

t cilatex

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR
ET DE RECHERCHE SCIENTIFIQUE
UNIVERSITÉ FERHAT ABBAS SÉTIF 1
FACULTÉ DES SCIENCES
DÉPARTEMENT DE MATHÉMATIQUES

MÉMOIRE

Présenté par

Samiha LARGUEM

Pour l'obtention du diplôme de

Magister en Mathématiques

Option : Mathématiques Appliquées

THÈME

Étude d'une méthode de Quasi Newton pour matrice diagonale

Soutenu le : Décembre 2013

à l'UNIVERSITÉ FERHAT ABBAS SÉTIF 1

devant le jury composé de :

Président : Pr. N. BENSALÉM Professeur Université Ferhat Abbas Sétif 1

Rapporteur : Dr. B. MERIKHI MCA Université Ferhat Abbas Sétif 1

Examineur : Pr. M. ACHACHE Professeur Université Ferhat Abbas Sétif 1

Résumé

Dans ce mémoire, on s'intéresse à l'adaptation de la technique de diagonalisation de la matrice Hessien dans la méthode de Quasi Newton pour la résolution des problèmes d'optimisation sans contraintes. Notre contribution consiste d'une part a changer l'ordre des étapes dans l'algorithme de *Toint* et d'autre part à utiliser la variante de \mathcal{DFP} au lieu de \mathcal{BFGS} . Les tests numériques programmés sont en faveur de notre variante.

Mots clés : optimisation sans contraintes, méthode Quasi-Newton, méthodes descendantes.

INTRODUCTION

L'idée fondamentale des méthodes de type Quasi-Newton est de générer une suite $(B^k)_{k \in \mathbb{N}}$ de matrices symétriques définies positives et de disposer, à chaque itération, d'une approximation de la matrice Hessienne réelle ou de son inverse. Malheureusement, dans la formule de Quasi-Newton de Broyden il n'y a aucune garantie que les matrices produites à chaque itération restent définies positives même si la fonction f est quadratique et que $B^0 = I$. On peut cependant noter l'intérêt des méthodes de mise à jour de rang 2 telle que la méthode de *DFP* et *BFGS*. L'objet de cet étude est de particulariser cette classe de méthodes au cas de figure dans lequel on désire également préserver une matrice diagonale à chaque itération. L'approximation de la matrice Hessienne par la méthode *DFP* est améliorée cependant par une procédure de diagonalisation, et ce pour réduire le coût calculatoire des opérations nécessaires pour exécuter l'algorithme en question.

Le mémoire est divisé en trois chapitres, en premier, On donne un rappel sur les notions fondamentales de l'analyse matricielle et l'analyse convexe. Le deuxième chapitre est réservé au développement d'un algorithme pour la résolution d'un problème d'optimisation sans contraintes basé essentiellement sur l'idée de diagonalisation de la matrice produite par la méthode *DFP*. Le chapitre 3, est réservé entièrement aux tests numériques. Cette étude est étayée par des expérimentations numériques concluantes. Notre travail ouvre des perspectives pour d'autres classes de problème d'optimisation.

Notations et terminologie

(P) : Problème d'optimisation.

Vecteurs

Les vecteurs sont désignés par des lettres minuscules. Si x est un vecteur de \mathbb{R}^n , on désigne par :

x^T : Le vecteur transposé de x ;

x_i : La i -ème composante de x ;

x^k : Le k -ème vecteur d'une suite de vecteurs.

Matrices

A^T : La matrice transposée de A ;

A^{-1} : L'inverse de la matrice A ;

I_n : La matrice identité de l'ordre n ;

Δ_k : La matrice correction de rang 2.

Opérations

$A \succ 0$: A est une matrice symétrique définie positive ;

$A \succeq 0$: A est une matrice symétrique sem-définie positive ;

$\langle x, y \rangle = x^T y$: Le produit scalaire de deux vecteurs avec $x, y \in \mathbb{R}^n$.

Ensembles

\mathbb{R}^n : L'espace euclidien des n -composantes réelles ;

$\mathbb{R}^{m \times n}$: L'espace des matrices réelles à m lignes et n colonnes ;

$\mathbb{R}^{n \times n}$: L'espace des matrices carrées d'ordre n .

Fonctions

Soit $f(\mathbb{R}^n \rightarrow \mathbb{R})$ une fonction différentiable à plusieurs variables $(x_1 \cdots x_n)$. Alors :

$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \dots, \frac{\partial f(x)}{\partial x_n} \right)^T$: gradient de f au point x ;

$\frac{\partial f(x)}{\partial x_i}$: Les dérivées partielles au point x ;

$H(x) : \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right)_{1 \leq i, j \leq n}$, La matrice Hessienne ;

C^1 : Espace vectoriel des fonctions continues et dérivables sur \mathbb{R}^n ;

C^2 : Espace vectoriel des fonctions deux fois continûment différentiables sur \mathbb{R}^n .

Chapitre 1

Rappels d'analyse matricielle et d'analyse convexe

1.1 Rappel sur le calcul matriciel

Soit \mathbb{R} désigne l'espace vectoriel.

1.1.1 Les valeurs propres et les vecteurs propres d'une matrice

Soit A une matrice carrée, $A \in \mathbb{R}^{n \times n}$. Les valeurs propres λ de cette matrice sont les racines de l'équation caractéristique :

$$\det(A - \lambda I) = 0.$$

L'ensemble des valeurs propres d'une matrice constitue son spectre. Les vecteurs propres v_i de cette matrice se calcul de l'équation :

$$Av_i = \lambda v_i$$

Comme le système est indéterminé, les vecteurs propres sont obtenus à une constante multiplicative près, c'est à dire que l'on détermine des directions propres.

Remarque 1 $\det(A - \lambda I) = 0$ est appelé *polynôme caractéristique de A*.

$\lambda I - A$ est appelée *la matrice caractéristique de A*.

Deux matrices sont dites *semblables* si elles ont le même *polynôme caractéristique*.

Propriétés

Si A est une matrice carrée, $A \in \mathbb{R}^{n \times n}$, ayant pour valeurs propres $\lambda_i, i = 1, \dots, n$:

$$\det(A) = \prod_{i=1}^n \lambda_i$$

$$\text{trace}(A) = \sum_{i=1}^n \lambda_i, \text{ désigne le trace de } A.$$

1.1.2 La définie positivité

On dit qu'une matrice $A \in \mathbb{R}^{n \times n}$ est symétrique définie positive, $A \succ 0$, si et seulement si la forme quadratique qui lui est associée est définie positive :

$$A \succ 0 \iff x^T A x > 0, \quad \forall x \neq 0.$$

A est symétrique semi définie positive, $A \succeq 0$ si et seulement si $x^T A x \geq 0, \forall x \in \mathbb{R}^n$.

A est symétrique définie négative, $A \prec 0$ si et seulement si $x^T A x < 0$ pour tout vecteur $x \neq 0$.

A est symétrique semi définie négative, $A \preceq 0$ si et seulement si $x^T A x \leq 0, \forall x \in \mathbb{R}^n$.

On parle, selon le cas, de positivité ou de négativité.

1.1.3 Le gradient d'une fonction scalaire

Soit une fonction multivariable à valeur réel, définie par :

$$\begin{aligned} f & : \mathbb{R}^n \longrightarrow \mathbb{R}. \\ x & \longmapsto f(x) \end{aligned}$$

$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_i} \right)_{i=1, \dots, n}^T = \left(\frac{\partial f(x)}{\partial x_1}, \dots, \frac{\partial f(x)}{\partial x_n} \right)^T.$$

ou $\frac{\partial f}{\partial x_i}$ sont les dérivés partielles de f en x_i .

1.1.4 La Hessienne d'une fonction scalaire

Étant donné une fonction scalaire $f(x)$, $f : \mathbb{R}^n \longrightarrow \mathbb{R}$. La matrice Hessienne de $f(x)$ relative au vecteur $x \in \mathbb{R}^n$ est définie comme la matrice symétrique suivante :

$$H(x) = \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right)_{n \times n} = \begin{pmatrix} \frac{\partial^2 f(x)}{\partial^2 x_1} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(x)}{\partial^2 x_n} \end{pmatrix}.$$

1.2 Rappels d'analyse convexe

1.2.1 Eléments d'analyse convexe

Ensembles convexes

Définition 1 *Un sous-ensemble C de \mathbb{R}^n est dit convexe si :*

$$\forall x, y \in C : (1 - \lambda)x + \lambda y \in C, \forall \lambda \in [0, 1].$$

Autrement dit si le segment de droite joignant deux points quelconques

$$x, y \in C : [x, y] = \{(1 - \lambda)x + \lambda y \in C, 0 \leq \lambda \leq 1\}$$

est entièrement inclus dans C .

Une conséquence immédiate de la définition dit que **toute intersection de convexes est convexe**. En particulier **toute intersection de convexes fermés est un ensemble convexe fermé**.

Définition 2 *C est dit affine si :*

$$\forall x, y \in C : (1 - \lambda)x + \lambda y \in C, \forall \lambda \in \mathbb{R}.$$

Fonctions convexes

Soit $f : \mathbb{R}^n \longrightarrow \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\} = [-\infty, +\infty]$. On associe à f les ensembles suivants :

Epigraphe de f : $\text{epi}(f) = \{(x, \lambda) \in \mathbb{R}^{n+1} : f(x) \leq \lambda\}$.

Epigraphe strict de f : $\widetilde{\text{epi}}(f) = \{(x, \lambda) \in \mathbb{R}^{n+1} : f(x) < \lambda\}$.

Courbe ou ensemble de niveau : $S_\lambda(f) = \{x \in \mathbb{R}^n : f(x) \leq \lambda\}$.

Courbe stricte : $\widetilde{S}_\lambda(f) = \{x \in \mathbb{R}^n : f(x) < \lambda\}, \lambda \in \mathbb{R}$.

Domaine effectif de f : $\text{dom}(f) = \{x \in \mathbb{R}^n : f(x) < +\infty\}$.

Définition 3 Une fonction $f : \mathbb{R}^n \longrightarrow \bar{\mathbb{R}}$ est dite *convexe* sur \mathbb{R}^n si :

$$f[(1 - \lambda)x + \lambda y] \leq (1 - \lambda)f(x) + \lambda f(y), \forall x, y \in \mathbb{R}^n, \forall \lambda \in [0, 1].$$

* f est dite **strictement convexe** sur \mathbb{R}^n si :

$$f[(1 - \lambda)x + \lambda y] < (1 - \lambda)f(x) + \lambda f(y), \forall x, y \in \mathbb{R}^n, x \neq y, \forall \lambda \in]0, 1[$$

* f est dite **concave** sur \mathbb{R}^n si $-f$ est convexe sur \mathbb{R}^n .

Définition 4 La fonction f est dite *affine* si :

$$f[(1 - \lambda)x + \lambda y] = (1 - \lambda)f(x) + \lambda f(y), \forall \lambda \in \mathbb{R}.$$

La fonction $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ est dite **semi-continue inférieurement** (s.c.i) en x^0 si :

$$\forall \lambda \in \mathbb{R}, \lambda < f(x^0) \implies \exists \alpha > 0 \text{ tel que } \|x - x^0\| \leq \alpha \implies f(x) > \lambda.$$

* f est dite **semi continue supérieurement** (s.c.s) en x^0 si $-f$ est semi continue inférieurement en x^0 .

* f est dite continue en x^0 (resp. sur C) si f est à la fois semi continue inférieurement et semi continue supérieurement en x^0 (resp. sur C).

Définition 5 Soit $f : \mathbb{R}^n \longrightarrow \bar{\mathbb{R}}$ une fonction propre et $x^0 \in \mathbb{R}^n$ tel que $f(x^0)$ est une valeur finie. $g \in \mathbb{R}^n$ est appelé **sous gradient** de f en x^0 si :

$$f(x) \geq f(x^0) + \langle x - x^0, g \rangle, \forall x \in \mathbb{R}^n.$$

L'ensemble de tous les sous gradients de f en x^0 est appelé **sous différentiel** de f en x^0 et noté $\partial f(x^0)$.

Propriétés

1. $\partial f(x^0)$ est un convexe fermé, (éventuellement vide).
2. $x^0 \notin \text{dom}(f) \implies \partial f(x^0) = \emptyset$ par définition.
3. Si $\partial f(x^0) \neq \emptyset$, f est dite **sous différentiable** en x^0 .
4. f est continue en $x^0 \implies \partial f(x^0) \neq \emptyset$ et compact (borné). La réciproque est fausse :
(une fonction sous-différentiable n'est pas nécessairement continue).
5. Si f est convexe, alors :
 - * f est différentiable en x^0 si et seulement si $\partial f(x^0)$ est un singleton, alors on a :
 $\partial f(x^0) = \{\nabla f(x^0)\}$.
 - * $\partial f(x^0) \neq \emptyset \implies f$ s.c.i en x^0 ($f(x^0) = \bar{f}(x^0)$). La réciproque est fausse.
 - * f s.c.i propre $\implies \partial f(x^0) \neq \emptyset \forall x^0 \in \text{ri}(\text{dom} f)$.
6. Si $x \in \text{int}(\text{dom} f)$ alors $\partial f(x)$ est convexe compact non vide.

Règles de calcul

- * $\partial(\lambda f)(x) = \lambda \partial f(x)$, $\forall \lambda > 0$.
- * $\partial(f + g)(x) \supseteq \partial f(x) + \partial g(x)$, $\forall x$. Si f et g sont convexes et propres, alors il y a égalité lorsque $\text{int}(\text{dom} f) \cap \text{int}(\text{dom} g) \neq \emptyset$.
- * Si $f(x) = \max_{i \in I} f_i(x)$, f_i convexe $\forall i \in I = \{1, \dots, m\}$ alors

$$\partial f(x) = \text{co} \{ \partial f_i(x), i \in I : f_i(x) = f(x) \}.$$

1.3 Rappels d'optimisation

Sous sa forme générale, un problème d'optimisation s'écrit comme suit :

$$\min_x [f(x) : x \in C]. \quad (\text{P})$$

La fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est appelée **fonction objectif**, l'ensemble $\emptyset \neq C \subset \mathbb{R}^n$ est appelé **l'ensemble des solutions réalisables**.

On appelle **solution optimale globale** de (P), tout point $\bar{x} \in C$ satisfaisant $f(\bar{x}) \leq f(x), \forall x \in C$. L'ensemble des points \bar{x} ainsi défini est noté $\arg \min_C f(x)$.

On dit qu'un point $\bar{x} \in C$ est **solution optimale locale** de (P), s'il existe un voisinage V de \bar{x} tel que $f(\bar{x}) \leq f(x), \forall x \in C \cap V$. L'ensemble des points \bar{x} ainsi défini est noté $\text{loc min}_C f(x)$.

Nous avons toujours $\arg \min_C f(x) \subseteq \text{loc min}_C f(x)$.

Rappelons que si f est semi continue inférieurement (s.c.i) et C compact (fermé et borné) alors $\arg \min_C f(x) \neq \emptyset$.

1.3.1 Conditions d'optimalité

A) Le cas sans contraintes

Si $C = \mathbb{R}^n$. Alors (P) devient

$$\min_x [f(x) : x \in \mathbb{R}^n]. \quad (1.1)$$

Conditions du premier ordre

Supposons que la fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est de classe C^1 sur \mathbb{R}^n .

Conditions nécessaires (CN1)

– Si f a un minimum local en \bar{x} alors $\nabla f(\bar{x}) = 0$.

Conditions suffisantes (CS1)

- Si f est convexe et $\nabla f(\bar{x}) = 0$ alors f a un minimum global en \bar{x} .

Conditions du second ordre

Supposons la fonction $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ de classe C^2 sur \mathbb{R}^n .

Conditions nécessaires (CN2)

- Si f a un minimum local en $\bar{x} \in \mathbb{R}^n$ alors $\nabla f(\bar{x}) = 0$ et la matrice Hessienne est sem-définie positive.

Conditions suffisantes (CS2)

- Si $\nabla f(\bar{x}) = 0$ et si la matrice $\nabla^2 f(\bar{x})$ est définie positive, alors f a un minimum local en \bar{x} .

Cas des fonctions convexes

Si f est convexe (non nécessairement différentiable), alors une condition nécessaire et suffisante (CNS) pour que x soit un minimum global est :

$$0 \in \partial f(\bar{x}).$$

B) Le cas avec contraintes

On considère le problème (P) :

$$\min_x [f(x) : x \in C]. \tag{1.2}$$

où C est un convexe de \mathbb{R}^n et f est C^1 sur un ouvert contenant C . On a les conditions suivantes d'optimalité en un point $\bar{x} \in C$:

- (CN1) Si f a un minimum local sur C en \bar{x} alors $\langle \nabla f(\bar{x}), x - \bar{x} \rangle \geq 0$ pour tout $x \in C$.
- (CS1) Si f est convexe et $\langle \nabla f(\bar{x}), x - \bar{x} \rangle \geq 0$ pour tout $x \in C$ alors f a un minimum global sur C en \bar{x} .

Notons que dans les deux cas nous avons $\nabla f(\bar{x}) \in \tilde{N}(\bar{x})$ on

$$\tilde{N}(\bar{x}) = \left\{ x \in \mathbb{R}^n : \langle x, y - \bar{x} \rangle \leq 0 \forall y \in \tilde{S}(\bar{x}) \right\}$$

est le cône normal à l'ensemble de niveau

$$\tilde{S}(\bar{x}) = \{y \in \mathbb{R}^n \text{ tel que } f(y) < f(\bar{x})\}.$$

On suppose maintenant que C est défini comme suit :

$$C = \{x \in \mathbb{R}^n : g_i(x) \leq 0, i = 1, \dots, p; h_j = 0, j = 1, \dots, q\}.$$

Les contraintes $g_i(x) \leq 0$ sont appelées **contraintes d'inégalités** et les contraintes $h_j(x) = 0$ **contraintes d'égalités**.

L'obtention des conditions nécessaires d'optimalité nécessite que certaines conditions appelées conditions de qualification des contraintes soient vérifiées.

Qualification des contraintes :

Sans rentrer dans des détails inutiles, nous rappelons ici les conditions suffisantes les plus utilisées pour la qualification des contraintes.

Condition de Slater

La condition de qualification des contraintes de Slater est comme suit :

- Les fonctions g_i sont convexes et les fonctions h_j sont affines.
- Il existe x^0 tel que $g_i(x^0) < 0$ et $h_j(x^0) = 0$ pour tout i, j .

Condition de Mangasarian-Fromowitz

La condition de qualification des contraintes de Mangasarian-Fromowitz en un point $\bar{x} \in C$ est comme suit :

- Les vecteurs $\nabla h_j(\bar{x})$ sont linéairement indépendants.
- Il existe \bar{d} tel que $\langle \nabla h_j(\bar{x}), \bar{d} \rangle = 0$ pour tout j et $\langle \nabla g_i(\bar{x}), \bar{d} \rangle < 0$ pour tout i .

Théorème 1 *Un ensemble $C \subset \mathbb{R}^n$ est compact si et seulement si il est fermé et borné.*

1.3.2 Direction de descente

- Un élément $d \in \mathbb{R}^n$ est appelé une direction de descente pour f en $x^0 \in C$ si : $\exists T > 0 : \forall \alpha \in]0, T[, f(x^0 + \alpha d) < f(x^0)$.
- Si la dérivée directionnelle de f en x^0 suivant d existe alors :

$$f'(x^0, d) < 0 \implies d \text{ une direction de descente pour } f \text{ en } x^0.$$

- En particulier si f est différentiable en x^0 , on a :

$$\nabla f(x^0)^T d < 0 \implies d \text{ une direction de descente pour } f \text{ en } x^0.$$

Voici maintenant deux résultats d'existences :

Théorème 2 (Weierstrass) [3]

Si f continue sur C compact} $\implies \arg \min_{x \in C} f(x) \neq \emptyset$

Corollaire 1 :

*Si f continue sur C fermé
et coercive ($\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty$) } $\implies \arg \min_C f(x) \neq \emptyset$*

Remarque 2 *L'unicité d'une solution éventuelle est en général une conséquence de la convexité de C et la stricte convexité de f .*

Théorème 3 *Soit $f : C \subset \mathbb{R}^n \rightarrow \mathbb{R}$ strictement convexe sur C convexe. Le minimum de f sur C s'il existe, il est unique.*

1.3.3 Méthodes de recherche linéaire

L'objectif d'une recherche linéaire est de trouver un pas de déplacement α^k pour que la fonction f ait une décroissance suffisante le long d'une direction de descente d^k . Cela se traduit souvent par une inégalité de la forme :

$$f(x^k + \alpha^k d^k) \leq f(x^k),$$

où x^k désigne l'iteration à l'ordre k .

Quand l'iteration est loin du minimum de f , prendre le pas de Newton peut ne pas diminuer la fonction ; il y a une possibilité de se déplacer trop loin pour que l'approximation quadratique soit valable. Un déplacement le long de la direction de Newton garantit seulement que f décroisse initialement. Ainsi, le but est de déplacer vers un nouveau point x^{k+1} pour lequel le modèle quadratique de f soit plus précis. En fait, si le pas est trop grand, il faut le réduire. Ceci se traduit par :

$$x^{k+1} = x^k + \alpha^k d^k, \quad 0 < \alpha^k \leq 1.$$

Pour simplifier les notations, nous définissons la restriction de f à la demi-droite

$$\{x^k + \alpha^k d^k : \alpha^k \geq 0\}$$

par la fonction

$$\varphi_k(\alpha) = f(x^k + \alpha d^k).$$

Comme nous cherchons à minimiser f , il semble naturel de chercher à minimiser le critère le long de d^k et donc de déterminer le pas de déplacement α^k comme solution du problème

$$\min_{\alpha \geq 0} \varphi_k(\alpha).$$

Jusqu'au milieu des années 60, la méthode classique était de choisir α^k pour que x^{k+1}

minimise exactement f dans la direction d^k . Une telle approche s'appelle une recherche linéaire exacte. Elle donne un meilleur déplacement le long de la direction de descente, mais elle est coûteuse à cause du nombre d'évaluations de la fonction. D'ailleurs, elle n'est plus nécessaire. En effet, puisque d^k est toujours la direction de Newton dans notre méthode, nous essayons d'abord $\alpha^k = 1$, ceci peut mener à la convergence quadratique si x^k est suffisamment proche de la solution. Cependant, si $f(x^{k+1})$ ne vérifie pas nos critères de décroissance suffisante, nous faisons un rebroussement le long de la direction de Newton, essayant une valeur plus petite de α^k , jusqu'à ce que nous trouvions un point approprié. Puisque la direction de Newton est une direction de descente pour f , il est assuré que f décroisse pour une valeur de α^k suffisamment petite. Une telle approche s'appelle une recherche linéaire inexacte. Dans ce cas, des conditions moins restrictives et plus facilement vérifiées seront utilisées. D'ailleurs, il n'y aura plus un unique pas (ou quelques pas) vérifiant ces conditions mais tout un intervalle de pas (ou plusieurs intervalles), ce qui rendra la recherche de tels pas plus aisée. Les sections ci-dessous présentent les conditions pour accepter un pas menant à une décroissance suffisante de f .

La règle d'Armijo

- Si $\varphi(\alpha) \leq \varphi(0) + m\varphi'(0)\alpha$, alors α convient.
- Si $\varphi(\alpha) > \varphi(0) + m\varphi'(0)\alpha$, alors α est trop grand.

où $m \in]0, 1[$

On peut noter que l'on a

$$\begin{aligned}\varphi(0) &= f(x^k), \\ \varphi'(0) &= \nabla f(x^k)^T d^k.\end{aligned}$$

La règle de Goldstein

En ajoutant une deuxième inégalité à la règle d'Armijo, on obtient la règle de Goldstein, où m_1 et m_2 sont deux constantes vérifiant $0 < m_1 < m_2$:

- Si $\varphi(\alpha) < \varphi(0) + m_2\varphi'(0)\alpha$, alors α est trop petit.
- Si $\varphi(\alpha) > \varphi(0) + m_1\varphi'(0)\alpha$, alors α est trop grand.
- Si $\varphi(0) + m_1\varphi'(0)\alpha \geq \varphi(\alpha) \geq \varphi(0) + m_2\varphi'(0)\alpha$, alors α convient.

Chapitre 2

Technique de diagonalisation dans les méthodes de Quasi-Newton

2.1 Méthode de Newton et Quasi Newton

2.1.1 Rappel

Rappelons brièvement la procédure des méthodes de type Quasi Newton. On envisage un problème de minimisation sans contraintes d'une fonction deux fois continûment dérivable :

$$\min_{x \in \mathbb{R}^n} f(x) \quad (2.1)$$

2.1.2 Méthode de Newton

La méthode de Newton consiste à remplacer, au voisinage du point courant x^k , la fonction par son approximation quadratique suivante :

$$q(x) = f(x^k) + \nabla f^T(x^k)(x - x^k) + \frac{1}{2}(x - x^k)^T \nabla^2 f(x^k)(x - x^k) \quad (2.2)$$

On prend alors comme nouveau candidat optimum, le point x^{k+1} minimum de $q(x)$ lorsqu'il existe. Ceci ne peut être le cas que si la matrice Hessienne $\nabla^2 f(x^k)$ est définie positive. Dans ce cas, la fonction $q(x)$ est strictement convexe et a un minimum unique x^{k+1} défini par la condition $\nabla q(x^{k+1}) = 0$. Soit en explicitant cette condition :

$$x^{k+1} = x^k - [\nabla^2 f(x^k)]^{-1} \cdot \nabla f(x^k) \quad (2.3)$$

où $[\nabla^2 f(x^k)]^{-1}$ est l'inverse de la matrice $\nabla^2 f(x^k)$. Pour remédier à ces difficultés, une première modification consiste à introduire une recherche linéaire le long de la direction de recherche :

$$d^k = -[\nabla^2 f(x^k)]^{-1} \cdot \nabla f(x^k). \quad (2.4)$$

On détermine le pas α^k afin de minimiser $\varphi(\alpha) = f(x^k + \alpha d^k)$ ou, tout au moins, de trouver un nouveau point tel que $f(x^k + \alpha^k d^k) < f(x^k)$.

2.1.3 Relation de Quasi Newton

Une méthode de Quasi Newton est une méthode du type :

$$\begin{cases} d^k = -B^k g^k, \\ x^{k+1} = x^k + \alpha^k d^k, \end{cases}$$

ou

$$\begin{cases} d^k = -H^{-1} g^k, \\ x^{k+1} = x^k + \alpha^k d^k, \end{cases}$$

où B^k (respectivement H^k) est une matrice destinée à approcher l'inverse du hessien de f (respectivement le hessien de f) en x^k . Il se pose donc un problème : quelle stratégie adopter pour faire cette approximation ? On peut démarer par exemple : $B^0 = I$, mais comment ensuite mettre à jour l'approximation B^k au cours des itérations. L'idée est la suivante : on sait qu'au point x^k , le gradient et le hessien de f vérifient la relation

$$g^{k+1} = g^k + \nabla^2 f(x^k)(x^{k+1} - x^k) + \epsilon(x^{k+1} - x^k).$$

Si on suppose que l'approximation quadratique est bonne, on peut alors négliger le terme $\epsilon(x^{k+1} - x^k)$ et considérer que l'on a :

$$g^{k+1} - g^k \approx \nabla^2 f(x^k)(x^{k+1} - x^k),$$

cela conduit à la notion de relation de Quasi-Newton :

Définition 6 *On dit que les matrices B^{k+1} (H^{k+1}) vérifient une relation de Quasi-Newton si :*

$$H^{k+1}(x^{k+1} - x^k) = \nabla f(x^{k+1}) - \nabla f(x^k) = g^{k+1} - g^k,$$

ou

$$x^{k+1} - x^k = B^{k+1}(\nabla f(x^{k+1}) - \nabla f(x^k)).$$

Il reste le problème à résoudre : comment mettre à jour B^k tout en assurant $B^k \succ 0$? C'est ce que nous allons voir maintenant.

2.2 Formules de mise à jour de l'approximation du hessien

Le principe de la mise à jour consiste, à une itération donnée de l'algorithme :

$$\begin{cases} d^k = -B^k g^k, \\ x^{k+1} = x^k + \alpha^k d^k, \end{cases}$$

à appliquer une formule du type

$$B^{k+1} = B^k + \Delta_k,$$

avec Δ_k symétrique, assurant la relation de Quasi Newton

$$x^{k+1} - x^k = B^{k+1}(g^{k+1} - g^k), \quad (\text{A})$$

ainsi que $B^{k+1} \succ 0$, sous l'hypothèse que $B^k \succ 0$.

La formule (A) permet d'utiliser, les nouvelles informations obtenues lors de l'étape k de l'algorithme, c'est à dire essentiellement le gradient $g^{k+1} = \nabla f(x^{k+1})$ au point x^{k+1} , obtenu par recherche linéaire (exacte ou approchée) dans la direction d^k . Il existe différentes formules du type (A). Suivant que Δ_k est de rang 1 ou 2, on parlera de correction de rang 2.

2.2.1 Formule de Broyden

On peut chercher à déterminer une formule de correction de rang 1 de la façon suivante. On écrit B^{k+1} sous la forme

$$B^{k+1} = B^k + vv^T,$$

et on cherche $v \in \mathbb{R}^n$ tel que la relation de Quasi Newton

$$B^{k+1}y^k = s^k,$$

on pose $y^k = g^{k+1} - g^k$ et $s^k = x^{k+1} - x^k$. On a donc

$$B^k y^k + vv^T y^k = s^k,$$

et en prenant le produit scalaire des deux membres de l'égalité précédente avec y^k on obtient

$$((y^k)^T v)^2 = (s^k - B^k y^k)^T y^k.$$

Si on utilise maintenant l'égalité

$$v v^T = \frac{v v^T y^k (v v^T y^k)^T}{(v^T y^k)^2},$$

alors on peut écrire, en remplaçant $v^T y^k$ par $s^k - B^k y^k$ et $(v^T y^k)^2$ par $(y^k)^T (s^k - B^k y^k)$, la formule de correction devient :

$$B^{k+1} = B^k + \frac{(s^k - B^k y^k)(s^k - B^k y^k)^T}{(s^k - B^k y^k)^T y^k},$$

connue sous le nom de *formule de Broyden*. La validité de cette formule provient du résultat suivant :

Théorème 4 [10] *Soit f une forme quadratique définie positive. Considérons la méthode itérative qui, partant d'un point x_0 arbitraire engendre successivement les points*

$$x^{k+1} = x^k + s^k,$$

les s^k sont des vecteurs linéairement indépendants. Alors la suite (B^k) générée par B^0 qu'est symétrique et donnée par la formule :

$$\begin{aligned} B^{k+1} &= B^k + \frac{(s^k - B^k y^k)(s^k - B^k y^k)^T}{(s^k - B^k y^k)^T y^k}, \\ \text{avec } y^k &= \nabla f(x^{k+1}) - \nabla f(x^k), \end{aligned}$$

en au plus n étapes vers A^{-1} , l'inverse du hessien de f .

Preuve. Puisque le hessien de f est constant et égal à A on a :

$$y^i = \nabla f(x^{i+1}) - \nabla f(x^i) = A(x^{i+1} - x^i), \forall i.$$

On a vu que B^{k+1} est construite de façon à ce que l'on ait :

$$B^{k+1}y^k = s^k,$$

montrons que l'on a aussi

$$B^{k+1}y^i = s^i, \quad i = 0, \dots, k-1.$$

On raisonne par récurrence en supposant que cette propriété est vraie pour B^k , à savoir

$$B^{k+1}y^i = s^i, \quad i = 0, \dots, k-2.$$

Soit donc $i \leq k-2$ quelconque. On a

$$B^{k+1}y^i = B^k y^i + \frac{(s^k - B^k y^k)((s^k)^T y^i - B^k (y^k)^T y^i)}{(s^k - B^k y^k)^T y^k}. \quad (\text{B})$$

Par l'hypothèse de récurrence on a $B^k y^i = s^i$ donc

$$(y^k)^T B^k y^i = (y^k)^T s^i,$$

mais comme $As^i = y^i, \forall i$, on obtient

$$(y^k)^T s^i = (s^k)^T As^i = (s^k)^T y^i,$$

donc dans (B) le numérateur est nul et on a $B^{k+1}y^i = B^k y^i = s^i$. On a donc

$$B^{k+1}y^i = s^i, \quad i = 0, \dots, k.$$

Au bout de n itérations on a donc

$$B^n y^i = s^i, \quad i = 0, \dots, n-1,$$

et puisque l'on a $y^i = As^i$ cette dernière formule s'écrit comme :

$$B^n As^i = s^i, \quad i = 0, \dots, n-1.$$

Comme les s^i constituent une base de \mathbb{R}^n on a $B^n A = I$ ou encore

$$B^n = A^{-1};$$

ce qui montre le résultat. ■

Inconvénient :

Le problème de la formule de Broyden est qu'il n'y a aucune garantie que les matrices B^k soient définies positives même si la fonction f est quadratique et si par exemple $B^0 = I$. On peut cependant recourir aux méthodes de rang 2 (DFP , $BFGS$).

2.2.2 Formule de Davidon, Fletcher et Powell (DFP)

La formule de mise à jour de Davidon, Fletcher et Powell est une formule de correction de rang 2 donnée par :

$$B^{k+1} = B^k + \frac{s^k (s^k)^T}{(s^k)^T y^k} - \frac{B^k y^k (y^k)^T B^k}{(y^k)^T B^k y^k}. \quad (C)$$

Le résultat suivant montre que sous certaines conditions, la formule (C) conserve la définie-positivité des matrices B^k .

Théorème 5 [10] *On considère la méthode définie par*

$$\begin{aligned} d^k &= -B^k g^k, \\ x^{k+1} &= x^k + \alpha^k d^k, \end{aligned}$$

α^k optimal, où $B^0 \succ 0$ est donnée ainsi que x^0 . Alors les matrices B^k sont définies positives, $\forall k > 0$.

Preuve. Soit x un vecteur de \mathbb{R}^n . On a :

$$\begin{aligned} x^T B^{k+1} x &= x^T B^k x + \frac{((s^k)^T x)^2}{(s^k)^T y^k} - \frac{((y^k)^T B^k x)^2}{(y^k)^T B^k y^k} \\ &= \frac{(y^k)^T B^k y^k x^T B^k x - ((y^k)^T B^k x)^2}{(y^k)^T B^k y^k} + \frac{((s^k)^T x)^2}{(s^k)^T y^k}, \end{aligned}$$

si on définit le produit scalaire $\langle x, y \rangle = x^T B^k y$ alors on a :

$$x^T B^{k+1} x = \frac{\langle y^k, y^k \rangle \langle x, x \rangle - \langle y^k, x \rangle^2}{\langle y^k, y^k \rangle} + \frac{((s^k)^T x)^2}{(s^k)^T y^k}. \quad (D)$$

Le premier terme du second membre est positif ou nul d'après l'inégalité de Cauchy-Schwartz. Quant au deuxième terme, on peut faire l'analyse suivante : puisque le pas est optimal, on a la relation :

$$(g^{k+1})^T d^k = 0,$$

donc

$$(s^k)^T y^k = \alpha^k (g^{k+1} - g^k)^T d^k = \alpha^k (g^k)^T B^k g^k > 0,$$

a donc $x^T B^{k+1} x \geq 0$. Les deux termes dans (D) étant positifs, cette quantité ne peut s'annuler que si les deux termes sont simultanément nuls. Le premier terme ne peut s'annuler que si $x = \lambda y^k$ pour un scalaire $\lambda \neq 0$. Dans ce cas le deuxième terme est non nul car

$(s^k)^T x = \lambda (s^k)^T y^k$. On a donc bien $B^{k+1} \succ 0$. ■

Remarque 3 La propriété $(s^k)^T y^k > 0$ est vérifiée également par des méthodes de recherche linéaire approchées comme par exemple la règle de Wolfe et Powell : en effet dans ce cas on détermine un point x^{k+1} tel que :

$$\varphi'(\alpha^k) = \nabla f(x^{k+1})^T d^k \geq m_2 \nabla f(x^k)^T d^k, 0 < m_2 < 1,$$

d'où

$$(g^{k+1})^T \frac{x^{k+1}}{\alpha^k} > (g^k)^T \frac{x^{k+1} - x^k}{\alpha^k},$$

et donc

$$(g^{k+1} - g^k)^T (x^{k+1} - x^k) > 0.$$

2.2.3 Algorithme de Davidon-Fletcher-Powel (DFP)

1. Choisir x^0 et B^0 définie positive quelconque (par exemple $B^0 = I$);
2. A l'itération k , calculer la direction de déplacement

$$d^k = -B^k \nabla f(x^k);$$
 déterminer le pas optimal α^k et poser

$$x^{k+1} = x^k + \alpha^k d^k;$$
3. Poser $s^k = \alpha^k d^k$ et $y^k = \nabla f(x^{k+1}) - \nabla f(x^k)$, puis calculer

$$B^{k+1} = B^k + \frac{s^k (s^k)^T}{(s^k)^T y^k} - \frac{B^k y^k (y^k)^T B^k}{(y^k)^T B^k y^k};$$
4. Faire $k \leftarrow k + 1$. Retourner en 1 sauf si le critère d'arrêt est vérifié.

Comme critère d'arrêt, on retiendra par exemple $\|g^{k+1}\| < \epsilon$.

Cet algorithme a un comportement remarquable dans le cas où f est une forme quadratique :

Théorème 6 [10] Appliqué à une forme quadratique f , l'algorithme DFP engendre des

directions s^0, \dots, s^k vérifiant

$$\langle s^i, s^j \rangle_A = (s^i)^T A s^j = 0, 0 \leq i < j \leq k+1, \quad (E)$$

$$B^{k+1} A s^i = s^i, 0 \leq i \leq k. \quad (2.5)$$

Preuve. En utilisant la formule (C), on a pour tout k

$$\begin{aligned} B^{k+1} A s^k &= B^{k+1} y^k, \\ &= s^k, \end{aligned}$$

par construction. Dans (D) est en particulier vérifiée pour $k = 0$, soit

$$B^1 A s^0 = s^0.$$

On a aussi

$$\begin{aligned} (s^0)^T A s^1 &= -\alpha^1 (s^0)^T A B^1 g^1, \\ &= -\alpha^1 (s^0)^T g^1, \\ &= 0, \end{aligned}$$

puisque $B^1 A s^0 = s^0$ et que x^1 est obtenu par un pas optimal dans la direction s^0 . Donc (E) est vérifiée pour $k = 0$.

Supposons maintenant que (E) et (F) sont vérifiées à l'ordre $k - 1$. On peut écrire d'une part pour $i = 0, \dots, k - 1$.

$$\begin{aligned} g^{k+1} - g^{i-1} &= y^{i+1} + y^i + \dots + y^k, \\ &= A(s^{i+1} + s^i + \dots + s^k) \end{aligned}$$

car f est une forme quadratique de hessien A . D'autre part, puisque x^{i+1} est obtenu par un pas optimal dans la direction s^i on a $(s^i)^T g^{i+1} = 0$ et donc

$$(s^i)^T (g^{k+1} - g^{i+1}) = (s^i)^T A (s^{i+1} + s^i + \dots + s^k), \quad i = 0, \dots, k-1,$$

donc en vertu de l'hypothèse de récurrence (conjugaison des s^i) on a :

$$(s^i)^T g^{k+1} = 0, \quad i = 0, \dots, k-1, \quad (G)$$

cette relation reste aussi valable pour $i = k$ puisque l'on a $(s^k)^T g^{k+1} = 0$ (pas optimal). La deuxième hypothèse de récurrence permet donc d'écrire, en remplaçant s^i par $B^{k+1} A s^i$ dans (G)

$$(s^i)^T A B^{k+1} g^{k+1} = 0, \quad i = 0, \dots, k$$

et donc, puisque $H^{k+1} g^{k+1} = -s^{k+1}/\alpha^{k+1}$,

$$(s^i)^T A s^{k+1} = 0, \quad i = 0 \dots k,$$

ce qui démontre donc la propriété (D) au rang k .

Montrons maintenant que

$$B^{k+1} A s^i = s^i, \quad i = 0 \dots k-1.$$

Cette relation est vraie pour $i = k$ comme on l'a déjà montré plus haut. On a :

$$B^{k+1} A s^i = B^k A s^i + \frac{s^k (s^k)^T A s^i}{(s^k)^T y^k} - \frac{B^k y^k (y^k)^T B^k A s^i}{(y^k)^T B^k y^k}.$$

Le deuxième terme du second membre est nul car $(s^k)^T A s^i = 0$. Si on note que par l'hypothèse de récurrence on a $B^k A s^i = s^i$ pour $i = 0, \dots, k-1$ et $(y^k)^T = (s^k)^T A$ le

numérateur du troisième terme est donné par :

$$B^k y^k (y^k)^T B^k A s^i = B^k y^k (s^k)^T A s^i = 0.$$

Par conséquent on a bien

$$B^{k+1} A s^i = s^i, \quad i = 0, \dots, k - 1,$$

ce qui démontre la propriété (F) à l'ordre k . ■

La méthode \mathcal{DFP} se comporte donc, dans le cas quadratique, comme une méthode de directions conjuguées. Dans ce cas l'algorithme converge en au plus n itérations. On peut aussi remarquer que l'on a pour $k = n - 1$, la relation :

$$B^n A s^i = s^i, \quad i = 0, \dots, n - 1,$$

et comme les s^i sont linéairement indépendants (car mutuellement conjugués) on en déduit que

$$B^n = A^{-1}.$$

Remarque 4 On peut montrer que dans le cas général (non quadratique), sous les mêmes réserves que pour la méthode de Fletcher-Reeves (réinitialisation périodique $d^k = -g^k$), cet algorithme permet de converger vers un minimum local \hat{x} de f , et que l'on a

$$\lim_{k \rightarrow \infty} B^k = \nabla^2 f(\hat{x})^{-1},$$

qui montre que près de l'optimum \hat{x} , et si la recherche linéaire est exacte, la méthode se comporte asymptotiquement comme la méthode de Newton classique. Cette remarque permet de justifier le choix d'une estimation du pas de déplacement donnée par $\alpha^k = 1$, dans les méthodes de recherche linéaire approchée.

2.2.4 Algorithme de Broyden, Fletcher, Goldfarb et Shanno (BFGS)

La forme de mise à jour de Broyden, Fletcher, Goldfarb et Shanno *BFGS* est une formule de correction de rang 2 qui s'obtient à partir de la formule de *DFP* en intervertissant les rôles de s^k et y^k .

La formule obtenue permet de mettre à jour une approximation H^k du hessien possédant les mêmes propriétés, à savoir $H^{k+1} \succ 0$ si $H^k \succ 0$ et vérifiant la relation de Quasi-Newton suivante :

$$y^k = H^k s^k.$$

La formule est donc :

$$H^{k+1} = H^k + \frac{y^k (y^k)^T}{(y^k)^T s^k} - \frac{H^k s^k (s^k)^T H^k}{(s^k)^T H^k s^k}.$$

L'algorithme associé est le suivant :

2.2.5 Algorithme de BFGS

1. Choisir x^0 et H^0 définie positive quelconque (par exemple $H^0 = I$) ;
2. A l'itération k , calculer la direction de déplacement

$$d^k = -(H^k)^{-1} \nabla f(x^k);$$
 déterminer le pas optimal α^k et poser

$$x^{k+1} = x^k + \alpha^k d^k;$$
3. Poser $s^k = \alpha^k d^k$ et $y^k = \nabla f(x^{k+1}) - \nabla f(x^k)$ puis calculer

$$H^{k+1} = H^k + \frac{y^k (y^k)^T}{(y^k)^T s^k} - \frac{H^k s^k (s^k)^T H^k}{(s^k)^T H^k s^k};$$
4. Faire $k \leftarrow k + 1$. Retourner en 2 sauf si le critère d'arrêt est vérifié.

Notons que la direction d^k est obtenue par résolution d'un système linéaire. En pratique la mise à jours de H^k est faite directement sur le facteur de Cholesky C_k où $H^k = C_k C_k^T$ ce qui ramène le calcul de d^k au même coût que pour la formule de *DFP*.

De plus, cette technique permet de contrôler précisément la définie positivité de H^k , qui peut se dégrader à cause des erreurs d'arrondi.

Remarque 5 [10] *La méthode de BFGS possède les mêmes propriétés que la méthode de DFP : dans le cas quadratique les directions engendrées sont conjuguées et on a $H^n = A$. Cette méthode est reconnue comme étant beaucoup moins sensible que la méthode DFP aux imprécisions dans la recherche linéaire, du point de vue de la vitesse de convergence. Elle est donc tout à fait adaptée quand la recherche linéaire est faite de façon économique, avec par exemple la règle de Goldstein ou la règle de Wolfe et Powell.*

2.3 La méthode de Gauss-Newton

Dans les problèmes de moindres carrés non linéaires, la fonction à minimiser prend en général la forme :

$$f(x) = \frac{1}{2} \sum_{i=1}^m f_i^2(x),$$

on doit calculer le Hessien de f , qui dans ce cas précis prend une forme particulière : on a d'une part

$$\nabla f(x) = \sum_{i=1}^m \nabla f_i(x) f_i(x),$$

et le hessien de f est donné par

$$\nabla^2 f(x) = \sum_{i=1}^m \nabla f_i(x) \nabla f_i(x)^T + \sum_{i=1}^m f_i(x) \nabla^2 f_i(x).$$

Si l'on se place près de l'optimum, où on supposera que les fonctions $f_i(x)$ sont négligables. D'où on aura :

$$H(x) = \sum_{i=1}^m \nabla f_i(x) \nabla f_i(x)^T,$$

La matrice obtenue possède une propriété intéressante : elle est symétrique sem-définie

positive pour tout x . De plus dans la plupart des cas $m \gg n$ et la matrice est dans la plupart des cas définie positive. La méthode originale que l'on obtient à partir de la méthode de Newton en remplaçant $\nabla^2 f(x)$ par $H(x)$ est la méthode de Gauss-Newton :

$$\begin{cases} x_0 & \text{donné,} \\ H^k = \sum_{i=1}^m \nabla f_i(x) \nabla f_i(x)^T, \\ x^{k+1} = x^k - (H^k)^{-1} \nabla f(x^k). \end{cases}$$

2.3.1 La méthode de Levenberg-Marquardt

Pour assurer la convergence globale de Gauss-Newton, on peut combiner l'algorithme précédent avec une recherche linéaire, et dans ce cas on peut alors faire les itérations suivantes :

$$\begin{cases} d^k = -(H^k)^{-1} \nabla f(x^k) \\ x^{k+1} = x^k + \alpha^k d^k, \end{cases}$$

Cependant, il n'y a aucune garantie pour que H^k reste définie positive, et en générale on fait appel à une méthode modifiée, qui est la méthode de Levenberg-Marquardt : l'idée consiste à remplacer, dans la méthode précédente, la matrice H^k par la matrice $H^k + \mu I$ où μ est un réel positif. Si μ est très grand, on retombe alors sur la méthode de gradient.

Méthode de Levenberg-Marquardt [10]

$$\begin{cases} x_0 & \text{donné,} \\ H^k = \sum_{i=1}^m \nabla f_i(x) \nabla f_i(x)^T, \\ d^k = -(H^k + \mu I)^{-1} \nabla f(x^k) \\ x^{k+1} = x^k + \alpha^k d^k. \end{cases}$$

2.4 Position du problème : génération d'une suite de matrices approximantes creuses où diagonales

Si les méthodes de Quasi-Newton sont applicables avec succès aux problèmes de taille modérée, il n'en va pas de même pour les problèmes de grande taille, puisque ceux-ci exigent, entre autres, la manipulation et la mémorisation de matrices de dimension $n \times n$ symétriques à chaque itération. De plus, les formules d'actualisation n'exploitent absolument pas les caractéristiques liées à la nature du problème. Ainsi, même si de nombreux problèmes de grande taille se caractérisent par une matrice Hessienne creuse, quasi diagonale ou à diagonale dominante, les formules de mise à jour classiques perdent ces propriétés intéressantes. Comme les mises à jour énoncées ci-dessus engendrent de rechercher des formules capables de préserver cette structure creuse. Cette tâche doit reposer sur des fondements mathématiques rigoureux. Grâce aux travaux de Toint (1977) et Thapa (1981), on dispose maintenant d'une théorie permettant de démontrer l'existence et la convergence des approximations de type Quasi-Newton préservant la structure creuse de la matrice Hessienne réelle. Nous avons adapté aux matrices diagonales le théorie générale établie pour des matrices creuses [7].

2.5 Méthodes de Quasi-Newton préservant la structure diagonale des matrices

Dans la suite de ce chapitre, on prend la convention de simplifier les notations. On omet l'indice d'itération k et on remplace l'indice d'itération $k + 1$ par $+$. On cherche à enrichir la matrice B , approximation du Hessien en vue de trouver une matrice B^+ qui soit à la fois symétrique définie positive, et qui satisfasse à l'équation de Quasi-

Newton :

$$B^+ s = y \quad (2.6)$$

$$\text{avec } s = x^+ - x \text{ et } y = g^+ - g. \quad (2.7)$$

heureusement, on sait que ces schémas classiques de mise à jour ne conservent pas la structure diagonale ou creuse de la matrice Hessienne originale. On sait aussi qu'il n'est pas automatique de construire une approximation qui satisfasse en même temps aux contraintes de positivité, de symétrie et à la condition de Quasi Newton. Une manière de procéder pour obtenir une mise à jour de type Quasi Newton qui répond à toutes ces contraintes est de recourir à une minimisation de normes.

2.5.1 Version modifiée des formules de mise à jour de Davidon-Fletcher-Powell (*DFP*)

Toint (1977) [7], a montré qu'il est possible d'obtenir une version modifiée des formules de mise à jour de *Davidon-Fletcher-Powell* pour matrices creuses qui satisfassent à l'équation de Quasi-Newton ainsi qu'aux conditions de symétrie et définie positivité. Celle-ci se définit en trouvant une matrice de correction E de telle sorte que la matrice remise à jour $B^+ = B + E$ soit la plus proche possible de la matrice originale B au sens de la norme de Frobenius, tout en possédant la même structure (creuse ou diagonale). Pour une matrice diagonale, le problème s'énonce mathématiquement comme suit :

$$\begin{cases} \min \|E\|_F \\ \text{avec } Es = y - Bs \\ E_{ij} = 0 \quad (i \neq j) \end{cases} \quad (2.8)$$

Toint a montré que la solution unique E du problème (2.6) est donnée par :

$$E_{ij} = \begin{cases} 0 & (i \neq j) \\ 2\lambda_i s_i & (i = j), \end{cases} \quad (2.9)$$

où le vecteur $\lambda = (\lambda_1, \dots, \lambda_n)^T$ est la solution du système linéaire diagonal :

$$Q\lambda = y - Bs, \quad (2.10)$$

dans lequel la matrice Q est définie par :

$$Q_{ij} = \begin{cases} 0 & (i \neq j) \\ 2s_i^2 & (i = j) \end{cases} \quad (2.11)$$

2.5.2 Généralisation de la mise à jour de la technique de diagonalisation à d'autres procédures de type Quasi-Newton

Les résultats de *Toint* peuvent être étendus à d'autres schémas de mise à jour symétriques. L'extraction de tels procédés de mise à jour, préservant la structure creuse de la matrice repose sur une série de théorèmes énoncés par *Thapa* (1981) [7]. La théorie des matrices diagonales n'est qu'une adaptation de théorie générale pour matrices creuses de *Thapa*. On utilise les notions suivantes afin de distinguer la partie diagonale et la partie non diagonale de la mise à jour :

$$\widehat{B}_D^+ = \begin{cases} \widehat{B}_{ii}^+ & (i = j) \\ 0 & (i \neq j) \end{cases} \quad et \quad \widehat{B}_{ND}^+ = \begin{cases} 0 & (i = j) \\ \widehat{B}_{ij}^+ & (i \neq j) \end{cases} \quad (2.12)$$

Soit $\widehat{B}^+ = \eta B + U$ où U , est une matrice de mise à jour symétrique mais généralement non diagonale et η est un nombre réel. On suppose de plus, que cette approximation \widehat{B}^+ satisfait à l'équation de Quasi Newton :

$$\widehat{B}^+ s = y \quad (2.13)$$

Dès lors, la matrice B^+ la plus proche de \widehat{B}_D^+ dans la norme de Frobenius et satisfaisant à la condition de Quasi Newton, est de la forme :

$$B^+ = \widehat{B}_D^+ + E \quad (2.14)$$

La matrice de correction E est la solution du problème de minimisation de norme :

$$\begin{cases} \min \|E\|_F \\ \text{avec } Es = \widehat{B}_{ND}^+ s \\ E_{ij} = 0 \quad (i \neq j) \end{cases} \quad (2.15)$$

La solution de ce problème est donnée par :

$$E_{ij} = \begin{cases} 0 & (i \neq j) \\ 2\lambda_i s_i & (i = j) \end{cases} \quad (2.16)$$

où le vecteur $\lambda = (\lambda_1, \dots, \lambda_n)^T$ est lui-même la solution du système linéaire :

$$Q\lambda = \widehat{B}_{ND}^+ s \quad (2.17)$$

dans lequel la matrice Q est une matrice diagonale définie par :

$$Q_{ij} = \begin{cases} 0 & (i \neq j) \\ 2s_i^2 & (i = j) \end{cases} \quad (2.18)$$

Il résulte du théorème précédent que la version de la méthode de Quasi-Newton pour matrices diagonales requiert la solution du système linéaire $Q\lambda = \widehat{B}_{ND}^+ s$. Toutefois, il est possible de montrer qu'il n'est pas nécessaire de calculer les éléments non diagonaux de la matrice de remise à jour U_{ND} . En effet, par simple manipulation algébrique, on montre facilement à partir de l'équation de Quasi-Newton (2.11) et des contraintes du problème (2.13), la propriété suivante :

$$\widehat{B}_{ND}^+ s = y - \eta B s - U_D s \quad (2.19)$$

On conclut que seuls les termes diagonaux de U_D doivent être calculés pour obtenir l'analogue de la remise à jour de type \mathcal{DFP} pour des matrices diagonales. On n'évalue donc que les termes :

$$U_{ii} = \frac{s_i^2}{s^T y} - \frac{(B_{ii} y^i)^2}{y^T B y}. \quad (2.20)$$

2.5.3 Caractère défini positif des approximations de la matrice Hessienne

La méthode de Quasi Newton pour matrices diagonales requiert la résolution de deux systèmes linéaires. La mise à jour de l'approximation de la Hessienne nécessite d'abord la résolution du système $Q\lambda = b$. Ensuite, la recherche de la direction de descente demande encore la solution du système $Bd = -g$. Malheureusement, le pro-

blème vient du fait que les matrices B et Q ne sont pas nécessairement définies positives.

Considérons d'abord le système $Q\lambda = b$. En examinant l'expression (2.16) de la matrice Q , il est facile de voir que la matrice Q est symétrique et définie positive si, et seulement si, chaque composante du vecteur s est différente de zéro : $s^i \neq 0, \forall i$. S'il existe au moins un $s^i = 0$, la matrice Q est singulière et le système peut ne pas admettre de solution. Dans ce cas de figure, *Toint* a proposé d'annuler les λ_i et b_i correspondant aux s^i nuls et de résoudre le système réduit obtenu en ignorant la ligne et la colonne correspondante. Toutefois, pour des matrices diagonales, on perd la garantie que la nouvelle approximation diagonale satisfasse à la condition de Quasi-Newton.

La résolution du second système linéaire $Bd = -g$, détermine la direction de progression. Si la matrice B n'est pas définie positive, il n'est plus garantie que d soit une direction de descente. Plusieurs techniques sont discutées par *Thapa* (1981) lorsque l'estimation du Hessien n'est pas définie positive. *Thapa* (1981) a notamment proposé, soit d'utiliser une version creuse de l'algorithme de factorisation de Cholesky soit de recourir à une procédure de Levenberg-Marquardt. On remplace B par $\bar{B} = B + P$, où P est une matrice diagonale dont les éléments sont tous positifs ou par $\bar{B} = B + \mu I$, où μ est le module de la valeur propre la plus négative. On résout ensuite le système modifié $\bar{B}d = -g$.

Un moyen d'obtenir des estimations définies positives de la Hessienne consiste à remplacer B par \bar{B} qui est également diagonale symétrique et définie positive mais qui ne satisfait peut être pas à la condition de Quasi-Newton. On espère ainsi que la nouvelle approximation \bar{B}^+ obtenue à partir de \bar{B} soit meilleure que B^+ obtenue à partir de B . Hélas, *Thapa* (1981) a constaté que dans ces cas difficiles, la mise à jour converge plus lentement [7].

Algorithme de Toint

- 1- Soit $x^0 \in \mathbb{R}^n$ et B^0 une matrice diagonale de dimension n définie positive ;
calculer g^0 ;
- 2- Si B^k n'est pas définie positive, employer une procédure de Levenberg-Marquart pour obtenir une matrice définie positive $\bar{B}^k = B^k + \mu I$;
- 3- Remplacer B^k par son substitut défini positif $B^k := \bar{B}^k$;
- 4- Calculer la direction de descente d^k en résolvant le système $B^k d^k = -g^k$;
- 5- Effectuer une recherche linéaire selon la direction d^k , c'est-à-dire calculer un pas de progression positif α^k le long de d^k de sorte que $f(x^k + \alpha^k d^k) > f(x^k)$;
- 6- $x^{k+1} := x^k + \alpha^k d^k$;
- 7- Calculer g^{k+1} ;
- 8- Calculer $s^k = x^{k+1} - x^k$ et $y^k = g^{k+1} - g^k$;
- 9- Si x^{k+1} ne satisfait pas au critère de convergence, calculer une nouvelle approximation de matrice hessienne comme suit :
 - Calculer les éléments diagonaux de la mise à jour de type \mathcal{BFGS} soit U_D^k :

$$U_{ii}^k = \frac{y_i^2}{s^k \cdot y^k} - \frac{(B_{ii}^k s_i^k)^2}{s^k B^k s^k}$$
 ;
 - Calculer les éléments de la matrice diagonale Q : $Q_{ii} = 2s_i^2$;
 - Pour obtenir le vecteur λ résoudre le système : $Q\lambda = y^k - B^k s^k - U_D^k s^k$,
 si Q n'est pas définie positive, parce qu'au moins un des $s^i = 0$,
 alors on impose $\lambda_i = 0$ et $(y^k - B^k s^k - U_D^k s^k)_i = 0$,
 ensuite résoudre le système réduit correspondant ;
 - Calculer les éléments de la matrice de correction diagonale E^k : $E_{ii}^k = 2\lambda_i s_i^k$;
 - Calculer la nouvelle approximation diagonale du Hessien $B^{k+1} = \eta^k B^k + U_D^k + E^k$;
- 10- Retour en 2.

Remarque 6 Dans l'algorithme de Toint on a déplacé les étapes (2 et 3) en les étapes (8 et 9) parce que : dans l'étape 2 de l'algorithme de \mathcal{DFP} on a déjà la matrice $B^0 = I$

qui est définie positive. De même d'autre côté on a utilisé la variante \mathcal{DFP} au lieu de \mathcal{BFGS} dans l'algorithme de Toint.

On propose la modification suivante :

Initialisation : Soit $\epsilon > 0$ est une précision donnée,

f une fonction différentiable, $x^0 \in \mathbb{R}^n$ un point de départ et B^0 une matrice identité ;

1- Soient $x^0 \in \mathbb{R}^n$ un point de départ et B^0 une matrice identité,

calculer $g^0 = \nabla f(x^0)$;

Tant que critère d'arrêt non satisfait faire

2- Calculer la direction de descente d^k en résolvant le système $B^k d^k = -g^k$;

3- Effectuer une recherche linéaire selon la direction d^k , c'est-à-dire calculer un pas de progression positif α^k le long de d^k de sorte que $f(x^k + \alpha^k d^k) > f(x^k)$;

4- $x^{k+1} := x^k + \alpha^k d^k$;

5- Calculer g^{k+1} ;

6- Calculer $s^k = x^{k+1} - x^k$ et $y^k = g^{k+1} - g^k$;

7- Calculer les éléments diagonaux de la mise à jour de type \mathcal{DFP} soit U_D^k :

$$U_{ii}^k = \frac{s_i^2}{s^k \cdot y^k} - \frac{(B_{ii}^k y_i^k)^2}{y^k B^k y^k};$$

- Calculer les éléments de la matrice diagonale Q : $Q_{ii} = 2s_i^2$;

- Pour obtenir le vecteur λ résoudre le système : $Q\lambda = y^k - B^k s^k - U_D^k s^k$,

si Q n'est pas définie positive, parce qu'au moins un des $s^i = 0$,

alors on impose $\lambda_i = 0$ et $(y^k - B^k s^k - U_D^k s^k)_i = 0$,

ensuite résoudre le système réduit correspondant ;

- Calculer les éléments de la matrice de correction diagonale E^k : $E_{ii}^k = 2\lambda_i s_i^k$;

- Calculer la nouvelle approximation diagonale du Hessien $B^{k+1} = B^k + U_D^k + E^k$;

8- Si B^{k+1} n'est pas définie positive, employer une procédure de Levenberg-Marquardt pour obtenir une matrice définie positive $\overline{B}^{k+1} = B^{k+1} + \mu I$;

9- Remplacer B^{k+1} par son substitut défini positif $B^{k+1} := \overline{B}^{k+1}$;

Fin tant que.

Chapitre 3

Tests numériques

3.1 Exemples

Les tests numériques sont effectués sur la station COMPAQ. Le pas de déplacement est calculé en utilisant la recherche linéaire de Goldstein.

Exemple 1

$$\min_{x \in \mathbb{R}^2} f(x) = x_1 x_2^2 \quad \text{avec la solution optimale est } (1, 0)^T$$

<i>Méthodes</i>	<i>Solution</i>	<i>Nombre d'itérations</i>	<i>Temps</i>
<i>DFP</i>	$(0.91581727, 0.00023888)^T$	10	1 ^s 54 ^{cent}
<i>DFP avec diagonalisation</i>	$(0.91581727, 0.00023888)^T$	10	1 ^s 26 ^{cent}

$\epsilon = 10^{-5}$, point de départ est $(1, 1)^T$.

Exemple 2

$$\min_{x \in \mathbb{R}^2} f(x) = x_1^2 + 2x_2^2 + x_1 x_2 - x_1 + x_2 \quad \text{avec la solution optimale est } \left(\frac{5}{7}, -\frac{3}{7}\right)^T$$

Méthodes	Solution	Nombre d'itérations	Temps
DFP	$(0.71429129, -0.42857360)^T$	28	4 ^s 69 ^{cent}
DFP avec diagonalisation	$(0.71429129, -0.42857360)^T$	28	3 ^s 88 ^{cent}

$\epsilon = 10^{-5}$, point de départ est $(1, 1)^T$.

Exemple 3

$$\min_{x \in \mathbb{R}^2} f(x) = 3x_1^2 \sqrt{x_2} \quad \text{avec la solution optimale est } (0, \frac{1}{2})^T$$

Méthodes	Solution	Nombre d'itérations	Temps
DFP	$(-0.00000115, 0.44899533)^T$	7	2 ^s 24 ^{cent}
DFP avec diagonalisation	$(-0.00000115, 0.44899533)^T$	7	1 ^s 69 ^{cent}

$\epsilon = 10^{-5}$, point de départ est $(1, 1)^T$.

Exemple 4

$$\min_{x \in \mathbb{R}^2} f(x) = x_1^2 x_2 + \ln(1 + x_2^2) \quad \text{avec la solution optimale est } (0, 0)^T$$

Méthodes	Solution	Nombre d'itérations	Temps
DFP	$(0.00003949, -0.00000003)^T$	24	4 ^s 28 ^{cent}
DFP avec diagonalisation	$(0.00003949, -0.00000003)^T$	24	3 ^s 17 ^{cent}

$\epsilon = 10^{-4}$, point de départ est $(1, 1)^T$.

Exemple 5

$$\min_{x \in \mathbb{R}^2} f(x) = x_1^3 + x_2^3 - 3x_1 x_2 \quad \text{avec la solution optimale est } (1, 1)^T$$

Méthodes	Solution	Nombre d'itérations	Temps
DFP	$(1.00000053, 1.00000005)^T$	18	3 ^s 13 ^{cent}
DFP avec diagonalisation	$(1.00000053, 1.00000005)^T$	18	2 ^s 53 ^{cent}

$\epsilon = 10^{-5}$, point de départ est $(0, 1)^T$.

3.2 Conclusion et perspectives

Conclusion 1 *Les simulations numériques testées montrent l'efficacité de notre approche (DFP avec technique de diagonalisation) en terme de temps relativement à la méthode de DFP classique.*

Perspectives

- Etendre cette approche à d'autres classes de problèmes d'optimisation.
- Répondre à la question qu'est qu'on perd en diagonalisant les matrices B^k .

Bibliographie

- [1] L. Armijo. "Minimization of functions having Lipschitz-continuous first partial derivatives". Pacific J. Math., 16 :1–3, (1966).
- [2] F., Jr. Ayres, " Matrices : cours et problèmes", McGraw Hill, New York, (1991).
- [3] S. Bazarra, H. D. Sherali and C. M. Shetty, "Nonlinear Programming, Theory and algorithms", Second edition (1993).
- [4] P. Borne, F. Rotella, "Théorie et pratique du calcul matriciel", Edition Technip, Paris (1995).
- [5] R. Bulirsh, J. Stoer, "Introduction to numerical Analysis", Springer-Verlag, New York (1993).
- [6] J.E. Dennis and J.J. Moré, "Quasi-Newton methods, motivation and theory". SIAM Rev.,19(1) : 46 -89, (1977).
- [7] P. Duysinx, "Optimisation Topologique : du milieu continu a la structure elastique". Thèse de doctorat en Sciences Appliquées. Université de Liège, (1995,1996).
- [8] A. Keraghel, "Analyse convexe : Théorie fondamentale et exercices", Editions Dar El'Houda, Ain Mila, Algérie (2001).
- [9] M. Minoux, "Programmation mathématique : Théorie et algorithmes", Tome I, Dunod , Paris (1983).
- [10] S. Mottelet, "RO04/TI07 - Optimisation non-linéaire", Université de Technologie de Compiègne (2003).

