	.173662698396828D+03
26	79	.173648240766576D+03
27	81	.173647414357017D+03
28	83	.173647226435056D+03
29	87	.173647226201125D+03
30	89	.173647226088431D+03
31	97	.173266006068004D+03
32	100	.110493335996789D+03
33	101	.454424793238800D+02
34	102	.433532455903307D+01
35	106	.316687907717057D+00
36	110	.142673855691412D+00
37	111	.289422143759619D-01

B. F. G. S

IT	NF	VALEUR DE F
38	54	.52233293419776 8D+03
39	55	.52228155339177 8D+03
40	57	.52227677674540 6D+03
41	61	.52032647273130 1D+03
42	64	.39351525574498 1D+03
43	66	.32489574543844 4D+03
44	67	.26575758190501 1D+03
45	68	.17016822948463 0D+03
46	69	.66296068438591 5D+02
47	70	.30572778995156 5D+02
48	71	.10518292172234 8D+02
49	72	.87798778305557 0D+01
50	73	.76936790043871 3D+01
51	75	.54172244489106 3D+01
52	76	.31566300498246 1D+01
53	77	.12393389103851 8D+01
54	78	.59057956539263 2D+00
55	79	.44634186213303 7D+00
56	80	.42600673492301 8D+00
57	81	.40152739066062 7D+00
58	82	.39308703595210 4D+00
59	84	.35267381421500 7D+00
60	86	.14181713694403 0D+00
61	87	.67194840260697 7D-01
62	88	.29424445904313 5D-01
63	89	.18351482929615 8D-01
64	90	.10533443094274 9D-01
65	91	.79765820897674 1D-02
66	92	.75629750177693 9D-02
67	93	.75116158307446 8D-02
68	94	.74563292338613 7D-02
69	96	.72245851530543 6D-02
70	98	.54533014973352 0D-02
71	99	.46898476702208 1D-02
72	100	.42576818269752 8D-02
73	101	.35565129008868 4D-02
74	102	.24549056351725 8D-02
75	103	.12315962160498 4D-02
76	104	.95235109894098 5D-03
77	105	.80715561703519 7D-03
78	107	.35348422215114 0D-03
79	108	.20468891967356 6D-03

Formule (III.5.1) avec

$\varepsilon = \varepsilon' - 1$, $\rho_k \cdot \delta_k = 1$

$\beta_k = \gamma_k = (0.9)^k$

IT	NF	VALEUR DE F
38	113	.65877535818354 2D-02
39	115	.40898621397186 5D-02
40	116	.11841341778767 1D-02
41	118	.38958337232360 3D-03
42	120	.94280508001332 6D-04
43	122	.10644072678828 6D-04
44	124	.14614840077321 5D-05
45	127	.19262209461729 0D-06
46	129	.65101455510766 1D-07
47	132	.38410193531033 5D-07
48	134	.36739998708647 7D-07
49	136	.36409423838479 2D-07
50	138	.36322906755645 3D-07
51	141	.23637207619307 2D-07
52	142	.69841684518432 0D-08
53	143	.42453700961621 4D-09
54	145	.14819866096367 2D-09
55	147	.86463615457135 1D-11
56	151	.14857452334895 9D-11
57	155	.13844592827594 2D-11
58	157	.13695396750400 5D-11
59	160	.13691869817176 6D-11
60	162	.13635648848946 0D-11
61	163	.13616738404917 3D-11
62	164	.13598722238121 7D-11
63	168	.13598626089372 4D-11
64	170	.13598617691501 5D-11
65	172	.13598617037363 8D-11
66	174	.13598616674105 7D-11
67	176	.13598610609679 5D-11
68	178	.13598580431904 4D-11
69	180	.13598372882435 7D-11
70	186	.30375426942623 2D-12
71	188	.28537807597262 4D-14
72	194	.26617814493663 9D-14
73	195	.23452829639071 0D-14
74	196	.18843692828867 5D-14
75	197	.11839079350268 8D-14
76	198	.56670778813455 8D-15
77	200	.41634941578134 6D-15
78	201	.24637894873117 7D-15
79	203	.18831207454629 1D-15

B.F.G.S

Formule (III 5.1) avec

$\varepsilon = \varepsilon' = -1$, $\alpha_k = \delta_k = 1$

$\beta_k = \gamma_k = (0.9)^k$

IT	NF	VALEUR DE F
80	109	.857262084066345D-04
81	110	.397682960846600D-04
82	111	.161323553941588D-04
83	112	.814066491240726D-05
84	113	.755826381339454D-05
85	115	.584000389762778D-05
86	117	.135561056584838D-05
87	118	.691598792738548D-06
88	119	.526433004069952D-06
89	120	.506564068256751D-06
90	122	.414503126519807D-06
91	123	.282006907315626D-06
92	124	.104229794756120D-06
93	125	.452414879634288D-07
94	126	.189715038602724D-07
95	127	.875604134911466D-08
96	128	.569488783713139D-08
97	129	.507375988994302D-08
98	130	.410460592869034D-08
99	131	.256677699157405D-08
100	132	.110003271642587D-08
101	133	.774652790096170D-09
102	134	.566489108302964D-09
103	135	.435399986677121D-09
104	136	.391333444869738D-09
105	137	.363110603562789D-09
106	140	.308955954196275D-09
107	141	.308279000314825D-09
108	143	.303838921749603D-09
109	145	.296053186525833D-09
110	147	.206681396717946D-09
111	148	.938228064074567D-10
112	149	.423286167069407D-10
113	150	.240738490951317D-10
114	151	.754418204038168D-11
115	152	.317036101161253D-11
116	153	.128728567214471D-11
117	154	.588061116779278D-12
118	155	.272938588658687D-12
119	156	.120509801198345D-12
120	157	.630842029064225D-13
121	158	.607099402718808D-13
122	160	.577881582997661D-13
123	162	.270035088814803D-13
124	163	.133243127472410D-13

IT	NF	VALEUR DE F
80	205	.171324249034324D-15
81	207	.498686969813889D-16
82	209	.189915243732074D-16
83	211	.660237177216640D-17
84	213	.279291799310536D-17
85	215	.210041213861322D-17
86	216	.122373594432581D-17
87	217	.224046797994179D-18
88	220	.180586018551056D-18
89	222	.172618229696238D-18
90	224	.113665540491920D-18
91	227	.108149942361741D-18
92	230	.107928750458877D-18
93	232	.107907872788752D-18
94	234	.107903564094697D-18
95	237	.107212506047254D-18
96	239	.104651674521902D-18
97	241	.664276174107568D-19
98	242	.187135531452295D-19
99	243	.148675231675175D-20
100	245	.321146605601181D-21
101	247	.162742521073153D-21
102	248	.367605750158937D-22
103	252	.362634123177177D-22

B.F.G.S

Formule (III.5.1) avec

$\varepsilon = \varepsilon' = -1 \qquad \alpha_k = \delta_k = 1$

$\beta_k = \gamma_k = (0.9)^k$

IT	NF	VALEUR DE F
125	164	.42277636294543 6D-14
126	165	.20273209319770 5D-14
127	166	.83480153535505 7D-15
128	167	.32000764319491 7D-15
129	168	.13112901393962 4D-15
130	169	.68596384526372 4D-16
131	170	.55494837253306 0D-16
132	171	.52691792638145 4D-16
133	172	.48157444606831 4D-16
134	173	.40683645382162 8D-16
135	174	.29517092807939 4D-16
136	175	.22016795674641 4D-16
137	176	.19484964471853 1D-16
138	177	.17504706048498 8D-16
139	178	.17017273423426 1D-16
140	179	.16951818857782 8D-16
141	180	.16021972234543 5D-16
142	181	.16839120827250 0D-16
143	183	.16716801546713 6D-16
144	186	.47150454169157 2D-17
145	187	.21418224239215 2D-17
146	188	.72144643245339 2D-18
147	189	.27032915029046 5D-18
148	190	.11065372476765 6D-18
149	191	.72308408155034 5D-19
150	192	.62462814831896 2D-19
151	193	.51520723364265 5D-19
152	194	.34327885102143 8D-19
153	195	.15365159414627 2D-19
154	196	.57856343391561 4D-20
155	197	.20768490543587 5D-20

NORME GRADIENT CARRE

.33581069835348 5D-25 .89071849693204 2D-25

TEMPS CALCUL

9 8

POINT OPTIMAL

X(1)= -.12278282408460 2D-03 X(1)= -.82193089016309 6D-04
X(2)= .99987401265035 8D+00 X(2)= .99991784180410 6D+00
X(3)= .99999359982486 3D+00 X(3)= .99999825856633 2D+00
X(4)= .99999999999990 8D+00 X(4)= .10000000000001 5D+01

Formule (III.5.1) avec

$\varepsilon = \varepsilon' = -1 \quad , \quad \alpha_k = \delta_k = 1 \qquad \beta_k = \gamma_k = (0.9995)^k$

0	1	.36742743335137 9D+09
1	2	.41140972624005 7D+08
2	4	.41679800144815 9D+05
3	13	.89094343462072 0D+03
4	18	.37537600064410 1D+03
5	20	.37515922068306 3D+03
6	21	.37494277516316 0D+03
7	22	.37494157117737 3D+03
8	24	.37493398521035 3D+03
9	30	.30249356112732 5D+03
10	32	.30237745816182 8D+03
11	35	.30237395418990 2D+03
12	37	.30237364679300 0D+03
13	42	.30024118126380 2D+03
14	45	.80809452910099 8D+02
15	47	.70967125160268 8D+02
16	48	.56254807508992 6D+02
17	52	.50672963716922 7D+02
18	54	.50599202465115 3D+02
19	55	.50559250018375 4D+02
20	57	.50551599195411 1D+02
21	58	.50550701413496 6D+02
22	60	.50550447577539 5D+02
23	61	.50550390893729 8D+02
24	63	.50549571360877 1D+02
25	65	.50543319890752 7D+02
26	69	.45177070487854 1D+02
27	71	.94691378526999 8D+01
28	72	.75236721752706 2D+01
29	73	.51546547370773 4D+01
30	74	.35944101107431 1D+01
31	75	.21866195930956 9D+01
32	76	.20449309447223 9D+01
33	77	.18095448130538 0D+01
34	79	.16465879959778 0D+01
35	80	.16330262998935 6D+01
36	81	.16307638786730 9D+01
37	82	.16300005781075 2D+01
38	83	.16295893020702 8D+01
39	84	.16295609430871 6D+01
40	85	.16295577757150 6D+01
41	86	.16295575989077 2D+01
42	87	.16295575868058 1D+01
43	88	.16295575861489 1D+01
44	89	.16295575860994 1D+01
45	91	.16295575860825 8D+01
46	92	.16295575860825 8D+01
47	93	.16295575860825 8D+01
48	94	.16295575860825 8D+01
49	108	.16295575860825 8D+01
50	122	.16295575860825 8D+01
50	131	.16295575860825 8D+01

NORME GRADIENT CARRE ,34383310810394 4D-14

TEMPS CALCUL 4

POINT OPTIMAL:

X(1)= -,73163517416041 2D+00 X(2)= -,29443298557111 5D+00
X(3)= -,60960611281821 2D+00 X(4)= ,19329688729942 6D+01

Formule (III.5.1) avec

$\varepsilon = \varepsilon' = -1$, $\alpha_k = \delta_k = 1$ $\beta_k = \gamma_k = 1/k^{1,25}$

IT	NF	VALEUR DE F
0	1	.367427433351379D+09
1	2	.411409726240057D+08
2	4	.110844664954110D+08
3	5	.254156613286145D+05
4	15	.217789399697535D+05
5	17	.209851033325879D+05
6	20	.209692106467895D+05
7	28	.209692104505310D+05
8	37	.668639198224696D+04
9	38	.880563189311531D+03
10	43	.594786625289408D+03
11	46	.588449724144498D+03
12	48	.586957132438238D+03
13	52	.586952671720913D+03
14	53	.586951286066996D+03
15	54	.586950050132576D+03
16	57	.586950036433419D+03
17	59	.586950034976026D+03
18	62	.586949825154123D+03
19	64	.586948338206001D+03
20	70	.375065362187083D+03
21	72	.276279780354080D+03
22	73	.225096102853886D+03
23	74	.184611493457409D+03
24	75	.128853293439327D+03
25	79	.108047238331032D+03
26	80	.103658108776898D+03
27	81	.961956567654353D+02
28	82	.868101276922010D+02
29	83	.798880349436750D+02
30	84	.734538580169883D+02
31	86	.719357444188173D+02
32	88	.716861232135869D+02
33	90	.716231179940294D+02
34	93	.716224177219339D+02
35	95	.716222352931615D+02
36	99	.716222351120578D+02
37	100	.716222348858189D+02
38	102	.716222315090477D+02
39	103	.716222258768552D+02
40	105	.716221911665957D+02
41	109	.715993888204009D+02
42	112	.680155131900031D+02
43	114	.479460998774861D+02
44	115	.185235393052960D+02
45	116	.783685555921482D+01
46	118	.182016693609674D+01
47	120	.117321702619135D+01
48	122	.102482609985251D+01
49	123	.834213300967215D+00
50	125	.747716108398077D+00
51	127	.726066389175903D+00
52	129	.714182799962664D+00
53	131	.710431451390114D+00
54	133	.710125779467443D+00
55	135	.710098585024824D+00
56	138	.710097131853116D+00
57	142	.710097128105114D+00
58	145	.710097128070370D+00
59	148	.710097128070300D+00
60	149	.710097128070200D+00
61	151	.710097128070176D+00
62	163	.710097128070176D+00
63	179	.710097128070176D+00
64	183	.710097128070176D+00
64	192	.710097128070176D+00

NURME GRADIENT CARRE .410111437608753D-14

TEMPS CALCUL 6

POINT OPTIMAL:

X(1)= .826245162503054D+00 X(2)= .289752503396840D+01
X(3)= .317116120583623D+01 X(4)= .539755312761817D+00

Formule (III.5.1) avec

$\varepsilon = \varepsilon' = -1$, $q_k = \delta_k = 1$ $\beta_k = \gamma_k = 1/k^{1,001}$

IT	NF	VALEUR DE F
0	1	.3674274333513790+09
1	2	.4114097262400570+08
2	4	.1108446649541100+03
3	5	.2541566133166550+05
4	15	.2177893991855300+05
5	17	.2098510319124460+05
6	20	.2096921042992120+05
7	27	.2096921010364210+05
8	29	.2096920670241860+05
9	33	.2096794903005180+05
10	38	.7603012974150200+04
11	40	.1346008826507310+04
12	43	.8809821706789150+03
13	46	.8405411697299600+03
14	49	.8393618667732900+03
15	53	.8393598676934970+03
16	56	.8393598557754600+03
17	58	.8393598530913800+03
18	65	.8382385178628170+03
19	68	.7008438170615430+03
20	69	.6267249962852480+03
21	70	.5254606037556550+03
22	73	.5150170963580700+03
23	75	.5109462010817000+03
24	77	.5087300798642310+03
25	80	.4021766285741820+03
26	82	.3862324122554290+03
27	84	.3854044486560060+03
28	86	.3851282689142910+03
29	88	.3850336096920660+03
30	90	.3850261864987370+03
31	93	.3850260614336230+03
32	94	.3850259050396070+03
33	96	.3850258686868680+03
34	98	.3850258593549140+03
35	102	.3850258593305950+03
36	107	.3850257273007760+03
37	113	.3670727153322820+03
38	116	.1830743743085100+03
39	121	.4702910475880380+02
40	124	.2307582385688450+02
41	127	.1920718126207140+02
42	128	.1500726388895510+02
43	129	.1155076990383830+02
44	130	.8855532465525260+01
45	133	.8746773858732210+01
46	136	.8742701624122940+01
47	138	.8742115994395930+01
48	140	.8741813113458180+01
49	142	.8741744142613820+01
50	147	.8741744125561480+01
51	149	.8741744124150600+01
52	152	.8741744124096290+01
53	154	.8741744124090100+01
54	159	.8741744124090100+01
55	162	.8741744124090100+01
56	183	.8741744124090100+01
56	193	.8741744124090100+01

NORME GRADIENT CARRE .4101114376087530-14

TEMPS CALCUL 5

POINT OPTIMAL:

X(1)= .1072624906992570+01 X(2)= .3960833438992900+01
X(3)= .4336103240127530+01 X(4)= -.1235424748572880+01

Formule (III 5.1) avec

$\varepsilon = \varepsilon' = -1$, $q_k = \delta_k = 1$ $\beta_k = \gamma_k = 1/k^{1,00005}$

IT	NF	VALEUR DE F
0	1	.3674274533513790D+09
1	2	.4114097262400570D+08
2	4	.110844664954110D+08
3	5	.2541566133168080D+05
4	15	.2177893991829170D+05
5	17	.2008510319051610D+05
6	20	.2006921042879760D+05
7	27	.2096921010125700D+05
8	29	.2096920654035290D+05
9	32	.200688641351610D+05
10	38	.6984232835463980D+04
11	40	.4720167697096620D+04
12	41	.1841964631588620D+04
13	43	.8731298933435180D+03
14	48	.8503360200757250D+03
15	50	.8393686722315630D+03
16	55	.8393599631635270D+03
17	58	.8393598548030220D+03
18	59	.8393596815197310D+03
19	61	.8393585163889180D+03
20	67	.7120531588931050D+03
21	68	.5787255460972410D+03
22	71	.5632060688662700D+03
23	72	.5398345108017360D+03
24	74	.5316001327624430D+03
25	76	.5301540902409650D+03
26	78	.5231687592605440D+03
27	79	.5161342701960500D+03
28	81	.4106047747707160D+03
29	83	.3831082919488280D+03
30	85	.3613639729692010D+03
31	87	.3565721682645570D+03
32	88	.3524925162332100D+03
33	90	.3513332991053370D+03
34	92	.3509938525078490D+03
35	94	.3509444692929330D+03
36	96	.3509418267873530D+03
37	98	.3509408152770070D+03
38	101	.3509407971248020D+03
39	103	.3509407950328150D+03
40	105	.3509407948461700D+03
41	108	.3509407879588440D+03
42	109	.3509407770018930D+03
43	110	.3509407599497840D+03
44	111	.3509407314075320D+03
45	113	.3509406026776190D+03
46	120	.2563421669340790D+03
47	121	.1603732432469460D+03
48	122	.1213308771094330D+03
49	123	.8870019875823960D+02
50	124	.6146447293607080D+02
51	128	.5981363899453960D+02
52	129	.5740939598628940D+02
53	130	.5426111472482650D+02
54	131	.4911859711172750D+02
55	133	.4472029263625260D+02
56	135	.4447833402422490D+02
57	137	.4435941825637200D+02
58	139	.4431641406495370D+02
59	141	.4431278745745300D+02
60	144	.4431269601116170D+02
61	147	.4431269297246360D+02
62	150	.4431269289227630D+02
63	152	.4431269286500100D+02
64	155	.4431269286369700D+02
65	160	.4431269286369690D+02
66	163	.4431269286369690D+02
67	173	.4431269286369690D+02
68	184	.4431269286369690D+02
69	196	.4431269286369690D+02
70	207	.4431269286369690D+02
70	218	.4431269286369690D+02

NORME GRADIENT CARRE ,4101114376087530-14

TEMPS CALCUL 6

POINT OPTIMAL:

X(1)= -,9295338389421270D+00 X(2)= -,1053564824845540D+01
X(3)= -,1512020462280290D+01 X(4)= ,7075861147956360D+01

Fonction test : Fonction étendue de Dennis [12]

$$f(x) = \sum_{i=1}^{n} i\, x_i^2 + \left(\sum_{i=1}^{n} x_i \right)^4$$

Point optimal théorique : $\overset{*}{x} = 0$.

Valeur optimale théorique de f = 0.

Point de départ $x_0(i) = 10 \quad i = 1, 2, \ldots, n$.

Valeur initiale de f (voir itération 0).

1) n = 10

Formule (III.5.1) avec

$\varepsilon = \varepsilon' = 1$

$\gamma_k = \delta_k = 1$

$\beta_k = \gamma_k = (0.999995)^k$

B. F. G. S

IT	NF	VALEUR DE F
0	1	.1000055000000000D+09
1	3	.1391236636511590D+08
2	4	.6047791485589350D+07
3	5	.1748049799206140D+07
4	6	.5933299107847610D+06
5	7	.1894489712260790D+06
6	8	.6194945356386790D+05
7	9	.2008071682344240D+05
8	10	.6535580626654500D+04
9	11	.2126307892523290D+04
10	12	.6932801556804780D+03
11	13	.2266234204323670D+03
12	14	.7442061384301980D+02
13	15	.2461148071010680D+02
14	16	.8223500522978090D+01
15	17	.2783870573834870D+01
16	18	.9537573638219000D+00
17	19	.3266150050210540D+00
18	20	.1079991412862240D+00
19	21	.3278089825369190D-01
20	22	.1025965103360300D-01
21	23	.6151196389647660D-02
22	24	.5920803145751660D-02
23	25	.5915795175876850D-02
24	26	.5907476659405100D-02
25	28	.5867212949677290D-02
26	31	.2267963966009920D-02
27	32	.9899914798463980D-03
28	35	.9660261657794860D-03
29	37	.7051176981587510D-03
30	38	.4414359716818820D-03
31	39	.2800560582129210D-03
32	40	.2797456346106130D-03
33	42	.2794960567076290D-03
34	46	.9419256404905900D-04
35	47	.9357327304132670D-04
36	49	.9319566855740320D-04
37	51	.9140933198576110D-04

IT	NF	VALEUR DE F
0	1	.100005500000000D+09
1	3	.139123663651159D+08
2	4	.272550508394879D+03
3	7	.379243509765453D+02
4	8	.164180716436392D+02
5	9	.494245950232913D+01
6	10	.175623571548254D+01
7	11	.5960973738547267D+00
8	12	.204693238881115D+00
9	13	.647516197747825D-01
10	14	.165037807903737D-01
11	15	.327518406654357D-02
12	16	.177232569793117D-02
13	17	.160964221297121D-02
14	19	.168585604124649D-02
15	22	.968230790533155D-03
16	23	.286882416133124D-03
17	25	.236014686916463D-03
18	27	.279406203071061D-03
19	30	.814314793546037D-04
20	32	.813347485161878D-04
21	36	.272752920128956D-04
22	37	.272272233950608D-04
23	39	.274750639170575D-04
24	41	.264778451044580D-04
25	43	.212199114521800D-04
26	45	.897307393121784D-05
27	50	.270409098279889D-05
28	51	.259157511857251D-05
29	53	.254973947297078D-05
30	57	.772436633208944D-06
31	58	.522122779552225D-06
32	61	.507375576671359D-06
33	63	.414590434527672D-06
34	65	.675374931193214D-07
35	70	.404099707433020D-08
36	73	.388988485322223D-08
37	75	.297633594674376D-08

B.F.G.S

Formule (III.S.1) avec

$\varepsilon = \varepsilon' = -1, \quad \varrho_k = \delta_k = 1$

$\beta_k = \gamma_k = (0.999995)^k$

IT	NF	VALEUR DE F
38	54	.3083330530456330D-04
39	55	.3073568771389950D-04
40	57	.3066564524526260D-04
41	59	.3015064533624960D-04
42	62	.8624612088518070D-05
43	65	.8483532773527530D-05
44	68	.2983960772208150D-05
45	69	.1823985585093860D-05
46	70	.1783189928111130D-05
47	73	.1744494311610980D-05
48	75	.1210565961487250D-05
49	76	.5553307194659700D-06
50	77	.2374569300345600D-06
51	78	.2303569633835030D-06
52	79	.2300931259336600D-06
53	81	.2297260152382670D-06
54	85	.1848093244086150D-07
55	86	.1375434711469950D-07
56	87	.1372858155334860D-07
57	89	.1369425453990350D-07
58	93	.6103393954282310D-25
59	94	.1841008250847020D-31

IT	NF	VALEUR DE F
38	80	.1429053770022930D-21
39	81	.5554281707150960D-27

NORME GRADIENT CARRE

.3207058290804610D-30   .1606071897348030D-25

TEMPS CALCUL

10   7

Valeur du point final

min $\lvert x_i \rvert$	max $\lvert x_i \rvert$	min $\lvert x_i \rvert$	max $\lvert x_i \rvert$
10^{-18}	10^{-17}	10^{-16}	10^{-15}

n = 20

B.F.G.S

Formule (III.S.1) avec

$\varepsilon = \varepsilon' = -1$

$\alpha_k = \delta_k = 1$

$\beta_k = \gamma_k = (0.9995)^k$

IT	NF	VALEUR DE F
0	1	.1600021000000000D+10
1	3	.2225789092607170D+09
2	4	.9675410054146740D+08
3	5	.2796360219329210D+08
4	6	.9490586869044080D+07
5	7	.3029737169970910D+07
6	8	.9903901563008070D+06
7	9	.3208454599831960D+06
8	10	.1043194023937760D+06
9	11	.3388092668300710D+05
10	12	.1101422277927160D+05
11	13	.3582484316734470D+04
12	14	.1166801041753640D+04
13	15	.3608319190363110D+03
14	16	.1247521685917860D+03
15	17	.4110855048863410D+02
16	18	.1367179568982510D+02
17	19	.4607421165206640D+01
18	20	.1577636451379870D+01
19	21	.5465529233095760D+00
20	22	.1870637432157500D+00
21	23	.5915105789842130D-01
22	24	.1500752321357730D-01
23	25	.2965693897144380D-02
24	26	.1454275360885620D-02
25	27	.1421225563688500D-02
26	29	.1420749219763170D-02
27	34	.2453113291394570D-03
28	35	.2451084400680650D-03
29	37	.2449556277377290D-03
30	39	.2429981892921390D-03
31	41	.2203010706463820D-03
32	43	.9717262595232050D-04
33	44	.7579983951461600D-04
34	45	.7510666676020790D-04
35	48	.7474927774666390D-04
36	50	.7187276594577280D-04
37	52	.3935355102233000D-04

0	1	.1600021000000000D+10
1	3	.2225789092607170D+09
2	5	.1348361920679550D+08
3	8	.1044922443180510D+07
4	11	.5040140799970470D+04
5	15	.3151054083583900D-03
6	17	.5985477849967840D-04
7	18	.5671951814136080D-06
8	19	.1929798151660310D-06
9	20	.1891102078234720D-06
10	21	.1890374116147690D-06
11	23	.1887817926700050D-06
12	25	.1869010289147760D-06
13	28	.3256703783245650D-07
14	33	.1336900619268210D-07
15	34	.1006193840068720D-07
16	35	.9976369153374430D-08
17	36	.9972095343213690D-08
18	38	.9970242252025510D-08
19	42	.6868978659778390D-08
20	43	.3942724347811150D-08
21	44	.3933677373684220D-08
22	45	.3933023630265750D-08
23	48	.3928096997974950D-08
24	52	.1762069029974290D-08
25	57	.1448024630377350D-08
26	58	.1057087929337340D-08
27	59	.6618125890940280D-09
28	60	.8382339650483190D-09
29	61	.8351453319696023D-09
30	62	.8346870242756750D-09
31	63	.8345938741650750D-09
32	65	.8344361270014910D-09
33	69	.5758323991462770D-09
34	70	.4194069676254600D-09
35	71	.3999783034723500D-09
36	72	.3974065226274360D-09
37	73	.3970357636179740D-09

B.F.G.S

Formule (III.S.1) avec
$\varepsilon = \varepsilon' = -1, \quad \alpha_k = \delta_k = 1$
$\beta_k = \gamma_k = (0.9995)^k$

IT	NF	VALEUR DE F	IT	NF	VALEUR DE F
38	53	.296407136375719D-04	38	74	.396969348287101D-09
39	58	.235502220108420D-04	39	75	.396944901663134D-09
40	59	.173287511400970D-04	40	77	.396845297913800D-09
41	60	.132834604142955D-04	41	81	.243727511485809D-09
42	65	.108794052454521D-04	42	82	.192998711489100D-09
43	66	.826773275688241D-05	43	83	.183779003855345D-09
44	67	.629347785131189D-05	44	84	.182016663396366D-09
45	72	.505825906511415D-05	45	85	.181659500269419D-09
46	73	.380206230469345D-05	46	86	.181579691732297D-09
47	74	.299385384512799D-05	47	87	.181557301161972D-09
48	79	.245701409206471D-05	48	88	.181545227258784D-09
49	80	.165700710333642D-05	49	90	.181488559328137D-09
50	81	.136941643914326D-05	50	94	.986409236743739D-10
51	86	.100791917711116D-05	51	95	.832051060242576D-10
52	87	.766270200181115D-06	52	96	.787555218736937D-10
53	88	.579533117644720D-06	53	97	.774371936853943D-10
54	93	.456567264484183D-06	54	98	.770298045117729D-10
55	94	.322283880060431D-06	55	99	.766978991485717D-10
56	95	.219446227102968D-06	56	100	.766523578647572D-10
57	100	.158921674469708D-06	57	101	.766348483618390D-10
58	101	.102346490930710D-06	58	102	.766205555933642D-10
59	102	.723640362328217D-07	59	103	.766204791570881D-10
60	107	.533431906464268D-07	60	105	.767925984441069D-10
61	108	.336264759291667D-07	61	109	.373290645388518D-10
62	109	.203491777143526D-07	62	110	.325916577184228D-10
63	114	.131857165806538D-07	63	111	.295355005956555D-10
64	115	.723664269742191D-08	64	112	.292931014798071D-10
65	116	.479371239758548D-08	65	113	.291020310703374D-10
66	121	.329666297824946D-08	66	114	.290956580643064D-10
67	122	.181745658489935D-08	67	115	.290913568521325D-10
68	123	.928083373071657D-09	68	116	.290890028443390D-10
69	128	.511553638132223D-09	69	118	.290712292932370D-10
70	129	.224622485661354D-09	70	122	.974400421991699D-11
71	130	.143947239992045D-09	71	123	.967814459236222D-11
72	135	.101667548496370D-09	72	124	.959825271908006D-11
73	136	.547288195190024D-10	73	125	.959607482285282D-11
74	137	.172156884012604D-10	74	126	.959370880516943D-11
75	142	.662986167298813D-11	75	128	.959093726960376D-11
76	143	.189785102140740D-11	76	132	.516234676702858D-11
77	144	.149192651099036D-11	77	133	.426511592450109D-11
78	145	.149157192546715D-11	78	134	.298446653222416D-11
79	147	.149098364905331D-11	79	136	.269824197046043D-11
80	151	.763152275981937D-12	80	141	.119086351431069D-11
81	152	.833874725166710D-13	81	143	.635822285923183D-12
82	155	.830556346209919D-13	82	145	.635665075466446D-12
83	158	.313863072516652D-13	83	149	.372515836219826D-12
84	159	.225941084567611D-14	84	150	.123885435669856D-12
85	164	.417165706866263D-15	85	152	.123120485798481D-12

B.F.G.S

Formule (III.S.1) avec

$\varepsilon = \varepsilon' = -1, \quad \alpha_k = \delta_k = 1$

$\beta_k = \gamma_k = (0.9995)^k$

IT	NF	VALEUR DE F
86	165	.106021187219966D-17
87	166	.151136968638212D-20
88	167	.232399674534898D-24
89	168	.194172057012564D-28

IT	NF	VALEUR DE F
86	157	.100876973444965D-12
87	159	.192266835625857D-13
88	160	.191141583773925D-13
89	162	.190997475961964D-13
90	164	.190877042118982D-13
91	166	.189615498739994D-13
92	169	.705582588704282D-14
93	170	.289610056930622D-14
94	172	.229447875845265D-14
95	173	.228566142909525D-14
96	175	.228433262068492D-14
97	177	.228269064556448D-14
98	179	.227416766139606D-14
99	182	.950082676035557D-15
100	183	.682606243158550D-15
101	184	.295686659594410D-15
102	186	.198072361776952D-15
103	187	.197959486053685D-15
104	188	.197935345764693D-15
105	189	.197906035378663D-15
106	191	.197703222976280D-15
107	195	.113536368054766D-16
108	196	.111276669985976D-16
109	198	.110670095553159D-16
110	200	.110575469665103D-16
111	202	.110132108428645D-16
112	205	.400386729754019D-17
113	207	.366585343947444D-18
114	208	.303182630773139D-18
115	209	.299894616254254D-18
116	210	.299858869473904D-18
117	212	.299639841096166D-18
118	216	.227817764983417D-19
119	218	.831776179974305D-22
120	220	.654135790152677D-25
121	222	.120709879553896D-31

Formule (III.5.1) avec
$\varepsilon = \varepsilon' = -1, \quad \alpha_k = \delta_k = 1$
$\beta_k = \gamma_k = (0.9995)^k$

B.F.G.S

NORME GRADIENT CARRE

.82710425646844490-27	.72852642917494390-30

TEMPS CALCUL

44	61

Valeur du point final

min $\lvert x_i \rvert$	max $\lvert x_i \rvert$	min $\lvert x_i \rvert$	max $\lvert x_i \rvert$
10^{-18}	10^{-16}	10^{-19}	10^{-17}

n = 30

B.F.G.S

Formule (III.5.1) avec

$\varepsilon = \varepsilon' = -1$

$\alpha_k = \delta_k = 1$

$\beta_k = \gamma_k = (0.9995)^k$

IT	NF	VALEUR DE F
0	1	.8100046500000000D+10
1	3	.1126788670244700D+10
2	4	.4898081600533960D+09
3	5	.1415622710136880D+09
4	6	.4804366328004960D+08
5	7	.1533674642093330D+08
6	8	.5013129502585420D+07
7	9	.1623877414382550D+07
8	10	.5278900977172550D+06
9	11	.1713940774630010D+06
10	12	.5568707353484570D+05
11	13	.1809541347254020D+05
12	14	.5883872509655590D+04
13	15	.1915022901312830D+04
14	16	.6243493661428080D+03
15	17	.2041432773083930D+03
16	18	.6707136283555420D+02
17	19	.2220824740172040D+02
18	20	.7440665486935500D+01
19	21	.2533016434452980D+01
20	22	.8768268766676800D+00
21	23	.3049928992891630D+00
22	24	.1020366544082970D+00
23	25	.2930561023265220D-01
24	26	.5704951185711020D-02
25	27	.8948355475546730D-03
26	28	.6206563637308050D-03
27	29	.6194335794865090D-03
28	31	.6191098601202160D-03
29	33	.6165689483138820D-03
30	36	.3255444879292810D-03
31	37	.1074276473049560D-03
32	39	.1073937495981920D-03
33	41	.1071374173610090D-03
34	45	.3341325756418630D-04
35	46	.3341205329946130D-04
36	48	.3340676021272460D-04
37	53	.1357997298786720D-04

0	1	.8100046500000000D+10
1	3	.1126788670244700D+10
2	5	.6825589665947160D+08
3	6	.5249621367700750D+07
4	11	.2536779839444640D+05
5	16	.6762591889696980D+01
6	17	.5510452053465950D-01
7	18	.3008876459094330D-01
8	19	.1571681645589840D-02
9	20	.3588657877242420D-05
10	21	.2051654648757900D-06
11	22	.1034485226126710D-06
12	23	.9904835675075020D-07
13	24	.9861371559662750D-07
14	25	.9879457056525370D-07
15	27	.9874659881164240D-07
16	29	.9815709588249410D-07
17	31	.9009568844193750D-07
18	33	.5343098937892260D-07
19	34	.2208993170133790D-07
20	35	.1756712292807550D-07
21	36	.1716162844689490D-07
22	37	.1712095856558460D-07
23	38	.1711607368465350D-07
24	39	.1711502076941690D-07
25	41	.1710748461883750D-07
26	43	.1702542596521820D-07
27	46	.1210469490197410D-07
28	47	.6527560103669360D-08
29	48	.5782902690721690D-08
30	49	.5491979256380670D-08
31	50	.5395700436023140D-08
32	51	.5325609892130810D-08
33	52	.5322670595966310D-08
34	53	.5319839428410980D-08
35	55	.5319260303341080D-08
36	60	.2161489620845350D-08
37	66	.1036373393433850D-08

B.F.G.S

Formule (III.5.1) avec

$\varepsilon = \varepsilon' = -1 \quad , \quad \alpha_k = \delta_k = 1$

$\beta_k = \gamma_k = (0.9995)^k$

IT	NF	VALEUR DE F
38	55	.1357753730361940D-04
39	57	.1355907096500690D-04
40	61	.6407881545768110D-05
41	64	.6398969723990200D-05
42	66	.6318029475275810D-05
43	69	.3300382715315660D-05
44	73	.3267702583820180D-05
45	75	.2737222661055490D-05
46	76	.1994516250259500D-05
47	77	.1788446388774140D-05
48	78	.1783720917233960D-05
49	79	.1783517074504590D-05
50	81	.1783355127518780D-05
51	85	.1546654201526090D-05
52	86	.1185363099121230D-05
53	87	.9891252980162000D-06
54	88	.9824814336032150D-06
55	89	.9820289005178530D-06
56	91	.9819547781947680D-06
57	95	.9257039666278720D-06
58	97	.5377172853534460D-06
59	103	.3160821211541160D-06
60	104	.2869692713583280D-06
61	105	.2865556086573400D-06
62	108	.2862020444985200D-06
63	110	.2826835713529530D-06
64	113	.1457852349785950D-06
65	119	.7144591376356710D-07
66	120	.6962999872256020D-07
67	121	.6962750062572540D-07
68	123	.6962165557235240D-07
69	128	.3079528996841190D-07
70	134	.1250824182525950D-07
71	135	.1248401447344200D-07
72	138	.1246717769100560D-07
73	141	.1062147461798820D-07
74	143	.4778113449224740D-08
75	144	.4605036281166020D-08
76	145	.4604661854997680D-08
77	147	.4604105003885920D-08
78	152	.1537436522141580D-08
79	158	.4626054750845700D-09
80	160	.4625325085020330D-09
81	165	.1248547837847660D-09
82	169	.1216656963718460D-09
83	171	.1010330962566090D-09
84	173	.3005929381713970D-10
85	179	.6413118574006190D-11

IT	NF	VALEUR DE F
38	67	.1021416520532700D-08
39	68	.1020133909087060D-08
40	69	.1019809682777390D-08
41	70	.1019713550148740D-08
42	71	.1019675640702730D-08
43	73	.1019002498627600D-08
44	78	.5251492666103790D-09
45	79	.5251310838707100D-09
46	80	.5251162702680070D-09
47	82	.5250894612075560D-09
48	87	.2637722330925710D-09
49	88	.2637601517256090D-09
50	93	.2633519359063030D-09
51	95	.1631725564479570D-09
52	96	.1590435238030750D-09
53	97	.1568850626653990D-09
54	99	.1562472500660020D-09
55	105	.1059894421140350D-09
56	106	.8564368730966680D-10
57	107	.8573430924721060D-10
58	108	.8558210820025560D-10
59	109	.8557168542076860D-10
60	110	.8555677419864340D-10
61	111	.8555432604826070D-10
62	112	.8555156952505550D-10
63	114	.8558353381363010D-10
64	119	.4559427106932540D-10
65	123	.4504504675277630D-10
66	125	.4178867863113550D-10
67	127	.2321822271911970D-10
68	129	.2319770711664650D-10
69	133	.2289332362824130D-10
70	135	.2103743762721930D-10
71	137	.1108046513792660D-10
72	139	.1107988603176610D-10
73	144	.6217232938700280D-11
74	145	.4981762919055490D-11
75	147	.4901681663197800D-11
76	149	.4901122862309530D-11
77	152	.4892568671350840D-11
78	154	.4846005366449300D-11
79	157	.1998029812860670D-11
80	159	.1987051887587140D-11
81	160	.1987001864384540D-11
82	162	.1986824189340520D-11
83	167	.8729176254936400D-12
84	169	.7344105156614480D-12
85	171	.7330357637344950D-12

B.F.G.S

Formule (III.5.1) avec

$\varepsilon = \varepsilon' = -1$, $\alpha_k = \delta_k = 1$

$\beta_k = \gamma_k = (0.9995)^k$

86	180	.641113699213165D-11
87	182	.640925734800355D-11
88	184	.639213222154626D-11
89	188	.120056156008856D-11
90	194	.195173453034620D-12
91	200	.458917046926901D-13
92	201	.273691625111077D-13
93	202	.271341970412995D-13
94	206	.269003636039939D-13
95	209	.318029997670642D-14
96	210	.316293033462715D-14
97	213	.315155386007179D-14
98	215	.306335860301356D-14
99	218	.300750539353971D-15
100	224	.224106562720451D-16
101	230	.128849289088689D-17
102	231	.122742393304250D-17
103	234	.122597988275817D-17
104	238	.439464456231330D-19
105	244	.781083562085985D-21
106	245	.771822377947901D-21
107	249	.764338707723298D-21
108	252	.837733313918892D-24
109	253	.407772632937512D-44

86	176	.567435745853016D-12
87	177	.327777766303959D-12
88	179	.246677173181288D-12
89	180	.245789748547017D-12
90	181	.244790597448568D-12
91	183	.244771512822804D-12
92	189	.753886791257295D-13
93	191	.736575046197574D-13
94	193	.736511272958711D-13
95	198	.202735709435732D-13
96	200	.198827184521678D-13
97	202	.198818402864163D-13
98	204	.196763321613451D-13
99	208	.148136191234950D-13
100	209	.787011210672066D-14
101	211	.495751570562676D-14
102	213	.478721706067230D-14
103	214	.478705261359965D-14
104	216	.478649714066088D-14
105	221	.100664523122612D-14
106	223	.102113788709604D-14
107	224	.102110414608565D-14
108	226	.102061947905193D-14
109	228	.101730616649465D-14
110	231	.533003626211474D-15
111	232	.278751213595982D-15
112	234	.192209592811796D-15
113	236	.191237773629958D-15
114	238	.191223078618862D-15
115	241	.189456431388626D-15
116	243	.164944835880242D-15
117	245	.747501303463784D-16
118	247	.311426588404170D-16
119	249	.310901910814584D-16
120	251	.310794614540156D-16
121	253	.310083696201314D-16
122	257	.430969754757259D-17
123	259	.432293918600965D-17
124	260	.432277084146717D-17
125	262	.432193072456182D-17
126	267	.529126626925075D-18
127	269	.503982392052526D-18
128	271	.503800512847289D-18
129	273	.502776557684047D-18
130	277	.160732446820164D-18
131	279	.487360764706681D-19
132	281	.479346346690268D-19
133	283	.479269880001788D-19
134	286	.477295201133556D-19

B.F.G.S

Formule (III.5.1) avec

$\varepsilon = \varepsilon' = -1$, $\alpha_k = \delta_k = 1$

$\beta_k = \gamma_k = (0.9995)^k$

135	288	.46403422431212 6D-19
136	290	.24866045258929 27D-19
137	291	.43319632087833 3D-20
138	293	.35768970932596 7D-20
139	295	.35714184586345 29D-20
140	298	.35640986457396 9D-20
141	301	.32593320869646 0D-20
142	303	.27137321734012 0D-21
143	305	.19619982169829 7D-21
144	307	.19562146388671 1D-21
145	308	.19561188264594 0D-21
146	310	.19555388838251 6D-21
147	314	.97568709631857 2D-22
148	315	.76702933733123 2D-23
149	317	.70061778512010 7D-23
150	319	.70044224612921 6D-23
151	321	.70030874571261 2D-23
152	324	.68131962149217 7D-23
153	326	.58664731656914 2D-23
154	328	.13392096794749 9D-24
155	330	.12359677404096 9D-24
156	332	.12307432752673 2D-24
157	335	.12287906382680 9D-24
158	338	.11336259479967 9D-24
159	340	.23948621909570 3D-25
160	342	.73136508441933 8D-28

NORME GRADIENT CARRE

.29477101700439 0D-42 .38341062267632 8D-26

TEMPS CALCUL

107 161

Valeur du point final

min $\lvert x_i \rvert$	max $\lvert x_i \rvert$	min $\lvert x_i \rvert$	max $\lvert x_i \rvert$
10^{-24}	10^{-23}	10^{-16}	10^{-15}

Fonction test : la fonction singulière de Powell [37]

$$f(x) = \sum_{i=1}^{n/4} [x_{4i-3} + 10_{4i-2}]^2 + 5[x_{4i-1} - x_{4i}]^2$$

$$+ [x_{4i-2} - 2\, x_{4i-1}]^2 + 10[x_{4i-3} - x_{4i}]^2$$

$n = 64.$

Point optimal théorique $\overset{*}{x} = 0.$

Valeur optimale théorique de $f = 0.$

Point de départ : pour $i = 1, 2, \ldots, n/4.$

$x_0(4i-3) = 6.$

$x_0(4i-2) = - 2.$

$x_0(4i-1) = 0.$

$x_0(4i) = 2.$

Valeur initiale de $f = 0.447\ 10^5.$

Formule (III.5.1) avec

B. F. G. S

$\varepsilon = \varepsilon' = 1$

$\gamma_k = \delta_k = 1$

$\beta_k = \gamma_k = (0.9995)^k$

IT	NF	VALEUR DE F
0	1	.446720000000000D+05
1	2	.237139605897796D+05
2	3	.658297249561873D+04
3	4	.220646460091831D+04
4	5	.108572675628177D+04
5	6	.614281064292015D+03
6	7	.367656692358526D+03
7	8	.258032960102283D+03
8	9	.194036367068212D+03
9	12	.752639514272167D+02
10	13	.672843962236824D+02
11	14	.569813168625903D+02
12	15	.490684851406722D+02
13	16	.363276573034367D+02
14	17	.178817639500839D+02
15	18	.605826294324176D+01
16	19	.199576062459185D+01
17	20	.631611068457799D+00
18	21	.257175384137407D+00
19	22	.896332823557583D-01
20	23	.270668701544809D-01
21	24	.909829993245667D-02
22	25	.316302406758893D-02
23	26	.100047477518579D-02
24	27	.322590846916447D-03
25	28	.108365548888252D-03
26	29	.351627848202713D-04
27	30	.112924574517552D-04
28	31	.371342699271980D-05
29	32	.121272160607308D-05
30	33	.391314125230542D-06
31	34	.127422956784402D-06
32	35	.415801159865643D-07
33	36	.134724378963521D-07
34	37	.437289960200508D-08
35	38	.142379453494561D-08
36	39	.462261569340335D-09
37	40	.149975811145751D-09

IT	NF	VALEUR DE F
0	1	.446720000000000D+05
1	2	.237139605897796D+05
2	3	.317885883610023D+04
3	4	.648285749843284D+03
4	5	.459504200670314D+03
5	7	.238013905969728D+03
6	8	.902205213251564D+02
7	10	.724386818272423D+02
8	12	.353062189279750D+02
9	13	.116385753049453D+02
10	14	.209909940132799D+01
11	15	.861448573338704D+00
12	16	.415670197658187D+00
13	17	.586904300376358D-01
14	18	.241251336950076D-01
15	19	.631777116775480D-02
16	20	.201358835181959D-02
17	21	.603892424989628D-03
18	22	.194924662611269D-03
19	23	.623371588155919D-04
20	24	.199703160325054D-04
21	25	.632503323648556D-05
22	26	.195744973990636D-05
23	27	.569401247698905D-06
24	28	.155297273343120D-06
25	29	.427908050659649D-07
26	30	.129229252465633D-07
27	31	.405713504022000D-08
28	32	.127454682587826D-08
29	33	.403629837089864D-09
30	34	.128167773420924D-09
31	35	.408042601169260D-10
32	36	.130074027667256D-10
33	37	.414969765588835D-11
34	38	.132434480619583D-11
35	39	.422682673059121D-12
36	40	.134892109179539D-12
37	41	.430382605435960D-13

Formule (III.5.1) avec

$\varepsilon = \varepsilon' = -1$

$\alpha_k = \delta_k = 1$

$\beta_k = \gamma_k = (0.9995)^k$

B.F.G.S

IT	NF	VALEUR DE F
38	41	.48751959481635900-10
39	42	.15837521688367000-10
40	43	.51400201502007300-11
41	44	.16694882186200300-11
42	45	.54233122013836400-12
43	46	.17607068168201000-12
44	47	.57171913384491300-13
45	48	.18569005939735100-13
46	49	.60295695265287000-14
47	50	.19577729449926000-14
48	51	.63578373630820900-15
49	52	.20645735611007000-15
50	53	.67037399188020800-16
51	54	.21768773139954300-16
52	55	.70689476392120400-17
53	56	.22953746471842800-17
54	57	.74534857954332400-18
55	58	.24203281922554200-18
56	59	.78592226136767300-19
57	60	.25520164636542200-19
58	61	.82869244116605100-20
59	62	.26909199281762400-20
60	63	.87378639419575600-21
61	64	.28373460152492900-21

IT	NF	VALEUR DE F
38	42	.13727423813419400-13
39	43	.43768474099326700-14
40	44	.13049497782255600-14
41	45	.44439411549979400-15
42	46	.14150921304488200-15
43	47	.45040435734017900-16
44	48	.14329113114567200-16
45	49	.45565136251304400-17
46	50	.14482432937296900-17
47	51	.46009224821096600-18
48	52	.14609734594807700-18
49	53	.46369598167588100-19
50	54	.14710169724079300-19

NORME GRADIENT CARRE

.31219354293697100-25 .99175234756704400-25

TEMPS CALCUL

224 186

valeur du point final

min $\lvert x_i \rvert$	max $\lvert x_i \rvert$	min $\lvert x_i \rvert$	max $\lvert x_i \rvert$
10^{-9}	10^{-6}	10^{-9}	10^{-6}

Formule (III.5.1) avec

B. F. G.S

$\varepsilon = \varepsilon' = 1$

$\varphi_k = \delta_k = 1$

$\beta_k = \gamma_k = (0.99)^k$

IT	NF	VALEUR DE F
0	1	.446720000000000D+05
1	2	.237139605897796D+05
2	3	.658297249561873D+04
3	4	.229646460091831D+04
4	5	.108572675628177D+04
5	6	.614281064292015D+03
6	7	.367656692358526D+03
7	8	.258032960102283D+03
8	9	.194036367068212D+03
9	12	.752639514272167D+02
10	13	.672843962236824D+02
11	14	.569813168625903D+02
12	15	.490684851406722D+02
13	16	.363276573034367D+02
14	17	.178817639500839D+02
15	18	.605826294324176D+01
16	19	.199576062459185D+01
17	20	.631611068457799D+00
18	21	.257175384137407D+00
19	22	.896332823557583D-01
20	23	.270668701544809D-01
21	24	.909829993245667D-02
22	25	.316302406758893D-02
23	26	.100047477518579D-02
24	27	.322590846916447D-03
25	28	.108365548888259D-03
26	29	.351627848202713D-04
27	30	.112924574517552D-04
28	31	.371342699271980D-05
29	32	.121272160607308D-05
30	33	.391314125230542D-06
31	34	.127422956784402D-06
32	35	.415801159865643D-07
33	36	.134724378963521D-07
34	37	.437289960200508D-08
35	38	.142379453494561D-08
36	39	.465261569340335D-09
37	40	.149975811145751D-09

IT	NF	VALEUR DE F
0	1	.446720000000000D+05
1	2	.237139605897796D+05
2	3	.532208929763180D+04
3	5	.327303280510881D+04
4	6	.128939648944371D+04
5	8	.577147588574103D+03
6	10	.474468134889050D+03
7	12	.461436519152498D+03
8	16	.414558839518806D+03
9	18	.304954370575469D+03
10	19	.126612844092587D+03
11	20	.254301263061565D+02
12	22	.117195447745841D+01
13	24	.295245569851328D+00
14	26	.174340418207727D+00
15	27	.341815835590567D-01
16	29	.240759550912768D-01
17	30	.937874937968970D-02
18	32	.783675031767472D-02
19	34	.120734419626312D-02
20	36	.685164313272134D-03
21	37	.344893458942962D-03
22	39	.180805268966639D-03
23	40	.421842233167221D-04
24	42	.324942268483655D-04
25	43	.236032569598168D-04
26	44	.102275376865218D-04
27	46	.418086191346329D-05
28	48	.252839848012968D-05
29	49	.623639165211138D-06
30	51	.429520861897136D-06
31	52	.220287102258996D-06
32	53	.119362059057732D-06
33	55	.772578162656364D-07
34	56	.238462413794540D-07
35	58	.620357884396043D-08
36	60	.117375937758020D-08
37	63	.108385831359043D-08

B.F.G.S

IT	NF	VALEUR DE F
38	41	.487519594816359D-10
39	42	.158375216883670D-10
40	43	.514002015020073D-11
41	44	.166948821862003D-11
42	45	.542331220138364D-12
43	46	.176070681682010D-12
44	47	.571719133844913D-13
45	48	.185690059397351D-13
46	49	.602956952652870D-14
47	50	.195777294499260D-14
48	51	.635783736308209D-15
49	52	.206457356110070D-15
50	53	.670373991880208D-16
51	54	.217687731399543D-16
52	55	.706894763921204D-17
53	56	.229537464718428D-17
54	57	.745348579543324D-18
55	58	.242032819225542D-18
56	59	.785922261367673D-19
57	60	.255201646365422D-19
58	61	.828692441166051D-20
59	62	.269091992817624D-20
60	63	.873786394195756D-21
61	64	.283734601524929D-21

Formule (III.5.1) avec

$\varepsilon = \varepsilon' = -1$, $\alpha_k = \delta_k = 1$

$\beta_k = \gamma_k = (0.99)^k$

IT	NF	VALEUR DE F
38	65	.212005993643695D-09
39	67	.160229038799098D-09
40	68	.103795612506017D-09
41	69	.770637788493017D-10
42	70	.536418419964519D-10
43	72	.418385767587031D-10
44	73	.225624761287406D-10
45	74	.369765797337368D-11
46	76	.134563016478295D-11
47	78	.555908914049999D-12
48	80	.413844708275389D-12
49	81	.188629894860651D-12
50	82	.368750636802585D-13
51	84	.105457238243737D-13
52	86	.878358751096924D-15
53	89	.833780276244592D-15
54	91	.244910382591701D-15
55	93	.182973434267200D-15
56	94	.940288305400702D-16
57	96	.554112497239780D-16
58	97	.121028457024554D-16
59	99	.109021784667066D-17
60	102	.485841636526295D-18
61	104	.169371616139819D-18
62	106	.393614213871845D-19
63	108	.282711084590927D-19
64	110	.235646023149669D-19
65	112	.660590863975766D-20
66	114	.371316459136132D-20
67	115	.529557358012652D-21
68	117	.342751013134452D-21
69	118	.137766577806390D-21
70	120	.410901582414719D-22
71	122	.268468713578719D-23
72	123	.267958637366470D-24
73	125	.553948282431036D-25
74	127	.104046322496145D-26

NORME GRADIENT CARRE

.31219354293697 1D-25	.42606013319692 9D-25

TEMPS CALCUL

224	276

valeur du point final

min $\|x_i\|$	max $\|x_i\|$	min $\|x_i\|$	max $\|x_i\|$
10^{-9}	10^{-6}	10^{-9}	10^{-8}

Fonction test : VAR [10] (100 variables)

$$f(x) = \sum_{i=1}^{100} x_i^2 + \left(\sum_{i=1}^{100} \sqrt{i}\, x_i\right)^2 + \left(\sum_{i=1}^{100} \sqrt{i}\, x_i\right)^4.$$

Point optimal $\overset{*}{x} = 0$.

Valeur optimale de f $f(\overset{*}{x}) = 0$.

Point de départ $x_0(i) = 6$, $i = 1, 2, \ldots, 100$.

Valeur initiale de f $f(x_0) = 0.263\ 10^{25}$.

Formule (III.5.1) avec

B.F.G.S

$$\varepsilon = \varepsilon' = -1$$

$$\alpha_k = \delta_k = 1$$

$$\beta_k = \gamma_k = (0.95)^k$$

IT	NF	VALEUR DE F
0	1	.263446987870664D+15
1	3	.401122895256115D+14
2	4	.169562382524287D+14
3	5	.495747982611523D+13
4	6	.167523025053336D+13
5	7	.535657149759592D+12
6	8	.174970244341636D+12
7	9	.566866611792210D+11
8	10	.184233530644922D+11
9	11	.598039079192815D+10
10	12	.194220159464205D+10
11	13	.630641530699459D+09
12	14	.204787705231623D+09
13	15	.664996534869396D+08
14	16	.215948879910290D+08
15	17	.701303317942402D+07
16	18	.227782268091360D+07
17	19	.740083208191698D+06
18	20	.240677351465511D+06
19	21	.784689052005560D+05
20	22	.257727623058172D+05
21	23	.864761673095603D+04
22	24	.307893270195028D+04
23	25	.126621252003842D+04
24	26	.675043554872912D+03
25	27	.4816297222626421D+03
26	28	.417998592283249D+03
27	29	.396866435174815D+03
28	30	.389738071213585D+03
29	31	.387274183348071D+03
30	32	.386393872213567D+03
31	33	.386070918736869D+03
32	34	.385958919040553D+03

IT	NF	VALEUR DE F
0	1	.263446987870664D+15
1	3	.401122895256115D+14
2	5	.723837725676494D+12
3	10	.879061700228656D+08
4	19	.152933624944772D+07
5	24	.468233362633005D+03
6	34	.386232739928536D+03
7	40	.385982264483198D+03
8	42	.385938853498067D+03
9	44	.385930878167749D+03
10	47	.385930769111969D+03
11	52	.385930769110502D+03
12	58	.385930568773559D+03
13	60	.385928757015543D+03
14	66	.374695086905191D+03
15	68	.227526523762802D+03
16	69	.298587993739779D+02
17	71	.696160074672679D+01
18	73	.336741960164611D+00
19	76	.194137664826863D+00
20	77	.230303601098094D-01
21	80	.135755629874739D-02
22	82	.483688267988993D-03
23	86	.482945422227193D-03
24	88	.478682367858495D-03
25	91	.573103192149539D-05
26	94	.283154486379375D-07
27	98	.977729723685937D-26
28	104	.520668020251950D-26

Formule ($\underline{\text{III}}$.5.1) avec
$\varepsilon = \varepsilon' = -1, \quad \varphi_k = \delta_k = 1$
$\beta_k = \gamma_k = (0.95)^k$

B.F.G.S

IT	NF	VALEUR DE F
33	35	.385932586521333D+03
34	36	.385930772246455D+03
35	37	.385930764704778D+03
36	40	.385930763584955D+03
37	49	.204090577546522D+03
38	50	.917733433820267D+02
39	51	.198294293060847D+01
40	52	.131276778513030D+01
41	53	.534445979664777D+00
42	54	.150551493052847D+00
43	55	.322774274267388D-01
44	56	.262503538335810D-02
45	57	.196824783758643D-04
46	58	.130800840844012D-06
47	59	.406551236610178D-08
48	60	.109025167488256D-11
49	61	.815246681092535D-15
50	62	.536482333799936D-20
51	63	.103406886490395D-24
52	64	.658838458343841D-26

	B.F.G.S	Formule (III.5.1)
NORME GRADIENT CARRE	.266647836465145D-25	.219726119975017D-25
TEMPS CALCUL	466	283

valeur du point final

	min $\lvert x_i \rvert$	max $\lvert x_i \rvert$
B.F.G.S	10^{-15}	10^{-14}
Formule (III.5.1)	10^{-16}	10^{-14}

Formule (III. 5.1) avec

B.F.G.S

$$\varepsilon = \varepsilon' = -1$$

$$\gamma_k = \delta_k = 1$$

$$\beta_k = \gamma_k = 1/k^{1,05}$$

IT		VALEUR DE F
0	1	.263446987870664D+15
1	3	.401122895256115D+14
2	4	.169562382524287D+14
3	5	.495747982611523D+13
4	6	.167523025053336D+13
5	7	.535657149759592D+12
6	8	.174970244341636D+12
7	9	.566866611792210D+11
8	10	.184233530644922D+11
9	11	.598039079192815D+10
10	12	.194220159464205D+10
11	13	.630641530699459D+09
12	14	.204787705231623D+09
13	15	.664996534869396D+08
14	16	.215948879910290D+08
15	17	.701303317942402D+07
16	18	.227782268091360D+07
17	19	.740083208191698D+06
18	20	.240677351465511D+06
19	21	.784680052005560D+05
20	22	.257727623058172D+05
21	23	.864761673095603D+04
22	24	.307893270195028D+04
23	25	.126621252003842D+04
24	26	.675043554872917D+03
25	27	.481629722626421D+03
26	28	.417998592283249D+03
27	29	.396866435174815D+03
28	30	.389738071213585D+03
29	31	.387274183348071D+03
30	32	.386393872213567D+03
31	33	.386070918736869D+03
32	34	.385958919040553D+03

IT	NF	VALEUR DE F
0	1	.263446987870664D+15
1	3	.401122895256115D+14
2	4	.169568382524287D+14
3	5	.171715066969063D+12
4	10	.165925884044961D+10
5	15	.259723662024190D+08
6	20	.389752725913606D+03
7	33	.385932483896001D+03
8	40	.385930768989712D+03
9	47	.385930768927696D+03
10	50	.385930768209635D+03
11	60	.285084698028790D-01
12	69	.111325418580689D-01
13	70	.326947771313988D-02
14	72	.158754631480986D-07
15	80	.540464906686875D-08
16	82	.170549477983154D-15
17	88	.685594757891580D-16
18	89	.163011762186919D-16
19	91	.392575195993931D-26

Formule $(\underline{\text{III}}.5.1)$ avec

$\varepsilon = \varepsilon' = -1$, $\alpha_k = \delta_k = 1$

$\beta_k = \gamma_k = 1/k^{1,05}$

B.F.G.S

IT	NF	VALEUR DE F
33	35	.385932586521333D+03
34	36	.385930772246455D+03
35	37	.385930764704778D+03
36	40	.385930763584955D+03
37	49	.204090577546522D+03
38	50	.917733433820267D+02
39	51	.198294293060847D+01
40	52	.131276778513030D+01
41	53	.534445979664777D+00
42	54	.150551493052847D+00
43	55	.322774274267388D-01
44	56	.262503538335810D-02
45	57	.196824783758643D-04
46	58	.130800840844012D-06
47	59	.406551236610178D-08
48	60	.109025167488256D-11
49	61	.815246681092535D-15
50	62	.536482333799936D-20
51	63	.103406886490395D-24
52	64	.658838458343841D-26

NORME GRADIENT CARRE

.266647836465145D-25 .372353497867205D-26

TEMPS CALCUL

466 202

valeur du point final

min $\lvert x_i \rvert$	max $\lvert x_i \rvert$	min $\lvert x_i \rvert$	max $\lvert x_i \rvert$
10^{-15}	10^{-14}	10^{-16}	10^{-14}

Formule (III.5.1) avec

B.F.G.S

$\varepsilon = \varepsilon' = -1$

$\alpha_k = \delta_k = 1$

$\beta_k = \gamma_k = (0,999)^k$

IT	NF	VALEUR DE F
0	1	.263446987870664D+15
1	3	.401122895256115D+14
2	4	.169562382524287D+14
3	5	.495747982611523D+13
4	6	.167523025053336D+13
5	7	.535657149759592D+12
6	8	.174970244341636D+12
7	9	.566866611792210D+11
8	10	.184233530644922D+11
9	11	.598039079192815D+10
10	12	.194220159464205D+10
11	13	.630641530699459D+09
12	14	.204787705231623D+09
13	15	.664996534869396D+08
14	16	.215948879910290D+08
15	17	.701303317942402D+07
16	18	.227782268091360D+07
17	19	.740083208191698D+06
18	20	.240677351465511D+06
19	21	.784680052005560D+05
20	22	.257727623058172D+05
21	23	.864761673095603D+04
22	24	.307893270195028D+04
23	25	.126621252003842D+04
24	26	.675043554872912D+03
25	27	.481629722626421D+03
26	28	.417998592283249D+03
27	29	.396866435174815D+03
28	30	.389738071213585D+03
29	31	.387274183348071D+03
30	32	.386393872213567D+03
31	33	.386070918736869D+03
32	34	.385958919040553D+03

IT	NF	VALEUR DE F
0	1	.263446987870664D+15
1	3	.401122895256115D+14
2	5	.723837725676500D+12
3	10	.879061700228802D+08
4	19	.152933624944581D+07
5	24	.468233362629727D+03
6	32	.385952410326777D+03
7	36	.385941692462842D+03
8	37	.385930778315439D+03
9	40	.385930769109027D+03
10	41	.385930769086724D+03
11	43	.385930768783171D+03
12	45	.385930766208317D+03
13	54	.247786306997415D+03
14	55	.913775027531600D+02
15	56	.237639109421243D+01
16	60	.114094331095203D+01
17	61	.959322322447393D-01
18	64	.164863761299194D-01
19	66	.612792442061894D-02
20	68	.193970182618483D-02
21	70	.473011053866231D-07
22	74	.259802930560497D-15
23	76	.150094343769229D-26
24	80	.572825917555887D-27
25		
26		
27		
28		
29		
30		
31		
32		

B.F.G.S

Formule (III.5.1) avec
$\varepsilon = \varepsilon' = -1$, $\alpha_k = \delta_k = 1$
$\beta_k = \gamma_k = (0.999)^k$

IT	NF	valeur de f
33	35	.385932586521333D+03
34	36	.385930772246455D+03
35	37	.385930764704778D+03
36	40	.385930763584955D+03
37	49	.204090577546522D+03
38	50	.917733433820267D+02
39	51	.198294293060847D+01
40	52	.131276778513030D+01
41	53	.534445979664777D+00
42	54	.150551493052847D+00
43	55	.322774274267388D-01
44	56	.262503538335810D-02
45	57	.196824783758643D-04
46	58	.130800840844012D-06
47	59	.406551236610178D-08
48	60	.109025167488256D-11
49	61	.815246681092535D-15
50	62	.536482333799936D-20
51	63	.103406886490395D-24
52	64	.658838458343841D-26

NORME GRADIENT CARRE

.266647836465145D-25	.229130367022356D-26

TEMPS CALCUL

.466	.242

valeur du point final

min $\|x_i\|$	max $\|x_i\|$	min $\|x_i\|$	max $\|x_i\|$
10^{-15}	10^{-14}	10^{-16}	10^{-15}