

Table des matières

Résumé	iii
Abstract	iv
Introduction	v
1 Optimisation sans contraintes. Notions de base	1
1.1 Quelques généralités sur les problèmes d'optimisation sans contraintes	1
1.2 Conditions d'optimalité	2
1.2.1 Conditions nécessaires d'optimalité	2
1.2.2 Condition suffisante d'optimalité	3
1.2.3 Conditions nécessaires et suffisantes d'optimalité dans le cas convexe	4
1.3 Convergence des algorithmes	4
1.3.1 Schemas general des algorithmes d'optimisation sans contraintes	4
1.3.2 Notion de convergence globale	5
1.3.3 Notion de vitesse de convergence	6
1.3.4 Les modes de convergence	6
1.4 Recherches linéaires	7
1.4.1 Principe des méthodes de descente	7
1.4.2 Recherche linéaire	12
1.4.3 Recherches linéaires inexactes	14
1.4.4 Convergence des méthodes à directions de descente . .	18

2	Méthodes quasi-Newtoniennes	24
2.1	Méthode de Newton	24
2.1.1	Avantages et inconvénients	25
2.1.2	Comment peut-on remédier les défauts de l'algorithme de Newton?	26
2.2	Méthodes de Quasi-Newton	27
2.2.1	Quelques propriétés des algorithmes de quasi-Newton	28
2.2.2	Formules de mise à jour de l'approximation du Hessien	28
2.2.3	Méthode de correction de rang un (SR1 [1960])	29
2.2.4	Méthode de Davidon-Fletcher-Powell (DFP [D-1959 et F.P-1963])	34
2.2.5	Méthode de Broyden-Fletcher-Goldfarb-Shanno (BFGS [1970])	40
2.2.6	Les méthodes de la classe Broyden	45
3	Rôle joué par la convexité dans la convergence de la méthode BFGS	47
3.1	Convergence globale de la méthode BFGS dans le cas convexe	47
3.2	Convergence globale de la méthode BFGS dans le cas non convexe	55
3.2.1	Méthode BFGS avec mise à jour appropriée	55
3.2.2	Convergence globale de Méthode BFGS avec mise à jour appropriée	60
3.3	Problèmes ouverts et conjectures	73
3.4	Tests numériques	75

Résumé

On s'intéresse dans ce mémoire aux problèmes de convergence de la méthode de BFGS avec des recherches linéaires inexactes.

On fait une synthèse sur les résultats de convergence des méthodes de quasi Newton et spécialement sur la méthode BFGS dans le cas convexe et le cas non convexe.

Mots clés : Optimisation non linéaire sans contraintes, Méthode de BFGS, Convergence globale.

Abstract

We are interested in this memory to the problems of convergence of the BFGS method with inexact line search.

We give a synthesis of the results of convergence of the Quasi-Newton methods and especially on the BFGS method in the convex case and the nonconvex case.

Key Words : unconstrained optimization, BFGS method, global convergence.

Introduction

Les problèmes d'optimisation différentiable se posent lorsque l'on cherche à déterminer la valeur optimale d'un nombre fini de paramètres. L'optimalité signifie ici la minimalité d'un critère donné. La différentiabilité supposée des fonctions qui définissent le problème écarte d'emblée de notre propos l'optimisation combinatoire (les paramètres à optimiser ne prennent que des valeurs entières ou discrètes) et l'optimisation non lisse (les fonctions ont des irrégularités).

L'optimisation est un sujet très ancien. Taylor [1685-1731], Newton [1643-1727], Lagrange [1736-1813] et Cauchy [1789-1857] ont élaboré les bases des développements limités. L'optimisation a connu un nouvel essor depuis l'apparition des ordinateurs et s'applique désormais dans de très nombreux domaines : économie, gestion, planification, logistique, automatique, robotique, conception optimale, science de l'ingénieur, traitement du signal, etc.

Les méthodes numériques de l'optimisation ont principalement été développées après la seconde guerre mondiale, en parallèle avec l'amélioration des ordinateurs, et n'ont cessé depuis de s'enrichir. En optimisation non linéaire, on peut ainsi distinguer plusieurs vagues : méthodes de pénalisation, méthode du lagrangien augmenté (1958), méthodes de quasi-Newton (1959), méthodes newtoniennes ou SQP (1976), algorithmes de points intérieurs (1984). Une vague n'efface pas la précédente mais permet d'apporter de meilleures réponses à certaines classes de problèmes, comme ce fut le cas pour les méthodes de points intérieurs en optimisation semi-définie positive (SDP). Une attention particulière sera portée aux algorithmes pouvant traiter les problèmes de grande taille, ceux qui se présentent dans les applications.

Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$. Considérons le problème d'optimisation sans contraintes (P) suivant :

$$(P) : \underset{x \in \mathbb{R}^n}{\text{Min}} f(x)$$

Les méthodes quasi-Newton, et particulièrement la méthode BFGS est parmi les méthodes les plus célèbres pour résoudre le problème (P). Cette méthode a été étudiée et développée par Broyden ([6]), Fletcher ([32]), Goldfarb ([9]) et Shanno ([10]).

On s'intéresse dans ce mémoire aux problèmes de convergence de la méthode BFGS dans le cas des recherches linéaires inexactes de Wolfe et d'Armijo. On essaye de mettre en lumière et d'étudier de façon approfondie les principaux résultats de convergence dans le cas où la fonction objectif f est convexe et dans le cas où f n'est pas convexe.

Dans le cas convexe, on a détaillé le travail de M.J.D. Powell ([25]). C'est l'un des plus beaux et plus importants résultats de convergence globale pour les problèmes d'optimisation sans contraintes. Il donne des conditions assurant la convergence globale de l'algorithme BFGS lorsque la fonction est convexe. La beauté du résultat tient au fait qu'il est rare de pouvoir montrer la convergence d'un algorithme quasi-Newtonien avec si peu d'hypothèses sur les objets générés par celui-ci. On suppose seulement que la suite $\{f(x_k)\}_{k \in \mathbb{N}}$ est bornée inférieurement.

Dans le cas non convexe, on présente de façon approfondie le travail de Dong-Hui Li et Masao Fukushima ([11]). Dans ce travail, les auteurs en question, étudient le problème de la convergence globale du problème (P). Pour cela les auteurs construisent une suite $\{x_k\}_{k \in \mathbb{N}}$ de la forme suivante :

$$x_{k+1} = x_k + \lambda_k p_k, \quad (1)$$

où p_k vérifie :

$$B_k p + g_k = 0, \quad g_k = \nabla f(x_k), \quad (2)$$

les matrices B_k sont calculées itérativement d'une façon semblable à celle de la méthode BFGS ([18]). Plus explicitement :

$$B_{k+1} = \begin{cases} B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k}, & \text{si } \frac{y_k^T s_k}{\|s_k\|^2} \geq \epsilon \|g_k\|^\alpha \\ B_k, & \text{sinon} \end{cases} \quad (3)$$

$$s_k = x_{k+1} - x_k \quad \text{et} \quad y_k = g_{k+1} - g_k,$$

α et ϵ sont des constantes positives. λ_k est un pas positif obtenu par une recherche linéaire de Wolfe ou d'armijoo.

Bien sur l'expression précédente est une modification de la formule classique de la méthode BFGS dont l'expression des matrices B_k est la suivante :

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k}, \quad (4)$$

Les auteurs réussissent à démontrer la convergence globale de la méthode définie par les relations (1), (2) et (3) sans la convexité.

Le problème de la convergence globale de la méthode BFGS définie par les relations (1), (2) et (4), en l'absence de convexité, demeure toujours un problème ouvert.

Le mémoire est divisé en 3 parties. On expose dans la première partie les notions de base pour étudier le problème (P). Plus particulièrement on y trouve les notions de conditions d'optimalité, de convergence globale, de vitesse de convergence. Un paragraphe important est réservé aux recherches linéaires, surtout inexactes.

Dans la seconde partie on donne un aperçu général sur les méthodes quasi-Newtoniennes surtout la méthode BFGS.

La partie essentielle du mémoire se trouve dans le chapitre 3 dans lequel on s'intéresse aux problèmes de convergence de la méthode BFGS et BFGS modifiée dans le cas des recherches linéaires inexactes de Wolfe et d'Armijo, pour le cas des fonctions convexes ou non convexes.

On trouve à la fin du mémoire une conclusion et des problèmes ouverts.

Chapitre 1

Optimisation sans contraintes. Notions de base

1.1 Quelques généralités sur les problèmes d'optimisation sans contraintes

Le problème que l'on étudie ici est celui de la recherche du minimum (maximum) d'une fonction réelle f de n variables réelles x_1, x_2, \dots, x_n .

Soit $f : IR^n \rightarrow IR$ qui à tout $x \in IR^n$, $x = (x_1, x_2, \dots, x_n)^T$ associe la valeur réelle :

$$f(x) = f(x_1, x_2, \dots, x_n)$$

On cherche à résoudre :

$$(P) \quad \underset{x \in IR^n}{\text{Min}} f(x) \quad (1.1)$$

Il s'agit donc de déterminer un point x^* de IR^n tel que :

$$\forall x \in IR^n : f(x^*) \leq f(x)$$

C'est-à-dire un **minimum global** de f Sur IR^n .

Néanmoins, il existe aussi bien des situations dans lesquelles le minimum n'est pas global, il faut alors se contenter d'**optimums locaux**, C'est-à-dire des points x^* vérifiant la condition suivante :

$$\exists V_\epsilon(x^*) \text{ tel que } \forall x \in V_\epsilon(x^*) : f(x^*) \leq f(x)$$

Dans d'autres situations moins favorables, on cherche des points x^* qui ne sont pas nécessairement des optimums locaux, mais qui vérifiant une propriété proche de l'optimalité. Citons par exemple des points x^* qui sont stationnaires, c'est à dire des points tels que $\nabla f(x^*) = 0$.

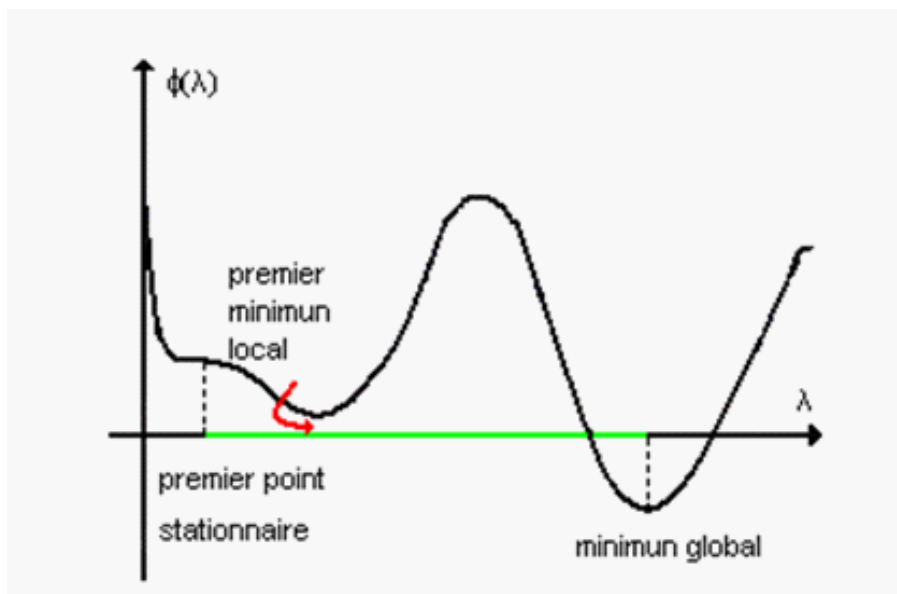


Fig.1.1 point stationnaire, optimum local, optimum global

1.2 Conditions d'optimalité

1.2.1 Conditions nécessaires d'optimalité

Définition 1.1 Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$. On dira qu'un vecteur $d \in \mathbb{R}^n$ est une direction de descente en \hat{x} s'il existe $\delta > 0$ tel que :

$$f(\hat{x} + \lambda d) < f(\hat{x}) : \quad \forall \lambda \in]0, \delta[.$$

Proposition ([22]). Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ et $\hat{x} \in \mathbb{R}^n$. On suppose que f est différentiable au point \hat{x} et qu'en \hat{x} et d f vérifie :

$$\nabla f(\hat{x})^T \cdot d < 0.$$

Alors d est une direction de descente en \hat{x} .

◇ **Condition nécessaire d'optimalité du premier ordre**

Théorème 1.2 ([22]) Soit $f : IR^n \rightarrow IR$ différentiable au point $\varkappa^* \in IR^n$. Si \varkappa^* est un minimum local de (P) alors $\nabla f(\varkappa^*) = 0$.

Preuve. Supposons le contraire c. a. d que $\nabla f(\varkappa^*) \neq 0$ ce qui implique que $d = -\nabla f(\varkappa^*)$ est une direction de descente. Donc $\exists \delta > 0$ tel que

$$\forall \lambda \in]0, \delta[: f(\varkappa^* + \lambda d) < f(\varkappa^*)$$

Ceci est contradiction avec le fait que \varkappa^* est une solution optimale locale de (P) . ■

◇ **Condition nécessaire d'optimalité du second ordre**

Théorème 1.3 ([22]) Soit $f : IR^n \rightarrow IR$ deux fois différentiable au point $\varkappa^* \in IR^n$. Si \varkappa^* est un minimum local de (P) alors $\nabla f(\varkappa^*) = 0$ et la matrice hessienne de f au point \varkappa^* , qu'on note $H(\varkappa^*)$, est semi définie positive.

1.2.2 Condition suffisante d'optimalité

Théorème 1.4 ([22]) Soit $f : IR^n \rightarrow IR$ deux fois différentiable au point $\varkappa^* \in IR^n$. Si $\nabla f(\varkappa^*) = 0$ et $H(\varkappa^*)$ est définie positive alors \varkappa^* est un minimum local strict de (P) .

1.2.3 Conditions nécessaires et suffisantes d'optimalité dans le cas convexe

Définition 1.2 soit $f : IR^n \rightarrow IR$.

a) f est dite convexe dans IR^n si :

$$f(t\kappa_1 + (1-t)\kappa_2) \leq tf(\kappa_1) + (1-t)f(\kappa_2)$$

pour tous points κ_1 et $\kappa_2 \in IR^n$ et pour tout $t \in [0, 1]$.

b) si l'inégalité précédente est stricte pour tous points κ_1 et κ_2 distincts et pour tout $t \in]0, 1[$, alors f est dite strictement convexe.

c) supposons que f soit différentiable, f est dite fortement convexe si :

$$f((1-t)\kappa_1 + t\kappa_2) \leq (1-t)f(\kappa_1) + tf(\kappa_2) - \frac{\alpha}{2}t(1-t)\|\kappa_1 - \kappa_2\|^2$$

pour tous points κ_1 et $\kappa_2 \in IR^n$ et pour tout $t \in [0, 1]$.

Remarque 1.1 Si f est convexe, alors tout minimum local est aussi global. De plus si f est strictement convexe, alors tout minimum local devient non seulement global mais aussi unique (voir [24])

Théorème 1.5 ([24]). Soit $f : IR^n \rightarrow IR$ et $\kappa^* \in IR^n$. Supposons que f est convexe dans IR^n et différentiable au point κ^* . Alors κ^* est un minimum global de f si et seulement si

$$\nabla f(\kappa^*) = 0.$$

Remarque 1.2 le théorème 1.5 demeure vrai si on remplace IR^n par un ouvert S de IR^n .

1.3 Convergence des algorithmes

1.3.1 Schemas general des algorithmes d'optimisation sans contraintes

La plupart des méthodes de résolution des problèmes d'optimisation qui seront exposées dans les chapitres qui suivent sont de nature itérative, c'est-à-dire qu'à partir d'un point initial donné κ_1 , ils engendrent une suite infinie

$\varkappa_1, \varkappa_2, \dots, \varkappa_k, \dots$ dont on espère qu'elle converge vers la solution optimale ou au moins vers un point stationnaire.

Un algorithme de résolution est un procédé qui permet, à partir de la donnée du point initial \varkappa_1 , d'engendrer la suite $\varkappa_1, \varkappa_2, \dots, \varkappa_k, \dots$. Un algorithme est parfaitement défini par la donnée de l'application A qui au point \varkappa_k associe le point \varkappa_{k+1} tel que $\varkappa_{k+1} \in A(\varkappa_k)$. Ceci permettra dans la suite de confondre un algorithme et l'application A qui lui est associée.

Pour construire des algorithmes de minimisation sans contraintes on fait appel à des processus itératifs du type

$$\varkappa_{k+1} = \varkappa_k + \lambda_k d_k, \quad \lambda_k > 0 \quad ((\text{mod}\blacksquare\text{le algorithme}))$$

où d_k détermine la direction de déplacement à partir du point \varkappa_k et λ_k est un facteur numérique dont le grandeur donne la longueur du pas dans la direction d_k

Le type d'algorithme permettant de résoudre le problème (1.1) sera déterminé dès qu'on définit les procédés de construction du vecteur d_k et de calcul de λ_k à chaque itération.

La façon avec laquelle on construit les vecteurs d_k et les scalaires λ_k déterminent directement les propriétés du processus et spécialement en ce qui concerne la convergence de la suite $\{\varkappa_k\}$, la vitesse de la convergence,...

Pour s'approcher de la solution optimale du problème (1.1), on se déplace naturellement à partir du point \varkappa_k dans la direction de la décroissance de la fonction f .

1.3.2 Notion de convergence globale

Définition 1.3 Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ différentiable. Supposons que l'on a construit une suite $\{\varkappa_k\}$, à l'aide d'un algorithme d'optimisation sans contraintes décrit dans le modèle (modèle algorithme). Nous dirons que l'algorithme converge globalement si on a :

$$\liminf_{k \rightarrow \infty} \|\nabla f(\varkappa_k)\| = 0$$

Remarque 1.3 : Des auteurs exigent parfois pour la même définition la relation plus forte suivante :

$$\lim_{k \rightarrow \infty} \|\nabla f(\mathcal{x}_k)\| = 0$$

1.3.3 Notion de vitesse de convergence

La convergence globale d'un algorithme ayant été établie, nous nous intéressons maintenant à l'évaluation de son *efficacité*. D'un point de vue pratique, l'efficacité d'un algorithme dépend du nombre d'itérations nécessaires pour obtenir une approximation à ε près (ε fixé à l'avance) de l'optimum \mathcal{x}^* .

Si l'on compare entre eux, plusieurs algorithmes, et si l'on admet que le temps de calcul par itération est sensiblement le même pour tous, le meilleur est celui qui nécessitera le plus petit nombre d'itérations.

Malheureusement, il se révèle impossible de dégager des conclusions générales de ce genre de comparaison.

Suivant le point de départ choisi, la nature de la fonction à optimiser, la valeur de la tolérance choisie, la hiérarchie des algorithmes peut varier considérablement.

Si l'on veut dégager un critère ayant une certaine valeur d'absolu, il faut par conséquent recourir à un autre type d'analyse : c'est l'objet de l'étude de la *convergence asymptotique*, c'est-à-dire du comportement de la suite $\{\mathcal{x}_k\}$ au voisinage du point limite \mathcal{x}^* .

Ceci conduit à attribuer à chaque algorithme un indice d'efficacité appelé sa *vitesse de convergence*.

Remarque 1.4 on est quelque fois amené à exprimer la vitesse de convergence de la suite $\{\mathcal{x}_k\}$ en étudiant, non pas la façon dont $\|\mathcal{x}_k - \mathcal{x}^*\|$ tend vers 0, mais la façon dont la suite $\{f(\mathcal{x}_k)\}$ tend vers $f(\mathcal{x}^*)$ où f la fonction que l'on minimise.

1.3.4 Les modes de convergence

Soit $\{\mathcal{x}_k\}_{k \in \mathbb{N}}$ une suite dans IR^n convergeant vers \mathcal{x}^*

1) si

$$\limsup_{k \rightarrow \infty} \frac{\|\mathcal{x}_{k+1} - \mathcal{x}^*\|}{\|\mathcal{x}_k - \mathcal{x}^*\|} = \alpha < 1,$$

on dit que la convergence est linéaire et α est le taux de convergence associé.

2) si

$$\limsup_{k \rightarrow \infty} \frac{\|\mathcal{X}_{k+1} - \mathcal{X}^*\|}{\|\mathcal{X}_k - \mathcal{X}^*\|^\gamma} < +\infty; \quad \gamma > 1,$$

On dit que la convergence est superlinéaire d'ordre γ .

3) si

$$\lim_{k \rightarrow \infty} \frac{\|\mathcal{X}_{k+1} - \mathcal{X}^*\|}{\|\mathcal{X}_k - \mathcal{X}^*\|} = 0,$$

On dit que $\{\mathcal{X}_k\}$ tend vers \mathcal{X}^* de façon superlinéaire.

1.4 Recherches linéaires

La recherche linéaire consiste à trouver λ_k de façon à diminuer la fonction f *Suffisamment* le long de cette direction.

Ce " *suffisamment* " sera quantifié dans la suite dans la description des conditions dites d' Armijo, Wolfe, Goldstein&Price(recherches linéaires inexactes).

1.4.1 Principe des méthodes de descente

Le principe d'une méthode de descente consiste à faire les itérations suivantes

$$\mathcal{X}_{k+1} = \mathcal{X}_k + \lambda_k d_k, \quad \lambda_k > 0 \tag{1.2}$$

tout en assurant la propriété

$$f(\mathcal{X}_{k+1}) < f(\mathcal{X}_k).$$

Le vecteur d_k est la direction de descente en \mathcal{X}_k . Le scalaire λ_k est appelé le pas de la méthode à l'itération k .

On peut caractériser les directions de descente en \mathcal{X}_k à l'aide du gradient :

Proposition 1.1 *Soit $d \in \mathbb{R}^n$ vérifiant*

$$\nabla f(\mathcal{X})^\top . d < 0$$

alors d est une direction de descente en \varkappa .

Preuve. ([22])

on a pour $\lambda > 0$

$$f(\varkappa + \lambda d) = f(\varkappa) + \lambda \nabla f(\varkappa)^\top d + \lambda \varepsilon(\lambda)$$

donc si on écrit

$$\frac{f(\varkappa + \lambda d) - f(\varkappa)}{\lambda} = \nabla f(\varkappa)^\top d + \varepsilon(\lambda)$$

on voit bien que pour λ suffisamment petit on aura

$$f(\varkappa + \lambda d) - f(\varkappa) < 0.$$

■

Ou encore que d fait avec l'opposé du gradient $-\nabla f(\varkappa)$ un angle θ strictement plus petit que 90° :

$$\theta := \arccos \frac{-\nabla f(\varkappa)^\top d}{\|\nabla f(\varkappa)\| \|d\|} \in \left[0, \frac{\pi}{2}\right[.$$

L'ensemble des directions de descente de f en \varkappa ,

$$\{d \in \mathbb{R}^n : \nabla^\top f(\varkappa) \cdot d < 0\}$$

forme un demi-espace ouvert de \mathbb{R}^n (illustration à la figure 1.2).

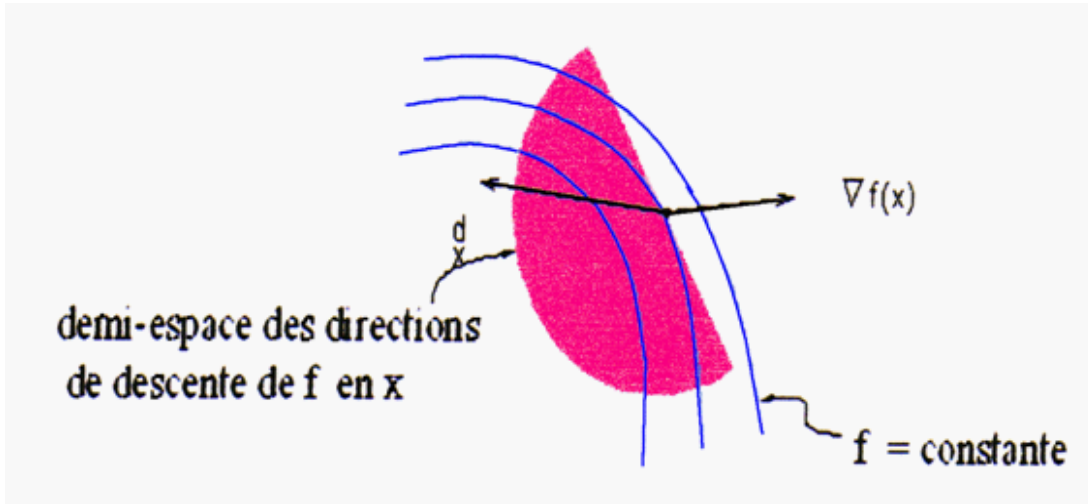


Fig.1.2 Demi-espace (translaté) des direction de descente d de f en x

De telles directions sont intéressantes en optimisation car, pour faire décroître f , il suffit de faire un déplacement le long de d .

Les méthodes à directions de descentes utilisent cette idée pour minimiser une fonction

Dans la méthode (1.2) le choix de λ_k est lié à la fonction

$$\varphi(\lambda) = f(\boldsymbol{x}_k + \lambda d_k)$$

Comme dans la méthode de la direction de descente, la trajectoire de la solution suit un modèle de zigzag (voir figure [1.4].) Si λ est choisi tels que $f(\boldsymbol{x}_k + \lambda d_k)$ soit le minimum dans chaque itération, alors les directions successives sont orthogonales.

En effet

si on note $g(\boldsymbol{x}) = \nabla f(\boldsymbol{x})$

$$\frac{\mathbf{d}f(\boldsymbol{x}_k + \lambda d_k)}{\mathbf{d}\lambda} = \sum_{i=1}^n \frac{\partial f(\boldsymbol{x}_k + \lambda d_k)}{\partial x_{ki}} \frac{\mathbf{d}(x_{ki} + \lambda d_{ki})}{\mathbf{d}\lambda}$$

$$\begin{aligned}
 &= \sum_{i=1}^n g_i(\boldsymbol{x}_k + \lambda d_k) d_{ki} \\
 &= g(\boldsymbol{x}_k + \lambda d_k)^\top d_k
 \end{aligned}$$

où $g(\boldsymbol{x}_k + \lambda d_k)$ est le gradient au point $\boldsymbol{x}_k + \lambda d_k$.

En particulier, une façon de choisir λ_k peut être de résoudre le problème d'optimisation (à une seule variable)

$$\min_{\lambda > 0} \varphi(\lambda) \tag{1.3}$$

Si le pas $\tilde{\lambda}_k$ obtenu ainsi s'appelle le pas optimal alors nous pouvons écrire

$$\varphi'(\tilde{\lambda}_k) = \nabla f(\boldsymbol{x}_k + \tilde{\lambda}_k d_k)^\top d_k = 0$$

c'est-à-dire

$$g(\boldsymbol{x}_k + \tilde{\lambda}_k d_k)^\top d_k = 0$$

ou bien

$$d_{k+1}^\top d_k = 0$$

où

$$d_{k+1} = -g(\boldsymbol{x}_k + \tilde{\lambda}_k d_k) = -g_{k+1}$$

est la direction de descente au point $\boldsymbol{x}_k + \tilde{\lambda}_k d_k$. Donc les directions successives d_k et d_{k+1} sont orthogonales comme représenté dans la figure (1.3).

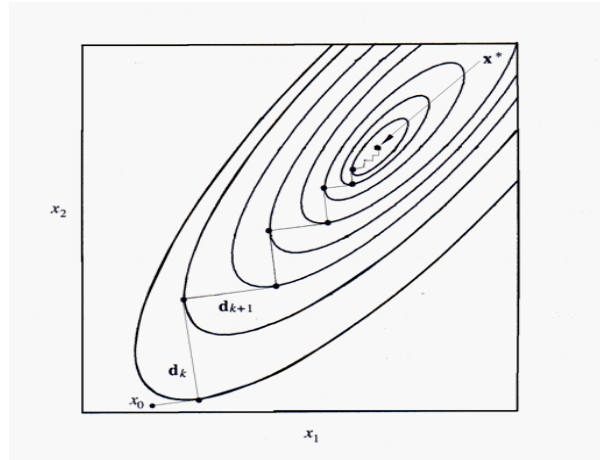


Fig.1.3 Trajectoire d'une solution typique d'une méthode à direction de descente

Pour définir une direction de descente il faut donc spécifier deux choses :

- ♣ Dire comment la direction d_k est calculée. Ce choix influe directement dans la nomination de l'algorithme.
- ♣ Dire comment on détermine le pas λ_k ; c'est ce que l'on appelle : la recherche linéaire.

Algorithme 1.1 (méthode à directions de descente – une itération)

Etape 0 : (initialisation)

On suppose qu'au début de l'itération k , on dispose d'un itéré

$$x_k \in \mathbb{R}^n$$

Etape 1 :

Test d'arrêt : si $\|\nabla f(x_k)\| \simeq 0$, d'arrêt de l'algorithme ;

Etape 2 :

Choix d'une direction de descente $d_k \in \mathbb{R}^n$;

Etape 3 :

Recherche linéaire : déterminer un pas $\lambda_k > 0$ le long de d_k de manière à "faire décroître f suffisamment" ;

Etape 4 :

Si la recherche linéaire est finie $x_{k+1} = x_k + \lambda_k d_k$;

remplacer k par $k + 1$ et aller à l'étape 1. \square

1.4.2 Recherche linéaire

Faire la recherche linéaire veut dire déterminer un pas λ_k le long d'une direction de descente d_k , autrement dit résoudre le problème unidimensionnel (1.3).

Notre intérêt pour la recherche linéaire ne vient pas seulement du fait que dans les application on rencontre, naturellement, des problèmes unidimensionnels, mais plutôt du fait que la recherche linéaire est un composant fondamental de toutes les méthodes traditionnelles d'optimisation multidimensionnelle

1.4.2.1 Objectifs à atteindre

il s'agit de réaliser deux objectifs

Le premier objectif

Consiste à faire décroître f suffisamment, cela se traduit le plus souvent par la réalisation d'une inégalité de la forme

$$f(\boldsymbol{x}_k + \lambda_k d_k) \leq f(\boldsymbol{x}_k) + \text{"un terme négatif"} \quad (1.4)$$

Le terme négatif, disons ν_k , joue un rôle-clé dans la convergence de l'algorithme utilisant cette recherche linéaire.

L'argument est le suivant.

Si $f(\boldsymbol{x}_k)$ est minorée ($\exists c$ telle que $f(\boldsymbol{x}_k) \geq c$ pour tout k), alors ν_k tend nécessairement vers zéro ($\nu_k \rightarrow 0$). c'est souvent à partir de la convergence vers zéro de cette suite que l'on parvient à montrer que le gradient lui-même doit tendre vers zéro. Le terme négatif devra prendre une forme bien particulière si on veut pouvoir en tirer de l'information.

En particulier, il ne suffit pas d'imposer $f(\boldsymbol{x}_k + \lambda_k d_k) < f(\boldsymbol{x}_k)$

Le second objectif

Consiste d'empêcher le pas $\lambda_k > 0$ d'être trop petit, trop proche de zéro.

Le premier objectif n'est en effet pas suffisant car l'inégalité (1.4) est en général satisfaite par des pas $\lambda_k > 0$ Arbitrairement petit.

Or ceci peut entraîner une "*fausse convergence*", c'est-à-dire la convergence des itérés vers un point non stationnaire.

1.4.2.2 Types de recherches linéaires

Il existe deux grandes classes de méthodes qui s'intéressent à l'optimisation unidimensionnelle :

- a) Les recherches linéaires exactes.
- b) Les recherches linéaires inexactes.

Les recherches linéaires exactes, malgré qu'elles n'aboutissent qu'à une solution optimale approchée, nécessitent beaucoup d'observations à chaque itération de l'algorithme principal.

Le mot exact prend sa signification dans le fait que si f est quadratique la solution de la recherche linéaire s'obtient d'une façon exacte est dans un nombre fini d'itérations.

Pour une fonction non linéaire (non quadratique) arbitraire,
- la détermination de ces pas demande en général beaucoup de temps de calcul et ne peut de toute façon pas être faite avec une précision infinie,
- l'efficacité supplémentaire éventuellement apportée à un algorithme par une recherche linéaire exacte ne permet pas, en général, de compenser le temps perdu à déterminer un tel pas,
- les résultats de convergence autorisent d'autres types de règles (recherche linéaire inexacte), moins gourmandes en temps de calcul.

On préfère imposer des conditions moins restrictives, plus facilement vérifiées, qui permettent toute fois de contribuer à la convergence des algorithmes. En particulier, 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 d'ailleurs leur recherche plus aisée. C'est ce que l'on fait avec les règles d'Armijo, de Goldstein et de Wolfe décrites dans la prochaine section.

1.4.3 Recherches linéaires inexactes

On considère la situation qui est typique pour l'application de la technique de recherche linéaire à l'intérieur de la méthode principale multidimensionnelle.

Sur une itération k de la dernière méthode nous avons l'itération courante $x_k \in \mathbb{R}^n$ et la direction de recherche $d_k \in \mathbb{R}^n$ qui est direction de descente pour notre objectif : $f : \mathbb{R}^n \rightarrow \mathbb{R}$:

$$\nabla f(x_k)^\top \cdot d_k < 0 \quad (1.5)$$

Le but est de réduire "de façon importante" la valeur de l'objectif par un pas $x_k \rightarrow x_{k+1} = x_k + \lambda_k d_k$ de x_k dans la direction d_k .

Pour cela de nombreux mathématiciens (Armijo, Goldstein, Wolfe, Albaali, Lemaréchal, Fletcher...) ont élaboré plusieurs règles.

L'objectif de cette section consiste à présenter les principales tests.

D'abord présentons le schéma d'une recherche linéaire inexacte. La règle d'Armijo, La règle de Goldstein, La règle de Wolfe)

La recherche linéaire inexacte d'Armijo

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $x \in \mathbb{R}^n$, $d \in \mathbb{R}^n$ tel que $\nabla f^T(x_k) d_k < 0$. Définissons la fonction $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ avec $\varphi(\lambda_k) = f(x_k + \lambda_k d_k)$ $\lambda_k \geq 0$

Notons que :

$$\begin{aligned} \varphi'(\lambda) &= \nabla f^T(x_k + d_k) d_k, \\ \varphi'(0) &= \nabla f^T(x_k) d_k < 0, \\ \varphi(0) &= f(x_k) \end{aligned}$$

L'équation de la tangente au point $(0, \varphi(0))$ est la suivante

$$\begin{aligned} \{\lambda, y\} &: y = \varphi(0) + \varphi'(0)(\lambda - 0). \\ \varphi_1(\lambda_k) &= f(x_k) + \nabla f^T(x_k) d_k \lambda_k \end{aligned}$$

Posons

$$\varphi_1(\lambda) = \varphi(0) + \varphi'(0) \lambda.$$

L'équation de la tangente devient :

$$\varphi_1(\lambda) = f(x_k) + \nabla f^T(x_k) d_k \lambda$$

Définissons maintenant la fonction $\hat{\varphi}(\lambda)$ comme suit :

$$\hat{\varphi}(\lambda) = f(x_k) + \varepsilon \lambda \hat{\varphi}'(0) = f(x_k) + \varepsilon \lambda \nabla f^T(x_k) d_k; \quad 0 < \varepsilon < 1 \quad (1.6)$$

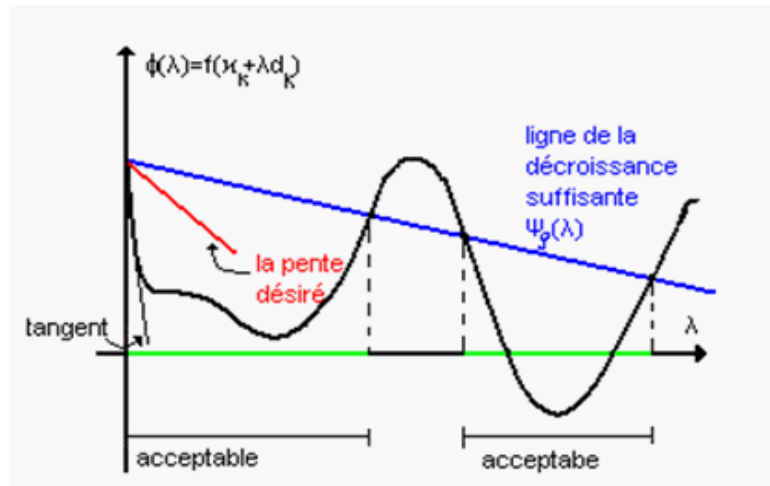


Fig1.4-Règle d'Amijo

On cherche $\bar{\lambda}_k$ tel que

$$\varphi(\bar{\lambda}_k) \leq \hat{\varphi}(\bar{\lambda}_k)$$

Remarques

1-La condition $\varphi(\bar{\lambda}_k) \leq \hat{\varphi}(\bar{\lambda}_k)$ implique la décroissance de la fonction f .
En effet

$$\begin{aligned} \varphi(\bar{\lambda}_k) &\leq \hat{\varphi}(\bar{\lambda}_k) \\ f(x_k + \bar{\lambda}_k d_k) &\leq f(x_k) + \varepsilon \bar{\lambda}_k \nabla f(x_k) d_k < f(x_k) \end{aligned}$$

car la direction d est une direction de descente.

2-Il ne faut pas prendre $\bar{\lambda}_k$ très proche de zero car celà va alterer la convergence et l'avitessse de convergence.

En effet

$$\begin{aligned}
 f(x_k + \bar{\lambda}_k d_k) &= f(x_k) + \bar{\lambda}_k \nabla f(x_k) d_k + \bar{\lambda}_k \alpha(x_k, \bar{\lambda}_k d_k) \\
 f(x_{k+1}) - f(x_k) &= \bar{\lambda}_k [\nabla f(x_k) d_k + \alpha(x_k, \bar{\lambda}_k d_k)] \\
 \text{si } \bar{\lambda}_k &\rightarrow 0 \quad \alpha(x_k, \bar{\lambda}_k d_k) \xrightarrow{\bar{\lambda}_k \rightarrow 0} 0 \quad \text{donc} \quad f(x_{k+1}) \simeq f(x_k)
 \end{aligned}$$

Algorithme d'Armijo

- λ^* initial
- Si $\varphi(\lambda^*) \leq \hat{\varphi}(\lambda^*)$, alors choisi le plus grand entier $t \geq 0$ tel que $\varphi(2^t \lambda^*) \leq \hat{\varphi}(2^t \lambda^*)$
 $\bar{\lambda}_k = 2^t \lambda^*$
- Sinon choisir le plus petit entier $t \geq 0$ tel que $\varphi(\frac{\lambda^*}{2^t}) < \hat{\varphi}(\frac{\lambda^*}{2^t})$
 $\bar{\lambda}_k = \frac{\lambda^*}{2^t}$

La recherche linéaire inexacte de Goldstein

En ajoutant une deuxième inégalité $\varphi(\lambda) \geq \varphi(0) + \sigma_2 \hat{\varphi}(0)$ a la règle d'Armijo on obtient la règle de Goldstein, avec σ_1, σ_2 sont deux constantes vérifiant $0 < \sigma_1 < \frac{1}{2} < \sigma_2 < 1$. Les deux inégalités de la règle de Goldstein sont donc :

$$\begin{aligned}
 \varphi(\bar{\lambda}_k) &\leq \hat{\varphi}(\bar{\lambda}_k) \quad \text{et} \\
 \varphi(\bar{\lambda}_k) &\geq \varphi(0) + \sigma_2 \hat{\varphi}(0).
 \end{aligned}$$

Soit en remplaçant

$$\begin{aligned}
 f(x_k + \bar{\lambda}_k d_k) &\leq f(x_k) + \sigma_1 \bar{\lambda}_k \nabla f(x_k) d_k \\
 &\quad \text{et} \\
 f(x_k + \bar{\lambda}_k d_k) &\geq f(x_k) + \sigma_2 \bar{\lambda}_k \nabla f(x_k) d_k
 \end{aligned}$$

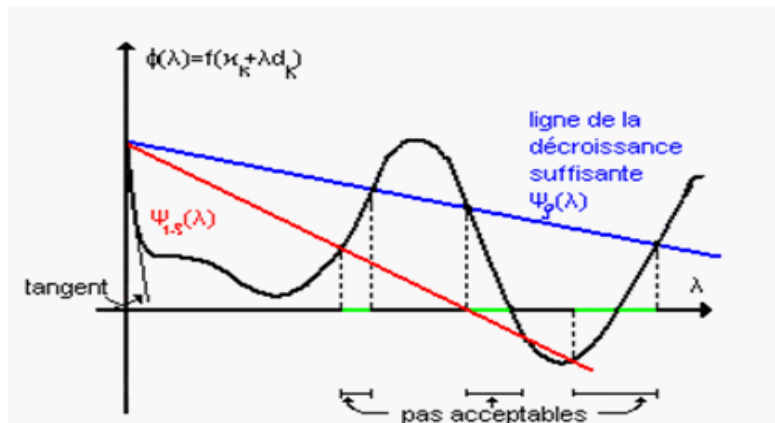


Fig1.5-Règle de Goldstein

La figure 1.5 montre, sur un exemple, l'ensemble des points satisfaisant les deux conditions Goldstein.

La recherche linéaire inexacte de Wolfe

La règle de Wolfe fait appel au calcul de $\varphi'(\lambda)$, elle est donc en théorie plus coûteuse que la règle de Goldstein. Cependant dans de nombreuses applications, le calcul du gradient $\nabla f(x)$ représente un faible coût additionnel en comparaison du coût d'évaluations de $f(x)$, c'est pourquoi cette règle est très utilisée.

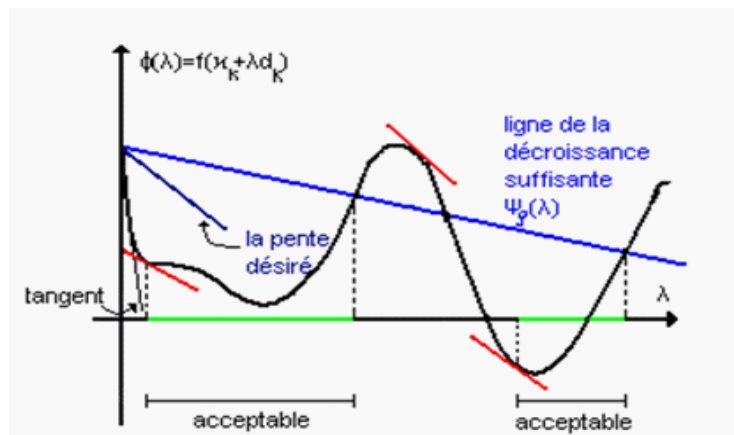


Fig1.3-Règle de Wolfe

Le figure 1.3 montre sur un exemple l'ensembles des points satisfaisant les conditions de Wolfe

$\sigma_1 = 0, 1, \sigma_2 = 0, 7$ (Lemaréchal 1980).

Les deux conditions de la règle Wolfe sont donc les suivantes :

$$f(x_k + \bar{\lambda}_k d_k) \leq f(x_k) + \sigma_1 \bar{\lambda}_k \nabla f(x_k)^T d_k \quad (\text{Wolfe1})$$

et

$$\nabla f(x_k + \bar{\lambda}_k d_k)^T d_k \geq \sigma_2 \nabla f(x_k)^T d_k \quad (\text{Wolfe2})$$

avec

$$0 < \sigma_1 < \sigma_2 < 1.$$

Le théorème suivant est important car il nous explicite les conditions suffisantes pour l'existence des scalaires $\bar{\lambda}_k$.

Théorème 1.6 [13]

Si d_k est une direction de descente, f est continument différentiable et bornée, alors il existe toujours un scalaire $\bar{\lambda}_k$ satisfait les conditions Wolfe1 et Wolfe2.

1.4.4 Convergence des méthodes à directions de descente

1.4 4 1 Condition de Zoutendijk

Dans cette section, on va étudier la *contribution* de la recherche linéaire inexacte dans la convergence des algorithmes à directions de descente. Ce n'est qu'une contribution, parce que la recherche linéaire ne peut à elle seule assurer la convergence des itérés. On comprend bien que le choix de la direction de descente joue aussi un rôle. Cela se traduit par une condition, dite de *Zoutendijk*, dont on peut tirer quelques informations qualitatives intéressantes.

On dit qu'une règle de recherche linéaire satisfait la *condition de Zoutendijk* s'il existe une constante $c > 0$ telle que pour tout indice $k \geq 1$ on ait

$$f(\mathcal{x}_{k+1}) \leq f(\mathcal{x}_k) - c \|\nabla f(\mathcal{x}_k)\|^2 \cos^2 \theta_k \quad (1.7)$$

où θ_k est l'angle que fait d_k avec $-\nabla f(\mathcal{x}_k)$, défini par

$$\cos\theta_k = \frac{-\nabla^\top f(\boldsymbol{x}_k)d\boldsymbol{x}}{\|\nabla f(\boldsymbol{x}_k)\| \|d\boldsymbol{x}\|} \quad (1.8)$$

Voici comment on se sert de la condition de Zoutendijk.

Proposition 1.2 *Si la suite $\{\boldsymbol{x}_k\}$ générée par un algorithme d'optimisation vérifie la condition de Zoutendijk (1.7) et si la suite $\{f(\boldsymbol{x}_k)\}$ est minorée, alors*

$$\sum_{k \geq 1} \|\nabla f(\boldsymbol{x}_k)\|^2 \cos^2\theta_k < \infty \quad (1.9)$$

Preuve. ([18])

En sommant les inégalités (1.7), on a

$$\sum_{k=1}^l \|\nabla f(\boldsymbol{x}_k)\|^2 \cos^2\theta_k \leq \frac{1}{c} (f(\boldsymbol{x}_1) - f(\boldsymbol{x}_{l+1}))$$

et puisque la suite $\{f(\boldsymbol{x}_k)\}$ est minorée, c'est-à-dire $\exists c' > 0$ telle que pour tout k , $f(\boldsymbol{x}_k) \geq c'$ alors

$$\sum_{k \geq 1} \|\nabla f(\boldsymbol{x}_k)\|^2 \cos^2\theta_k < \infty \text{ quand } l \rightarrow \infty.$$

■

Les deux propositions suivantes précisent les circonstances dans lesquelles la condition de Zoutendijk (1.7) est vérifiée avec les règles d'Armijo et de Wolfe.

Proposition 1.3 *Soit $f : IR^n \rightarrow IR$ une fonction $C^{1,1}$ (différentiable et sa dérivée vérifie pour une constante L et pour tout \boldsymbol{x} et \boldsymbol{y} : $\|\nabla f(\boldsymbol{y}) - \nabla f(\boldsymbol{x})\| \leq L \|\boldsymbol{y} - \boldsymbol{x}\|$) dans un voisinage de $\Gamma = \{\boldsymbol{x} \in IR^n : f(\boldsymbol{x}) \leq f(\boldsymbol{x}_1)\}$.*

On considère un algorithme à directions de descente d_k ; qui génère une suite $\{\boldsymbol{x}_k\}$ en utilisant la recherche linéaire d'Armijo avec $\lambda_1 > 0$.

Alors il existe une constante $c > 0$ telle que, pour tout $k \geq 1$, l'une des conditions

$$f(\boldsymbol{x}_{k+1}) \leq f(\boldsymbol{x}_k) - c \nabla^\top f(\boldsymbol{x}_k)d_k \quad (1.10)$$

ou

$$f(\boldsymbol{x}_{k+1}) \leq f(\boldsymbol{x}_k) - c \|\nabla f(\boldsymbol{x}_k)\|^2 \cos^2 \theta_k \quad (1.7)$$

est vérifiée.

Preuve. ([18])

Si le pas $\lambda_k = \lambda_1$ est accepté, on a (1.10), en effet

d_k direction de descente implique que $f(\boldsymbol{x}_{k+1}) < f(\boldsymbol{x}_k)$ et car λ_1 est uniformément positif

alors

$$f(\boldsymbol{x}_{k+1}) \leq f(\boldsymbol{x}_k) - c \nabla^\top f(\boldsymbol{x}_k) d_k.$$

Dans le cas contraire ($\lambda_k \neq \lambda_1$), la condition d'Armijo (wolfe1) n'est pas vérifiée avec un pas $\lambda'_k \leq \frac{\lambda_k}{\tau}$ c'est-à-dire $f(\boldsymbol{x}_k + \lambda'_k d_k) > f(\boldsymbol{x}_k) + \rho \lambda'_k \nabla^\top f(\boldsymbol{x}_k) d_k$

Comme f est $C^{1,1}$, on a pour tout $\lambda_k > 0$:

$$\begin{aligned} f(\boldsymbol{x}_k + \lambda_k d_k) &= f(\boldsymbol{x}_k) + \lambda_k \nabla^\top f(\boldsymbol{x}_k) d_k \\ &\quad + \int_0^1 [\nabla f(\boldsymbol{x}_k + t \lambda_k d_k) - \nabla f(\boldsymbol{x}_k)]^\top \lambda_k d_k dt \end{aligned}$$

$$\Rightarrow f(\boldsymbol{x}_k + \lambda_k d_k) \leq f(\boldsymbol{x}_k) + \lambda_k \nabla^\top f(\boldsymbol{x}_k) d_k + c \lambda_k^2 \|d_k\|^2$$

où $c > 0$ est une constante. Avec l'inégalité précédente, et le fait que , on obtient :

$$\begin{cases} f(\boldsymbol{x}_k + \lambda'_k d_k) - f(\boldsymbol{x}_k) > \rho \lambda'_k \nabla^\top f(\boldsymbol{x}_k) d_k \\ f(\boldsymbol{x}_k + \lambda'_k d_k) - f(\boldsymbol{x}_k) \leq \lambda'_k \nabla^\top f(\boldsymbol{x}_k) d_k + c \lambda_k'^2 \|d_k\|^2 \end{cases}$$

$$\begin{aligned} &\Rightarrow \rho \lambda'_k \nabla^\top f(\boldsymbol{x}_k) d_k \leq \lambda'_k \nabla^\top f(\boldsymbol{x}_k) d_k + c \lambda_k'^2 \|d_k\|^2 \\ &\Rightarrow -c \lambda_k'^2 \|d_k\|^2 \leq (1 - \rho) \lambda'_k \nabla^\top f(\boldsymbol{x}_k) d_k \end{aligned}$$

or

$$\rho < 1 \Rightarrow 0 < 1 - \rho < 1 \Rightarrow \frac{1}{1 - \rho} > 1$$

d'où

$$\begin{aligned} \nabla^\top f(\mathbf{x}_k)d_k \geq \frac{-c}{1-\rho} \lambda'_k \|d_k\|^2 &\Rightarrow -\nabla^\top f(\mathbf{x}_k)d_k \leq \frac{c}{1-\rho} \lambda'_k \|d_k\|^2 \\ \Rightarrow |\nabla^\top f(\mathbf{x}_k)d_k| &= \|\nabla^\top f(\mathbf{x}_k)\| \|d_k\| \cos\theta_k \leq \frac{c}{1-\rho} \lambda'_k \|d_k\|^2 \end{aligned}$$

ce qui permet de minorer $\lambda'_k \|d_k\|$ et donc aussi $\lambda_k \|d_k\|$ par une constante fois $\|\nabla^\top f(\mathbf{x}_k)\| \cos\theta_k$. Cette minoration et l'expression suivante de la condition d'Armijo (wolfe1)

$$f(\mathbf{x}_k + \lambda_k d_k) \leq f(\mathbf{x}_k) - \frac{\rho}{1-\rho} c \lambda_k \|\nabla^\top f(\mathbf{x}_k)\| \|d_k\| \cos\theta_k$$

car $\frac{\rho}{1-\rho} < 1$ on obtient

$$f(\mathbf{x}_k + \lambda_k d_k) \leq f(\mathbf{x}_k) - c \|\nabla^\top f(\mathbf{x}_k)\|^2 \cos^2\theta_k.$$

■

Proposition 1.4 Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction $C^{1,1}$ (différentiable et sa dérivée vérifie pour une constante c et pour tout \mathbf{x} et \mathbf{y} : $\|\nabla f(\mathbf{y}) - \nabla f(\mathbf{x})\| \leq c \|\mathbf{y} - \mathbf{x}\|$) dans un voisinage de

$$\Gamma = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \leq f(\mathbf{x}_1)\}.$$

On considère un algorithme à directions de descente d_k ; qui génère une suite $\{\mathbf{x}_k\}$ en utilisant la recherche linéaire de Wolfe (wolfe1)-(wolfe2) avec $\lambda_1 > 0$.

Alors il existe une constante $c > 0$ telle que, pour tout $k \geq 1$, la condition de Zoutendijk (1.7) est vérifiée.

Preuve. ([18])

D'après (wolfe1)

$$\begin{aligned} \nabla^\top f(\mathbf{x}_k + \lambda_k d_k)d_k &\geq \sigma \nabla^\top f(\mathbf{x}_k)d_k \\ \Rightarrow (\nabla f(\mathbf{x}_k + \lambda_k d_k) - \nabla f(\mathbf{x}_k))^\top d_k &\geq (\sigma - 1) \nabla^\top f(\mathbf{x}_k)d_k \\ = -(1 - \sigma) \nabla^\top f(\mathbf{x}_k)d_k &= (1 - \sigma) |\nabla^\top f(\mathbf{x}_k)d_k| \\ \Leftrightarrow (1 - \sigma) |\nabla^\top f(\mathbf{x}_k)d_k| &\leq (\nabla f(\mathbf{x}_k + \lambda_k d_k) - \nabla f(\mathbf{x}_k))^\top d_k \end{aligned}$$

et du fait que f est $C^{1,1}$:

$$\begin{aligned}
 (1 - \sigma) \left| \nabla^\top f(\mathbf{z}_k) d_k \right| &= (1 - \sigma) \left\| \nabla^\top f(\mathbf{z}_k) \right\| \|d_k\| \cos \theta_k \\
 &\leq \left\| \nabla f(\mathbf{z}_k + \lambda_k d_k) - \nabla f(\mathbf{z}_k) \right\| \|d_k\| \\
 &\Rightarrow (1 - \sigma) \left\| \nabla^\top f(\mathbf{z}_k) \right\| \cos \theta_k \leq L \lambda_k \|d_k\| \\
 &\Rightarrow \lambda_k \|d_k\| \leq \frac{(1 - \sigma)}{L} \left\| \nabla^\top f(\mathbf{z}_k) \right\| \cos \theta_k
 \end{aligned}$$

en utilisant (wolfe1), on aura :

$$\begin{aligned}
 f(\mathbf{z}_k + \lambda_k d_k) &\leq f(\mathbf{z}_k) + \rho \lambda_k \nabla^\top f(\mathbf{z}_k) d_k \\
 \Rightarrow f(\mathbf{z}_k + \lambda d_k) &\leq f(\mathbf{z}_k) + \rho \lambda \nabla^\top f(\mathbf{z}_k) d_k \leq f(\mathbf{z}_k) + \left| \rho \lambda \nabla^\top f(\mathbf{z}_k) d_k \right| \\
 \Rightarrow f(\mathbf{z}_k + \lambda d_k) &\leq f(\mathbf{z}_k) + \rho \lambda \left| \nabla^\top f(\mathbf{z}_k) d_k \right| \leq f(\mathbf{z}_k) - \rho \lambda \nabla^\top f(\mathbf{z}_k) d_k \\
 \Rightarrow f(\mathbf{z}_k + \lambda d_k) &\leq f(\mathbf{z}_k) - \rho \lambda \left\| \nabla^\top f(\mathbf{z}_k) \right\| \|d_k\| \cos \theta_k \\
 \Rightarrow f(\mathbf{z}_k + \lambda d_k) &\leq f(\mathbf{z}_k) - \frac{\rho(1 - \sigma)}{L} \left\| \nabla^\top f(\mathbf{z}_k) \right\|^2 \cos^2 \theta_k
 \end{aligned}$$

On en déduit (1.7). ■

L'inégalité (1.7), que nous appelons la condition de Zoutendijk, implique

$$\left\| \nabla^\top f(\mathbf{z}_k) \right\|^2 \cos^2 \theta_k \xrightarrow[k \rightarrow \infty]{} 0 \tag{1.11}$$

Cette limite peut être utilisée alternativement pour extraire des résultats de convergence globale pour les algorithmes de la recherche linéaire.

En effet si on s'assure dans le choix de la direction d_k que l'angle θ_k défini par (1.8) est tel que :

$$\exists \delta > 0 \text{ tel que } \cos \theta_k \geq \delta \text{ pour tout } k$$

alors en utilisant (1.11), on aura :

$$\lim_{k \rightarrow \infty} \left\| \nabla f(\mathbf{z}_k) \right\| = 0$$

En d'autres termes, nous pouvons être sûrs que les normes du gradient $\left\| \nabla f(\mathbf{z}_k) \right\|$ convergent vers zéro, si les directions de descente ne sont pas trop proches de l'orthogonalité avec le gradient.

1.4. RECHERCHES LINÉAIRES

En particulier, une méthode à direction de descente d_k , d_k faisant un angle nul avec le gradient génère une suite \varkappa_k vérifiant

$$\lim_{k \rightarrow \infty} \|\nabla f(\varkappa_k)\| = 0,$$

à condition bien sur d'utiliser une recherche linéaire de Wolfe ou de Goldstein.

Chapitre 2

Méthodes quasi-Newtoniennes

2.1 Méthode de Newton

Considérons le problème d'optimisation sans contraintes (P)

$$(P) : \underset{x \in \mathbb{R}^n}{\text{Min}} f(x) \quad (2.1)$$

où $f : \mathbb{R}^n \rightarrow \mathbb{R}$

Le principe de la méthode de Newton consiste à minimiser successivement les approximations du second ordre de f , plus précisément si

$$f(x) = f(x_k) + \nabla f(x_k)^t (x - x_k) + (x - x_k)^t H(x_k) (x - x_k) + o\|x - x_k\|^2,$$

posons

$$q(x) = f(x_k) + \nabla f(x_k)^t (x - x_k) + (x - x_k)^t H(x_k) (x - x_k).$$

Soit x_{k+1} l'optimum de q , alors il vérifie $\nabla q(x_{k+1}) = 0$, soit en remplaçant :

$$\nabla f(x_k) + H(x_k) (x_{k+1} - x_k) = 0,$$

ou encore

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

donc

$$x_{k+1} = x_k - [H(x_k)]^{-1} \nabla f(x_k)$$

Donc pour obtenir une méthode qui converge superlinéairement, il est nécessaire d'approximer l'étape de Newton asymptotiquement. C'est le principe de Dennis et Moré. Comment peut-on y aboutir sans évaluer la matrice Hessienne dans chaque itération ?

La réponse a été découverte par Davidon en 1959 ([35]) et a été développée et popularisée par Fletcher et Powell en 1963 ([31]). Elle consiste à commencer par n'importe quelle approximation de la matrice Hessienne et à chaque itération, on améliore la matrice en introduisant la courbure du problème mesuré tous au long de l'étape. Si cette amélioration est faite correctement, on obtient quelques méthodes remarquablement robustes et efficaces, qu'on appelle les méthodes de la variable métrique ou quasi Newton. Ils ont libéré l'optimisation non linéaire en procurant une alternative à la méthode de Newton, qui est très coûteuse pour plusieurs applications.

Il y a plusieurs méthodes de variable métrique, on s'étalera particulièrement sur les trois plus importantes, la méthode de correction de rang un, la méthode DFP (Davidon, Fletcher,

Powell), et la méthode BFGS (Broyden, Fletcher, Goldfarb, Shanno).

Algorithme 2.1 (méthode de Newton)

Étape initiale :

Soit $\varepsilon > 0$, critère d'arrêt. Choisir \varkappa_1 point initial, poser $k = 1$ et aller à l'étape principale.

Étape principale :

Si $\|\nabla f(\varkappa_k)\| \leq \varepsilon$ STOP, sinon poser $\varkappa_{k+1} = \varkappa_k - [H(\varkappa_k)]^{-1} \nabla f(\varkappa_k)$ remplacer k par $k + 1$ et aller à l'étape principale.

2.1.1 Avantages et inconvénients

♣ **Avantages**

- Si le point \varkappa_0 est assez proche de la solution optimale locale \varkappa^* telle que $H(\varkappa^*)$ soit définie positive, alors l'algorithme de Newton converge de façon quadratique vers la solution \varkappa^*

c'est-à-dire que l'on a

$$\|\varkappa_{k+1} - \varkappa^*\| \leq \gamma \|\varkappa_k - \varkappa^*\|^2 \quad \gamma \geq 0$$

- La méthode de Newton fonctionne bien si $1 < n < 10$ car il est facile de déterminer $\nabla^2 f(\boldsymbol{x}_k)$.
- La méthode de Newton converge en une seule itération lorsqu'elle est appliquée à une fonction quadratique strictement convexe.

♣ Inconvénients

- Il faut calculer les dérivées secondes de f qui peut être très coûteux en temps de calcul.
- L'algorithme n'est pas globalement convergent si le premier itéré \boldsymbol{x}_0 est éloigné de \boldsymbol{x}^* .
- Le successeur \boldsymbol{x}_{k+1} de \boldsymbol{x}_k n'est pas toujours bien défini, si $\nabla^2 f(\boldsymbol{x}) = H(\boldsymbol{x})$ est singulière.
- Si f n'est pas strictement convexe, l'algorithme ne génère par nécessairement des directions de descente de f .
- Un système linéaire d'ordre n doit être résolu à chaque itération, et on dispose pas de place mémoire pour stocker les $\mathcal{O}(n^2)$ éléments d'une matrice.

2.1.2 Comment peut-on remédier les défauts de l'algorithme de Newton ?

Pour obtenir une méthode qui converge superlinéairement, il est nécessaire d'approximer l'étape de Newton asymptotiquement qui est le principe de Dennis et Moré.

Comment peut-on y aboutir sans évaluer la matrice jacobienne $F'(\boldsymbol{x}_k)$ dans la résolution d'équation non linéaires ou le hessien $H(\boldsymbol{x}_k)$ de f en optimisation ?

La réponse a été découverte par Davidon en 1959 ([35]) et a été développée et popularisée par Fletcher et Powell en 1963 ([31]). Elle consiste à commencer par n'importe quelle approximation de la matrice hessienne et à chaque itération, on améliore la matrice en introduisant la courbure du problème mesuré tout au long de l'étape.

Si cette amélioration est faite correctement, on obtient quelques méthodes remarquablement robustes et efficaces, qu'on appelle les méthodes de *la variable métrique* ou *quasi-Newton*, ils ont libéré l'optimisation non linéaire en procurant une alternative à la méthode de Newton, qui est très coûteuse pour plusieurs applications.

2.2 Méthodes de Quasi-Newton

La classe des méthodes quasi-newtoniennes n'exigent pas des expressions explicites pour les deuxièmes dérivés. Celles-ci désignés parfois sous le nom des méthodes métriques variables. Car le nom implique que la base de ces méthodes est la méthode classique de Newton. Le principe de base dans des méthodes quasi-newtoniennes est que la direction de recherche est basée sur une matrice M_k , qui atteint le même but de l'inverse du Hessien H_k , dans la méthode de Newton.

Les algorithmes de quasi-Newton prennent donc en charge la génération de la suite des itérés $\{\varkappa_k\}_{k \geq 0}$, mais aussi celle de la suite des matrices B_k approchant $F'(\varkappa_k)$ ou H_k . La première suite générée par la récurrence

$$\varkappa_{k+1} = \varkappa_k + s_k$$

où $s_k = \lambda_k d_k$ dans les algorithmes avec recherche linéaire ($\lambda_k > 0$ est le pas et d_k vérifie $B_k d_k = -F'(\varkappa)$ ou $B_k d_k = -g_k$ selon le problème) et s_k est la solution d'un problème d'optimisation quadratique avec une borne sur la norme du déplacement dans les algorithmes avec régions de confiance.

Donc une méthode de quasi Newton est une méthode de type :

$$\begin{cases} \varkappa_{k+1} = \varkappa_k + \lambda_k d_k \\ d_k = -M_k g_k \end{cases} \quad (2.2)$$

ou bien

$$\begin{cases} \varkappa_{k+1} = \varkappa_k + \lambda_k d_k \\ d_k = -B_k^{-1} g_k \end{cases} \quad (2.3)$$

où M_k (respectivement B_k) est une matrice destinée à approcher l'inverse du hessien de f (respectivement le hessien) de f en \varkappa_k .

Si l'on veut que M_{k+1} approche le hessien $H_{k+1} = \nabla^2 f(\varkappa_{k+1})$, sans calculer ce dernier, la variation du gradient d'un itéré au suivant qu'il faut tirer de l'information. Puisque l'on a

$$y_k = g_{k+1} - g_k = \left(\int_0^1 \nabla^2 f(\varkappa_k + t s_k) dt \right) s_k,$$

il est naturel d'imposer à la matrice M_{k+1} de vérifier l'équation

$$s_k = M_{k+1} y_k \quad (2.4)$$

Cette équation porte le nom d'*équation de quasi-Newton*.

Il est également normal d'imposer à M_{k+1} d'être symétrique (puisque $[H(\boldsymbol{x}_k)]^{-1}$ l'est).

2.2.1 Quelques propriétés des algorithmes de quasi-Newton

- Des méthodes quasi-newtoniennes, comme la plupart des autres méthodes, sont développées pour le problème quadratique convexe et puis sont prolongées au problème général. Elles sont considérées comme étant parmi les méthodes les plus efficaces et sont employées très intensivement dans de nombreuses applications. Plusieurs méthodes quasi-newtoniennes distinctes ont évolué ces dernières années. Dans ce chapitre, nous discuterons en détail les quatre méthodes les plus importantes de cette classe qui sont :

- 1- *Méthode de correction de rang un* (SR1 [1960])
- 2- *Méthode de Davidon-Fletcher-Powell* (DFP [D-1959 et F.P-1963])
- 3- *Méthode de Broyden-Fletcher-Goldfarb-Shanno* (BFGS [1970])

- les algorithmes quasi-newtoniens convergent moins rapidement que l'algorithme de Newton. Dans les implémentations correctes, les itérés convergent toutefois q-superlinéairement.

- Chaque itération demande moins de calcul au simulateur que dans l'algorithme de Newton : il n'est pas nécessaire d'évaluer les dérivées secondes.

- Dans leur version standard, ces algorithmes peuvent être utilisés pour un nombre de variables qui n'est pas trop grand, disons $n \leq 500$ pour fixer les idées, cependant on a développé ces dernières années (à partir de 1990) des algorithmes qui sont des variantes de la méthode BFGS et qui prennent en charge les problèmes de grande taille (citons particulièrement la méthode BFGS à mémoire limitée, La LBFGS).

2.2.2 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

$$\begin{cases} \varkappa_{k+1} = \varkappa_k + \lambda_k d_k \\ d_k = -M_k g_k \end{cases}$$

Nous cherchons M_k sous la forme suivante :

$$M_{k+1} = M_k + \Delta_k \tag{2.5}$$

avec Δ_k symétrique, assurant la relation de quasi-Newton. On exige aussi que M_{k+1} soit définie positive, sous l'hypothèse que M_k est définie positive.

La formule (2.5) permet d'utiliser les nouvelles informations obtenues lors de l'étape k de l'algorithme, c'est-à-dire essentiellement le gradient $g_{k+1} = \nabla f(\varkappa_{k+1})$ au point \varkappa_{k+1} , obtenu par une recherche linéaire (exacte ou approchée) dans la direction d_k . Il existe différentes formules de type (2.2). Suivant que Δ_k est de rang un ou deux, on parlera de correction de rang un (SR1) ou de rang deux (DFP, BFGS).

2.2.3 Méthode de correction de rang un (SR1 [1960])

Cette méthode a été proposée par Broyden ([6]), Davidon ([35]), Fiacco et McCormick ([4]), Murtagh et Sargent ([5]), et Wolfe ([28]).

Etant donné que $[H(\varkappa_k)]^{-1}$ est symétrique, la formule de mise à jour de l'approximation du Hessien est la suivante :

$$M_{k+1} = M_k + a_k u_k u_k^\top \tag{2.6}$$

où u_k est un vecteur de IR^n et a_k constante. La matrice de correction $\Delta_k = a_k u_k u_k^\top$ est symétrique et de rang un.

Des équations (2.4)-(2.6) on obtient :

$$s_k = M_k y_k + a_k u_k u_k^\top y_k \tag{2.7}$$

et par conséquent

$$\begin{aligned} y_k^\top (s_k - M_k y_k) &= a_k y_k^\top u_k u_k^\top y_k \\ &= a_k (u_k^\top y_k)^2 \end{aligned} \tag{2.8}$$

Alternativement, de l'équation (2.7) on a :

$$(s_k - M_k y_k) = a_k u_k u_k^\top y_k = a_k (u_k^\top y_k) u_k$$

$$(s_k - M_k y_k)^\top = a_k y_k^\top u_k u_k^\top = a_k (u_k^\top y_k) u_k^\top$$

Si on prend

$$u_k = (s_k - M_k y_k),$$

alors

$$a_k (u_k^\top y_k) = 1.$$

Soit en remplaçant :

$$a_k \cdot (s_k - M_k y_k)^\top \cdot y_k = 1$$

ou encore :

$$a_k = \frac{1}{(s_k - M_k y_k)^\top \cdot y_k},$$

et (2.6) devient :

$$M_{k+1} = M_k + \frac{(s_k - M_k y_k) \cdot (s_k - M_k y_k)^\top}{(s_k - M_k y_k)^\top \cdot y_k} \quad (2.9)$$

Cette formule porte le nom de formule SR1 (*système de rang 1*).

Algorithme 2.2 (Méthode de SR1)

Etape initiale :

Soit $\varepsilon > 0$ (critère d'arrêt). Choisir un point initial \varkappa_1 et une matrice symétrique définie positive M_1 quelconque (par exemple $M_1 = I$) poser $k = 1$, et aller aux étapes principales

Etapes principales.

Etape 1 :

Si $\|\nabla f(\varkappa_k)\| < \varepsilon$ stop ; sinon, poser $d_k = -M_k g_k$ et déterminer le pas optimal λ_k solution optimale du problème $\min f(\varkappa_k + \lambda d_k)$, $\lambda \geq 0$. et poser $\varkappa_{k+1} = \varkappa_k + \lambda_k d_k$

Etape 2 :

Construire M_{k+1} comme suit :

$$M_{k+1} = M_k + \frac{(s_k - M_k y_k) \cdot (s_k - M_k y_k)^\top}{(s_k - M_k y_k)^\top \cdot y_k}$$

avec

$$\begin{aligned} s_k &= \varkappa_{k+1} - \varkappa_k \\ y_k &= \nabla f(\varkappa_{k+1}) - \nabla f(\varkappa_k) \end{aligned}$$

Remplacer k par $k + 1$ et aller à l'étape 1. \square

Cet algorithme a un comportement remarquable dans le cas où f est une fonction quadratique.

Pour un problème quadratique convexe, cette méthode converge au plus dans $(n + 1)$ itérations et $(M_{k+1})^{-1} = H$ à condition que $s_i = M_{k+1}y_i$ pour $0 \leq i \leq k$. C'est en effet le cas, comme sera démontré par le théorème suivant.

Théorème 2.1. ([1]) *Soit f une fonction quadratique, de matrice Hessienne H définie positive. Considérons la méthode itérative qui, partant d'un point \varkappa_0 arbitraire engendre les points*

$$\varkappa_{k+1} = \varkappa_k + s_k$$

où les s_k sont des vecteurs lin. indépendants. Pour toute matrice symétrique initiale M_0 , considérons la suite des matrices $\{M_k\}$ définis par :

$$M_{k+1} = M_k + \frac{(s_k - M_k y_k)(s_k - M_k y_k)^\top}{s_k^\top (s_k - M_k y_k)}.$$

avec

$$y_k = \nabla f(\varkappa_{k+1}) - \nabla f(\varkappa_k).$$

Alors on a :

$$s_i = M_{k+1}y_i \quad \text{pour } 0 \leq i \leq k$$

et dans au plus n étapes, la suite $\{\varkappa_k\}$ converge vers la solution optimale, et M_k converge vers l'inverse du Hessien H^{-1} ($M_{k+1} = H^{-1}$)

Preuve. ([1])

Supposons que

$$s_i = M_k y_i \quad \text{pour } 0 \leq i \leq k - 1 \tag{2.10}$$

est vraie et montrons qu'elle est vraie pour k , si $0 \leq i \leq k - 1$ de l'équation (2.9)

$$M_{k+1}y_i = M_k y_i + u_k (s_k - M_k y_k)^\top y_i$$

où

$$u_k = \frac{s_k - M_k y_k}{y_k^\top (s_k - M_k y_k)}$$

Puisque M_k est symétrique, nous pouvons écrire

$$M_{k+1} y_i = M_k y_i + u_k (s_k^\top y_i - y_k^\top M_k y_i)$$

et de (2.10)

$$M_{k+1} y_i = s_i + u_k (s_k^\top y_i - y_k^\top s_i) \quad (2.11)$$

pour $0 \leq i \leq k$ on a

$$y_i = H s_i$$

et

$$y_k^\top = s_k^\top H$$

Par conséquent pour $0 \leq i \leq k-1$, nous prenons

$$s_k^\top y_i - y_k^\top s_i = s_k^\top H s_i - s_k^\top H s_i = 0$$

et de l'équation (2.11) on obtient

$$s_i = M_{k+1} y_i \quad \text{pour } 0 \leq i \leq k-1 \quad (2.12)$$

on a (voir (2.4))

$$s_k = M_{k+1} y_k$$

et avec (2.12), nous pouvons écrire

$$s_i = M_{k+1} y_i \quad \text{pour } 0 \leq i \leq k \quad (2.13)$$

Pour accomplir l'induction, nous notons

$$s_i = M_1 y_i = M_1 H s_i, \quad \text{pour } 0 \leq i \leq 0$$

et d'après (2.13)-(2.10) nous pouvons écrire

$$s_i = M_2 y_i = M_2 H s_i, \quad \text{pour } 0 \leq i \leq 1$$

$$\begin{array}{rcl}
 s_i = M_3 y_i = M_3 H s_i, & \text{pour } 0 \leq i \leq 2 \\
 \vdots & & \vdots \\
 s_i = M_{k+1} y_i = M_{k+1} H s_i, & \text{pour } 0 \leq i \leq k
 \end{array}$$

et comme les vecteurs s_i sont indépendants, c'est-à-dire l'ensemble $\{s_0, s_1, \dots, s_k\}$ forme une base de \mathbb{R}^{k+1} alors,

$$M_{k+1} H = I$$

ou encore,

$$M_{k+1} = H^{-1}.$$

■

Remarque 2.1 si $M_{k+1} = H^{-1}$ alors on revient à la méthode de Newton qui converge en une seule itération puisqu'elle est appliquée à une fonction quadratique strictement convexe ; et aussi le théorème précédant permet d'obtenir l'inverse d'une matrice avec une méthode itérative.

Avantages et inconvénients de la méthode de SR1

♣ Avantages

- Cette méthode présente l'avantage, que le point x_{k+1} n'a pas besoin d'être choisi comme le minimum exact, c'est à dire qu'on n'a pas besoin d'effectuer des recherches linéaire exactes.

♣ Inconvénients

- Même si la fonction est quadratique, et même si son Hessien est défini positif, la matrice M_k n'est pas forcément définie positive.
- Le dénominateur $(s_k - M_k y_k)^\top y_k$ peut devenir nul ou très petit, ce qui rend le procédé instable c'est-à-dire, la méthode n'est pas bien définie.

Remarque 2.2 la méthode de SR1 est souvent utilisée lorsqu'il n'est pas possible ou qu'il n'est pas nécessaire que M_k soit définie positive.

2.2.4 Méthode de Davidon-Fletcher-Powell (DFP [D-1959 et F.P-1963])

Cette méthode a été proposée par Davidon en 1959 ([35]) et développé plus tard en 1963 par Fletcher et Powell ([31]). La formule de mise à jour de DFP est une formule de correction de rang deux. De façon plus précise construisons M_{k+1} en fonction de M_k de la forme :

$$M_{k+1} = M_k + A_k + \Delta_k \quad (2.14)$$

Bien sur le choix de A_k et Δ_k doit prendre en considération l'équation de quasi-Newton

$$M_{k+1}y_k = s_k \quad (2.14 \text{ bis})$$

Si on écrit A_k sous la forme $A_k = a_k u_k u_k^\top$ et $\Delta_k = b_k v_k v_k^\top$, alors (2.14 bis) prend la forme :

$$(a_k u_k u_k^\top) \cdot y_k + (b_k v_k v_k^\top) \cdot y_k = s_k - M_k \cdot y_k.$$

Comme dans le cas de la formule SR1, un bon choix de a_k , b_k , u_k et v_k conduit à la formule de correction de rang 2 suivante :

$$M_{k+1} = M_k + \frac{s_k s_k^\top}{s_k^\top y_k} - \frac{M_k y_k y_k^\top M_k}{y_k^\top M_k y_k} \quad (2.15)$$

Algorithme 2.3 (Méthode de DFP)

Etape initiale :

Soit $\varepsilon > 0$, déterminer le critère d'arrêt. Choisir un point initial \mathbf{x}_1 et une matrice symétrique définie positive M_1 quelconque

(par exemple $M_1 = I$) poser $k = 1$, et aller aux étapes principales

Etapes principales.

Etape 1 :

Si $\|\nabla f(\mathbf{x}_k)\| < \varepsilon$ STOP; sinon, poser $d_k = -M_k g_k$ et déterminer le pas optimal λ_k solution optimale du problème

$\min f(\mathbf{x}_k + \lambda d_k)$, $\lambda \geq 0$. et poser $\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda_k d_k$

Etape 2 :

Construire M_{k+1} comme suit :

$$M_{k+1} = M_k + \frac{s_k s_k^T}{s_k^T y_k} - \frac{M_k y_k y_k^T M_k}{y_k^T M_k y_k}$$

avec

$$\begin{aligned} s_k &= \varkappa_{k+1} - \varkappa_k \\ y_k &= \nabla f(\varkappa_{k+1}) - \nabla f(\varkappa_k) \end{aligned}$$

Remplacer k par $k + 1$ et aller à l'étape 1. \square

Cet algorithme a un comportement remarquable dans le cas où f est une fonction quadratique.

Théorème 2.2. *La formule (2.15) conserve la définie positivité des matrices M_k , c'est-à-dire, si M_k est définie positive alors M_{k+1} est également définie positive.*

Preuve. Soit \varkappa un vecteur de \mathbb{R}^n on a

$$\begin{aligned} \varkappa^\top M_{k+1} \varkappa &= \varkappa^\top M_k \varkappa + \frac{(s_k^\top \varkappa)^2}{s_k^\top y_k} - \frac{(y_k^\top M_k \varkappa)^2}{y_k^\top M_k y_k} \\ &= \frac{y_k^\top M_k y_k^\top \varkappa M_k \varkappa - (y_k^\top M_k \varkappa)^2}{y_k^\top M_k y_k} + \frac{(s_k^\top \varkappa)^2}{s_k^\top y_k} \end{aligned}$$

Si on définit le produit scalaire $\langle \varkappa, y \rangle = \varkappa^\top M_k y_k$, alors on a

$$\varkappa^\top M_{k+1} \varkappa = \frac{\langle y_k, y_k \rangle \langle \varkappa, \varkappa \rangle - \langle y_k, \varkappa \rangle^2}{\langle y_k, y_k \rangle} + \frac{(s_k^\top \varkappa)^2}{s_k^\top y_k} \quad (2.16)$$

Le premier terme du second membre est positive ou nul d'après l'inégalité de Cauchy Schwartz. et pour le deuxième terme on peut faire l'analyse suivante :

Puisque le pas est optimal on a la relation d'après l'orthogonalité des directions successives (voir section 1.4.1)

$$g_{k+1}^\top d_k = 0$$

et donc

$$\begin{aligned} s_k^\top y_k &= \lambda_k (g_{k+1} - g_k)^\top d_k \\ &= \lambda_k g_k^\top M_k g_k > 0 \end{aligned}$$

On a donc

$$\varkappa^\top M_{k+1} \varkappa \geq 0.$$

Les deux termes dans (2.16) 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 $\varkappa = \lambda y_k$ pour un scalaire $\lambda \neq 0$, dans ce cas le deuxième terme est non nul car $s_k^\top \varkappa = \lambda s_k^\top y_k$. on a donc bien

$$\varkappa^\top M_{k+1} \varkappa > 0 \quad \text{pour } \varkappa \neq 0.$$

■

Remarque 2.3

• Il est important de noter que le résultat ci-dessus est vérifié pour n'importe quel $\lambda_k > 0$ pour lequel

$$s_k^\top y_k = s_k^\top g_{k+1} - s_k^\top g_k > 0 \tag{2.17}$$

c'est-à-dire si $s_k^\top y_k > 0$, alors la définie positivité de M_{k+1} est préservée même dans le cas où la minimisation de $f(\varkappa_k + \lambda d_k)$ est inexacte et donc l'équation de quasi-newton a toujours une solution M_{k+1} .

• Si f est strictement convexe alors l'inégalité (2.17) est vérifiée pour tous \varkappa_k et \varkappa_{k+1}

$$\left(\begin{array}{l} \text{car : } f \text{ est strictement convexe} \Leftrightarrow \forall \varkappa_k, \varkappa_{k+1} \in IR^n; \varkappa_k^\top H(\varkappa_{k+1}) \varkappa_k > 0 \\ \Rightarrow s_k^\top M_{k+1} s_k = s_k^\top y_k > 0 \end{array} \right)$$

Si f est non convexe alors l'inégalité (2.17) n'est pas forcément vérifiée.

Dans ce cas on utilise la recherche linéaire de Wolfe ou de Wolfe forte pour vérifier la condition (2.17). En effet

de (wolfe2) on a

$$g_{k+1}^\top s_k \geq \sigma g_k^\top s_k \Rightarrow y_k^\top s_k \geq (\sigma - 1) \lambda_k g_k^\top d_k$$

Puisque $\sigma < 1$ et $g_k^\top d_k < 0$, on peut écrire

$$y_k^\top s_k \geq (\sigma - 1) \lambda_k g_k^\top d_k > 0.$$

Théorème 2.3. ([1]) **(Directions conjuguées engendrées par méthode de DFP)**

a) Si la recherche linéaire utilisée dans la méthode de DFP est exacte et si la fonction objective $f(\varkappa)$ est une fonction convexe et quadratique, alors l'algorithme DFP décrit par la relation

$$M_{k+1} = M_k + \frac{s_k s_k^\top}{s_k^\top y_k} - \frac{M_k y_k y_k^\top M_k}{y_k^\top M_k y_k} \tag{2.15}$$

2.2. MÉTHODES DE QUASI-NEWTON

engendre un ensemble des directions conjuguées s_0, s_1, \dots, s_k c'est-à-dire,

$$s_i^\top H s_j = 0 \quad \text{pour } 0 \leq i < j \leq k \quad (2.18)$$

b) Si

$$y_i = H s_i \quad \text{pour } 0 \leq i \leq k \quad (2.19)$$

alors

$$s_i = M_{k+1} y_i \quad (2.20)$$

Preuve. ([1])

Supposons que

$$s_i^\top H s_j = 0 \quad \text{pour } 0 \leq i < j \leq k-1 \quad (2.21)$$

$$s_i = M_k y_i \quad \text{pour } 0 \leq i \leq k-1 \quad (2.22)$$

puis nous montrons l'équations (2.18) et (2.20)

a) on a

$$y_k = H s_k$$

où

$$s_k = \mathcal{X}_{k+1} - \mathcal{X}_k$$

$$y_k = g_{k+1} - g_k$$

donc nous pouvons écrire

$$\begin{aligned} g_k &= g_{k-1} + H s_{k-1} \\ &= g_{k-2} + H s_{k-2} + H s_{k-1} \\ &= g_{k-3} + H s_{k-3} + H s_{k-2} + H s_{k-1} \\ &\vdots \\ &= g_{i+1} + H(s_{i+1} + s_{i+2} + \dots + s_{k-1}) \end{aligned}$$

Ainsi pour $0 \leq i \leq k-1$, nous prenons

$$s_i^\top g_k = s_i^\top g_{i+1} + s_i^\top H(s_{i+1} + s_{i+2} + \dots + s_{k-1}) \quad (2.23)$$

et quand la recherche linéaire utilisée dans la méthode de DFP est exacte, alors $f(\varkappa)$ est minimisée exactement au point \varkappa_{i+1} et ainsi

$$g_{i+1}^\top d_i = 0 \quad (2.24)$$

implique

$$\lambda_i g_{i+1}^\top d_i = g_{i+1}^\top \lambda_i d_i = s_i^\top g_{i+1} = 0$$

Maintenant pour $0 \leq i \leq k-1$, l'équation (2.21) donne

$$s_i^\top H(s_{i+1} + s_{i+2} + \dots + s_{k-1}) = 0 \quad (2.25)$$

et de (2.23)-(2.25), nous obtenons

$$s_i^\top g_k = 0$$

Alternativement, d'équations. (2.22)-(2.19) nous pouvons écrire

$$s_i^\top g_k = (M_k y_i)^\top g_k = (M_k H s_i)^\top g_k = s_i^\top H M_k g_k = 0$$

De plus, pour éliminer $M_k g_k$ on utilise l'équation

$$s_k = \lambda_k d_k = -\lambda_k M_k g_k$$

ainsi

$$s_i^\top g_k = -\frac{1}{\lambda_k} s_i^\top H s_k = 0$$

et depuis $\lambda_k > 0$, nous avons

$$s_i^\top H s_k = 0 \quad \text{pour } 0 \leq i \leq k-1 \quad (2.26)$$

Maintenant, par la combinaison de. (2.21) et (2.26)

$$s_i^\top H s_j = 0 \quad \text{pour } 0 \leq i < j \leq k \quad (2.27)$$

Pour accomplir l'induction, nous pouvons écrire

$$s_0^\top g_1 = (M_1 y_0)^\top g_1 = (M_1 H s_0)^\top g_1 = s_0^\top H M_1 g_1 = -\frac{1}{\lambda_1} s_0^\top H s_1$$

et depuis $f(x)$ est minimisée exactement au point \varkappa_1 , nous avons $s_0^\top g_1 = 0$ et

$$s_i^\top H s_j = 0 \quad \text{pour } 0 \leq i < j \leq 1$$

2.2. MÉTHODES DE QUASI-NEWTON

Depuis (2.27) est vérifiée si (2.21) est vérifiée, nous pouvons écrire

$$\begin{aligned} s_i^\top H s_j &= 0 && \text{pour } 0 \leq i < j \leq 2 \\ s_i^\top H s_j &= 0 && \text{pour } 0 \leq i < j \leq 3 \\ &\vdots && \vdots \\ s_i^\top H s_j &= 0 && \text{pour } 0 \leq i < j \leq k \end{aligned}$$

c'est-à-dire, les directions s_0, s_1, \dots, s_k , forment un ensemble des directions conjuguées.

b) de (2.22)

$$y_k^\top s_i = y_k^\top M_k y_i \quad \text{pour } 0 \leq i \leq k-1 \quad (2.28)$$

et d'autre part de (2.19) on a

$$y_k^\top s_i = s_k^\top H s_i \quad \text{pour } 0 \leq i \leq k-1 \quad (2.29)$$

Depuis s_0, s_1, \dots, s_k , forme un ensemble des directions conjuguées (de la partie (a)), et de (2.28)-(2.29)

$$y_k^\top s_i = y_k^\top M_k y_i = s_k^\top H s_i \quad \text{pour } 0 \leq i \leq k-1 \quad (2.30)$$

En notant

$$s_k^\top = y_k^\top M_{k+1} \quad \text{et} \quad H s_i = y_i$$

(2.30) peut être exprimé

$$y_k^\top s_i = y_k^\top M_k y_i = y_k^\top M_{k+1} y_i = 0 \quad \text{pour } 0 \leq i \leq k-1$$

et, en conséquence,

$$s_i = M_k y_i = M_{k+1} y_i \quad \text{pour } 0 \leq i \leq k-1 \quad (2.31)$$

Maintenant, de (2.4)

$$s_k = M_{k+1} y_k$$

et par la combinaison de. (2.31) et (2.4), nous obtenons

$$s_i = M_{k+1} y_i \quad \text{pour } 0 \leq i \leq k.$$

■

Remarque 2.4 Pour $k = n - 1$, des équations (2.19) et (2.20) on peut écrire

$$[M_n H - \alpha I] s_i = 0 \quad \text{pour } 0 \leq i \leq n - 1$$

avec $\alpha = 1$. En effet, les vecteurs s_i sont des vecteurs propres qui correspondent à la valeur propre d'unité pour la matrice $M_n H$, et puisqu'ils sont linéairement indépendants, nous avons

$$M_n = H^{-1}$$

c'est-à-dire, dans un problème quadratique M_{k+1} devient le Hessien à l'itération $n - 1$.

Remarque 2.5 Une méthode de directions conjuguées est une méthode itérative qui appliquée à une fonction quadratique de n variables conduit à l'optimum en n étapes au plus; autrement dit, une méthode de directions conjuguées converge donc de façon finie en au plus n étapes.

Avantages et inconvénients de la méthode de DFP

♣ Avantages

- 1- Pour des fonctions quadratiques (avec une recherche linéaire exacte) :
 - L'algorithme converge dans au plus n étapes avec $M_{n+1} = H^{-1}$.
 - Elles engendrent des directions conjuguées.
- 2- Pour les fonctions quelconques :
 - la matrice M_k reste définie positive, ce qui est nécessaire pour que la direction soit une direction de descente.

2.2.5 Méthode de Broyden-Fletcher-Goldfarb-Shanno (BFGS [1970])

La méthode de BFGS est une méthode quasi-newtonienne bien connue pour résoudre des problèmes d'optimisation sans contraintes, elle a été proposée par : Broyden ([6]), Fletcher ([30]), Goldfarb ([9]), et Shanno ([10]). Elle est devenue une méthode de choix pour les ingénieurs et les mathématiciens qui sont intéressés à résoudre des problèmes d'optimisation.

Une autre manière d'imposer à M_{k+1} d'être proche de M_k est de minimiser l'"écart" entre M_{k+1} et M_k , toujours en requérant que M_{k+1} soit symétrique

2.2. MÉTHODES DE QUASI-NEWTON

et vérifie l'équation de quasi-Newton (2.4). On est donc conduit à considérer le problème en la variable matricielle $M \in IR^{n \times n}$ suivant :

$$\begin{cases} \min \text{ "écart" } (M, M_k) \\ y_k = Ms_k \\ M = M^\top \end{cases} ; M \in IR^{n \times n} \quad (2.32)$$

On dit alors que la matrice est obtenue par une *approche variationnelle*.

Notations : Pour indiquer qu'une matrice carrée M est symétrique semi-définie positive (resp. définie positive), on notera $M \geq 0$ (resp. $M > 0$). L'ensemble des matrices symétriques d'ordre n est noté S^n , $S_+^n := \{M \in S^n : M \geq 0\}$ et $S_{++}^n := \{M \in S^n : M > 0\}$.

Il est souvent intéressant d'imposer également la définie positivité des matrices M_k , car pour que $d_k = -M_k^{-1}g_k$ soit une direction de descente, on a besoin que M_k soit une matrice symétrique, définie positive (en effet : on sait que, toute matrice symétrique réelle et définie positive est inversible, et son inverse est elle aussi définie positive, c'est-à-dire on peut écrire :

$$g_k^\top d_k = -g_k^\top M_k^{-1}g_k < 0 \text{ donc, } d_k \text{ est une direction de descente}).$$

Cette condition définissant un ensemble ouvert ne peut être utilisée directement comme contrainte dans le problème définissant M_{k+1} ; et dans ce but, on commence par introduire une fonction : $\psi : S^n \rightarrow IR$ dont le domaine est S_{++}^n et qui forme une "barrière" au bord du cône S_{++}^n (elle tend vers l'infini lorsque son argument se rapproche du bord de S_{++}^n) ainsi qu'à l'infini :

$$\psi(\Upsilon) = tr \Upsilon + ld \Upsilon \quad (2.33)$$

où la fonction log-déterminant $ld : S^n \rightarrow IR \cup \{+\infty\}$ est définie en $\Upsilon \in S^n$ par

$$ld(\Upsilon) = \begin{cases} -\log \det \Upsilon & \text{si } \Upsilon \in S_{++}^n \\ +\infty & \text{sinon} \end{cases}$$

Les propriétés annoncées de ψ peuvent se voir sur son expression suivante. Si on note $\{\alpha_i\}$ les valeurs propres de Υ , on a $tr \Upsilon = \sum_{i=1}^n \alpha_i$, $\det \Upsilon = \prod_{i=1}^n \alpha_i$ et donc

$$\psi(\Upsilon) = \sum_{i=1}^n (\alpha_i - \log \alpha_i), \quad \text{si } \Upsilon \in S_{++}^n \quad (2.34)$$

Etant donné l'allure de la fonction $t \in IR_{++} \mapsto t - \log t$, $\psi(\Upsilon)$ tend vers l'infini si l'une des valeurs propres de Υ tend vers zéro ou vers l'infini.

$$(\exists j \in \{1, \dots, n\}; \lim_{\alpha_j \rightarrow 0 \text{ ou } \infty} \psi(\Upsilon) = \infty).$$

La formule (2.34) montre aussi que l'unique minimiseur de ψ est $\Upsilon = I$ la matrice identité ($\alpha_i = 1$ pour tout i).

Si M_k est une matrice réelle symétrique alors :

la matrice M_k est définie positive si et seulement s'il existe une matrice définie positive A_k telle que $A_k^2 = M_k$; dans ce cas, la matrice définie positive A_k et on peut la noter $A_k = M_k^{\frac{1}{2}}$ est unique.

en effet :

Si M_k est une matrice réelle symétrique alors on peut écrire

$$U^\top M_k U = \Lambda$$

où U est vérifiée

$$U^\top U = U U^\top = I$$

et Λ est une matrice diagonale où les éléments du diagonale sont les valeurs propres de M_k qui sont strictement positives car M_k est définie positive. Donc on peut écrire

$$M_k = U \Lambda U^\top = U \Lambda^{\frac{1}{2}} \Lambda^{\frac{1}{2}} U^\top = \left(U \Lambda^{\frac{1}{2}} U^\top \right) \left(U \Lambda^{\frac{1}{2}} U^\top \right) = M_k^{\frac{1}{2}} M_k^{\frac{1}{2}}. \quad \square$$

On veut dire, qu'il faut trouver une matrice M symétrique et définie positive être proche de M_k c'est-à-dire

$$M \cong M_k^{\frac{1}{2}} M_k^{\frac{1}{2}}$$

implique

$$M_k^{-\frac{1}{2}} M M_k^{-\frac{1}{2}} \cong I$$

Donc afin de minimiser l'écart entre M et M_k , on va chercher à ce que $M_k^{-\frac{1}{2}} M M_k^{-\frac{1}{2}}$ soit proche de I ; et ceci peut être obtenu en minimisant $\psi \left(M_k^{-\frac{1}{2}} M M_k^{-\frac{1}{2}} \right)$, si bien que l'on aura M_{k+1} proche de M_k en résolvant :

$$\begin{cases} \min \psi \left(M_k^{-\frac{1}{2}} M M_k^{-\frac{1}{2}} \right) \\ y_k = M s_k \\ M \in S_{++}^n \quad (\text{contrainte implicite}). \end{cases} \quad (2.35)$$

Si $s_k = 0$ alors $\begin{cases} \text{si } y_k \neq 0 \text{ alors (2.35) n'a pas de solutions} \\ \text{ou bien} \\ \text{si } y_k = 0 \text{ alors la solution du (2.35) est } M = M_k \end{cases}$;

Sinon ; le cas non trivial où $s_k \neq 0$ est examiné dans le proposition suivante.

Proposition 2.1 *Supposons que M_k soit symétrique définie positive et que $s_k \neq 0$. Alors, le problème (2.35) a une solution si et seulement si $y_k^\top s_k > 0$.*

Sous cette condition la solution M_{k+1} de (2.35) est unique et est donnée par l'une des formules suivantes :

$$M_{k+1} = M_k + \frac{y_k y_k^\top}{y_k^\top s_k} - \frac{M_k s_k s_k^\top M_k}{s_k^\top M_k s_k} \quad (2.36)$$

$$B_{k+1} = \left(I - \frac{s_k y_k^\top}{y_k^\top s_k} \right) B_k \left(I - \frac{y_k s_k^\top}{y_k^\top s_k} \right) + \left(I - \frac{s_k s_k^\top}{y_k^\top s_k} \right) \quad (2.37)$$

où on a noté $B_k := M_k^{-1}$ et $B_{k+1} := M_{k+1}^{-1}$.

Algorithme 2.4 (Méthode de BFGS)

Étape initiale :

Soit $\varepsilon > 0$, déterminer le critère d'arrêt. Choisir \varkappa_1 point initial et M_1 définie positive quelconque (par exemple $M_1 = I$).

Poser $k = 1$ et aller aux étapes principales

Étapes principales.

Étape 1 :

Si $\|\nabla f(\varkappa_k)\| < \varepsilon$ STOP ; sinon, poser $d_k = -M_k g_k$ et déterminer le pas

optimal λ_k solution optimale du problème $\min f(\varkappa_k + \lambda d_k)$, $\lambda \geq 0$ et poser

$$\varkappa_{k+1} = \varkappa_k + \lambda_k d_k$$

Étape 2 :

Construire M_{k+1} comme suit :

$$M_{k+1} = M_k + \frac{y_k y_k^\top}{y_k^\top s_k} - \frac{M_k s_k s_k^\top M_k}{s_k^\top M_k s_k}$$

avec

$$\begin{aligned} s_k &= \mathcal{X}_{k+1} - \mathcal{X}_k \\ y_k &= \nabla f(\mathcal{X}_{k+1}) - \nabla f(\mathcal{X}_k) \end{aligned}$$

Remplacer k par $k + 1$ et aller à l'étape 1. \square

Proposition 2.2 ([18]) (*Propriété algébrique*) Soit H une matrice symétrique et s_1, s_2, \dots, s_p des directions non nulles, conjuguées par rapport à cette matrice (c'est-à-dire : $s_i^\top H s_j = 0$, pour $i \neq j$ et $s_i^\top H s_i > 0$). On se donne une matrice M_1 symétrique définie positive et on définit

$$M_{i+1} = M_i + \frac{y_i y_i^\top}{y_i^\top s_i} - \frac{M_i s_i s_i^\top M_i}{s_i^\top M_i s_i}, \quad \text{pour } i = 1, \dots, p$$

Alors,

- (i) $M_{k+1} s_i = H s_i$, pour $k = 1, \dots, p$ et $i = 1, \dots, k$,
- (ii) si $p = n$, on a $M_{n+1} = H$.

Preuve. ([18])

(i) La première propriété se démontre par récurrence.

Elle est vérifiée pour $k = 1$: $M_2 s_1 = H s_1$ (c'est l'équation de quasi-Newton vérifiée par M_2 où $y_1 = H s_1$). Supposons qu'elle le soit pour un $k = 1, \dots, l - 1$, avec $2 \leq l \leq p$, et démontrons la pour $k = l$.

Par conjugaison et récurrence, $s_l^\top H s_i = 0$ et $s_l^\top M_l s_i = s_l^\top H s_i = 0$.

Dès lors, la formule de BFGS (2.36) donne $M_{l+1} s_i = M_l s_i = H s_i$. Si $i = l$, on a $M_{l+1} s_l = H s_l$ (c'est l'équation de quasi-Newton vérifiée par M_{l+1} où $y_l = H s_l$).

Donc on a $M_{k+1} s_i = H s_i$, pour $k = 1, \dots, p$ et $i = 1, \dots, k$.

(ii) Si $p = n$, alors M_{n+1} prend la même valeur que H sur les n vecteurs linéairement indépendants s_1, s_2, \dots, s_n , c'est-à-dire de (i) on peut écrire

$$M_{n+1} s_1 = H s_1$$

$$M_{n+1} s_2 = H s_2$$

⋮

$$M_{n+1}s_n = Hs_n$$

et puisque s_1, s_2, \dots, s_n sont linéairement indépendant, nous avons

$$M_{n+1} = H.$$

■

Propriétés

- La condition $y_k^\top s_k > 0$ assure la défini-positivité de M_k .
- La condition $y_k^\top s_k > 0$ est satisfaite lorsqu' on utilise la recherche linéaire de Wolfe.
- Les directions s_1, s_2, \dots, s_n engendrées par la formule de BFGS sont conjuguées.
- Dans le cas quadratique la méthode de BFGS possède les mêmes propriétés que DFP.
- Dans le cas non quadratique, il faut procéder à des réinitialisations périodiques pour assurer la convergence globale.

2.2.6 Les méthodes de la classe Broyden

Une méthode de classe Broyden est une méthode de quasi-Newton où l'approximation de l'inverse du Hessien prendre la formule suivant :

$$M_{k+1} = M_k - \frac{M_k s_k s_k^\top M_k}{s_k^\top M_k s_k} + \frac{y_k y_k^\top}{y_k^\top s_k} + \phi (s_k^\top M_k s_k) v_k v_k^\top \quad (2.38)$$

avec $\phi \in [0, 1]$, $v_k^\top = \left[\frac{y_k}{y_k^\top s_k} - \frac{M_k s_k}{s_k^\top M_k s_k} \right]$

- ◆ Si $\phi = 0$ on obtient la méthode BFGS.
- ◆ Si $\phi = 1$ on obtient la méthode DFP.
- ◆ Si $\phi = \frac{s_k^\top y_k}{s_k^\top y_k - y_k^\top M_k y_k}$ on obtient la méthode de SR1.

Propriétés des méthodes de la classe de broyden

• Nous pouvons donc récrire (2.38) comme "*combinaison linéaire*" de deux méthodes, c'est-à-dire,

$$M_{k+1} = (1 - \phi)M_{k+1}^{BFGS} + \phi M_{k+1}^{DFP},$$

Cette relation indique que tous les membres de la classe de Broyden satisfont l'équation de quasi-newton (2.4), puisque les matrices de BFGS et de DFP elles-mêmes satisfont cette équation et aussi elle conserve la définie positivité des matrices M_k .

Théorème 2.4. ([1]) Si une méthode de Broyden est appliquée à une fonction quadratique convexe avec une recherche linéaire exacte, elle converge après $m \leq n$ itérations. Les propriétés suivantes s'appliquent pour tout

$k = 0, 1, \dots, m$:

- (a) $s_i = M_{k+1}y_i$ pour $0 \leq i \leq k$.
- (b) $s_i^\top H s_j = 0$ pour $0 \leq i < j \leq k$.
- (c) Si $m = n - 1$, alors $M_m = H^{-1}$.

Chapitre 3

Rôle joué par la convexité dans la convergence de la méthode BFGS

Si on utilise une recherche linéaire exacte et dans le cas où la fonction objectif est convexe, la théorie de la convergence de la méthode de BFGS a été bien établie par les auteurs ([33, 34, 20, 2, 23, 25, 26]).

Le but de cette partie est de voir le rôle joué par la convexité dans la convergence de la méthode BFGS, dans le cas de recherches linéaires inexactes..

Cette partie est divisée en deux sections principales ; La première section sera consacrée aux résultats de convergence pour la méthode BFGS dans le cas convexe. Dans la dernière section, on s'intéresse aux résultats de convergence de cette méthode mais dans le cas non convexe.

3.1 Convergence globale de la méthode BFGS dans le cas convexe

Un des plus beaux résultats de convergence en optimisation sans contrainte est le théorème 3.1 ci-dessous. Il est du à M.J.D. Powell ([25]) en 1976. Il donne des conditions assurant la convergence globale de l'algorithme de BFGS lorsque la fonction à minimiser est convexe. La beauté du résultat tient au fait qu'il est rare de pouvoir montrer la convergence d'un algorithme

CHAPITRE 3. RÔLE JOUÉ PAR LA CONVEXITÉ DANS LA
CONVERGENCE DE LA MÉTHODE BFGS

quasi-newtonien, avec si peu d'hypothèses sur les objets générés par celui-ci. On suppose seulement que l'algorithme génère une suite $\{\mathcal{x}_k\}$ et que la suite $\{f(x_k)\}$ est bornée inférieurement (il existe une constante c telle que pour tout indice k , $f(\mathcal{x}_k) \geq c$). On ne fait aucune hypothèse sur la suite des matrices $\{M_k\}$.

Commençons cette section d'abord par une hypothèse et deux formules qui nous seront utiles pour étudier la convergence de l'algorithme BFGS.

Dans tout ce que suit nous supposons que la fonction f est de classe $C^{1,1}$ (elle est différentiable et sa dérivée vérifie pour \mathcal{L} et pour tout \mathcal{x} et y : $\|\nabla f(y) - \nabla f(\mathcal{x})\| \leq \mathcal{L} \|y - \mathcal{x}\|$). Notons par $H(\mathcal{x})$ la matrice hessienne au point \mathcal{x} , \mathcal{x}_1 est un point initial de l'algorithme. Notons par D l'ensemble suivant :

$$D = \{\mathcal{x} \in IR^n : f(\mathcal{x}) \leq f(\mathcal{x}_1)\}$$

On suppose dans cette section que f est convexe dans D . On supposera vraie l'hypothèse 3.1 suivante :

Hypothèse 3.1

$\exists c, C$ deux constantes positives tels que :

$$c \|z\|^2 \leq z^\top H(\mathcal{x}) z \leq C \|z\|^2 \quad \forall z \in IR^n, \forall \mathcal{x} \in D \quad (3.1)$$

Conséquences

a) Soit \bar{H}_k défini comme suit

$$\bar{H}_k = \int_0^1 \nabla^2 f(\mathcal{x}_k + \tau s_k) d\tau = \int_0^1 H(\mathcal{x}_k + \tau s_k) d\tau,$$

alors nous avons

$$\bar{H}_k = \int_0^1 H(\mathcal{x}_k + \tau s_k) d\tau = \frac{1}{s_k} [g(\mathcal{x}_k + s_k) - g(\mathcal{x}_k)] = \frac{1}{s_k} y_k$$

donc

$$y_k = \bar{H}_k s_k$$

ce qui implique

$$y_k^\top s_k = s_k^\top \bar{H}_k s_k \quad (3.2)$$

3.1. CONVERGENCE GLOBALE DE LA MÉTHODE BFGS DANS LE
CAS CONVEXE

de (3.1) et (3.2) on obtient

$$\frac{y_k^\top s_k}{\|s_k\|^2} = \frac{s_k^\top \bar{H}_k s_k}{\|s_k\|^2} \geq c > 0 \quad (3.3)$$

De (3.3) l'hypothèse 3.1 implique que \bar{H}_k soit définie positive et car \bar{H}_k est une matrice symétrique (puisque le hessien est symétrique), alors \bar{H}_k admet une unique racine carrée $\bar{H}_k^{\frac{1}{2}}$ telle que

$$\bar{H}_k = \bar{H}_k^{\frac{1}{2}} \bar{H}_k^{\frac{1}{2}}$$

par conséquent, si on définit $z_k = \bar{H}_k^{\frac{1}{2}} s_k$ alors de (3.2) et (3.1) on obtient

$$\frac{y_k^\top s_k}{\|y_k\|^2} = \frac{\left(\bar{H}_k^{\frac{1}{2}} s_k\right)^\top \bar{H}_k^{\frac{1}{2}} s_k}{\left(\bar{H}_k^{\frac{1}{2}} s_k\right)^\top \bar{H}_k \bar{H}_k^{\frac{1}{2}} s_k} = \frac{\|z_k\|^2}{z_k^\top \bar{H}_k z_k} \geq \frac{1}{C}$$

ce qui implique

$$\frac{\|y_k\|^2}{y_k^\top s_k} \leq C \quad (3.4)$$

Proposition 3.1 *Si M_{k+1} et M_k sont reliées par la formule de BFGS (2.36) avec M_k définie positive et $y_k^\top s_k > 0$, on a*

$$Tr(M_{k+1}) = Tr(M_k) + \frac{\|y_k\|^2}{y_k^\top s_k} - \frac{\|M_k s_k\|^2}{s_k^\top M_k s_k} \quad (3.5)$$

et

$$\det(M_{k+1}) = \det(M_k) \cdot \left(\frac{y_k^\top s_k}{s_k^\top M_k s_k}\right) \quad (3.6)$$

Preuve. ([18])

CHAPITRE 3. RÔLE JOUÉ PAR LA CONVEXITÉ DANS LA
CONVERGENCE DE LA MÉTHODE BFGS

La formule de la trace se déduit directement de la relation BFGS (2.36) suivante :

$$M_{k+1} = M_k + \frac{\|y_k\|^2}{y_k^\top s_k} - \frac{\|M_k s_k\|^2}{s_k^\top M_k s_k} \quad (BFGS),$$

celle du déterminant utilise la formule suivante :

$$\det(M_{k+1}) = \det(M_k) \cdot \det\left(I + \frac{M_k^{-1} y_k y_k^\top}{y_k^\top s_k} - \frac{s_k s_k^\top M_k}{s_k^\top M_k s_k}\right)$$

avec $M_k^{-1} y_k = s_k$, implique

$$\det(M_{k+1}) = \det(M_k) \cdot \det\left(\frac{y_k^\top s_k s_k^\top M_k s_k + s_k y_k^\top s_k^\top M_k s_k - y_k^\top s_k s_k s_k^\top M_k}{y_k^\top s_k s_k^\top M_k s_k}\right)$$

implique

$$\det(M_{k+1}) = \det(M_k) \cdot \det\left(\frac{y_k^\top s_k y_k^\top s_k + s_k y_k^\top y_k^\top s_k - y_k^\top s_k s_k y_k^\top}{y_k^\top s_k s_k^\top M_k s_k}\right)$$

donc on trouve

$$\det(M_{k+1}) = \det(M_k) \cdot \left(\frac{y_k^\top s_k}{s_k^\top M_k s_k}\right).$$

■

Théorème 3.1 ([22]) *Supposons que $f : IR^n \rightarrow IR$ soit convexe et $C^{1,1}$ dans un voisinage convexe de $D = \{\varkappa \in IR^n : f(\varkappa) \leq f(\varkappa_1)\}$, où $\varkappa_1 \in IR^n$. On considère l'algorithme de BFGS démarrant en \varkappa_1 avec une matrice M_1 symétrique définie positive et on suppose que celui-ci génère une suite $\{\varkappa_k\}$ telle que $\{f(\varkappa_k)\}$ soit bornée inférieurement. Alors, $\liminf_{k \rightarrow \infty} \|g_k\| = 0$.*

Preuve. ([22])

On note par $c_1, c_2, c_3, c_4, c_5, c_6, c_7$ des constantes strictement positives. L'idée de la démonstration est de contrôler le comportement des matrices $\{M_k\}$ en obtenant une majoration de leur trace et une minoration de leur déterminant, ce qui permettra d'avoir une borne inférieure sur le pas λ_k . On note

$$q_k := \frac{s_k^\top M_k s_k}{\|s_k\|^2} \text{ et } \cos \theta_k = \frac{-g_k^\top d_k}{\|g_k\| \|d_k\|} = \frac{s_k^\top M_k s_k}{\|M_k s_k\| \|s_k\|}$$

3.1. CONVERGENCE GLOBALE DE LA MÉTHODE BFGS DANS LE CAS CONVEXE

Le quotient de Rayleigh de M_k dans la direction s_k et le cosinus de l'angle entre d_k et l'opposé du gradient.

✂ **Premier étape :**

On utilise la formule de la trace pour montrer que pour tout $k \geq 1$:

$$\text{Tr}(M_{k+1}) \leq c_1 k \text{ et } \sum_{i=1}^{i=k} \frac{\|M_i s_i\|_2^2}{s_i^\top M_i s_i} = \sum_{i=1}^{i=k} \frac{q_i}{\cos \theta_i} \leq c_1 k \quad (3.7)$$

En effet :

Comme $f \in C^{1,1}$, de (3.4) on a

$$\frac{\|y_k\|^2}{y_k^\top s_k} \leq c_2, \quad c_2 > 0,$$

et d'après la proposition (3.1) nous avons

$$\begin{aligned} \text{Tr}(M_{k+1}) &\leq \text{Tr}(M_k) + c_2 - \frac{\|M_k s_k\|^2}{s_k^\top M_k s_k} \\ &\leq \text{Tr}(M_1) + k c_2 - \sum_{i=1}^k \frac{\|M_i s_i\|^2}{s_i^\top M_i s_i} \end{aligned}$$

On obtient la première inégalité de (3.7) en posant

$$c_1 = \text{Tr}(M_1) + c_2 \leq k c_1.$$

D'autre part les matrices M_i étant définies positives, alors $s_i^\top M_i s_i > 0$, donc

$$\sum_{i=1}^{i=k} \frac{\|M_i s_i\|^2}{s_i^\top M_i s_i} \geq 0$$

D'autre part nous avons

$$q_k \|s_k\|^2 = s_k^\top M_k s_k \text{ et } \|M_k s_k\| \|s_k\| \cos \theta_k = s_k^\top M_k s_k$$

d'où

$$\frac{\|M_k s_k\|^2}{s_k^\top M_k s_k} = \frac{s_k^\top M_k s_k}{\|s_k\|^2 \cos^2 \theta_k} = \frac{q_k}{\cos^2 \theta_k}$$

donc

$$\sum_{i=1}^{i=k} \frac{\|M_i s_i\|^2}{s_i^\top M_i s_i} = \sum_{i=1}^{i=k} \frac{q_i}{\cos^2 \theta_i} \leq c_1 k$$

CHAPITRE 3. RÔLE JOUÉ PAR LA CONVEXITÉ DANS LA
CONVERGENCE DE LA MÉTHODE BFGS

car :

M_{k+1} est définie positive implique que, toutes les valeurs propres de M_{k+1} sont strictement positives

c'est-à-dire $\text{Tr}(M_{k+1}) > 0$ et Comme $\sum_{i=1}^{i=k} \frac{\|M_i s_i\|^2}{s_i^\top M_i s_i} \geq 0$

donc

$$\text{Tr}(M_{k+1}) + \sum_{i=1}^k \frac{\|M_i s_i\|^2}{s_i^\top M_i s_i} \leq c_1 k \Rightarrow \sum_{i=1}^k \frac{\|M_i s_i\|^2}{s_i^\top M_i s_i} \leq c_1 k$$

✂ *Deuxième étape :*

En utilisant la majoration de la trace (3.7) et une minoration du déterminant, On obtient une minoration de la plus petite valeur propre de M_{k+1} , donc une minoration du pas.

On montre en fait que cela conduit à la borne inférieure suivante sur le pas moyen : pour tout $k \geq 1$,

$$\left(\prod_{i=1}^k \lambda_i \right)^{\frac{1}{k}} \geq c_3, \quad c_3 > 0, \quad (3.8)$$

En effet : Par la formule de déterminant (3.6) et la première condition de wolfe nous avons

$$y_k^\top s_k \geq (1 - \sigma_2) \frac{s_k^\top M_k s_k}{\lambda_k},$$

ou encore

$$\frac{y_k^\top s_k}{s_k^\top M_k s_k} \geq \frac{1 - \sigma_2}{\lambda_k},$$

3.1. CONVERGENCE GLOBALE DE LA MÉTHODE BFGS DANS LE CAS CONVEXE

donc

$$\begin{aligned}
 \det(M_{k+1}) &= \frac{y_k^\top s_k}{s_k^\top M_k s_k} \det(M_k) \\
 &\geq \frac{(1 - \sigma_2)}{\lambda_k} \det(M_k) \\
 &\geq \frac{(1 - \sigma_2)^2}{\lambda_k \lambda_{k-1}} \det(M_{k-1}) \\
 &\geq \frac{(1 - \sigma_2)^k}{\prod_{i=1}^k \lambda_i} \det(M_1) \\
 &\geq \frac{c_4^k}{\prod_{i=1}^k \lambda_i}
 \end{aligned}$$

où $c_4 = (1 - \sigma_2) \min(1, \det M_1)$.

D'autre part, en utilisant l'inégalité des moyennes (3.7) et le fait que $k^n \leq (e^n)^k$ (où $k^n = e^{n \ln k}$ et $\ln k \leq k$) $\Rightarrow k^n \leq (e^n)^k$, on a

$$\det(M_{k+1}) \leq \left(\frac{1}{n} \text{Tr}(M_{k+1}) \right)^n \leq \left(\frac{c_1}{n} k \right)^n \leq \left(\frac{c_1}{n} \right)^n (e^n)^k.$$

Maintenant posons $c_5 = \max\left(\frac{c_1}{n}, 1\right)^n e^n$, on a :

$$\det(M_{k+1}) \leq c_5^k$$

Finalement on a

$$\frac{1}{\det(M_{k+1})} \geq \frac{1}{c_5^k}.$$

D'autre part

$$\prod_{i=1}^k \lambda_i \geq \frac{c_4^k}{\det(M_{k+1})} \geq \frac{c_4^k}{c_5^k} = \left(\frac{c_4}{c_5} \right)^k$$

d'où

$$\left(\prod_{i=1}^k \lambda_i \right)^{\frac{1}{k}} \geq \frac{c_4}{c_5}.$$

Si on pose $c_3 = \frac{c_4}{c_5}$ alors :

$$\left(\prod_{i=1}^k \lambda_i \right)^{\frac{1}{k}} \geq c_3$$

CHAPITRE 3. RÔLE JOUÉ PAR LA CONVEXITÉ DANS LA
CONVERGENCE DE LA MÉTHODE BFGS

✠ *Troisième étape :*

En exploitant la deuxième estimation de (3.7), on obtient une majoration du pas moyen pondéré. On se défera de la pondération dans le raisonnement par l'absurde de la quatrième étape, pour tout $k \geq 1$

$$\left(\prod_{i=1}^k \lambda_i \|g_i\|_2^2 \right)^{\frac{1}{k}} \leq \frac{c_6}{k}, \quad c_6 > 0 \quad (3.9)$$

En effet :

On a :

$$\sum_{i=1}^k \frac{\|M_i s_i\|_2^2}{s_i^\top M_i s_i} = \sum_{i=1}^k \frac{\lambda_i \|g_i\|_2^2}{-g_i^\top s_i} \leq c_1.$$

On utilise ensuite deux fois l'inégalité des moyennes et la première condition de Wolfe

donc

$$\lambda_i \|g_i\|_2^2 \leq c_1 (-g_i^\top s_i)$$

implique

$$\begin{aligned} \prod_{i=1}^k \lambda_i \|g_i\|_2^2 &\leq c_1^k \prod_{i=1}^k (-g_i^\top s_i) \\ &\leq \left(\frac{c_1}{k} \sum_{i=1}^k (-g_i^\top s_i) \right)^k \\ &\leq \left(\frac{c_1}{\sigma_1 k} \sum_{i=1}^k (f(\mathcal{x}_i) - f(\mathcal{x}_{i+1})) \right)^k \\ &\leq \left(\frac{c_1}{\sigma_1 k} (f(\mathcal{x}_1) - f(\mathcal{x}_{k+1})) \right)^k \\ &\leq \frac{c_6}{k} \end{aligned}$$

On en déduit (3.9), avec $c_6 = c_1 (f(\mathcal{x}_1) - f_{\min}) / \sigma_1$ où $f_{\min} \in \mathbb{R}$ est un minorant de $\{f(\mathcal{x}_k)\}$.

✠ *Quatrième étape :*

On conclut par un raisonnement par l'absurde. Supposons que le résultat ne soit pas vrai, alors

$$\|g_k\| > c_7 > 0$$

et (3.9) donne

$$\left(\prod_{i=1}^k \lambda_i \right)^{\frac{1}{k}} \leq \frac{c_6}{k c_7^2}$$

qui est contradiction avec (3.8) donc $\liminf_{k \rightarrow \infty} \|g_k\| = 0$. ■

3.2 Convergence globale de la méthode BFGS dans le cas non convexe

Comme nous l'avons signalé au paragraphe précédent, l'étude de la convergence globale de la méthode BFGS a bien progressé dans le cas où la fonction économique est convexe. De façon plus précise, Il a été démontré (voir ([25]), ([18])), que dans le cas où la fonction économique f est convexe et dans le cas de recherches linéaires exactes ou inexates, alors les itérés $\{x_k\}$ générés par l'algorithme BFGS, sont globalement convergentes. Malheureusement dans le cas où la fonction économique f n'est pas convexe, on a peu de résultats de convergence sur la méthode BFGS jusqu'à présent. Non seulement personne n'a pu prouver la convergence globale dans le cas non convexe, mais aussi personne n'a pu trouver un contre exemple qui prouve le contraire.

Récemment, Dong-Hui Li et Masao Fukushima ([8]) ont proposé une méthode modifiée BFGS modifiée qu'ils ont appelée Méthode BFGS avec mise à jour appropriée et ont établi sa convergence globale pour des problèmes d'optimisation non convexes.

Le but de cette partie est mettre toute la lumière sur cette étude et d'essayer de voir pourquoi il est difficile d'aborder le cas non convexe. Des problèmes ouverts et des conjectures seront formulés à la fin du chapitre.

3.2.1 Méthode BFGS avec mise à jour appropriée

On a vu qu'une des propriétés de la formule de BFGS (2.36) est que la matrice M_{k+1} hérite de la définie positivité de M_k si la condition $y_k^\top s_k > 0$ est vérifiée. Remarquons que si on utilise une recherche linéaire exacte ou une recherche linéaire inexacte de Wolfe, alors la condition $y_k^\top s_k > 0$ est vérifiée. Par contre, la recherche linéaire d'Armijo n'assure pas cette condition, et par conséquent M_{k+1} n'est pas nécessairement définie positive même si M_k est définie positive. Afin d'assurer la définie positivité de M_{k+1} , la condition $y_k^\top s_k > 0$ est parfois utilisée pour décider si M_{k+1} est une mise à jour ou non, c'est-à-dire on pose

$$M_{k+1} = \begin{cases} M_k + \frac{y_k y_k^\top}{y_k^\top s_k} - \frac{M_k s_k s_k^\top M_k}{s_k^\top M_k s_k}, & \text{si } y_k^\top s_k > 0 \\ M_k & \text{, ailleurs} \end{cases} \quad (3.10)$$

La condition $y_k^\top s_k > 0$ est souvent remplacée par $y_k^\top s_k > \eta$ où $\eta > 0$ est une petite constante. Dans ([8]), Dong-Hui Li et Masao Fukushima proposent une mise à jour appropriée à la méthode BFGS semblable à ce qui précède et patant de celà ils établissent un théorème de convergence globale pour des problèmes non convexes. Avant de décrire cette mise à jour appropriée, voyons d'abord un lemme important dû à Powell ([22]) et qui nous sera utile par la suite.

Lemme 3.1 (Powell ([22])) *Si la méthode de BFGS avec la recherche linéaire de Wolfe (wolfe1)-(wolfe2) est appliquée à une fonction f continument différentiable; et s'il existe une constante $c > 0$ telle que :*

$$\frac{\|y_k\|^2}{y_k^\top s_k} \leq c \quad \text{pour tout } k \quad (3.11)$$

Alors

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0. \quad \square \quad (3.12)$$

Preuve. (Powell ([22]))

Avec les expressions (3.5) et (3.6), nous déduisons une limite sur le produit

$$\prod_{j=1}^k \frac{\|g_j\|}{\|s_j\| \cos \theta_j}$$

où $\theta_j \in]0, \frac{\pi}{2}[$ est l'angle entre la direction de recherche d_j et la direction de descente $-g_j$. Cette limite montre une contradiction si la limite (3.12) n'est pas obtenue. De (2.36) et (3.11) on peut écrire

$$\begin{aligned} \text{Tr}(M_{k+1}) &< \text{Tr}(M_1) + \sum_{i=1}^k \frac{\|y_i\|^2}{y_i^\top s_i} \\ &\leq \text{Tr}(M_1) + ck \end{aligned}$$

3.2. CONVERGENCE GLOBALE DE LA MÉTHODE BFGS DANS LE CAS NON CONVEXE

$$\leq c_1 k, \quad k = 1, 2, 3, \dots \quad (3.13')$$

où c_1 est la constante $\{Tr(M_1) + c\}$. Puisque la matrice M_{k+1} est définie positive, nous pouvons appliquer l'inégalité moyenne géométrique/arithmétique à ses valeurs propres pour obtenir la relation

$$\det(M_{k+1}) < (c_1 k/n)^n, \quad k = 1, 2, 3, \dots \quad (3.14')$$

Puisque la trace de M_{k+1} est positive, nous pouvons également déduire de l'équation (2.36) l'inégalité

$$\begin{aligned} \sum_{i=1}^k \frac{\|M_i s_j\|^2}{s_j^\top M_i s_j} &< Tr(M_1) + \sum_{i=1}^k \frac{\|y_i\|^2}{y_j^\top s_j} \\ &\leq c_1 k, \quad k = 1, 2, 3, \dots \end{aligned} \quad (3.15')$$

Encore nous appliquons l'inégalité moyenne géométrique/arithmétique, et trouvons ainsi la limite

$$\prod_{j=1}^k \frac{\|M_i s_j\|^2}{s_j^\top M_i s_j} < c_1^k, \quad k = 1, 2, 3, \dots \quad (3.16')$$

D'ailleurs, parce que la formule (2.36) donne la relation de récurrence

$$\det(M_{k+1}) = \frac{y_k^\top s_k \det(M_k)}{s_k^\top M_k s_k} \quad (3.17')$$

(Pearson, 1969), nous avons l'équation

$$\prod_{j=1}^k \frac{y_j^\top s_j}{s_j^\top M_i s_j} = \frac{\det(M_{k+1})}{\det(M_1)} \quad (3.18')$$

Pour obtenir l'inégalité principale de la preuve, nous multiplions les côtés gauches et droits des inégalités (3.14'), (3.16') et (3.18') puis décommandant le terme $\det(M_{k+1})$, qui donne l'expression

$$\begin{aligned} \prod_{j=1}^k \frac{\|M_i s_j\|^2 y_j^\top s_j}{(s_j^\top M_i s_j)^2} &< \frac{(c_1 k/n)^n c_1^k}{\det(M_1)} \\ &\leq c_2^k, \quad k = 1, 2, 3, \dots \end{aligned} \quad (3.19')$$

CHAPITRE 3. RÔLE JOUÉ PAR LA CONVEXITÉ DANS LA
CONVERGENCE DE LA MÉTHODE BFGS

où c_2 est une constante appropriée. Nous éliminons M_i du Côté gauche. Egalement nous enlevons le terme $y_j^\top s_j$ du dénominateur, par la déduction de l'expression (wolfe2) on a

$$\frac{y_j^\top s_j}{-g_j^\top s_j} \geq 1 - \sigma \quad (3.20')$$

Ainsi la relation (3.19') donne la condition

$$\prod_{j=1}^k \frac{\|g_j\|^2}{-g_j^\top s_j} < \left(\frac{c_2}{1 - \sigma} \right)^k \quad (3.21')$$

Par conséquent l'inégalité

$$\prod_{j=1}^k \frac{\|g_j\|}{\|s_j\| \cos \theta_j} < c_3^k, \quad k = 1, 2, 3, \dots, \quad (3.22')$$

est satisfaisante, où c_3 est la constante $c_2/1 - \sigma$.

Puisque $f(x)$ est inférieurement bornée, la condition (wolfe1) de Wolfe implique que la somme

$$\sum_{k=1}^{\infty} -\lambda_k g_k^\top d_k = \sum_{k=1}^{\infty} \|g_k\| \|s_k\| \cos \theta_k \quad (3.23')$$

est convergente. Par conséquent, si $\|g_k\|$ est bornée à partir de zéro, nous avons la limite

$$\|s_k\| \cos \theta_k \rightarrow 0 \quad (3.24')$$

Dans ce cas le terme $\|g_j\| / (\|s_j\| \cos \theta_j)$ diverge à l'infini quand j tend à l'infini. Par conséquent la limite (3.22') donne une contradiction. Elle suit que la limite (3.12) est obtenue, qui accomplit la preuve du lemme 3.1. ■

Remarque 3.1 :

Si f est deux fois continument différentiable et strictement convexe, alors on obtient toujours l'inégalité (3.11) (voir section 3.1), mais, dans le cas où f est non convexe, il est difficile de garantir l'inégalité (3.11). Peut être c'est une des raisons pour laquelle la convergence globale de la méthode de BFGS n'a pas été prouvée.

Nous allons maintenant présenter la méthode BFGS avec mise à jour appropriée et montrer plus tard qu'elle est globalement convergente sans que

3.2. CONVERGENCE GLOBALE DE LA MÉTHODE BFGS DANS LE CAS NON CONVEXE

la fonction économique soit convexe. Pour être plus précis, déterminons M_{k+1} en fonction de M_k par :

$$M_{k+1} = \begin{cases} M_k + \frac{y_k y_k^\top}{y_k^\top s_k} - \frac{M_k s_k s_k^\top M_k}{s_k^\top M_k s_k}, & \text{si } \frac{y_k^\top s_k}{\|s_k\|^2} \geq \varepsilon \|g_k\|^\alpha \\ M_k, & \text{sinon} \end{cases} \quad (3.13)$$

où ε et α sont des constantes positives.

Algorithme 3.1 (Méthode BFGS avec mise à jour appropriée)

Étape 0 choisir un point initial $\varkappa_0 \in IR^n$ et une matrice initiale

$M_0 \in IR^{n \times n}$ symétrique, définie positive; choisir les constantes

$0 < \sigma_1 < \sigma_2 > 1$, $\alpha > 0$ et $\varepsilon > 0$.

soit $k = 0$

Étape 1 résoudre l'équation linéaire $M_k d_k + g_k = 0$ pour obtenir d_k .

Étape 2 déterminer le pas $\lambda_k > 0$ par la recherche linéaire inexacte de Wolfe ou d'Armijo.

Étape 3 Calculer $\varkappa_{k+1} := \varkappa_k + \lambda_k d_k$.

Étape 4 déterminer M_{k+1} par

$$M_{k+1} = \begin{cases} M_k + \frac{y_k y_k^\top}{y_k^\top s_k} - \frac{M_k s_k s_k^\top M_k}{s_k^\top M_k s_k}, & \text{si } \frac{y_k^\top s_k}{\|s_k\|^2} \geq \varepsilon \|g_k\|^\alpha \\ M_k, & \text{sinon} \end{cases}$$

avec

$$\begin{aligned} s_k &= \varkappa_{k+1} - \varkappa_k \\ y_k &= \nabla f(\varkappa_{k+1}) - \nabla f(\varkappa_k) \end{aligned}$$

Étape 5 Remplacer k par $k + 1$ et aller à l'étape 1. □

Remarque 3.2 il n'est pas difficile de voir que les matrices M_k générés par l'algorithme 3.1 sont symétriques et défini-positives pour tout k . Ceci implique qu'avec l'utilisation d'une recherche linéaire inexacte de Wolfe ou d'Armijo, qu' on obtienne que la suite $\{f(\varkappa_k)\}_{k \in \mathbb{N}}$ est décroissante.

Nous avons aussi en prenant en considération (wolfe1) ou (wolfe1)-(wolfe2) et si f est bornée inférieurement :

$$-\sum_{k=0}^{\infty} g_k^\top s_k < \infty. \quad (3.14)$$

Ceci implique que

$$-\lim_{k \rightarrow \infty} (-g_k^\top s_k) = 0 \quad (3.15)$$

et puisque

$$s_k = \varkappa_{k+1} - \varkappa_k = \lambda_k d_k.$$

Alors

$$-\lim_{k \rightarrow \infty} (\lambda_k g_k^\top d_k) = 0$$

3.2.2 Convergence globale de Méthode BFGS avec mise à jour appropriée

Dans cette section, nous prouvons la convergence globale de l'algorithme 3.1 sous l'hypothèse suivante :

Hypothèse 3.2

Considérons l'ensemble Ω suivant :

$$\Omega = \{\varkappa \in IR^n / f(\varkappa) \leq f(\varkappa_0)\}.$$

Nous supposons que Ω est contenu dans un ensemble convexe et bornée D et que la fonction économique f est continument différentiable sur D et qu'il existe une constante $l > 0$ telle que :

$$\|g(\varkappa) - g(y)\| \leq l \|\varkappa - y\|, \quad \forall \varkappa, y \in D \quad (3.16)$$

c'est-à-dire $f \in C^{1,1}(D)$.

Comme la suite $\{f(\varkappa_k)\}$ est décroissante, il est clair que la suite $\{\varkappa_k\}$ générée par l'algorithme 3.1 est contenue dans Ω (car $f(\varkappa_{k+1}) < f(\varkappa_k) < \dots < f(\varkappa_1) < f(\varkappa_0)$)

Pour des raisons de convenance et pour être plus pratique, définissons les ensembles d'indices suivants :

$$\bar{K} = \left\{ i \mid \frac{y_i^\top s_i}{\|s_i\|^2} \geq \varepsilon \|g_i\|^\alpha \right\} \text{ et } \bar{K}_k = \left\{ i \in \bar{K} \mid i \leq k \right\}. \quad (3.17)$$

3.2. CONVERGENCE GLOBALE DE LA MÉTHODE BFGS DANS LE CAS NON CONVEXE

Notons par i_k , l'ensemble des indices $i \in \bar{K}_k$.

Au moyen de \bar{K} , on peut récrire (3.13) sous la forme :

$$M_{k+1} = \begin{cases} M_k + \frac{y_k y_k^\top}{y_k^\top s_k} - \frac{M_k s_k s_k^\top M_k}{s_k^\top M_k s_k}, & \text{si } k \in \bar{K} \\ M_k & , \text{ sinon} \end{cases} \quad (3.18)$$

et en considérant la trace des deux membres de (3.18), on peut écrire pour n'importe quel k

$$Tr(M_{k+1}) = Tr(M_1) + \sum_{i \in \bar{K}_k} \frac{\|y_i\|^2}{y_i^\top s_i} - \sum_{i \in \bar{K}_k} \frac{\|M_i s_i\|^2}{s_i^\top M_i s_i} \quad (3.19)$$

Maintenant, nous sommes en mesure de prouver la convergence globale de l'algorithme **3.1**. Tout d'abord nous prouvons la convergence globale de cet algorithme dans le cas où \bar{K} est un ensemble fini :

Théorème 3.2 ([11]) *On suppose vraie l'hypothèse 3.2 et soit $\{\varkappa_k\}$ la suite générée par l'algorithme 3.1.*

Si \bar{K} est un ensemble fini, alors

$$\lim_{k \rightarrow \infty} \|g_k\| = 0. \quad (3.20)$$

Preuve. ([11])

si \bar{K} est un ensemble fini alors il existe un indice k_0 tel que

$$M_k = M_{k_0} \triangleq M \quad \text{pour tout } k \geq k_0$$

Par la défini- positivité de M , il existe des constantes positives $c_1 \leq C_1$ telles que

$$\begin{aligned} c_1 \|d\|^2 &\leq d^\top M d \leq C_1 \|d\|^2, \\ c_1 \|d\|^2 &\leq d^\top M^{-1} d \leq C_1 \|d\|^2 \quad \forall d \in IR^n \end{aligned} \quad (3.21)$$

✠ *Si la recherche linéaire inexacte utilisée est de Wolfe, on obtient de (3.16) et (wole2)*

$$l \|s_k\|^2 \stackrel{(3.16)}{\geq} \|(g^\top(\varkappa_{k+1}) - g^\top(\varkappa_k)) s_k\| \geq y_k^\top s_k$$

CHAPITRE 3. RÔLE JOUÉ PAR LA CONVEXITÉ DANS LA
CONVERGENCE DE LA MÉTHODE BFGS

$$\begin{aligned}
 & \stackrel{(1.12)}{\geq} \sigma_2 g_k^\top s_k - g_k^\top s_k \quad (\text{où } s_k = \lambda_k d_k) \\
 & \geq (1 - \sigma_2) \lambda_k^\top s_k^\top M^{-1} s_k \quad (\text{où } d_k^\top = \lambda_k^{-1} s_k^\top \text{ et } g_k^\top = -d_k^\top M^{-1}) \\
 & \stackrel{(3.21)}{\geq} (1 - \sigma_2) \lambda_k^\top c_1 \|s_k\|^2, \quad \forall k \geq k_0
 \end{aligned}$$

ceci implique :

$$\lambda_k \geq (1 - \sigma_2) c_1 l^{-1}, \quad \forall k \geq k_0$$

par conséquent on obtient de (3.15)

$$\lim_{k \rightarrow \infty} \lambda_k g^\top d_k = \lim_{k \rightarrow \infty} (-g_k^\top s_k) = 0$$

avec $\lambda_k \geq (1 - \sigma_2) c_1 l^{-1} > 0$, $\forall k \geq k_0$, on peut écrire

$$g_k^\top M^{-1} g_k = -g_k^\top d_k \rightarrow 0$$

et de (3.21) on a :

$$c_1 \|g_k\|^2 \leq g_k^\top M^{-1} g_k \rightarrow 0$$

c'est-à-dire

$$c_1 \|g_k\|^2 \rightarrow 0$$

avec $c_1 \neq 0$, donc

$$\lim_{k \rightarrow \infty} \|g_k\| = 0. \quad \square$$

✠ Si la recherche linéaire inexacte utilisée est d'Armijo. Notons par :

$$\bar{\lambda} = \limsup_{k \rightarrow \infty} \lambda_k$$

- Si $\bar{\lambda} > 0$, alors en utilisant un argument semblable à ce qui précède, on retrouve (3.20).

- Supposons maintenant que $\bar{\lambda} = 0$, ceci signifie que :

$$\lim_{k \rightarrow \infty} \lambda_k = 0.$$

Soit \bar{x} un point d'accumulation arbitraire de la suite $\{x_k\}$ et $\{x_k\}_{k \in K}$ une sous-suite convergeant vers \bar{x} . Comme $d_k = -M^{-1}g_k$ pour tout $k \geq k_0$, alors on peut écrire :

$$\{d_k\}_{k \in K} \xrightarrow[k \rightarrow \infty]{} \bar{d} \triangleq -M^{-1}g(\bar{x}).$$

3.2. CONVERGENCE GLOBALE DE LA MÉTHODE BFGS DANS LE CAS NON CONVEXE

D'après la règle d'Armijo, on obtient pour k suffisamment grand,

$$\lambda'_k \triangleq \lambda_k / \rho$$

ne satisfait pas (wolfe1). Par conséquent on a :

$$f(\boldsymbol{x}_k + \lambda'_k d_k) - f(\boldsymbol{x}_k) - \sigma \lambda'_k g^\top(\boldsymbol{x}_k) d_k \geq 0 \quad (3.22)$$

Divisons maintenant les deux membres de (3.22) par λ'_k puis passons à la limite quand $k \rightarrow \infty$ ($\lambda'_k \rightarrow 0$), on obtient

$$\underbrace{\frac{f(\boldsymbol{x}_k + \lambda'_k d_k) - f(\boldsymbol{x}_k)}{\lambda'_k}}_{\text{La dérivée directionnelle}} - \sigma g^\top(\boldsymbol{x}_k) d_k \geq 0 \stackrel{(\lambda'_k \rightarrow 0)}{\Rightarrow} g^\top(\bar{\boldsymbol{x}}) \bar{d} - \sigma g^\top(\bar{\boldsymbol{x}}) \bar{d} \geq 0$$

Ceci implique que

$$(1 - \sigma) g^\top(\bar{\boldsymbol{x}}) \bar{d} \geq 0$$

Ce qui donne aussi

$$g^\top(\bar{\boldsymbol{x}}) \bar{d} \geq 0 \quad (\text{car } \sigma \in (0, 1))$$

et finalement on obtient :

$$-g^\top(\bar{\boldsymbol{x}}) M^{-1} g(\bar{\boldsymbol{x}}) \geq 0 \quad (\text{avec } \bar{d} \triangleq -M^{-1} g(\bar{\boldsymbol{x}})).$$

Comme M^{-1} est définie positive (car M est définie positive), on obtient

$$g(\bar{\boldsymbol{x}}) = 0$$

et la preuve est complète. \square ■

Maintenant nous allons prouver la convergence globale de l'algorithme **3.1** dans le cas où \bar{K} est un ensemble infini.

Si on suppose le contraire, alors il existe une constante $\delta > 0$ telle que

$$\|g_k\| \geq \delta \quad \text{pour tout } k \quad (3.23)$$

et nous verrons que cela produit une contradiction. Avant d'établir le théorème de convergence globale de l'algorithme **3.1**, nous montrons d'abord quelques lemmes utiles.

Lemme 3.2 ([11]) *Supposons que l'hypothèse 3.2 soit vraie et que $\{\varkappa_k\}$ est une suite générée par l'algorithme 3.1.*

Supposons aussi que la relation (3.23) soit vraie pour tout k . alors il existe une constante $c_2 > 0$ telle que :

$$\text{Tr}(M_{k+1}) \leq c_2 i_k \quad (3.24)$$

et

$$\sum_{i \in \bar{K}_k} \frac{\|M_i s_i\|^2}{s_i^\top M_i s_i} \leq c_2 i_k \quad (3.25)$$

pour k suffisamment grand.

Preuve. ([11])

de (3.17) et (3.23) on a

$$y_i^\top s_i \stackrel{(3.17)}{\geq} \varepsilon \|g_i\|^\alpha \|s_i\|^2 \stackrel{(3.23)}{\geq} \varepsilon \delta^\alpha \|s_i\|^2 \quad \text{pour tout } i \in \bar{K}$$

implique

$$y_i^\top s_i \geq \varepsilon \delta^\alpha \|s_i\|^2 \quad \text{pour tout } i \in \bar{K} \quad (3.26)$$

Ceci avec (3.16) donne pour tout $i \in \bar{K}$

$$\|g_{i+1} - g_i\| \leq l \|\varkappa_{i+1} - \varkappa_i\|$$

c'est-à-dire

$$\|y_i\|^2 \leq l^2 \|s_i\|^2$$

et quand

$$\frac{1}{y_i^\top s_i} \leq \frac{1}{\varepsilon \delta^\alpha \|s_i\|^2}$$

on obtient

$$\frac{\|y_i\|^2}{y_i^\top s_i} \leq \frac{l^2}{\varepsilon \delta^\alpha} \triangleq c'_2 \quad (3.27)$$

Ceci avec la proposition (3.1) donne

$$\begin{aligned} \text{Tr}(M_{k+1}) &= \text{Tr}(M_1) + \sum_{i \in \bar{K}} \frac{\|y_i\|^2}{y_i^\top s_i} - \underbrace{\sum_{i \in \bar{K}} \frac{\|M_i s_i\|^2}{s_i^\top M_i s_i}}_{\text{terme positive}} \leq \text{Tr}(M_1) + \sum_{i \in \bar{K}} \frac{\|y_i\|^2}{y_i^\top s_i} \\ &\leq i_k (c_0 + c'_2) = i_k c_2 \end{aligned}$$

3.2. CONVERGENCE GLOBALE DE LA MÉTHODE BFGS DANS LE CAS NON CONVEXE

où on pose $c_2 = \max(c_0, c'_2)$. On obtient donc :

$$\text{Tr}(M_{k+1}) \leq i_k c_2.$$

Puisque $\text{Tr}(M_{k+1}) > 0$ pour n'importe quel k , on obtient, de (3.19) et (3.27),

$$0 < \text{Tr}(M_1) + \sum_{i \in \bar{K}_k} \frac{\|y_i\|^2}{y_i^\top s_i} - \sum_{i \in \bar{K}_k} \frac{\|M_i s_i\|^2}{s_i^\top M_i s_i}$$

ceci implique

$$\sum_{i \in \bar{K}_k} \frac{\|M_i s_i\|^2}{s_i^\top M_i s_i} < \text{Tr}(M_1) + i_k c'_2 < i_k c_2$$

c'est-à-dire

$$\sum_{i \in \bar{K}_k} \frac{\|M_i s_i\|^2}{s_i^\top M_i s_i} \leq c_2 i_k.$$

■

Lemme 3.3 ([11]) *On suppose vraie l'hypothèse 3.2. Si on a (3.23) pour tout k , alors il existe des constantes positives β_1, β_2 et β_3 tels que pour tout $k > 1$, il y a au moins $[i_k/2]$ indices $i \in \bar{K}_k$ tels que :*

$$\|M_i s_i\| \leq \beta_1 \|s_i\| \quad \text{et} \quad \beta_2 \|s_i\|^2 \leq s_i^\top M_i s_i \leq \beta_3 \|s_i\|^2 \quad (3.29)$$

Preuve. Remarquons que (3.23) implique que (3.26) et (3.27) ont lieu pour tout $i \in \bar{K}$.

Par conséquent, en procédant de manière semblable à la preuve du théorème 2.1 dans ([33]), on peut conclure que (3.29) a lieu.

En effet :

de (2.33), (3.5) et (3.6) on a

$$\begin{aligned} \psi(M_{k+1}) &= \psi(M_k) + \frac{\|M_k s_k\|^2}{s_k^\top M_k s_k} + \frac{\|y_k\|^2}{y_k^\top s_k} - \log \left(\frac{y_k^\top s_k}{s_k^\top M_k s_k} \right) \\ &= \psi(M_k) - \left[\frac{\|M_k s_k\| \|s_k\|}{s_k^\top M_k s_k} \right]^2 \frac{s_k^\top M_k s_k}{s_k^\top s_k} + \frac{\|y_k\|^2}{y_k^\top s_k} \\ &\quad - \log \left(\frac{y_k^\top s_k}{s_k^\top s_k} \frac{s_k^\top s_k}{s_k^\top M_k s_k} \right) \end{aligned}$$

CHAPITRE 3. RÔLE JOUÉ PAR LA CONVEXITÉ DANS LA
CONVERGENCE DE LA MÉTHODE BFGS

Si on note

$$q_k := \frac{s_k^\top M_k s_k}{\|s_k\|^2} \text{ et } \cos \theta_k = \frac{-g_k^\top d_k}{\|g_k\| \|d_k\|} = \frac{s_k^\top M_k s_k}{\|M_k s_k\| \|s_k\|}$$

Le quotient de Rayleigh de M_k dans la direction s_k et le cosinus de l'angle entre d_k et l'opposé du gradient. Alors on peut écrire

$$\begin{aligned} \psi(M_{k+1}) &= \psi(M_k) + \frac{\|y_k\|^2}{y_k^\top s_k} - \log \frac{y_k^\top s_k}{s_k^\top s_k} - \frac{q_k}{\cos^2 \theta_k} + \log q_k \\ &= \psi(M_k) + \frac{\|y_k\|^2}{y_k^\top s_k} - 1 - \log \frac{y_k^\top s_k}{s_k^\top s_k} + [\log \cos^2 \theta_k] \\ &\quad + \left[1 - \frac{q_k}{\cos^2 \theta_k} + \log \frac{q_k}{\cos^2 \theta_k} \right] \end{aligned} \quad (3.30')$$

De (3.30'), (3.26) et (3.27) (supposant $\varepsilon \delta^\alpha = c$ et $\frac{l^2}{\varepsilon \delta^\alpha} = C$)

$$\begin{aligned} \psi(M_{k+1}) &\leq \psi(M_1) + (C - 1 - \log c) i_k \\ &\quad + \sum_{i \in \bar{K}} \left(\log \cos^2 \theta_i + 1 - \frac{q_i}{\cos^2 \theta_i} + \log \frac{q_i}{\cos^2 \theta_i} \right) \end{aligned}$$

Définissons près $\eta_i \geq 0$ par

$$\eta_i = -\log \cos^2 \theta_i - \left[1 - \frac{q_i}{\cos^2 \theta_i} + \log \frac{q_i}{\cos^2 \theta_i} \right] \quad (3.31')$$

Puisque $\psi(M_{k+1}) \geq 0$ nous avons

$$\frac{1}{i_k} \sum_{i \in \bar{K}} \eta_i < \frac{\psi(M_1)}{i_k} + (C - 1 - \log c) \quad (3.32')$$

Maintenant définissons I_k pour être un ensemble comprenant les indices $[i_k/2]$ correspondant aux plus petites valeurs $[i_k/2]$ du η_i , pour $i \in \bar{K}_k$, et η_{m_k} la plus grande valeur du η_i pour $i \in I_k$. alors

$$\frac{1}{i_k} \sum_{i \in \bar{K}} \eta_i \geq \frac{1}{i_k} \left[\eta_{m_k} + \sum_{\substack{i \in \bar{K} \\ i \notin I_k}} \eta_i \right] \geq \eta_{m_k} \left(1 - \underbrace{\frac{1}{2}}_{p \text{ dans } [1]} \right)$$

3.2. CONVERGENCE GLOBALE DE LA MÉTHODE BFGS DANS LE CAS NON CONVEXE

Ainsi de (3.32') nous avons, pour tout $i \in I_k$,

$$\eta_i < \frac{1}{\left(1 - \frac{1}{2}\right)} [\psi(M_1) + C - 1 - \log c] \equiv \beta_0 \quad (3.33')$$

Puisque le terme à l'intérieur des parenthèses dans (3.31') est inférieur ou égal à zéro, nous concluons de (3.33') et (3.31'), pour tout $i \in I_k$

$$-\log \cos^2 \theta_i < \beta_0$$

Par conséquent

$$\cos \theta_i > e^{-\frac{\beta_0}{2}} \equiv \bar{\beta}_0 \quad (3.34')$$

De même, de (3.33') et (3.31') nous avons, pour tout $i \in I_k$

$$1 - \frac{q_i}{\cos^2 \theta_i} + \log \frac{q_i}{\cos^2 \theta_i} > -\beta_0$$

Notons que la fonction

$$u(t) = 1 - t + \log t$$

est non positif pour tout $t > 0$, réalise sa valeur maximum à $t = 1$, et satisfait $u(t) \rightarrow -\infty$ comme $t \rightarrow 0$ ou $t \rightarrow \infty$. Par conséquent, pour tout $i \in I_k$

$$0 < \bar{\beta}_1 \leq \frac{q_i}{\cos^2 \theta_i} \leq \tilde{\beta}_1 \quad (3.35')$$

Par conséquent, de (3.34') on obtient

$$\frac{\|M_i s_i\|}{\|s_i\|} = \frac{q_i}{\cos \theta_i} \leq \frac{\tilde{\beta}_1}{\bar{\beta}_0} \equiv \beta_1$$

c'est-à-dire $\|M_i s_i\| \leq \beta_1 \|s_i\|$.

En conclusion, de (3.35') et (3.34')

$$\beta_2 \equiv \left(\bar{\beta}_0\right)^2 \bar{\beta}_1 \leq q_i \leq \left(\bar{\beta}_0\right)^2 \tilde{\beta}_1 \equiv \beta_3$$

c'est-à-dire $\beta_2 \|s_i\|^2 \leq s_i^\top M_i s_i \leq \beta_3 \|s_i\|^2$. \square ■

Conséquence immédiate : Le lemme 3.3 prouve que lorsque \bar{K} est infini et que si on a (3.23), alors il existe un ensemble d'indices infini $\tilde{K} \subset \bar{K}$ tel que

$$\|g_i\| = \|M_i d_i\| \leq \beta_1 \|d_i\|, \quad \forall i \in \tilde{K} \quad (3.30)$$

et

$$\|d_i\|^2 \leq \beta_2^{-1} d_i^\top M_i d_i = -\beta_2^{-1} g_i^\top d_i \leq \beta_2^{-1} \|g_i\| \|d_i\|, \quad \forall i \in \tilde{K} \quad (3.31)$$

et par conséquent

$$\|d_i\| \leq \beta_2^{-1} \|g_i\|, \quad \forall i \in \tilde{K} \quad (3.32)$$

De plus, (3.29) et (3.30) impliquent

$$\|g_i\|^2 \leq \beta_1^2 \|d_i\|^2 \leq -\beta_1^2 \beta_2^{-1} g_i^\top d_i, \quad \forall i \in \tilde{K} \quad (3.33)$$

Convergence globale de l'algorithme 3.1 avec la recherche linéaire d'Armijo

Théorème 3.3 ([11]) *Supposons que l'hypothèse 3.2 soit vraie. Soit $\{z_k\}_{k \in \mathbb{N}}$ une suite générée par l'algorithme 3.1 avec λ_k déterminé par la recherche linéaire d'Armijo (wolfe1). Alors*

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0 \quad (3.34)$$

Preuve. ([11])

En prenant en considération le théorème 3.2, il suffit de vérifier (3.34) dans le cas où \bar{K} est un ensemble infini.

On suppose que (3.34) n'est pas vraie, alors il existe une constante $\delta > 0$ tel que $\|g_k\| \geq \delta$ pour tout k . Considérons l'ensemble $\tilde{K} \subset \bar{K}$ comme spécifié dans le paragraphe précédent. \tilde{K} contient donc une infinité d'indices. Notons par

$$\bar{\lambda} = \limsup_{\substack{k \rightarrow \infty \\ k \in \tilde{K}}} \lambda_k = \lim_{\substack{k \rightarrow \infty \\ k \in K'}} \lambda_k, \quad \text{où } K' \in \tilde{K}.$$

3.2. CONVERGENCE GLOBALE DE LA MÉTHODE BFGS DANS LE CAS NON CONVEXE

Comme la suite $\{\boldsymbol{x}_k\}_{k \in K'}$ est bornée, on obtient de (3.32) que la suite $\{d_i\}_{k \in K'}$ est aussi bornée. Sans perte de généralité, nous supposons que les suites $\{\boldsymbol{x}_k\}_{k \in K'}$ et $\{d_i\}_{k \in K'}$ convergent vers $\bar{\boldsymbol{x}}$ et \bar{d} respectivement. Alors de (3.15) on a

$$\bar{\lambda} g^\top(\bar{\boldsymbol{x}}) \bar{d} = 0.$$

En prenant en considération (3.33), il suffit de montrer que $\bar{\lambda} > 0$.

Supposons le contraire, $\bar{\lambda} = 0$, alors d'après la règle d'Armijo et pour $k \in K'$ suffisamment grand, $\lambda'_k \triangleq \lambda_k / \rho$ ne satisfait pas (wolfe1). Ce qui signifie que :

$$f(\boldsymbol{x}_k + \lambda'_k d_k) - f(\boldsymbol{x}_k) > \sigma \lambda'_k g^\top(\boldsymbol{x}_k) d_k \quad (3.35)$$

Par le théorème des accroissements finis, il existe $\theta_k \in (0, 1)$ tel que

$$f(\boldsymbol{x}_k + \lambda'_k d_k) - f(\boldsymbol{x}_k) = \lambda'_k g^\top(\boldsymbol{x}_k + \theta_k \lambda'_k d_k) d_k$$

Appliquons cela à (3.35), on obtient :

$$g^\top(\boldsymbol{x}_k + \theta_k \lambda'_k d_k) d_k > \sigma g^\top(\boldsymbol{x}_k) d_k$$

qui donne

$$(g^\top(\boldsymbol{x}_k + \theta_k \lambda'_k d_k) - g^\top(\boldsymbol{x}_k)) d_k > -(1 - \sigma) g^\top(\boldsymbol{x}_k) d_k$$

et on a

$$l \lambda'_k \|d_k\|^2 \stackrel{\text{car } \theta_k \in (0,1)}{\geq} l \theta_k \lambda'_k \|d_k\|^2 \stackrel{(3.16)}{\geq} (g^\top(\boldsymbol{x}_k + \theta_k \lambda'_k d_k) - g^\top(\boldsymbol{x}_k)) d_k$$

donc :

$$\begin{aligned} l \lambda'_k \|d_k\|^2 &> -(1 - \sigma) g_k^\top d_k \\ &= (1 - \sigma) d_k^\top M_k d_k \\ &\stackrel{(3.31)}{\geq} (1 - \sigma) \beta_2 \|d_k\|^2 \end{aligned}$$

ceci implique que

$$l \lambda'_k > (1 - \sigma) \beta_2$$

Cette inégalité contredit l'hypothèse, $\bar{\lambda} = 0$

(car : $((1 - \sigma) > 0, \beta_2 > 0, \text{ et } l > 0) \Rightarrow \bar{\lambda} > 0$). ■

Convergence globale de l'algorithme 3.1 avec la recherche linéaire de Wolfe

Pour cet effet, nous prouvons d'abord un lemme semblable au lemme 3.2 dans ([34]).

Lemme 3.4 *Supposons que l'hypothèse 3.2 soit vraie. Soit $\{\mathbf{x}_k\}_{k \in \mathbb{N}}$ une suite générée par l'algorithme 3.1 avec λ_k déterminé par la recherche linéaire de Wolfe (wolfe1)-(wolfe2). Si on a (3.23) pour tout k , alors il existe une constante $c_3 > 0$ tel que pour tout k assez grand on ait :*

$$\prod_{i \in \tilde{K}_k} \lambda_i \geq c_3^{i_k} \quad (3.36)$$

Preuve. la formule (3.18) donne la relation de récurrence suivante :

$$\det(M_{i+1}) = \det(M_i) \left(\frac{y_i^\top s_i}{s_i^\top M_i s_i} \right) \quad \forall i \in \bar{K} \quad (3.37)$$

et

$$\det(M_{i+1}) = \det(M_i) \quad \forall i \notin \bar{K} \quad (3.38)$$

Si on note par n_k le plus grand indice dans l'ensemble \bar{K}_k , alors on peut écrire

$$\det(M_{n_k+1}) = \det(M_1) \prod_{i \in \bar{K}_k} \frac{y_i^\top s_i}{s_i^\top M_i s_i} \quad (3.39)$$

D'autre part, de (wolfe2) on a :

$$g^\top (\mathbf{x}_i + \lambda_i d_i) d_i \geq \sigma_2 g^\top (\mathbf{x}_i) d_i$$

implique

$$y_i^\top s_i = (g_{i+1} + g_i)^\top s_i \geq \sigma_2 g_i^\top s_i - g_i^\top s_i$$

ceci implique :

$$y_i^\top s_i \geq -(1 - \sigma_2) g_i^\top s_i = (1 - \sigma_2) \lambda_i^{-1} s_i^\top M_i s_i,$$

avec

$$g^\top = -d_i^\top M_i = -\lambda_i^{-1} s_i^\top M_i$$

3.2. CONVERGENCE GLOBALE DE LA MÉTHODE BFGS DANS LE CAS NON CONVEXE

Ceci étant et en procédant d'une manière semblable à la preuve du lemme 3.2 dans ([34]), on obtient (3.36) en appliquant la dernière inégalité et (3.24) à (3.39)

En effet : de la dernière inégalité on peut écrire

$$\prod_{i \in \bar{K}_k} \frac{y_i^\top s_i}{s_i^\top M_i s_i} \geq \prod_{i \in \bar{K}_k} \frac{1 - \sigma_2}{\lambda_i}$$

avec (3.39) on obtient

$$\det(M_{n_{k+1}}) \geq \det(M_1) \prod_{i \in \bar{K}_k} \frac{1 - \sigma_2}{\lambda_i}$$

or

$$\det(M_{n_{k+1}}) \leq \left[\frac{\text{Tr}(M_{n_{k+1}})}{n} \right]^n,$$

par conséquent, d'après (3.24)

$$\prod_{i \in \bar{K}_k} \frac{1 - \sigma_2}{\lambda_i} \leq \frac{\det(M_{n_{k+1}})}{\det(M_1)} \leq \frac{1}{\det(M_1)} \left[\frac{\text{Tr}(M_{n_{k+1}})}{n} \right]^n \stackrel{(3.24)}{\leq} \frac{1}{\det(M_1)} \left[\frac{c_2 i_k}{n} \right]^n$$

Donc

$$\begin{aligned} \prod_{i \in \bar{K}_k} \frac{1 - \sigma_2}{\lambda_i} &\leq \frac{1}{\det(M_1)} \left[\frac{c_2 i_k}{n} \right]^n \\ \prod_{i \in \bar{K}_k} \frac{1 - \sigma_2}{\lambda_i} &\leq \frac{1}{\det(M_1) n^n} [c_2^{i_k}]^n \\ &\leq \frac{1}{\det(M_1) n^n} [c_2^n]^{i_k} \end{aligned}$$

donc il existe une constante c_3 telle que

$$\prod_{i \in \tilde{K}_k} \lambda_i \geq c_3^{i_k}. \square$$

■

Maintenant nous sommes en mesure de prouver **la convergence globale** de l'algorithme 3.1 avec la **recherche linéaire de Wolfe**.

CHAPITRE 3. RÔLE JOUÉ PAR LA CONVEXITÉ DANS LA
CONVERGENCE DE LA MÉTHODE BFGS

Théorème 3.4 *Supposons que l'hypothèse 3.2 soit vraie. Soit $\{z_k\}_{k \in \mathbb{N}}$ une suite générée par l'algorithme 3.1 avec λ_k déterminé par la recherche linéaire de Wolfe (wolfe1)-(wolfe2). Alors on a*

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0$$

Preuve. ([11])

En prenant en considération le théorème 3.2, il suffit de vérifier (3.34) dans le cas où \bar{K} est infini. Notons \bar{K} par

$$\bar{K} = \{k_1 < k_2 < \dots\}.$$

Remarquons que (3.14) donne particulièrement

$$-\sum_{j=0}^{\infty} g_{k_j}^\top s_{k_j} < \infty$$

et comme, $M_{k_j} s_{k_j} = -\lambda_{k_j} g_{k_j}$, alors

$$\sum_{j=1}^{\infty} \|g_{k_j}\|^2 \lambda_{k_j} \frac{s_{k_j}^\top M_{k_j} s_{k_j}}{\|M_{k_j} s_{k_j}\|^2} = -\sum_{j=0}^{\infty} g_{k_j}^\top s_{k_j} < \infty \quad (3.40)$$

car

$$\begin{aligned} -g_{k_j}^\top s_{k_j} &= -\|g_{k_j}\|^2 \lambda_{k_j}^2 \frac{g_{k_j}^\top s_{k_j}}{\|g_{k_j}\|^2 \lambda_{k_j}^2} = \|g_{k_j}\|^2 \lambda_{k_j} \frac{-\lambda_{k_j} g_{k_j}^\top s_{k_j}}{\|\lambda_{k_j} g_{k_j}\|^2} \\ &= \|g_{k_j}\|^2 \lambda_{k_j} \frac{s_{k_j}^\top M_{k_j} s_{k_j}}{\|M_{k_j} s_{k_j}\|^2}. \end{aligned}$$

Si (3.34) n'est pas satisfaite, alors il existe une constante $\delta > 0$ tel que $\|g_k\| \geq \delta$ pour tout k . Ainsi (3.40) implique

$$\sum_{j=1}^{\infty} \lambda_{k_j} \frac{s_{k_j}^\top M_{k_j} s_{k_j}}{\|M_{k_j} s_{k_j}\|^2} < \infty. \quad (3.41)$$

Par conséquent, pour n'importe quel $\xi > 0$, il existe un entier $j_0 > 0$, tel que pour tout entier positif q , on a

$$\left(\prod_{j=j_0+1}^{j_0+q} \lambda_{k_j} \frac{s_{k_j}^\top M_{k_j} s_{k_j}}{\|M_{k_j} s_{k_j}\|^2} \right)^{\frac{1}{q}} \leq \frac{1}{q} \sum_{j=j_0+1}^{j_0+q} \lambda_{k_j} \frac{s_{k_j}^\top M_{k_j} s_{k_j}}{\|M_{k_j} s_{k_j}\|^2} \leq \frac{\xi}{q}$$

ceci implique que

$$\begin{aligned}
 \left(\prod_{j=j_0+1}^{j_0+q} \lambda_{k_j} \right)^{\frac{1}{q}} &\leq \frac{\xi}{q} \left(\prod_{j=j_0+1}^{j_0+q} \frac{\|M_{k_j} s_{k_j}\|^2}{s_{k_j}^\top M_{k_j} s_{k_j}} \right)^{\frac{1}{q}} \\
 &\leq \frac{\xi}{q^2} \sum_{j=j_0+1}^{j_0+q} \frac{\|M_{k_j} s_{k_j}\|^2}{s_{k_j}^\top M_{k_j} s_{k_j}} \\
 &\leq \frac{\xi}{q^2} \sum_{j=0}^{j_0+q} \frac{\|M_{k_j} s_{k_j}\|^2}{s_{k_j}^\top M_{k_j} s_{k_j}} \\
 &\stackrel{((3.25) \text{ où } i_k=j_0+q+1)}{\leq} \frac{\xi (j_0 + q + 1)}{q^2} c_2.
 \end{aligned}$$

Si $q \rightarrow \infty$, alors on aboutit à une contradiction, car le lemme 3.4 certifie que le terme à gauche de l'inégalité ci-dessus est plus grand qu'une constante positive. ■. ■

3.3 Problèmes ouverts et conjectures

Nous avons vu que l'un des plus beaux résultats de convergence en optimisation sans contraintes est le théorème 3.1, du à M.J.D. Powell ([25]). Il donne des conditions assurant la convergence globale de l'algorithme BFGS lorsque la fonction est convexe. La beauté du résultat tient au fait qu'il est rare de pouvoir montrer la convergence d'un algorithme quasi-Newtonien avec si peu d'hypothèses sur les objets générés par celui-ci. On suppose seulement que la suite $\{f(x_k)\}_{k \in \mathbb{N}}$ est bornée inférieurement. On ne fait aucune hypothèse sur la suite des matrices $\{B_k\}$. Rappelons ce théorème

Théorème 3.1 ([25]) *Supposons que $f : \mathbb{R}^n \rightarrow \mathbb{R}$ soit convexe et $C^{1,1}$ dans un voisinage convexe de $D = \{x \in \mathbb{R}^n : f(x) \leq f(x_1)\}$, où $x_1 \in \mathbb{R}^n$. On considère l'algorithme de BFGS démarrant en x_1 avec une matrice B_1 symétrique définie positive et on suppose que celui-ci génère une suite $\{x_k\}$ telle que $\{f(x_k)\}$ soit bornée inférieurement. Alors, $\liminf_{k \rightarrow \infty} \|\nabla f(x_k)\| = 0$.*

Problème ouvert I *Considérons la méthode BFGS. Le résultat de convergence du théorème 3.1 dépend du fait que f est convexe. On se pose donc de façon légitime le problème ouvert suivant : Considérons la méthode BFGS avec une recherche linéaire inexacte satisfaisant les deux conditions de Wolfe. Supposons que f continûment différentiable et bornée inférieurement. Supposons que l'algorithme BFGS démarre d'un point initial x_1 quelconque et d'une matrice B_1 définie positive et génère une suite $\{x_k\}$. Est ce qu'on a comme dans le cas convexe la convergence globale de la suite $\{x_k\}$, c'est à dire a t'on*

$$\liminf \|\nabla f(x_k)\| = 0. \quad (\text{prob1})$$

Conjecture 1 Tous les tests numériques effectués sur des fonctions non convexes nous ont montré que la suite générée par l'algorithme BFGS converge vers des points stationnaires. Celà n'a pas été prouvé mais pesonne n'est parvenu à trouver un contre exemple d'une fonction non convexe pour laquelle la méthode BFGS diverge. nous conjecturons que la relation (prob1) est vraie.

Avant de poser le deuxième problème ouvert, donnons d'abord la notion de fonction quas-convexe.

Définition (*fonction quasi-convexe*) Soit $S \subset \mathbb{R}^n$ un ensemble convexe et $f : S \rightarrow \mathbb{R}$. Soit $\alpha \in \mathbb{R}$. Notons par S_α les ensembles suivants :

$$S_\alpha = \{x \in S : f(x) \leq \alpha\}$$

On dit que f est quasi convexe dans S si et seulement si pour tout $\alpha \in \mathbb{R}$, les ensembles S_α sont convexes.

Remarque : *Il est facile de voir que si f est convexe dans S , alors f est aussi quasi-convexe dans S . La réciproque est en général fausse. La notion de quasi-convexité est plus faible que la notion de convexité.*

Problème ouvert II Considérons dans ce cas aussi la méthode BFGS et au lieu de considerer des fonctions non convexes comme dans le problème ouvert 1, nous prenons des fonctions quasi-convexes, c'est à dire qu'on affaiblit l'hypothèse de convexité sans la supprimer totalement. En effect la notion de fonction quas-convexe est plus faible que la notion de convexité.

Nous conjecturons le résultat suivant :

Conjecture II *Supposons que $f : \mathbb{R}^n \rightarrow \mathbb{R}$ soit quasi-convexe et $C^{1,1}$ dans un voisinage convexe de $D = \{x \in \mathbb{R}^n : f(x) \leq f(x_1)\}$, où $x_1 \in \mathbb{R}^n$. On considère l'algorithme*

BFGS démarrant en x_1 avec une matrice B_1 symétrique définie positive et on suppose que celui-ci génère une suite $\{x_k\}$ telle que $\{f(x_k)\}$ soit bornée inférieurement. Alors,

$$\liminf_{k \rightarrow \infty} \|\nabla f(x_k)\| = 0.$$

3.4 Tests numériques

Cette section contient une étude numérique faite avec l'algorithme 3.1. Nous avons testé l'algorithme 3.1 sur les trois problèmes suivants pris de ([16]).

Problème 1 Extension de la fonction singulière de Powell

$$\begin{aligned} f(\mathcal{x}) &= \sum_{i=1}^{n/4} \{ (\mathcal{x}_{4i-3} + 10\mathcal{x}_{4i-2})^2 + 5(\mathcal{x}_{4i-1} - \mathcal{x}_{4i})^2 \\ &\quad + (\mathcal{x}_{4i-2} - 2\mathcal{x}_{4i-1})^4 + 10(\mathcal{x}_{4i-1} - \mathcal{x}_{4i})^4 \}, \\ \mathcal{x}^* &= (0, 0, \dots, 0)^\top. \end{aligned}$$

Problème 2 Fonction de Rosenbrock (Extension)

$$\begin{aligned} f(\mathcal{x}) &= \sum_{i=1}^{n/2} \{ 100(\mathcal{x}_{2i} - \mathcal{x}_{2i-1}^2) + (1 - \mathcal{x}_{2i-1})^2 \}, \\ \mathcal{x}^* &= (1, 1, \dots, 1)^\top. \end{aligned}$$

Problème 3 Fonction de Wood (Extension)

$$\begin{aligned} f(\mathcal{x}) &= \sum_{i=1}^{n/4} \{ 100(\mathcal{x}_{4i-2} - \mathcal{x}_{4i-3}^2)^2 + (1 - \mathcal{x}_{4i-3})^2 + 90(\mathcal{x}_{4i} - \mathcal{x}_{4i-1}^2)^2 \\ &\quad + (1 - \mathcal{x}_{4i-1})^2 + 10(\mathcal{x}_{4i-2} + \mathcal{x}_{4i} - 2)^2 + 0.1(\mathcal{x}_{4i-2} - \mathcal{x}_{4i})^2 \}, \\ \mathcal{x}^* &= (1, 1, \dots, 1)^\top. \end{aligned}$$

CHAPITRE 3. RÔLE JOUÉ PAR LA CONVEXITÉ DANS LA
CONVERGENCE DE LA MÉTHODE BFGS

Nous avons appliqué l'algorithme 3.1, qu'on a appelé la méthode BFGSA (A : abreviation de appropriée), avec la recherche linéaire de Wolfe ou d'Armijo à ces problèmes et nous l'avons comparé avec la méthode ordinaire de BFGS. Nous avons utilisé la condition $\max \{\|g(\mathcal{x}_k)\|, \|\mathcal{x}_k - \mathcal{x}^*\|\} \leq 10^{-5}$ comme critère d'arrêt, et pour chaque problème, nous avons choisi différents points de départ (points initiaux) mais avec la même matrice initiale $M_0 = I$ (matrice d'unité). Pour chaque problème, les paramètres communs aux deux méthodes ont été placés identiquement. Spécialement nous avons choisi les paramètres comme suit $\sigma_1 = 0.1$ et $\sigma_2 = 0.49$ dans la recherche linéaire de Wolfe (wolfe1)-(wolfe2) et $\sigma = 0.1$ dans la recherche linéaire d'Armijo (wolfe1) et $\alpha = \begin{cases} 0.01, & \text{si } \|g_k\| \geq 1 \\ 3, & \text{si } \|g_k\| < 1 \end{cases}$, $\varepsilon = 0.1$, dans la mise à jour appropriée (3.13). Il n'est pas difficile de voir que le théorème de la convergence dans la section 3.2.2 est vrai si le choix de α est comme ci-dessus ; en général, si $\alpha \in [\mu_1, \mu_2]$ avec $\mu_1 > 0$, tous les théorèmes dans la section 3.2.2 sont vrais. Les résultats sont donnés dans les tableaux 1-4, où P1, P2 et P3 représentent les problèmes 1, 2 et 3, respectivement, et k représente le nombre d'itérations. Pour la méthode de BFGSA, "off" dénote le nombre de fois où la condition dans la mise à jour prudente n'a pas été réunie, c'est-à-dire, le nombre des k où $\frac{y_k^\top s_k}{\|s_k\|^2} < \varepsilon \|g_k\|^\alpha$. Pour la méthode de BFGS, "off" dénote le nombre des k où $y_k^\top s_k < 10^{-17}$, et pour la méthode de BFGS avec la recherche linéaire de Wolfe, "off" est égale à zéro comme $y_k^\top s_k$ est toujours positive à cause de l'inégalité de (wolfe2).

Dans les tableaux "init" présente le point initial. Les points initiaux sont $\mathcal{x}_0 = (0, 0, \dots, 0)^\top$, $\mathcal{x}_1 = (1, 1, \dots, 1)^\top$, $\mathcal{x}_2 = (10, 10, \dots, 10)^\top$, $\mathcal{x}_3 = (100, 100, \dots, 100)^\top$, $\mathcal{x}_4 = -\mathcal{x}_2$, $\mathcal{x}_5 = -\mathcal{x}_3$, $\mathcal{x}_6 = (0, 100, 0, 100, \dots)^\top$ et $\mathcal{x}_7 = -\mathcal{x}_6$.

Dans la suite, pour voir l'influence de α et ε dans la méthode de BFGSA, on fait la résolution du problème 1 avec des valeurs différent de α et ε partant du point initial $\mathcal{x}_2 = (10, 10, \dots, 10)^\top$.

Les résultats sont dans le tableau 5, où

$$\alpha_1 = \begin{cases} 0.0001, & \text{si } \|g_k\| \geq 1 \\ 4, & \text{si } \|g_k\| < 1 \end{cases}, \alpha_2 = \begin{cases} 0.001, & \text{si } \|g_k\| \geq 1 \\ 4, & \text{si } \|g_k\| < 1 \end{cases},$$

$$\alpha_3 = \begin{cases} 0.01, & \text{si } \|g_k\| \geq 1 \\ 4, & \text{si } \|g_k\| < 1 \end{cases}, \alpha_4 = \begin{cases} 0.1, & \text{si } \|g_k\| \geq 1 \\ 4, & \text{si } \|g_k\| < 1 \end{cases},$$

$$\alpha_5 = \begin{cases} 1, & \text{si } \|g_k\| \geq 1 \\ 4, & \text{si } \|g_k\| < 1 \end{cases}, \alpha_6 = 1, \alpha_7 = 2.$$

Nous avons constaté que la condition de la mise à jour appropriée à été habituellement satisfaite. Cependant, quand elle est violée, la méthode de BFGSA est aussi mauvaise que la méthode BFGS.

Les résultats prouvent aussi que le choix des paramètres α et ε sont numériquement importants. Si nous choisissons α et ε convenablement, alors la condition de la mise à jour appropriée est presque toujours satisfaite et la méthode de BFGSA se réduit à la méthode de BFGS classique.

CHAPITRE 3. RÔLE JOUÉ PAR LA CONVEXITÉ DANS LA
CONVERGENCE DE LA MÉTHODE BFGS

Tableau 1
tests numériques pour la méthode de BFGSA (recherche de Wolfe)

Problem	n	Init	k	$\ \nabla f_k\ $	$\ x_k - x^*\ $	off
P1	4	x^1	31	1.0E-7	1.3E-3	0
P1	4	x^2	35	1.0E-6	3.6E-3	0
P1	4	x^3	39	1.0E-6	4.6E-3	0
P1	4	x^4	35	1.0E-6	3.6E-3	0
P1	4	x^5	39	1.0E-6	4.6E-3	0
P1	4	x^6	43	1.0E-6	1.7E-3	0
P1	4	x^7	43	1.0E-6	1.7E-3	0
P2	2	x^0	17	1.1E-4	1.0E-5	0
P2	2	x^2	61	1.8E-4	4.0E-6	0
P2	2	x^4	70	1.1E-4	7.0E-6	0
P2	2	x^6	44	1.1E-5	1.0E-6	0
P2	2	x^7	25	1.0E-7	4.3E-5	0
P2	10	x^0	37	1.7E-4	8.0E-6	0
P2	10	x^2	134	1.0E-6	1.5E-5	0
P2	10	x^4	144	1.4E-5	2.0E-6	0
P2	10	x^6	99	7.7E-5	6.0E-6	0
P2	10	x^7	70	3.0E-6	1.5E-5	0
P2	100	x^0	93	4.7E-5	6.0E-6	0
P2	100	x^2	354	8.0E-6	1.3E-5	0
P2	100	x^6	308	5.1E-5	7.0E-6	0
P2	100	x^7	277	3.6E-5	9.0E-6	0
P3	4	x^0	19	3.4E-5	4.0E-6	0
P3	4	x^2	29	4.9E-5	1.0E-7	0
P3	4	x^3	45	2.4E-4	6.0E-6	0
P3	4	x^6	31	4.0E-4	1.0E-6	0
P3	4	x^7	38	7.2E-5	1.0E-7	0
P3	40	x^0	51	6.9E-5	4.0E-6	0
P3	40	x^2	185	7.0E-4	8.0E-6	0
P3	40	x^3	499	1.3E-4	6.0E-6	0
P3	40	x^4	530	7.4E-5	3.0E-6	0
P3	40	x^5	524	8.0E-6	1.0E-6	0
P3	40	x^6	168	1.8E-4	5.0E-6	0
P3	40	x^7	148	2.0E-5	1.0E-5	0
P3	100	x^0	75	1.8E-4	6.0E-6	0
P3	100	x^2	352	1.0E-4	8.0E-6	0
P3	100	x^3	758	2.1E-5	5.0E-6	0
P3	100	x^4	875	1.6E-4	5.0E-6	0
P3	100	x^5	1239	1.1E-4	7.0E-6	0
P3	100	x^6	289	1.0E-4	5.0E-6	0
P3	100	x^7	233	3.4E-4	7.0E-6	0

Tableau 2

tests numériques pour la méthode de BFGSA (recherche de Wolfe)

Problem	n	Init	k	$\ \nabla f_k\ $	$\ x_k - x^*\ $	off
P1	4	x^1	31	1.0E-7	1.3E-3	0
P1	4	x^2	35	1.0E-6	3.6E-3	0
P1	4	x^3	39	1.0E-6	4.6E-3	0
P1	4	x^4	35	1.0E-6	3.6E-3	0
P1	4	x^5	39	1.0E-6	4.6E-3	0
P1	4	x^6	43	1.0E-6	1.7E-3	0
P1	4	x^7	43	1.0E-6	1.7E-3	0
P2	2	x^0	17	1.1E-4	1.0E-5	0
P2	2	x^2	61	1.8E-4	4.0E-6	0
P2	2	x^4	70	1.1E-4	7.0E-6	0
P2	2	x^6	44	1.1E-5	1.0E-6	0
P2	2	x^7	25	1.0E-7	4.3E-5	0
P2	10	x^0	37	1.7E-4	8.0E-6	0
P2	10	x^2	134	1.0E-6	1.5E-5	0
P2	10	x^4	144	1.4E-5	2.0E-6	0
P2	10	x^6	99	7.7E-5	6.0E-6	0
P2	10	x^7	70	3.0E-6	1.5E-5	0
P2	100	x^0	93	4.7E-5	6.0E-6	0
P2	100	x^2	354	8.0E-6	1.3E-5	0
P2	100	x^6	308	5.1E-5	7.0E-6	0
P2	100	x^7	277	3.6E-5	9.0E-6	0
P3	4	x^0	19	3.4E-5	4.0E-6	0
P3	4	x^2	29	4.9E-5	1.0E-7	0
P3	4	x^3	45	2.4E-4	6.0E-6	0
P3	4	x^6	31	4.0E-4	1.0E-6	0
P3	4	x^7	38	7.2E-5	1.0E-7	0
P3	40	x^0	51	6.9E-5	4.0E-6	0
P3	40	x^2	185	7.0E-4	8.0E-6	0
P3	40	x^3	499	1.3E-4	6.0E-6	0
P3	40	x^4	530	7.4E-5	3.0E-6	0
P3	40	x^5	524	8.0E-6	1.0E-6	0
P3	40	x^6	168	1.8E-4	5.0E-6	0
P3	40	x^7	148	2.0E-5	1.0E-5	0
P3	100	x^0	75	1.8E-4	6.0E-6	0
P3	100	x^2	352	1.0E-4	8.0E-6	0
P3	100	x^3	758	2.1E-5	5.0E-6	0
P3	100	x^4	875	1.6E-4	5.0E-6	0
P3	100	x^5	1239	1.1E-4	7.0E-6	0
P3	100	x^6	289	1.0E-4	5.0E-6	0
P3	100	x^7	233	3.4E-4	7.0E-6	0

CHAPITRE 3. RÔLE JOUÉ PAR LA CONVEXITÉ DANS LA
CONVERGENCE DE LA MÉTHODE BFGS

Tableau 3
tests numériques pour la méthode de BFGSA (recherche d'Armijo)

Problem	n	Init	k	$\ \nabla f_k\ $	$\ x_k - x^*\ $	off
P1	4	x^1	41	1.0E-7	1.7E-3	0
P1	4	x^2	43	1.0E-6	2.7E-3	0
P1	4	x^3	61	1.0E-7	2.4E-3	0
P1	4	x^4	43	1.0E-6	2.7E-3	0
P1	4	x^5	61	1.0E-7	2.4E-3	0
P1	4	x^6	70	1.0E-7	3.9E-4	0
P1	4	x^7	70	1.0E-7	3.9E-4	0
P2	2	x^0	21	1.8E-5	1.0E-6	0
P2	2	x^2	45	6.0E-6	2.8E-5	1
P2	2	x^3	65	2.1E-5	1.0E-6	2
P2	2	x^4	42	1.6E-5	2.0E-6	3
P2	2	x^5	68	1.4E-4	5.0E-6	1
P2	2	x^6	36	6.0E-6	1.9E-5	0
P2	2	x^7	70	1.9E-6	2.0E-6	3
P2	10	x^0	53	2.6E-5	5.0E-6	0
P2	10	x^2	118	2.0E-5	4.0E-6	0
P2	10	x^3	168	1.1E-5	1.0E-6	3
P2	10	x^4	124	1.2E-5	4.0E-6	1
P2	100	x^0	127	5.1E-5	9.0E-6	0
P2	100	x^2	364	8.0E-6	2.4E-5	0
P2	100	x^3	468	1.2E-5	1.0E-5	1
P2	100	x^4	322	1.9E-5	1.0E-5	0
P3	4	x^0	45	7.4E-5	2.0E-6	1
P3	4	x^2	68	7.9E-4	8.0E-6	1
P3	4	x^3	93	1.8E-5	1.0E-7	0
P3	4	x^4	34	4.2E-3	8.0E-6	0
P3	4	x^5	75	5.3E-5	1.0E-6	0
P3	4	x^6	49	2.8E-5	1.0E-6	0
P3	4	x^7	71	6.3E-4	4.0E-6	0
P3	40	x^0	158	3.1E-4	9.0E-6	0
P3	40	x^2	467	5.1E-5	3.0E-6	0
P3	40	x^3	293	1.8E-5	8.0E-6	0
P3	40	x^4	163	3.7E-5	9.0E-6	0
P3	40	x^5	671	1.3E-4	9.0E-6	0
P3	40	x^6	209	2.5E-5	9.0E-6	0
P3	40	x^7	288	8.5E-5	9.0E-6	0
P3	100	x^0	280	2.0E-4	1.0E-5	0
P3	100	x^2	612	2.6E-5	9.0E-6	0
P3	100	x^3	520	9.2E-5	7.0E-6	0
P3	100	x^4	303	9.7E-5	7.0E-6	0
P3	100	x^5	960	1.5E-5	9.0E-6	0
P3	100	x^6	416	6.2E-5	9.0E-6	0
P3	100	x^7	482	3.0E-5	8.0E-6	0

Tableau 4

tests numériques pour la méthode de BFGS (recherche d'Armijo)

Problem	n	Init	k	$\ \nabla f_k\ $	$\ x_k - x^*\ $	off
P1	4	x^1	41	1.0E-7	1.7E-3	0
P1	4	x^2	43	1.0E-6	2.7E-3	0
P1	4	x^3	61	1.0E-7	2.4E-3	0
P1	4	x^4	43	1.0E-6	2.7E-3	0
P1	4	x^5	61	1.0E-7	2.4E-3	0
P1	4	x^6	70	1.0E-7	3.9E-4	0
P1	4	x^7	70	1.0E-7	3.9E-4	0
P2	2	x^0	21	1.8E-5	1.0E-6	0
P2	2	x^2	45	6.0E-6	2.8E-5	1
P2	2	x^3	62	1.2E-4	2.0E-6	0
P2	2	x^4	43	2.0E-6	1.0E-7	2
P2	2	x^5	68	1.4E-4	5.0E-6	1
P2	2	x^6	36	6.0E-6	1.9E-5	0
P2	2	x^7	71	1.1E-5	1.0E-7	1
P2	10	x^0	53	2.6E-5	5.0E-6	0
P2	10	x^2	118	2.0E-5	4.0E-6	0
P2	10	x^3	169	7.0E-6	2.9E-5	0
P2	10	x^4	124	1.2E-5	4.0E-6	1
P2	100	x^0	127	5.1E-5	9.0E-6	0
P2	100	x^2	364	8.0E-6	2.4E-5	0
P2	100	x^3	435	1.1E-5	9.0E-6	0
P2	100	x^4	322	1.9E-5	1.0E-5	0
P3	4	x^0	45	7.4E-5	2.0E-6	1
P3	4	x^2	68	7.9E-4	8.0E-6	1
P3	4	x^3	93	1.8E-5	1.0E-7	0
P3	4	x^4	34	4.2E-3	8.0E-6	0
P3	4	x^5	75	5.3E-5	1.0E-6	0
P3	4	x^6	49	2.8E-5	1.0E-6	0
P3	4	x^7	71	6.3E-4	4.0E-6	0
P3	40	x^0	158	3.1E-4	9.0E-6	0
P3	40	x^2	467	5.1E-5	3.0E-6	0
P3	40	x^3	293	1.8E-5	8.0E-6	0
P3	40	x^4	163	3.7E-5	9.0E-6	0
P3	40	x^5	671	1.3E-4	9.0E-6	0
P3	40	x^6	209	2.5E-5	9.0E-6	0
P3	40	x^7	288	8.5E-5	9.0E-6	0
P3	100	x^0	280	2.0E-4	1.0E-5	0
P3	100	x^2	612	2.6E-5	9.0E-6	0
P3	100	x^3	520	9.2E-5	7.0E-6	0
P3	100	x^4	303	9.7E-5	7.0E-6	0
P3	100	x^5	960	1.5E-5	9.0E-6	0
P3	100	x^6	416	6.2E-5	9.0E-6	0
P3	100	x^7	482	3.0E-5	8.0E-6	0

CHAPITRE 3. RÔLE JOUÉ PAR LA CONVEXITÉ DANS LA
CONVERGENCE DE LA MÉTHODE BFGS

Tableau 5 (recherche de Wolfe) pour la fonction de Powell
Différentes choix des paramètres dans la méthode de BFGA

n	α	ϵ	k	$\ \nabla f_k\ $	$\ x_k - x^*\ $	off
4	α_1	1	36	8.0E-6	3.9E-3	11
4	α_1	10^{-1}	32	9.0E-6	7.2E-3	2
4	α_1	10^{-3}	32	9.0E-6	8.3E-3	0
4	α_1	10^{-5}	32	9.0E-6	8.3E-3	0
4	α_2	1	36	8.0E-6	3.9E-3	11
4	α_2	10^{-1}	32	9.0E-6	7.2E-3	2
4	α_2	10^{-3}	32	9.0E-6	8.3E-3	0
4	α_2	10^{-5}	32	9.0E-6	8.3E-3	0
4	α_3	1	36	8.0E-6	3.9E-3	11
4	α_3	10^{-1}	32	9.0E-6	7.2E-3	2
4	α_3	10^{-3}	32	9.0E-6	8.3E-3	0
4	α_3	10^{-5}	32	9.0E-6	8.3E-3	0
4	α_4	1	36	8.0E-6	3.9E-3	11
4	α_4	10^{-1}	32	9.0E-6	7.2E-3	2
4	α_4	10^{-3}	32	9.0E-6	8.3E-3	0
4	α_4	10^{-5}	32	9.0E-6	8.3E-3	0
4	α_5	10^{-1}	36	4.0E-6	5.4E-3	5
4	α_5	10^{-3}	32	9.0E-6	8.3E-3	0
4	α_5	10^{-5}	32	9.0E-6	8.3E-3	0
4	α_6	10^{-1}	1056	1.0E-5	1.1E-2	1043
4	α_6	10^{-3}	187	6.0E-6	4.8E-3	164
4	α_6	10^{-5}	32	9.0E-6	8.3E-3	0
4	α_7	10^{-3}	13	8.0E-6	2.4E-3	0
4	α_7	10^{-5}	2	1.0E-5	2.4E-3	0
40	α_1	10^{-1}	243	9.0E-6	1.7E-2	9
40	α_1	10^{-3}	238	9.0E-6	1.6E-2	0
40	α_1	10^{-5}	238	9.0E-6	1.6E-2	0
40	α_2	10^{-1}	243	9.0E-6	1.7E-2	9
40	α_2	10^{-3}	238	9.0E-6	1.6E-2	0
40	α_2	10^{-5}	238	9.0E-6	1.6E-2	0
40	α_3	10^{-1}	243	9.0E-6	1.7E-2	9
40	α_3	10^{-3}	238	9.0E-6	1.6E-2	0
40	α_3	10^{-5}	238	9.0E-6	1.6E-2	0
40	α_4	10^{-1}	243	9.0E-6	1.7E-2	9
40	α_4	10^{-3}	238	9.0E-6	1.6E-2	0
40	α_4	10^{-5}	238	9.0E-6	1.6E-2	0
40	α_5	10^{-3}	238	9.0E-6	1.6E-2	0
40	α_5	10^{-5}	238	9.0E-6	1.6E-2	0
40	α_7	10^{-3}	82	3.0E-6	3.1E-3	61
40	α_7	10^{-5}	1	7.0E-6	3.1E-3	0

Bibliographie

- [1] **Andreas Antoniou et Wu-Sheng Lu** "PRACTICAL OPTIMIZATION Algorithms and Engineering Applications" (2007 Springer Science+Business Media, LLC).
- [2] **A. Griewank**, "The global convergence of partitioned BFGS on problems with convex decompositions and Lipschitzian gradients", (Mathematical Programming, 50 (1991) 141-175).
- [3] **A.A. Goldstein** and **J.F. Price** (1969), "An effective algorithm for minimization", (Num.Math., 10, pp. 184-189).
- [4] **A. V. Fiacco** and **G. P. McCormick**, (1968) "Nonlinear Programming", (Wiley, New York).
- [5] **B. A. Murtagh** and **R. W. H. Sargent** (1969), "A constrained minimization method with quadratic convergence", (Optimization, ed. R. Fletcher, pp. 215-246, Academic Press, London).
- [6] **C. G. Broyden**, (1965) "Quasi-Newton methods and their application to function minimization", (Maths. Comput., vol. 21, pp. 368–381).
- [7] **D.H. Li** and **M. Fukushima**, "A globally and superlinearly convergent Gauss-Newton based BFGS method for symmetric nonlinear equations", (to appear in SIAM Journal on Numerical Analysis).
- [8] **D.H. Li** and **M. Fukushima**, "A modified BFGS method and its global convergence in nonconvex minimization", (Technical Report 98003, Department of Applied Mathematics and Physics, Kyoto University, January 1998)
- [9] **D. Goldfarb** (1970), "A family of variable metric methods derived by variational means", (Maths. Comput., vol. 24, pp. 23–26).
- [10] **D. F. Shanno** (1970), "Conditioning of quasi-Newton methods for function minimization", (Maths. Comput., vol. 24, pp. 647–656).

-
- [11] **D.H. Li** and **M. Fukushima**, "on the global convergence of BFGS method for nonconvex unconstrained optimization problems", (April 7, 1999).
- [12] **G.Allaire** "L'ouvrage complet est disponible auprès des Editions de l'Ecole Polytechnique" paris, le 4 Janvier 2005.
- [13] **J.E.Dennis, Jr.** and **J.J.More**, "A characterization of superlinear convergence and its application to quasi-Newton methods", *Mathematics of Computation*, (28 (1974) 549-560).
- [14] **J.E.Dennis** and **J.J.More**, "Quasi-Newton methods, motivation and theory", (*SIAM Review*, 19 (1977) 46-89).
- [15] **Jorge Nocedal et Stephen J.Wright** "Numerical Optimization with 85 illustrations" (1999)
- [16] **J.J. Moré, B.S. Garbow** and **K.E. Hillstrome**, "Testing unconstrained optimization software", (*ACM Trans. Math. Software*, 7 (1981),17-41).
- [17] **J. Nocedal**, "Theory of algorithms for unconstrained optimization, *Acta Numerica*, 1 (1992) 199-242).
- [18] **J. C. Gilbert** (2007), "Eléments d'Optimisation Différentiable : Théorie et Algorithmes", (Notes de cours, école Nationale Supérieure de Techniques Avancées, Paris).
- [19] **J. Charles Gilbert** (19 février 2008), "Optimisation Différentiable".
- [20] **L.C.W.Dixon**, "Variable metric algorithms : necessary and sufficient conditions for identical behavior on nonquadratic functions", (*Journal of Optimization Theory and Applications*, 10 (1972) 34-40).
- [21] **L. Armijo** (1966), "Minimization of function having lipschitz continuous first partial derivatives", (*Pacific Journal of Mathematics*, Vol. 16(1), pp.1-3).
- [22] **Michel Minoux** (1983), "Programmation Mathématique, Théorie et Algorithmes, (tome 1 , Dunod).
- [23] **M.J.D. Powell**, "On the convergence of the variable metric algorithm", (*Journal of the Institute of Mathematics and its Applications*, 7 (1971) 21-36).
- [24] **M. S. Bazaraa, H. D. Sherali**, and **C. M. Shetty** (1993), "Nonlinear Programming, Theory and Algorithms", (Wiley-Interscience).

-
- [25] **M.J.D. Powell**, "Some global convergence properties of a variable metric algorithm for minimization without exact line searches", (in Nonlinear Programming, SIAM-AMS Proceedings, Vol. IX, R.W. Cottle, and C.E. Lemke, eds., SIAM, 1976a, pp.53-72).
- [26] **Ph.L. Toint**, "Global convergence of the partitioned BFGS algorithm for convex partially separable optimization", (Mathematical Programming, 36 (1986), 290-306).
- [27] **P. Wolfe** (1969), "Convergence conditions for ascent methods", (SIAM Review, 11, pp. 226-235).
- [28] **P. Wolfe** (1967), "Methods of nonlinear programming", (Nonlinear Programming, ed. J. Abadie, pp. 97-131, Interscience, Wiley, New York).
- [29] **R. Fletcher**, "Practical Methods of Optimization", (Second Edition, John Wiley & Sons, Chichester, 1987).
- [30] **R. Fletcher**, "An overview of unconstrained optimization, in Algorithms for Continuous Optimization : The State of the Art", (E. Spedicato, ed., Kluwer Academic Publishers, Boston, 1994, pp.109-143).
- [31] **R. Fletcher** and **M. J. D Powell** (1963) "A rapidly convergent descent method for minimization". (Comput. J. 6 ? 163-168).
- [32] **R. Fletcher** (1970), "A new approach to variable metric algorithms", (Computer J., vol. 13, pp. 317-322).
- [33] **R. Byrd** and **J. Nocedal**, "A tool for the analysis of quasi-Newton methods with application to unconstrained minimization", (SIAM Journal on Numerical Analysis, 26 (1989) 727-739).
- [34] **R. Byrd**, **J. Nocedal** and **Y. Yuan**, "Global convergence of a class of quasi-Newton methods on convex problems", (SIAM Journal on Numerical Analysis, 24 (1987) 1171-1189).
- [35] **W. C. Davidon** (1959), "Variable Metric Methods For Minimization", (Argonne National Lab Report (Argonne, IL)).