



وزارة التعليم العالي والبحث العلمي



**BADJI MOKHTAR -ANNABA
UNIVERSITY
UNIVERSITE BADJI MOKHTAR
ANNABA**

**جامعة باجي مختار
- عنابة -**

Faculté des Sciences

Année : 2018

Département de Mathématiques

THÈSE

Présenté en vue de l'obtention du diplôme de
Doctorat en Sciences

**LA CONVERGENCE DES METHODES A DIRECTION DE
DESCENTE**

**Option
Mathématiques appliquées**

Par
BOUAZIZ Khelifa

DIRECTEUR DE THESE: LASKRI Yamina Prof E.S.T.I.ANNABA

Devant le jury

**PREDIDENT: BOURAS Med Cherif MCA U.B.M ANNABA
EXAMINATEUR: AMIAR Rachida MCA U.B.M ANNABA
EXAMINATEUR: SALLAMI Badreddine MCA U.M.C.M SOUK AHRAS**

Remerciements

Je tiens à adresser mes remerciements les plus chaleureux et ma profonde gratitude à mon encadreur madame LASKRI Yamina, Professeur à l'Ecole Supérieure des Technologies Industrielles Annaba, pour m'avoir proposé le sujet de cette thèse. C'est grâce à sa grande disponibilité, ses conseils, ses orientations, et ses encouragements que j'ai pu mener à bien ce travail.

Mes remerciements vont également à Mr BOURAS Med Cherif MCA à l'Université Badji mokhtar Annaba, pour avoir bien voulu me faire l'honneur d'accepter de présider le jury.

De même je remercie Mme AMIAR Rachida, MCA à l'Université Badji mokhtar Annaba et Mr SALLAMI Badreddine, MCA à l'Université Mohamed Cherif Messaadia Souk-Ahras, pour l'honneur qu'ils m'ont fait de bien vouloir accepter de faire partie du jury.

Je remercie également tout ceux qui ont contribué de près ou de loin à l'élaboration de cette thèse, en particulier, ma femme pour ses encouragements.

Enfin, Merci à toute ma famille, en particulier, ma mère, mon père, mes enfants Douaa, Ghaith, Taha, ma soeur et mes frères, et à tous mes amis et tous mes collègues sans exception.

ملخص

تمت دراسة خصائص التقارب الكلي لطريقة التدرج السلمي من قبل العديد من الباحثين لتحقيق بعض استراتيجيات البحث الخطي. لكن غالباً ما تشكل عبئاً للمشكلات واسعة النطاق لأنها تنطوي على حساب أكبر لقيم الدوال وتدرجاتها.

في هذه الأطروحة ، نبرهن تقارب طريقة النسب الهابط لعائلة من وسيطين لطرق التدرج السلمي ، باستخدام صيغة طول الخطوة الجديدة التي اقترحها وي كينغ [59] ، بالإضافة إلى النتائج العددية التي تؤكد فعالية الصيغة المقترحة ومقارنتها ببيانها مع طول الخطوة المقترحة من قبل صان وزملائه [7, 55]

الكلمات المفتاحية: التدرج السلمي، خوارزمية، التقارب الكلي، البحث الخطي غير الدقيق، معيار أرميجو، معيار ولف، طول الخطوة، الاتجاه الهابط، وسيطين، طريقة هستنس ستفان، طريقة فلتشر ريفز، طريقة بولاك ربيار بولاياكا، طريقة الهبوط السلمي، طريقة داي يوان.

Résumé

Les propriétés de la convergence globale de la méthode du gradient conjugué ont été étudiées par de nombreux chercheurs, généralement le pas vérifié certaines stratégies de recherche linéaire, mais la recherche par linéaire devient souvent un fardeau important pour les problèmes à grande échelle car elle implique un plus grand calcul des valeurs des fonctions et des gradients.

Dans cette thèse, nous démontrons la convergence de la méthode de descente d'une famille du deux paramètres de méthodes du gradient conjugué, en utilisant une nouvelle formule de pas suggérée par Wu [59]. Les résultats numériques sont présentés pour confirmer l'efficacité de la formule proposée et sa comparaison avec un pas suggérés par Sun et ses collègues [7,55].

Mots clés: Gradient conjugué, Algorithme, Convergence globale, Recherche linéaire inexacte, Règle d'Armijo, Règle de Wolfe, Taille de pas, Direction de descente, Méthode de Hestenes-Stiefel, Méthode de Fletcher-Reeves, Méthode de Polak-Ribière-Polyak, Méthode de la descente conjuguée, Méthode de Dai-Yuan.

Abstract

The global convergence properties of conjugate gradient method have been studied by many researchers, it is usually required that the step-size satisfy some line search strategies. But the line search often becomes a significant burden for large-scale problems for it involves more computation of functions values and gradients.

In this thesis, we prove the convergence of the method of descent of a family of two parameters of conjugate gradient methods, using a new step-size formula that was suggested by Wu [59]. Numerical results are presented to confirm the effectiveness of the proposed step-size by comparing with the step-size suggested by Sun and his colleagues [7, 55].

Key words: Conjugate gradient, Algorithm, Global convergence, Inexact line search, Armijo line search, Wolfe line search, Step-size, Directions of descent, Two parameters, Hestenes-Stiefel Method, Fletcher-Reeves Method, Polak-Ribière-Polyak Method, Conjugate descent Method, Dai-Yuan Method.

Table des matières

Introduction	iv
1 Optimisation sans contraintes	1
1.1 Concepts fondamentaux	1
1.1.1 Produit scalaire et normes vectorielles.	1
1.1.2 Normes matricielles	2
1.1.3 Orthogonalité dans \mathbb{R}^n	3
1.1.4 Dérivée directionnelle	4
1.2 Ensembles et fonctions convexes	4
1.2.1 Minima locaux et globaux	6
1.2.2 Théorèmes d'existence et d'unicité	7
1.3 Direction de descente	8
1.4 Théorèmes d'optimalité	10
1.4.1 Les conditions nécessaires	10
1.4.2 Les conditions suffisantes	11
1.5 Les problèmes de minimisation sans contraintes et leurs algo- rithmes	12
1.5.1 Convergence des algorithmes	13
1.5.2 Les modes de convergence	13
2 Méthodes à directions de descente	15
2.1 Stratégies de recherche d'un pas de descente	15
2.1.1 Les techniques utilisées	18
2.2 Description des méthodes à directions de descente	26
2.3 Convergence des algorithmes à directions de descente	27
2.3.1 Condition de Zoutendijk	27

2.3.2	Recherche linéaire d'Armijo	28
2.3.3	Recherche linéaire de Wolfe	30
2.4	Direction du gradient	31
2.4.1	Principe des méthodes du gradient	31
2.5	Direction du gradient conjugué	32
2.5.1	Le principe général d'une méthode à directions conjuguées	33
2.5.2	La méthode du gradient conjugué dans le cas quadratique	34
2.6	Direction de Newton	39
2.7	Direction de quasi Newton	41
2.7.1	Méthode de correction de rang un	42
2.7.2	Méthode de Davidon Fletcher Powell (DFP)	43
2.7.3	Méthode de Broyden, Fletcher, Goldfarb et Shanno (BFGS)	45
2.8	Direction de Gauss-Newton	47
2.8.1	Problèmes de moindres carrés	47
2.8.2	Algorithme de Gauss-Newton	49
2.8.3	Algorithme de Levenberg-Marquardt	51
3	Classes itératives à directions de descente	53
3.1	Méthodes du gradient conjugué non linéaire	53
3.2	Méthode de Fletcher-Reeves	55
3.2.1	Algorithme de la méthode de FR	55
3.2.2	La propriété de descente de la méthode de FR	56
3.2.3	Convergence de la méthode de FR	58
3.3	Méthode de Polak-Ribière-Polyak	61
3.3.1	Convergence de la méthode de PRP	61
3.4	Méthode de la descente conjuguée	63
3.4.1	La propriété de descente de la méthode de la DC	64
3.4.2	Convergence de la méthode de la DC	65
3.5	Méthode de Dai-Yuan	66
3.5.1	La propriété de descente de la méthode de DY	67
3.5.2	Convergence de la méthode de DY	68
3.6	Méthode de Hestenes et Stiefel HS	70
3.6.1	Convergence de la méthode HS	71
3.7	Méthode de Liu et Storey LS	72
3.8	Combinaison unifiée	72

4	La convergence de la méthode de descente d'une famille à deux paramètres des méthodes du gradient conjugué avec une formule fixe du pas	76
4.1	Introduction	76
4.2	Préliminaires	79
4.3	La convergence globale de la famille à deux paramètres	87
4.4	Tests numériques et discussions	93

Introduction

En optimisation différentiable, qui est une discipline d'analyse numérique en mathématiques étudiant en particulier les algorithmes minimisant des fonctions différentiables sur des ensembles, une direction de descente est une direction le long de laquelle la fonction à minimiser a une dérivée directionnelle strictement négative. Ces directions sont utilisées par les méthodes à directions de descente. C'est le long de ces directions qu'un déplacement est effectué afin de trouver l'itération suivant, en lequel la fonction à minimiser prend une valeur inférieure à celle qu'elle a en l'itération courant. Des directions de descente peuvent être calculées par de nombreuses techniques.

On peut citer la méthode du Gradient conjugué. Cette méthode est surtout utilisée pour les problèmes de grande taille.

Cette méthode a été découverte en 1952 par Hestenes et Steifel, pour la minimiser des fonctions quadratiques strictement convexes.

Plusieurs mathématiciens ont étendu cette méthode pour le cas non linéaire. Ceci a été réalisé pour la première fois, en 1964 par Fletcher et Reeves puis en 1969 par Polak, Ribière et Ployak. Une autre variante a été étudiée en 1987 par Fletcher. Une nouvelle variante a été proposée en 1991 par Liu et Storey. Et enfin une dernière variante qui a été étudiée en 1999 par Dai et Yuan.

Toutes ces méthodes génèrent une suite $\{x_k\}_{k \in \mathbb{N}}$ de la façon suivante :

$$x_{k+1} = x_k + \alpha_k d_k. \quad (0.1)$$

Le pas $\alpha_k \in \mathbb{R}$ est déterminé par une optimisation unidimensionnelle.

Les directions d_k sont calculées de façon récurrente par les formules suivantes :

$$d_k = \begin{cases} -g_1 & \text{si } k = 1 \\ -g_k + \beta_k d_{k-1} & \text{si } k \geq 2 \end{cases} \quad (0.2)$$

$g_k = \nabla f(x_k)$ et $\beta_k \in \mathbb{R}$.

Les différentes valeurs attribuées à β_k définissent les différentes formes du gradient conjugué.

Si on note $y_{k-1} = g_k - g_{k-1}$, on obtient les variantes suivantes :

$$\beta_k^{HS} = \frac{g_k^T (g_k - g_{k-1})}{d_{k-1}^T (g_k - g_{k-1})}, \quad \text{gradient conjugué variante Hestenes-Stiefel.} \quad (0.3)$$

$$\beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2}, \quad \text{gradient conjugué variante Fletcher- Reeves.} \quad (0.4)$$

$$\beta_k^{PRP} = \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2}, \quad \text{gradient conjugué variante Polak-Ribière-Polyak.} \quad (0.5)$$

$$\beta_k^{CD} = \frac{\|g_k\|^2}{-d_{k-1}^T g_{k-1}}, \quad \text{gradient conjugué variante descente conjugué.} \quad (0.6)$$

$$\beta_k^{LS} = \frac{g_k^T y_{k-1}}{-d_{k-1}^T g_{k-1}}, \quad \text{gradient conjugué variante de Liu-Storey.} \quad (0.7)$$

$$\beta_k^{DY} = \frac{\|g_k\|^2}{d_{k-1}^T y_{k-1}}, \quad \text{gradient conjugué variante de Dai-Yuan.} \quad (0.8)$$

Les propriétés de la convergence globale de la méthode du gradient conjugué ont été étudiées par de nombreux chercheurs. Pour établir les résultats de la convergence, il est généralement nécessaire que le pas α_k est vérifié certaines stratégies de recherche linéaire. Mais la recherche linéaire devient souvent un fardeau important pour les problèmes à grande échelle car elle implique un plus grand calcul des valeurs des fonctions et des gradients, par ailleurs Powell [49] a montré que la recherche linéaire exacte pourrait entraîner une séquence non convergente dans le cas de méthode de PR et méthode de HS. Récemment, un nouveau type de méthode de gradient conjugué est développé par Sun et Zhang [55]. Dans leurs méthode, le pas est calculé par une formule fixe au lieu de la recherche linéaire. La formule est donnée par

$$\alpha_k = -\delta g_k^T d_k / \|d_k\|_{Q_k}^2, \quad (0.9)$$

où $\|d_k\|_{Q_k} = \sqrt{d_k^T Q_k d_k}$, $\delta \in (0, \nu_{\min}/\tau)$, τ est une constante de Lipschitz de f et $\{Q_k\}$ est une matrices définies positives satisfaisant pour des constantes positives ν_{\min} et ν_{\max}

$$\nu_{\min} d^T d \leq d^T Q_k d \leq \nu_{\max} d^T d, \quad \forall k, \quad \forall d \in \mathbb{R}^n.$$

Ils ont prouvé que la formule ci-dessus pour α_k peut assurer la convergence globale pour la méthode de FR, la méthode de PR, la méthode de HS, la méthode de DY et la méthode CD. Chen et Sun [8, 55], ont montré qu'une famille de deux paramètres du gradient conjugué lequel α_k est calculé par la formule (0.9) est globalement convergent. Sun et Zhang [55], ont suggéré que Q_k puisse être mis à jour par une formule de quasi-Newton, comme BFGS et DFP, mais cela peut coûter plus de temps et de mémoire puisque le scalaire $d^T Q_k d$ est utilisé pour calculer le pas à chaque itération.

Dernièrement, Wu [59] a réussi à obtenir une formule de pas α_k sans matrice qui utilise à la fois la valeur de fonction disponible et l'information de gradient. La méthode est basée sur les méthodes de quasi-Newton dans [63] et [64]. Sous des hypothèses appropriées, cette formule est

$$\alpha_k = \frac{-\delta g_k^T d_k}{(\bar{g}_{k+1} - g_k)^T d_k + \gamma \theta_k}, \quad (0.10)$$

où

$$\theta_k = 6(f_k - \bar{f}_{k+1}) + 3(g_k + \bar{g}_{k+1})^T d_k,$$

f_k, g_k, \bar{f}_{k+1} et \bar{g}_{k+1} désignent $f(x_k), \nabla f(x_k), f(x_k + d_k)$, et $\nabla f(x_k + d_k)$, respectivement,

δ et γ sont des paramètres satisfaisant

$$\delta \in (0, \kappa/\tau), \text{ et } \gamma \geq 0 \text{ si } \tau = \kappa, \text{ ou } \gamma \in \left(0, \frac{\kappa - \delta\tau}{3(\tau - \kappa)}\right) \text{ si } \tau > \kappa.$$

Le but de cette thèse c'est l'utilisation de la nouvelle formule de pas (0.10) pour étudier la convergence globale d'une famille de deux paramètres :

$$\beta_k^{\mu_k, \omega_k} = \frac{g_k^T y_{k-1}}{(1 - \mu_k - \omega_k) \|g_{k-1}\|^2 + \mu_k d_{k-1}^T y_{k-1} - \omega_k d_{k-1}^T g_{k-1}},$$

sous les conditions de la direction de descente d_k telle que

$$g_k^T d_k = -c \|g_k\|^2,$$

où $0 < c < 1 - \mu_k - \omega_k$, et $\mu_k \in [0, 1[$ et $\omega_k \in [0, 1 - \mu_k[$.

Cette thèse comporte quatre chapitres.

Le premier chapitre est un rappel de notions préliminaires de base concernant l'optimisation sans contraintes.

Le second chapitre est une présentation générale des méthodes de descente lesquelles font intervenir deux principaux éléments : la direction de descente et le pas cherché de façon optimal.

Dans le troisième chapitre, toutes les classes de la méthode du gradient conjugué sont citées ainsi que toutes leurs propriétés.

Le quatrième chapitre est la partie principale de cette thèse, elle comporte un résultat original sur la convergence d'une famille de deux paramètres du gradient conjugué avec une nouvelle formule du pas, plus des résultats de calcul pour confirmer l'efficacité de cette formule.

Chapitre 1

Optimisation sans contraintes

Dans ce chapitre, on définit et on introduit les outils fonctionnels de base nécessaires pour l'optimisation sans contraintes.

1.1 Concepts fondamentaux

1.1.1 Produit scalaire et normes vectorielles.

Le produit scalaire de deux vecteurs x et y est

$$x^T y = \sum_{i=1}^{i=n} x_i y_i. \quad (1.1)$$

Dans le cas de vecteurs complexes, le produit scalaire hermitien est défini par

$$(x, y) \in \mathbb{C}^n \times \mathbb{C}^n \rightarrow x^T y = \sum_{i=1}^{i=n} \bar{x}_i y_i, \quad (1.2)$$

ou le surlignage d'une grandeur indique qu'on en considère le conjugué. Il est possible de définir plusieurs normes dans l'espace vectoriel \mathbb{R}^n .

Une norme d'un espace vectoriel E est une application $\|\cdot\|$ de E dans \mathbb{R}^+ qui vérifie les propriétés suivantes :

$$\begin{aligned} \forall x \in E, \|x\| &= 0 \Leftrightarrow x = 0 \\ \forall \lambda \in \mathbb{R}, x \in E, \|\lambda x\| &= |\lambda| \|x\| \\ \forall x, y \in E, \|x + y\| &\leq \|x\| + \|y\|. \end{aligned}$$

Dans \mathbb{R}^+ , les trois normes les plus courantes sont la norme infinie, la norme 1 et la norme euclidienne.

norme infinie :

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

norme 1 :

$$\|x\|_1 = \sum_{i=1}^{i=n} |x_i|$$

norme 2 ou norme euclidienne :

$$\|x\|_2 = \sqrt{x^T x} = \left(\sum_{i=1}^{i=n} x_i^2 \right)^{\frac{1}{2}},$$

ou $x = (x_1, \dots, x_n)^T$.

La norme euclidienne est donc définie par le produit scalaire $x^T y$.

1.1.2 Normes matricielles

Définition 1.1.1 Soit $\|\cdot\|$ une norme quelconque sur \mathbb{R}^n , on appelle norme matricielle subordonnée de la matrice carrée A d'ordre n

$$\|A\| = \max_{x \in \mathbb{R}^n, \|x\| \neq 0} \frac{\|Ax\|}{\|x\|} = \max_{x \in \mathbb{R}^n, \|x\|=1} \|Ax\| \quad (1.3)$$

Proposition 1.1.1 a) Une norme matricielle subordonnée vérifie les propriétés suivantes :

$$\begin{aligned} \|A\| &= 0 \Leftrightarrow A = 0 \text{ pour tout } A \\ \|\alpha A\| &= |\alpha| \|A\| \text{ pour tout } A, \alpha \in \mathbb{R} \\ \|A + B\| &\leq \|A\| + \|B\| \text{ pour tout } A, B \\ \|AB\| &\leq \|A\| \|B\| \text{ pour tout } A, B \end{aligned}$$

b) Soit A une matrice carrée réelle, pour les normes vectorielles usuelles on a les normes matricielles subordonnées.

$$\begin{aligned} \|A\|_\infty &= \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| = \text{maximum sur les lignes} \\ \|A\|_1 &= \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}| = \text{maximum sur les colonnes} \\ \|A\|_2 &= \sqrt{\rho(A^T A)} = \|A^T\|_2, \end{aligned}$$

où $\rho(A^T A)$ est le rayon spectral de la matrice symétrique $A^T A$

c) Si O et P sont des matrices orthogonales (i.e. $O^T = O^{-1}$ et $P^T = P^{-1}$), alors $\|OAP\|_2 = \|A\|_2$.

Si $AA^T = A^T A$ (i.e. A est normale), alors $\|A\|_2 = \rho(A)$

1.1.3 Orthogonalité dans \mathbb{R}^n

Définition 1.1.2

$$x \perp y \Leftrightarrow x^T y = 0 \quad (1.4)$$

Définition 1.1.3

$$\cos\langle x, y \rangle = \frac{x^T y}{\|x\|_2 \|y\|_2} \quad (1.5)$$

Proposition 1.1.2 *Théorème de Pythagore :*

$$x \perp y \Leftrightarrow \|x + y\|_2^2 = \|x\|_2^2 + \|y\|_2^2 \quad (1.6)$$

Définition 1.1.4 Soit S un sous-espace vectoriel de \mathbb{R}^n . L'orthogonal de S est S^\perp défini par :

$$S^\perp = \{y : y^T x = 0, \forall x \in S\}$$

Proposition 1.1.3 *Inégalité de Cauchy-Schwarz :*

$$\forall (x, y) \in (\mathbb{R}^n)^2 : |x^T y| \leq \|x\|_2 \|y\|_2, \quad (1.7)$$

l'égalité a lieu si et seulement si les vecteurs x et y sont liés

Définie positivité :

◇ **Matrice définie positive :**

$$A = A^T > 0 \Leftrightarrow x^T A x > 0, \forall x \neq 0. \quad (1.8)$$

Propriétés des matrices définies positives :

$$A = A^T > 0,$$

avec $A = \begin{bmatrix} a_{11} & \cdot & \cdot & \cdot & a_{1n} \\ \cdot & \cdot & & & \cdot \\ \cdot & & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot \\ a_{n1} & \cdot & \cdot & \cdot & a_{nn} \end{bmatrix} \Leftrightarrow$

1)

$$\lambda_i \{A\} > 0, i = 0, \dots, n$$

2)

$$\det \{A_k\} > 0, k = 0, \dots, n$$

$$\text{tq} : A_1 = a_{11}, A_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \dots, A_n = A$$

3) $A = LL^T$ par factorisation de Cholesky avec L inversible et triangulaire.

◇ **Matrices semi-définies positives :**

$$A = A^T \geq 0 \Leftrightarrow x^T A x \geq 0, \forall x \neq 0 \quad (1.9)$$

1.1.4 Dérivée directionnelle

Définition 1.1.5 [13]

On appelle *dérivée directionnelle* de f dans la direction d au point x , notée $\delta f(x, d)$, la limite (éventuellement $\pm\infty$) du rapport :

$$\frac{f(x + hd) - f(x)}{h} \text{ lorsque } h \text{ tend vers } 0.$$

Autrement dit :

$$\delta f(x, d) = \lim_{h \rightarrow 0} \frac{f(x + hd) - f(x)}{h} = \nabla^T f(x) d \quad (1.10)$$

○ Si $\|d\| = 1$: la *dérivée directionnelle* est le *d'accroissement* de f dans la direction d au point x .

Remarque 1.1.1 ◇ *Le taux d'accroissement est maximal dans la direction du gradient.*

 ◇ *Le gradient indique la direction de la plus grande pente.*

1.2 Ensembles et fonctions convexes

Définition 1.2.1 (*Ensemble convexe*).

Soit L l'ensemble $C \subseteq \mathbb{R}^n$. C est convexe si et seulement si

$$\forall (x, y) \in C^2, \forall \lambda \in [0, 1], \lambda x + (1 - \lambda)y \in C \quad (1.11)$$

Remarque 1.2.1 Soit $x_1, x_2, \dots, x_k \in \mathbb{R}^n$ et $\lambda_j \geq 0$ et $\sum_{j=1}^k \lambda_j = 1$. Toute expression de la forme $\sum_{j=1}^k \lambda_j x_j$. S' appelle combinaison convexe des points x_j ou barycentre.

Définition 1.2.2 (Fonction convexe).

Soit $C \subseteq \mathbb{R}^n$ un ensemble convexe non vide.

Une fonction $f : C \rightarrow \mathbb{R}$ est convexe si et seulement si

$$\forall (x, y) \in C^2, \forall \lambda \in [0, 1] : f(\lambda x + (1 - \lambda) y) \leq \lambda f(x) + (1 - \lambda) f(y). \quad (1.12)$$

On dit que f est une fonction concave si $-f$ est convexe. On dira que f est strictement convexe dans C si et seulement si

$$\forall (x, y) \in C^2, x \neq y, \forall \lambda \in [0, 1] : f(\lambda x + (1 - \lambda) y) < \lambda f(x) + (1 - \lambda) f(y) \quad (1.13)$$

Définition 1.2.3 (Fonction fortement ou uniformément convexe de module $v > 0$).

Soit $C \subseteq \mathbb{R}^n$ un ensemble convexe non vide. Une fonction $f : C \rightarrow \mathbb{R}$ est fortement ou uniformément convexe de module $v > 0$ si

$$f(\lambda x + (1 - \lambda) y) \leq \lambda f(x) + (1 - \lambda) f(y) - \frac{v}{2} \lambda (1 - \lambda) \|x - y\|^2, \forall x, y \in C^2, \forall \lambda \in [0, 1] \quad (1.14)$$

Définition 1.2.4 (Fonction convexe différentiable).

Soit $C \subseteq \mathbb{R}^n$, $f : \mathbb{R}^n \rightarrow \mathbb{R}$ et $\hat{x} \in \text{int}(C)$, f est dite différentiable au point \hat{x} , s'il existe un vecteur $A \in \mathbb{R}^n$ est une fonction $\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$ telle que

$$f(x) = f(\hat{x}) + A^T(x - \hat{x}) + \|x - \hat{x}\| \alpha(\hat{x}, x - \hat{x}), \quad (1.15)$$

où : $\alpha(\hat{x}, x - \hat{x}) \xrightarrow{x \rightarrow \hat{x}} 0$. On peut noter le vecteur A comme suit :

$$A = \nabla f(\hat{x}) = \left(\frac{\partial f(\hat{x})}{\partial x_1}, \dots, \frac{\partial f(\hat{x})}{\partial x_n} \right)^T. \quad (1.16)$$

Définition 1.2.5 (Fonction convexe deux fois différentiable).

Soit $C \subseteq \mathbb{R}^n$ non vide et $f : \mathbb{R}^n \rightarrow \mathbb{R}$, est dite deux fois différentiable ou point $\hat{x} \in \text{int}(C)$ s'il existe un vecteur $\nabla f(\hat{x})$ et une matrice symétrique $H(\hat{x})$ d'ordre (n,n) appelée matrice hessienne, et une fonction $\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$ telle que

$$\forall x \in C : f(x) = f(\hat{x}) + \nabla f(\hat{x})^T (x - \hat{x}) + \frac{1}{2} (x - \hat{x})^T H(\hat{x}) (x - \hat{x}) + \|x - \hat{x}\|^2 \alpha(\hat{x}, x - \hat{x}),$$

où : $\alpha(\hat{x}, x - \hat{x}) \xrightarrow{x \rightarrow \hat{x}} 0$. On peut écrire :

$$H(\hat{x}) = \begin{bmatrix} \frac{\partial^2 f(\hat{x})}{\partial x_1 \partial x_1} & \frac{\partial^2 f(\hat{x})}{\partial x_1 \partial x_2} & \dots & \dots & \frac{\partial^2 f(\hat{x})}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(\hat{x})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\hat{x})}{\partial x_2 \partial x_2} & \dots & \dots & \frac{\partial^2 f(\hat{x})}{\partial x_2 \partial x_n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \frac{\partial^2 f(\hat{x})}{\partial x_n \partial x_1} & \frac{\partial^2 f(\hat{x})}{\partial x_n \partial x_2} & \dots & \dots & \frac{\partial^2 f(\hat{x})}{\partial x_n \partial x_n} \end{bmatrix}$$

1.2.1 Minima locaux et globaux

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$. On appelle problème de minimisation sans contraintes le problème suivant :

$$\min \{ f(x) : x \in \mathbb{R}^n \} \tag{1.17}$$

Définition 1.2.6 1) $\hat{x} \in \mathbb{R}^n$ est un minimum local de (1.17) si et seulement si il existe un voisinage $V_\varepsilon(\hat{x})$ tel que

$$f(\hat{x}) \leq f(x) : \forall x \in V_\varepsilon(\hat{x}) \tag{1.18}$$

2) $\hat{x} \in \mathbb{R}^n$ est un minimum local strict de (1.17) si et seulement si il existe un voisinage $V_\varepsilon(\hat{x})$ tel que

$$f(\hat{x}) < f(x) : \forall x \in V_\varepsilon(\hat{x}), x \neq \hat{x} \tag{1.19}$$

3) $\hat{x} \in \mathbb{R}^n$ est un minimum global de (1.17) si et seulement si

$$f(\hat{x}) \leq f(x) : \forall x \in \mathbb{R}^n \tag{1.20}$$

Remarque 1.2.2 Dans le cas d'une fonction objectif convexe, il n'y a pas de distinction entre minimum local et global : tout minimum local est également global, comme l'établit le théorème suivant.

Théorème 1.2.1 Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction convexe définie sur l'ensemble convexe X . Alors, tout minimum local de f est également un minimum global. Si f est strictement convexe, alors il existe au plus un minimum global de f .

1.2.2 Théorèmes d'existence et d'unicité

Définition 1.2.7 Un ensemble $C \subset \mathbb{R}^n$ est dit compact si, de toute suite $\{x_k\}$, où $x_k \in C, \forall k$, on peut extraire une sous-suite convergente.

Théorème 1.2.2 Un ensemble $C \subset \mathbb{R}^n$ est compact si et seulement si il est fermé et borné .

Théorème 1.2.3 Si $f : C \subset \mathbb{R}^n \rightarrow \mathbb{R}$ est continue et si de plus C est un ensemble compact, alors le problème (1.17) admet une solution optimale $\hat{x} \in C$, qui vérifie donc

$$f(\hat{x}) \leq f(x), \forall x \in C.$$

Le second résultat est moins général car il considère le cas particulier $C = \mathbb{R}^n$:

Théorème 1.2.4 Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction continue sur \mathbb{R}^n . Si

$$\lim_{\|x\| \rightarrow \infty} f(x) = \infty,$$

alors (1.17) admet une solution optimale \hat{x} .

Preuve. Soit $x_0 \in \mathbb{R}^n$. Puisque $\lim_{\|x\| \rightarrow \infty} f(x) = \infty$ il existe $M > 0$ tel que $\|x\| > M \implies f(x) > f(x_0)$, donc

$$\exists M > 0, f(x) \leq f(x_0) \implies \|x\| \leq M.$$

Puisque \hat{x} est caractérisé par $f(\hat{x}) \leq f(x), \forall x \in \mathbb{R}^n$, on a donc forcément $\|\hat{x}\| \leq M$. Donc \hat{x} est solution du problème

$$\min_{\|x\| \leq M} f(x),$$

et le théorème précédent s'applique la boule $\{x \in \mathbb{R}^n, \|x\| \leq M\}$ étant compact ■

Théorème 1.2.5 Soit $f : C \subset \mathbb{R}^n \rightarrow \mathbb{R}$ strictement convexe sur C convexe. Le minimum de f sur C , s'il existe, est unique.

Preuve. Soit donc $\hat{x} \in C$ tel que $f(\hat{x}) \leq f(x), \forall x \in C$. Supposons qu'il existe $\hat{y} \neq \hat{x}$ tel que $f(\hat{y}) \leq f(x), \forall x \in C$. Formons pour $\lambda \in]0, 1[$ le vecteur

$$\mu = \lambda\hat{y} + (1 - \lambda)\hat{x}.$$

D'après la stricte convexité de f est puisque nécessairement

$$f(\hat{y}) = f(\hat{x}),$$

on a

$$f(\mu) < \lambda f(\hat{y}) + (1 - \lambda)f(\hat{x}) = f(\hat{x}),$$

ce qui contredit le fait que \hat{x} soit un minimum. On a donc $\hat{x} = \hat{y}$. ■

1.3 Direction de descente

Définition 1.3.1 Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$, d un vecteur non nul de \mathbb{R}^n est dit une direction de descente, s'il existe un réel $\delta > 0$ tel que pour tout λ dans l'intervalle $]0, \delta[$, on a

$$f(x + \lambda d) < f(x)$$

Théorème 1.3.1 Si f est différentiable au point $x \in \mathbb{R}^n$ et $d \in \mathbb{R}^n$ telle que

$$\nabla f(x)^T d < 0.$$

Alors d est une direction de descente en x

Remarque 1.3.1 d fait avec l'opposé du gradient $-\nabla f(x)$ un angle θ strictement plus petit que 90° :

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

L'ensemble des directions de descente de f en x , $\{d \in \mathbb{R}^n : \nabla f(x)^T \cdot d < 0\}$ forme un demi-espace ouvert de \mathbb{R}^n .

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

Théorème 1.3.2 *Soit f une fonction C^1 de \mathbb{R}^n dans \mathbb{R} . On suppose qu'il existe $\alpha > 0$ tel que*

$$\forall (x; y) \in \mathbb{R}^n \times \mathbb{R}^n : (\nabla f(x) - \nabla f(y), x - y) \geq \alpha \|x - y\|^2 \quad (1.21)$$

Alors f est strictement convexe et coercive, en particulier le problème (1.17) admet une solution unique.

Preuve. La condition (1.21) implique que ∇f est monotone et que f est convexe. De plus on a la stricte convexité de f . Enfin f est coercive : en effet, appliquons la formule de Taylor avec reste intégral

$$f(y) = f(x) + \int_0^1 \frac{d}{dt} f(x + t(y-x)) dt = f(x) + \int_0^1 (\nabla f(x + t'(y-x)), y-x) dt.$$

Donc

$$f(y) = f(x) + (\nabla f(x), y-x) + \int_0^1 (\nabla f(x + t(y-x)) - \nabla f(x), y-x) dt. \quad (1.22)$$

D'après (1.21) on obtien

$$f(y) \geq f(x) + (\nabla f(x), y-x) + \int_0^1 t\alpha \|x-y\|^2 dt.$$

Finalement

$$f(y) \geq f(x) - \|\nabla f(x)\| \cdot \|y-x\| + \frac{\alpha}{2} \|x-y\|^2.$$

Fixons $x = 0$ par exemple, il est alors clair que f est coercive. Par conséquent, f admet un minimum unique x^* sur \mathbb{R}^n caractérisé par $\nabla f(x^*) = 0$.

■

1.4 Théorèmes d'optimalité

1.4.1 Les conditions nécessaires

Théorème 1.4.1 *Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ telle que f est différentiable au point $\hat{x} \in \mathbb{R}^n$. Soit $d \in \mathbb{R}^n$ telle que $\nabla f(\hat{x})^T d < 0$. Alors il existe $\delta > 0$ tel que $f(\hat{x} + \alpha d) < f(\hat{x})$ pour tout $\alpha \in]0, \delta[$. La direction d s'appelle dans ce cas direction de descente.*

Preuve. [41] Comme f est différentiable en \hat{x} alors

$$f(\hat{x} + \alpha d) = f(\hat{x}) + \alpha \nabla f(\hat{x})^T d + \alpha \|d\| \lambda(\hat{x}, \alpha d),$$

où $\lambda(\hat{x}, \alpha d) \rightarrow 0$ pour $\alpha \rightarrow 0$. Ceci implique :

$$\frac{f(\hat{x} + \alpha d) - f(\hat{x})}{\alpha} = \nabla f(\hat{x})^T d + \|d\| \lambda(\hat{x}, \alpha d), \quad \alpha \neq 0,$$

et comme $\nabla f(\hat{x})^T d < 0$ et $\lambda(\hat{x}, \alpha d) \rightarrow 0$ pour $\alpha \rightarrow 0$, il existe $\delta > 0$ tel que

$$\nabla f(\hat{x})^T d + \|d\| \lambda(\hat{x}, \alpha d) < 0 \text{ pour tout } \alpha \in]0, \delta[,$$

et par conséquent on obtient :

$$f(\hat{x} + \alpha d) < f(\hat{x}) \text{ pour tout } \alpha \in]0, \delta[\tag{1.23}$$

■

Théorème 1.4.2 *Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ différentiable en \hat{x} , si \hat{x} est un minimum local alors $\nabla f(\hat{x}) = 0$.*

Preuve. Par contre, on suppose que $\nabla f(\hat{x}) \neq 0$. Si on pose $d = -\nabla f(\hat{x})$, on obtient :

$$\nabla f(\hat{x})^T d = -\|\nabla f(\hat{x})\|^2 < 0, \tag{1.24}$$

et par le théorème précédent, il existe $\delta > 0$ tel que :

$$f(\hat{x} + \alpha d) < f(\hat{x}) \text{ pour tout } \alpha \in]0, \delta[,$$

mais ceci est contradictoire avec le fait que \hat{x} est un minimum local, d'où $\nabla f(\hat{x}) = 0$ ■

Théorème 1.4.3 Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ deux fois différentiable en \hat{x} . Supposons que \hat{x} soit minimum local. Alors $\nabla f(\hat{x}) = 0$ et $H(\hat{x})$ est semi définie positive.

Preuve. Le théorème ci-dessus montre la première proposition, pour la deuxième proposition on a :

$$f(\hat{x} + \alpha d) = f(\hat{x}) + \alpha \nabla f(\hat{x})^T \cdot d + \frac{1}{2} \alpha^2 d^T H(\hat{x}) d + \alpha^2 \|d\|^2 \lambda(\hat{x}, \alpha d),$$

où $\lambda(\hat{x}, \alpha d) \rightarrow 0$ pour $\alpha \rightarrow 0$. Ceci implique :

$$\frac{f(\hat{x} + \alpha d) - f(\hat{x})}{\alpha^2} = \frac{1}{2} d^T H(\hat{x}) d + \|d\|^2 \lambda(\hat{x}, \alpha d), \alpha \neq 0. \quad (1.25)$$

Comme \hat{x} est un minimum local alors $f(\hat{x} + \alpha d) > f(\hat{x})$ pour α suffisamment petit, d'où

$$\frac{1}{2} d^T H(\hat{x}) d + \|d\|^2 \lambda(\hat{x}, \alpha d) \geq 0 \text{ pour } \alpha \text{ petit.}$$

En passant à la limite quand $\alpha \rightarrow 0$, on obtient que $d^T H(\hat{x}) d \geq 0$, d'où $H(\hat{x})$ est semi définie positive. ■

1.4.2 Les conditions suffisantes

Théorème 1.4.4 Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ deux fois différentiable en \hat{x} . Si $\nabla f(\hat{x}) = 0$ et $H(\hat{x})$ est définie positive, alors \hat{x} est minimum local strict.

Preuve. f est deux fois différentiable au point \hat{x} . Alors $\forall x \in \mathbb{R}^n$, on obtient :

$$f(x) = f(\hat{x}) + \nabla f(\hat{x})^T (x - \hat{x}) + \frac{1}{2} (x - \hat{x})^T H(\hat{x}) (x - \hat{x}) + \|x - \hat{x}\|^2 \lambda(\hat{x}, x - \hat{x}),$$

où $\lambda(\hat{x}, x - \hat{x}) \rightarrow 0$ pour $x \rightarrow \hat{x}$. Supposons que \hat{x} n'est pas un minimum local strict. Donc il existe une suite $\{x_k\}$ convergente vers \hat{x} telle que $f(x_k) \leq f(\hat{x})$, $x_k \neq \hat{x}$, $\forall k$.

Posons $d_k = (x_k - \hat{x}) / \|x_k - \hat{x}\|$. Donc $\|d_k\| = 1$ et on obtient à partir le théorème (1.4.2) :

$$\frac{f(x_k) - f(\hat{x})}{\|x_k - \hat{x}\|^2} = \frac{1}{2} d_k^T H(\hat{x}) d_k + \lambda(\hat{x}, x_k - \hat{x}) \leq 0, \forall k$$

et comme $\|d_k\| = 1, \forall k$ alors $\exists \{d_k\}_{k \in N_1 \subset \mathbb{N}}$ telle que $d_k \rightarrow d$ pour $k \rightarrow \infty$ et $k \in N_1$. On a bien sûr $\|d\| = 1$.

Considérons donc $\{d_k\}_{k \in N_1}$ et le fait que $\lambda(\hat{x}, x_k - \hat{x}) \rightarrow 0$ pour $k \rightarrow \infty$ et $k \in N_1$.

Alors le théorème (1.4.3) donne : $d_k^T H(\hat{x}) d_k \leq 0$, ce qui contredit le fait que $H(\hat{x})$ est définie positive car $\|d\| = 1$ (donc $d \neq 0$). Donc \hat{x} est un minimum local strict. ■

1.5 Les problèmes de minimisation sans contraintes et leurs algorithmes

L'étude de problème : $\min \{f(x) : x \in \mathbb{R}^n\}$ est importante pour des raisons diverses. Beaucoup des problèmes d'optimisation avec contraintes sont transformés en des suites de problèmes d'optimisation sans contraintes (multiplicateur de Lagrange, méthodes des pénalités, ...). L'étude des problèmes d'optimisation sans contraintes trouve aussi des applications dans la résolution des systèmes non linéaires. Les méthodes numériques de résolution de divers problèmes de minimisation sans contraintes ont pris ces dernières années un bel essor si bien que la bibliographie correspondante contient des centaines d'ouvrages et d'articles. Cet intérêt n'est nullement fortuit, il reflète le rôle de premier plan que les problèmes d'optimisation jouent dans les applications. La construction d'algorithmes consacrés à la recherche efficace de minimum d'une fonction sans contraintes est un problème complexe, car il ne suffit pas d'élaborer un algorithme il faut montrer de plus qu'il emporte sur ceux connus. On compare des algorithmes en se basant sur des plusieurs critères, par exemple, la précision du résultat, le nombre d'évaluations fonctionnelles et celui d'évaluations du gradient, la vitesse de la convergence, le temps de calcul, l'occupation de la mémoire nécessaire,

Même en se fixant des critères de comparaison, on ne peut pas classer les algorithmes ni en indiquant lequel est le meilleur ou le pire. Le fait est qu'on obtient les estimations de l'un des critères précédents pour des classes des problèmes, et un algorithme mauvais pour une vaste classe peut s'avérer efficace pour une autre, plus restreinte. L'utilisateur doit posséder tout un arsenal d'algorithmes pour être en mesure de faire face à chaque problème posé.

1.5.1 Convergence des algorithmes

Définition 1.5.1 (L’algorithme en optimisation)

Un algorithme de résolution est un procédé qui permet, à partir de la donnée du point initial x_0 , d’engendrer la suite $x_0, x_1, \dots, x_k, \dots$

Un algorithme est parfaitement défini par la donnée de l’application A qui à x_k associe à $x_{k+1} = A(x_k)$. Ceci permettra de confondre un algorithme et l’application A qui lui est associée.

Définition 1.5.2 (Fonction multivoque)

Une application multivoque est une application A qui à $x \in \mathbb{R}^n$ fait correspondre un sous ensemble $A(x)$ de \mathbb{R}^n .

Etant donné un point x_k : En appliquant les instructions d’un certain algorithme, on obtient un nouveau point x_{k+1} . Cette procédure peut être décrite par une application multivoque A appelée application algorithmique.

Donc étant donné un point $x_0 \in \mathbb{R}^n$, l’application algorithmique génère une suite x_0, x_1, \dots , où $x_{k+1} \in \mathbb{R}^n$ pour tout k .

Définition 1.5.3 (Convergence globale)

Nous dirons qu’un algorithme décrit par une application multivoque A , est globalement convergent si, quelque soit le point de départ x_0 choisi, la suite $\{x_k\}$ engendrée par $x_{k+1} \in A(x_k)$ (ou une sous suite) converge vers un point satisfaisant les conditions nécessaires d’optimalité (ou solution optimale).

1.5.2 Les modes de convergence

Définition 1.5.4 Soit $\{x_k\}_{k \in \mathbb{N}}$ une suite dans \mathbb{R}^n convergente vers x_* .

◆ Si

$$\limsup_{k \rightarrow \infty} \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} = \alpha < 1.$$

On dit que $\{x_k\}_{k \in \mathbb{N}}$ converge vers x_* linéairement avec le taux α .

◆ Si

$$\frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} \rightarrow 0 \text{ quand } k \rightarrow \infty.$$

On dit que la convergence est superlinéaire.

Plus précisément si $\exists \gamma > 1$ tel que :

$$\limsup_{k \rightarrow \infty} \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|^\gamma} < +\infty.$$

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

En particulier si

$$\limsup_{k \rightarrow \infty} \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|^2} < +\infty.$$

On dit que la convergence est quadratique (superlinéaire d'ordre 2).

Chapitre 2

Méthodes à directions de descente

Ce chapitre introduit une classe importante d'algorithmes de résolution des problèmes d'optimisation sans contraintes. Le concept central est celui de direction de descente. On le retrouvera dans des contextes variés, également pour résoudre des problèmes avec contraintes. Tous les algorithmes d'optimisation n'entrent pas dans ce cadre. Après avoir décrit comment fonctionne un algorithme à directions de descente, nous donnons quelques exemples d'algorithmes de ce type .

2.1 Stratégies de recherche d'un pas de descente

Considérons le problème d'optimisation sans contraintes P .

$$(P) : \min_{x \in \mathbb{R}^n} f(x),$$

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

Les algorithmes qu'on étudie par la suite suivant les schémas généraux suivants.

$$x_{k+1} = x_k + \lambda_k d_k, \tag{2.1}$$

où λ_k est solution optimale du problème d'optimisation unidimensionnel suivant :

$$\min_{\lambda > 0} f(x_k + \lambda d_k),$$

c'est -à- dire que λ_k vérifie

$$f(x_k + \lambda_k d_k) \leq f(x_k + \lambda d_k), \forall \lambda > 0,$$

x_k, d_k sont fixes et la fonction à minimiser est une fonction d'une variable réelle définie comme suit :

$$\lambda \rightarrow \varphi(\lambda) = f(x_k + \lambda d_k).$$

Il faut noter que dans les problèmes d'optimisation sans contraintes on a besoin de résoudre à chaque itération x_k , un problème d'optimisation dans \mathbb{R} .

Objectifs à atteindre

Il s'agit de réaliser deux objectifs :

Le premier objectif, faire décroître f suffisamment, et pour cela on cherche à vérifier l'inégalité :

$$f(x_k + \lambda_k d_k) \leq f(x_k) + \text{“un terme négatif”}. \quad (2.2)$$

Le terme négatif joue un rôle-clé dans la convergence de l'algorithme.

Le second objectif, on choisit le pas $\lambda_k > 0$ d'être trop petit, pour assurer la convergence d'algorithme au point stationnaire.

Le premier objectif n'est en effet pas suffisant car l'inégalité (2.2) 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, comme le montre l'observation suivante.

Si on prend

$$0 < \lambda_k \leq \frac{\varepsilon}{2^k \|d_k\|},$$

la suite $\{x_k\}$ générée par (2.2) est de Cauchy, puisque pour $1 \leq l < k$ on a

$$\|x_k - x_l\| = \left\| \sum_{i=l}^{k-1} \lambda_i d_i \right\| \leq \sum_{i=l}^{k-1} \frac{\varepsilon}{2^i} \rightarrow 0,$$

lorsque $l \rightarrow \infty$.

Donc $\{x_k\}$ converge, disons vers un point \bar{x} . En prenant $l = 1$ et $k \rightarrow \infty$ dans l'estimation ci-dessus, on voit que $\bar{x} \in \overline{B}(x_1, \epsilon)$ et donc \bar{x} ne saurait être solution s'il n'y a pas de solution dans $\overline{B}(x_1, \epsilon)$. On a donc arbitrairement forcé la convergence de $\{x_k\}$ en prenant des pas très petits.

Pour simplifier les notations, on définit la restriction de f à la droite $\{x_k + \lambda d_k : \lambda \in \mathbb{R}\}$ comme la fonction

$$\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}; \quad \lambda \longmapsto \varphi(\lambda) = f(x_k + \lambda d_k).$$

Schéma des 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 : $f : \mathbb{R}^n \rightarrow \mathbb{R}$

$$\nabla^T f(x_k) d_k < 0.$$

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.

Elles reviennent à déterminer, par tâtonnement un intervalle $[a, b]$, où $\lambda^* \in [a, b]$, dans lequel :

$$f(x_k + \lambda_k d_k) < f(x_k).$$

Le schéma de l'algorithme est donc :

Algorithme (Schéma général des recherches linéaires inexactes)

Etape 0 : (initialisation)

$a_1 = b_1 = 0$, choisir $\lambda_1 > 0$, poser $k = 1$ et aller à l'étape 1 :

Etape 1 :

Si λ_k est satisfaisant (suivant un certain critère) : STOP ($\lambda^* = \lambda_k$).

Si λ_k est trop petit (suivant un certain critère) : nouvel intervalle : $[a_{k+1} = \lambda_k, b_{k+1} = b]$, et aller à l'étape 2.

Si λ_k est trop grand (suivant un certain critère) : nouvel intervalle : $[a_{k+1} = a, b_{k+1} = \lambda_k]$,

et aller à l'étape 2.

Etape 2 :

Si $b_{k+1} = 0$ déterminer $\lambda_{k+1} \in]a_{k+1}, +\infty[$.

Si $a_{k+1} \neq 0$ déterminer $\lambda_{k+1} \in]a_{k+1}, b_{k+1}[$,
remplacer k par $k + 1$ et aller à l'étape 1.

Il nous reste donc à décider selon quel(s) critère(s) est trop petit ou trop grand ou satisfaisant.

2.1.1 Les techniques utilisées

Définition 2.1.1 *On dit que $[a, b]$ est un intervalle de sécurité s'il permet de classer les valeurs de λ de la façon suivante :*

- ◆ Si $\lambda < a$ alors λ est considéré trop petit.
- ◆ Si $a \leq \lambda \leq b$ alors λ est satisfaisant.
- ◆ Si $\lambda > b$ alors λ est considéré trop grand.

Le problème est de traduire de façon numérique sur φ les trois conditions précédentes, ainsi que de trouver un algorithme permettant de déterminer a et b .

Algorithme de base

Etape 0 : (initialisation)

$a = b = 0$, choisir $\lambda_1 > 0$, poser $k = 1$ et aller à l'étape 1 ;

Etape 1 :

Si λ_k convient, poser $\lambda^* = \lambda_k$ et on s'arrête.

Si λ_k est trop petit on prend $a_{k+1} = \lambda_k$, $b_{k+1} = b$,
et on va à l'étape 2.

Si λ_k est trop grand on prend $b_{k+1} = \lambda_k$, $a_{k+1} = a$,
et on va à l'étape 2.

Etape 2 :

Si $b_{k+1} = 0$ déterminer $\lambda_{k+1} \in]a_{k+1}, +\infty[$.

Si $b_{k+1} \neq 0$ déterminer $\lambda_{k+1} \in]a_{k+1}, b_{k+1}[$,
remplacer k par $k + 1$ et aller à l'étape 1.

Il faut maintenant préciser quelles sont les relations sur φ qui vont nous permettre de caractériser les valeurs de λ convenables, ainsi que les techniques utilisées pour réduire l'intervalle.

La règle d'Armijo

Une condition naturelle est de demander que f décroisse autant qu'une portion $\omega_1 \in]0, 1[$ de ce que ferait le modèle linéaire de f en x_k . Cela conduit à l'inégalité suivante, parfois appelée condition d'Armijo ou condition de décroissance linéaire :

$$f(x_k + \lambda_k d_k) \leq f(x_k) + \omega_1 \lambda_k \nabla^T f(x_k) d_k. \quad (2.3)$$

Elle est de la forme (2.1), car ω_1 devra être choisi dans $]0, 1[$.

$$\lambda \rightarrow f(x_k) + \omega_1 \lambda \nabla^T f(x_k) d_k. \quad (2.4)$$

Règle d'Armijo

- ◆ Si $\varphi(\lambda) \leq \varphi(0) + \omega_1 \varphi'(0)\lambda$, alors λ convient.
- ◆ Si $\varphi(\lambda) > \varphi(0) + \omega_1 \varphi'(0)\lambda$, alors λ est trop grand.

On peut noter que l'on a

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

Algorithme (Règle d'Armijo)

Étape 0 : (initialisation)

$a_1 = b_1 = 0$, choisir $\lambda_1 > 0$, $\omega_1 \in]0, 1[$ poser $k = 1$ et aller à l'étape 1.

Étape 1 :

Si $\varphi(\lambda_k) \leq \varphi(0) + \omega_1 \varphi'(0)\lambda_k$: STOP ($\lambda^* = \lambda_k$).

Si $\varphi(\lambda_k) > \varphi(0) + \omega_1 \varphi'(0)\lambda_k$, alors

$b_{k+1} = b$, $a_{k+1} = \lambda_k$ et aller à l'étape 2.

Étape 2 :

Si $b_{k+1} = 0$ déterminer $\lambda_{k+1} \in] a_{k+1}, +\infty [$.

Si $b_{k+1} \neq 0$ déterminer $\lambda_{k+1} \in] a_{k+1}, b_{k+1}[$,

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

Remarque 2.1.1 1) *En pratique, la constante ω_1 est prise très petite, de manière à satisfaire (2.3) le plus facilement possible. Typiquement, $\omega_1 = 10^{-4}$. Notons que cette constante ne doit pas être adaptée aux données du problème et donc que l'on ne se trouve pas devant un choix de valeur délicat.*

2) Dans certains algorithmes, il est important de prendre $\omega_1 < \frac{1}{2}$ pour que le pas λ_k soit accepté lorsque x_k est proche d'une solution.

Dans le théorème suivant on va assurer l'existence du pas d'Armijo en posant quelques conditions sur la fonction φ

Théorème 2.1.1 *Si $\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}$, définie par $\varphi(\lambda) = f(x_k + \lambda d_k)$ est continue et bornée inférieurement, si d_k est une direction de descente en x_k ($\varphi'(\lambda) < 0$) et si $\omega_1 \in]0, 1[$, alors l'ensemble des pas vérifiant la règle d'Armijo est non vide.*

Preuve. [39] on a

$$\begin{aligned}\varphi(\lambda) &= f(x_k + \lambda d_k) \\ \psi(\lambda) &= f(x_k) + \omega_1 \lambda \nabla^T f(x_k) d_k.\end{aligned}$$

Le développement de Taylor-Yong en $\lambda = 0$ de φ est :

$$\varphi(\lambda) = f(x_k + \lambda d_k) = f(x_k) + \omega_1 \lambda \nabla^T f(x_k) d_k + \lambda \xi(\lambda) \quad \text{où } \xi(\lambda) \rightarrow 0, \lambda \rightarrow 0,$$

et comme $\omega_1 \in]0, 1[$ et $\varphi'(0) = \nabla^T f(x_k) d_k < 0$ on déduit :

$$f(x_k) + \lambda \nabla^T f(x_k) d_k \leq f(x_k) + \omega_1 \lambda \nabla^T f(x_k) d_k \quad \text{pour } \lambda > 0.$$

On voit que pour $\lambda > 0$ assez petit on a :

$$\varphi(\lambda) < \psi(\lambda).$$

De ce qui précède et du fait que φ est bornée inférieurement, et $\psi(\lambda) \rightarrow -\infty, \lambda \rightarrow +\infty$, on déduit que la fonction $\psi(\lambda) - \varphi(\lambda)$ a la propriété :

$$\begin{cases} \psi(\lambda) - \varphi(\lambda) > 0 & \text{pour } \lambda \text{ assez petit,} \\ \psi(\lambda) - \varphi(\lambda) < 0 & \text{pour } \lambda \text{ assez grand,} \end{cases}$$

donc s'annule au moins une fois pour $\lambda > 0$.

En choisissant le plus petit de ces zéros on voit qu'il existe $\bar{\lambda} > 0$ tel que

$$\varphi(\bar{\lambda}) = \psi(\bar{\lambda}) \quad \text{et } \varphi(\lambda) < \psi(\lambda) \quad \text{pour } 0 < \lambda < \bar{\lambda}.$$

Ce qui achève la preuve. ■

La règle de Goldstein

La règle de Goldstein remédie à cet inconvénient (le pas λ_k doit être trop petit).

Dans celle-ci, en ajoutant une deuxième inégalité à la règle d'Armijo on obtient la règle de Goldstein.

$$f(x_k) + \omega_1 \lambda_k \nabla^T f(x_k) d_k \geq f(x_k + \lambda_k d_k) \geq f(x_k) + \omega_2 \lambda_k \nabla^T f(x_k) d_k, \quad (2.5)$$

où ω_1 et ω_2 sont deux constantes vérifiant $0 < \omega_1 < \omega_2 < 1$, cette inégalité qui empêche le pas d'être trop petit.

Règle de Goldstein

- ◆ Si $\varphi(\lambda) < \varphi(0) + \omega_2 \varphi'(0)\lambda$, alors λ est trop petit.
- ◆ Si $\varphi(\lambda) > \varphi(0) + \omega_1 \varphi'(0)\lambda$, alors λ est trop grand.
- ◆ Si $\varphi(0) + \omega_1 \varphi'(0)\lambda \geq \varphi(\lambda) \geq \varphi(0) + \omega_2 \varphi'(0)\lambda$, alors λ convient.

On peut noter que l'on a

$$\begin{aligned} \varphi(\lambda) &= f(x_k + \lambda d_k) \\ \varphi(0) &= f(x_k). \end{aligned}$$

Algorithme (Règle de Goldstein & Price)

Étape 0 :

$a_1 = b_1 = 0$, choisir $\lambda_1 > 0$, $\omega_1 \in]0, 1[$, $\omega_2 \in]\omega_1, 1[$, poser $k = 1$ et aller à l'étape 1.

Étape 1 :

Si $\varphi(0) + \omega_2 \varphi'(0)\lambda \leq \varphi(\lambda_k) \leq \varphi(0) + \omega_1 \varphi'(0)\lambda_k$: STOP ($\lambda^* = \lambda_k$).

Si $\varphi(\lambda_k) > \varphi(0) + \omega_1 \varphi'(0)\lambda_k$, alors

$b_{k+1} = \lambda_k$, $a_{k+1} = a_k$, et aller à l'étape 2.

Si $\varphi(\lambda_k) < \varphi(0) + \omega_2 \varphi'(0)\lambda_k$, alors

$b_{k+1} = b_k$, $a_{k+1} = \lambda_k$, et aller à l'étape 2.

Étape 2 :

Si $b_{k+1} = 0$ déterminer $\lambda_{k+1} \in]b_{k+1}, +\infty[$

Si $b_{k+1} \neq 0$ déterminer $\lambda_{k+1} \in]a_{k+1}, b_{k+1}[$.

Théorème 2.1.2 *Si $\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}$, définie par $\varphi(\lambda) = f(x_k + \lambda d_k)$ est continue et bornée inférieurement, si d_k est une direction de descente en x_k et si $\omega_1 \in]0, 1[$, $\omega_2 \in]\omega_1, 1[$, alors l'ensemble des pas vérifiant la règle de Goldstein & Price est non vide.*

Preuve. on a

$$\begin{aligned}\varphi(\lambda) &= f(x_k + \lambda d_k) \\ \psi_{\omega_1}(\lambda) &= f(x_k) + \omega_1 \lambda \nabla^T f(x_k) d_k \\ \psi_{\omega_2}(\lambda) &= f(x_k) + \omega_2 \lambda \nabla^T f(x_k) d_k.\end{aligned}$$

Le développement de Taylor-Yong en $\lambda = 0$ de φ est :

$$\varphi(\lambda) = f(x_k + \lambda d_k) = f(x_k) + \omega_1 \lambda \nabla^T f(x_k) d_k + \lambda \xi(\lambda) \quad \text{où } \xi(\lambda) \rightarrow 0, \lambda \rightarrow 0,$$

et comme $\omega_1 \in]0, 1[$ et $\varphi'(0) = \nabla^T f(x_k) d_k < 0$ on déduit :

$$f(x_k) + \lambda \nabla^T f(x_k) d_k < f(x_k) + \omega_2 \lambda \nabla^T f(x_k) d_k < f(x_k) + \omega_1 \lambda \nabla^T f(x_k) d_k \quad \text{pour } \lambda > 0.$$

On voit que pour $\lambda > 0$ assez petit on a :

$$\varphi(\lambda) < \psi_{\omega_2}(\lambda) < \psi_{\omega_1}(\lambda).$$

De ce qui précède et du fait que φ est bornée inférieurement, et $\psi_{\omega_1}(\lambda) \rightarrow -\infty, \lambda \rightarrow +\infty$, on déduit que la fonction $\psi_{\omega_1}(\lambda) - \varphi(\lambda)$ a la propriété :

$$\begin{cases} \psi_{\omega_1}(\lambda) - \varphi(\lambda) > 0 \text{ pour } \lambda \text{ assez petit,} \\ \psi_{\omega_1}(\lambda) - \varphi(\lambda) < 0 \text{ pour } \lambda \text{ assez grand,} \end{cases}$$

donc s'annule au moins une fois pour $\lambda > 0$.

En choisissant le plus petit de ces zéros on voit qu'il existe $\bar{\lambda} > 0$ tel que

$$\varphi(\bar{\lambda}) = \psi_{\omega_1}(\bar{\lambda}) \quad \text{et } \varphi(\lambda) < \psi_{\omega_1}(\lambda) \quad \text{pour } 0 < \lambda < \bar{\lambda}.$$

De la même manière, il existe $\hat{\lambda} > 0$ tel que :

$$\varphi(\hat{\lambda}) = \psi_{\omega_2}(\hat{\lambda}) \quad \text{et } \varphi(\lambda) < \psi_{\omega_2}(\lambda) \quad \text{pour } 0 < \lambda < \hat{\lambda},$$

et comme $\psi_{\omega_2}(\lambda) < \psi_{\omega_1}(\lambda)$ pour $\lambda > 0$, forcément $\hat{\lambda} < \bar{\lambda}$ et $\lambda = \bar{\lambda}$ satisfait (2.5)

$$\begin{aligned}\psi_{\omega_2}(\hat{\lambda}) = \varphi(\hat{\lambda}) &< \psi_{\omega_1}(\lambda), \text{ n'est autre que} \\ f(x_k) + \omega_2 \hat{\lambda} \nabla^T f(x_k) d_k &= f(x_k + \hat{\lambda} d_k) < f(x_k) + \omega_1 \hat{\lambda} \nabla^T f(x_k) d_k.\end{aligned}$$

Ce qu'il fallait démontrer ■

La règle 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'évaluation de $f(x)$, c'est pourquoi cette règle est très utilisée.

Nous allons présenter les conditions de Wolfe faibles sur $\lambda > 0$:

Règle de wolfe faible :

$$f(x_k + \lambda d_k) \leq f(x_k) + \omega_1 \lambda \nabla f(x_k)^T . d_k, \quad (W1) \quad (2.6)$$

$$\nabla f(x_k + \lambda d_k)^T . d_k \geq \omega_2 \nabla f(x_k)^T . d_k, \quad (W2) \quad (2.7)$$

avec $0 < \omega_1 < \omega_2 < 1$.

Règle de Wolfe faible

- ◆ Si $\varphi(\lambda) \leq \varphi(0) + \omega_1 \varphi'(0)\lambda$ et $\varphi'(\lambda) \geq \omega_2 \varphi'(0)$, alors λ convient.
- ◆ Si $\varphi(\lambda) > \varphi(0) + \omega_1 \varphi'(0)\lambda$, alors λ est trop grand.
- ◆ Si $\varphi'(\lambda) < \omega_2 \varphi'(0)$, alors λ est trop petit.

Algorithme (Règle de Wolfe)

Etape 0 : (initialisation)

$a_1 = b_1 = 0$, choisir $\lambda_1 > 0$, $\omega_1 \in]0, 1[$, $\omega_2 \in]\omega_1, 1[$, poser $k = 1$ et aller à l'étape 1.

Etape 1 :

Si $\varphi(\lambda_k) \leq \varphi(0) + \omega_1 \varphi'(0)\lambda_k$ et $\varphi'(\lambda) \geq \omega_2 \varphi'(0)$: STOP($\lambda^* = \lambda_k$).

Si $\varphi(\lambda_k) > \varphi(0) + \omega_1 \varphi'(0)\lambda_k$, alors

$b_{k+1} = \lambda_k$, $a_{k+1} = a_k$, et aller à l'étape 2.

Si $\varphi'(\lambda) < \omega_2 \varphi'(0)$, alors

$b_{k+1} = b_k$, $a_{k+1} = \lambda_k$, et aller à l'étape 2.

Etape 2 :

Si $b_{k+1} = 0$ déterminer $\lambda_{k+1} \in]a_{k+1}, +\infty[$.

Si $b_{k+1} \neq 0$ déterminer $\lambda_{k+1} \in]a_{k+1}, b_{k+1}[$.

Règle de Wolfe Forte : On obtient des contraintes plus fortes si l'on remplace (W2) par

$$|\nabla f(x_k + \lambda d_k)^T \cdot d_k| \leq -\omega_2 \nabla f(x_k)^T \cdot d_k, \quad (W3). \quad (2.8)$$

Les (W1) et (W3) sont les conditions de Wolfe fortes. La contrainte (W3) entraîne que $\omega_2 \varphi'(0) \leq \varphi'(\lambda) \leq -\omega_2 \varphi'(0)$ c-à-d. $\varphi'(\lambda)$ n'est pas "trop" positif.

Remarque 2.1.2 *On voit bien que les conditions de Wolfe fortes impliquent les conditions de Wolfe faibles, en effet :*

$$\begin{aligned} |\nabla^T f(x_k + \lambda_k d_k) d_k| &\leq -\omega_2 \nabla^T f(x_k) d_k \\ \Leftrightarrow \omega_2 \nabla^T f(x_k) d_k &\leq \nabla^T f(x_k + \lambda_k d_k) d_k \leq -\omega_2 \nabla^T f(x_k) d_k \\ \Rightarrow \omega_2 \nabla^T f(x_k) d_k &\leq \nabla^T f(x_k + \lambda_k d_k) d_k. \end{aligned} \quad (W2)$$

La règle de Wolfe relaxée : Proposée par *Dai* et *Yuan* [1996], cette règle consiste à choisir le pas satisfaisant aux conditions :

$$f(x_k + \lambda d_k) \leq f(x_k) + \omega_1 \lambda \nabla^T f(x_k) d_k, \quad (W4) \quad (2.9)$$

$$\omega_2 \nabla^T f(x_k) d_k \leq \nabla^T f(x_k + \lambda_k d_k) d_k \leq -\omega_2 \nabla^T f(x_k) d_k, \quad (W5) \quad (2.10)$$

où $0 < \omega_1 < \omega_2 < 1$ et $\omega_2 > 0$.

Remarque 2.1.3 1) *On voit bien que les conditions de Wolfe relaxées impliquent les conditions de Wolfe fortes. Effectivement (W4) est équivalente à (W1), tandis que pour le cas particulier $\omega_2 = \omega_2 = \omega_2$, (W5) est équivalente à (W3). En effet :*

$$\begin{aligned} \omega_2 \nabla^T f(x_k) d_k &\leq \nabla^T f(x_k + \lambda_k d_k) d_k \leq -\omega_2 \nabla^T f(x_k) d_k \\ \Rightarrow \omega_2 \nabla^T f(x_k) d_k &\leq \nabla^T f(x_k + \lambda_k d_k) d_k \leq -\omega_2 \nabla^T f(x_k) d_k \\ \Rightarrow |\nabla^T f(x_k + \lambda_k d_k) d_k| &\leq -\omega_2 \nabla^T f(x_k) d_k. \end{aligned} \quad (W3)$$

2) *Les conditions de Wolfe relaxées impliquent les conditions de Wolfe faibles. Effectivement (W4) est équivalente à (W1), tandis que pour le cas particulier $\omega_2 = \omega_2$ et $\omega_2 = +\infty$, (W5) est équivalente à (W2). En effet :*

$$\begin{aligned} \omega_2 \nabla^T f(x_k) d_k &\leq \nabla^T f(x_k + \lambda_k d_k) d_k \leq -\omega_2 \nabla^T f(x_k) d_k \\ \Rightarrow \omega_2 \nabla^T f(x_k) d_k &\leq \nabla^T f(x_k + \lambda_k d_k) d_k. \end{aligned} \quad (W2)$$

Dans le théorème suivant on va assurer l'existence du pas de Wolfe en posant quelques conditions sur la fonction φ :

Théorème 2.1.3 *Si $\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}$, définie par $\varphi(\lambda) = f(x_k + \lambda d_k)$ est dérivable et bornée inférieurement, si d_k est une direction de descente en x_k et si $\omega_1 \in]0, 1[$, $\omega_2 \in]\omega_1, 1[$, alors l'ensemble des pas vérifiant la règle de Wolfe (faible) est non vide.*

Preuve. on a

$$\begin{aligned}\varphi(\lambda) &= f(x_k + \lambda d_k) \\ \psi_{\omega_1}(\lambda) &= f(x_k) + \omega_1 \lambda \nabla^T f(x_k) d_k.\end{aligned}$$

Le développement de Taylor-Yong en $\lambda = 0$ de φ est :

$$\varphi(\lambda) = f(x_k + \lambda d_k) = f(x_k) + \omega_1 \lambda \nabla^T f(x_k) d_k + \lambda \xi(\lambda) \quad \text{où } \xi(\lambda) \rightarrow 0, \lambda \rightarrow 0,$$

et comme $\omega_1 \in]0, 1[$ et $\varphi'(0) = \nabla^T f(x_k) d_k < 0$ on déduit :

$$f(x_k) + \lambda \nabla^T f(x_k) d_k < f(x_k) + \omega_1 \lambda \nabla^T f(x_k) d_k \quad \text{pour } \lambda > 0.$$

On voit que pour $\lambda > 0$ assez petit on a :

$$\varphi(\lambda) < \psi_{\omega_1}(\lambda).$$

De ce qui précède et du fait que φ est bornée inférieurement, et $\psi_{\omega_1}(\lambda) \rightarrow -\infty, \lambda \rightarrow +\infty$, on déduit que la fonction $\psi_{\omega_1}(\lambda) - \varphi(\lambda)$ a la propriété :

$$\begin{cases} \psi_{\omega_1}(\lambda) - \varphi(\lambda) > 0 \text{ pour } \lambda \text{ assez petit,} \\ \psi_{\omega_1}(\lambda) - \varphi(\lambda) < 0 \text{ pour } \lambda \text{ assez grand,} \end{cases}$$

donc s'annule au moins une fois pour $\lambda > 0$.

En choisissant le plus petit de ces zéros on voit qu'il existe $\bar{\lambda} > 0$ tel que

$$\varphi(\bar{\lambda}) = \psi_{\omega_1}(\bar{\lambda}) \quad \text{et } \varphi(\lambda) < \psi_{\omega_1}(\lambda) \quad \text{pour } 0 < \lambda < \bar{\lambda}. \quad (2.11)$$

La formule des accroissements finis fournit alors un nombre $\hat{\lambda}$, $0 < \hat{\lambda} < \bar{\lambda}$ tel que

$$\begin{aligned}\varphi(\bar{\lambda}) - \varphi(0) &= \bar{\lambda} \varphi'(\hat{\lambda}) = \bar{\lambda} \nabla^T f(x_k + \hat{\lambda} d_k) d_k \\ &\Rightarrow \omega_1 \bar{\lambda} \nabla^T f(x_k) d_k = \bar{\lambda} \nabla^T f(x_k + \hat{\lambda} d_k) d_k \\ &\Rightarrow \nabla^T f(x_k + \hat{\lambda} d_k) d_k = \omega_1 \nabla^T f(x_k) d_k \geq \omega_2 \nabla^T f(x_k) d_k,\end{aligned}$$

car $0 < \omega_1 < \omega_2 < 1$ et $\nabla^T f(x_k)d_k < 0$.

Donc $\hat{\lambda}$ satisfait (2.10). D'autre part, $\lambda = \hat{\lambda}$ satisfait (2.9), en effet, $\hat{\lambda}$ satisfait (2.11)

$$\begin{aligned} \varphi(\hat{\lambda}) &< \psi_{\omega_1}(\lambda), \text{ n'est autre que} \\ f(x_k + \hat{\lambda}d_k) &< f(x_k) + \omega_1 \hat{\lambda} \nabla^T f(x_k)d_k. \end{aligned}$$

Ce qu'il fallait démontrer ■

2.2 Description des méthodes à directions de descente

Les méthodes à directions de descentes utilisent l'idée suivante pour minimiser une fonction [7]. Elles construisent la suite des itérés $\{x_k\}_{k \geq 1}$, approchant une solution x_k de (1.17) par la récurrence :

$$x_{k+1} = x_k + \alpha_k d_k, \text{ pour } k \geq 1.$$

Où α_k est appelé le pas et d_k la direction de descente de f en x_k .

Pour définir une méthode à directions de descente il faut donc spécifier deux choses :

◇ dire comment la direction d_k est calculée ; la manière de procéder donne le nom à l'algorithme ;

◇ dire comment on détermine le pas α_k ; c'est ce que l'on appelle : la recherche linéaire.

Décrivons cette classe d'algorithmes de manière précise.

Algorithme (méthode à directions de descente- une itération)

Etape 0 : (initialisation)

On suppose qu'au début de l'itération k , on dispose $x_k \in \mathbb{R}^n$

Etape 1 :

Test d'arrêt : si $\|\nabla f(x_k)\| \simeq 0$, 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 $\alpha_k > 0$ le long de d_k de manière à "faire décroître f suffisamment" ;

Etape 4 :

Si la recherche linéaire réussit : $x_{k+1} = x_k + \alpha_k d_k$; remplacer k par $k + 1$ et aller à l'étape 1.

Exemples de méthodes à directions de descente

Supposons que d_k soit une direction de descente au point x_k . Ceci nous permet de considérer le point x_{k+1} , successeur de x_k de la manière suivante :

$$x_{k+1} = x_k + \alpha_k d_k, \quad \alpha_k \in]0, +\delta[.$$

Vu la définition de direction de descente, on est assuré que

$$f(x_{k+1}) = f(x_k + \alpha_k d_k) < f(x_k).$$

Un bon choix de d_k et de α_k permet ainsi de construire une multitude d'algorithmes d'optimisation

Exemples de choix de directions de descente

par exemple si on choisit $d_k = -\nabla f(x_k)$ et si $\nabla f(x_k) \neq 0$, on obtient la méthode du gradient.

Bien sur $d_k = -\nabla f(x_k)$ est une direction de descente puisque $\nabla f(x_k)^T \cdot d_k = -\nabla f(x_k)^T \nabla f(x_k) = -\|\nabla f(x_k)\|^2 < 0$.

La méthode de Newton correspond à $d_k = -(H(x_k))^{-1} \cdot \nabla f(x_k)$, d_k direction de descente si la matrice hessienne $H(x_k)$ est définie positive.

Exemples de choix de pas α_k

On choisit en général α_k de façon optimale, c'est à dire que α_k doit vérifier

$$f(x_k + \alpha_k d_k) \leq f(x_k + \alpha d_k) : \forall \alpha \in [0, +\infty[.$$

En d'autres termes on est ramené à étudier à chaque itération un problème de minimisation d'une variable réelle. C'est ce qu'on appelle recherche linéaire.

2.3 Convergence des algorithmes à directions de descente

2.3.1 Condition de Zoutendijk

Dans cette section on va étudier la contribution de la recherche linéaire inexacte à 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

$$f(x_{k+1}) \leq f(x_k) - C \|\nabla f(x_k)\|^2 \cos^2 \theta_k, \quad (2.12)$$

où θ_k est l'angle que fait d_k avec $-\nabla f(x_k)$, défini par

$$\cos \theta_k = \frac{-\nabla^T f(x_k) d_k}{\|\nabla f(x_k)\| \|d_k\|}.$$

Voici comment on se sert de la condition de Zoutendijk.

Proposition 2.3.1 [31] *Si la suite $\{x_k\}$ générée par un algorithme d'optimisation vérifie la condition de Zoutendijk (2.12) et si la suite $\{f(x_k)\}$ est minorée, alors*

$$\sum_{k \geq 1} \|\nabla f(x_k)\|^2 \cos^2 \theta_k < \infty. \quad (2.13)$$

Preuve. En sommant les inégalités (2.12), on a

$$\sum_{k \geq 1} \|\nabla f(x_k)\|^2 \cos^2 \theta_k \leq \frac{1}{C} (f(x_1) - f(x_{l+1})).$$

La série est donc convergente puisqu'il existe une constante C' telle que pour tout k , $f(x_k) \geq C'$. ■

Il ya des propositions précisent les circonstances dans lesquelles la condition de Zoutendijk (2.12) est vérifiée avec les règles de la recherche linéaire inexacte (Armijo, Wolfe).

2.3.2 Recherche linéaire d'Armijo

Proposition 2.3.2 [31] *Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction continuellement différentiable dans un voisinage de $\Gamma = \{x \in \mathbb{R}^n : f(x) \leq f(x_1)\}$.*

CHAPITRE 2. MÉTHODES À DIRECTIONS DE DESCENTE

On considère un algorithme à directions de descente d_k , qui génère une suite $\{x_k\}$ en utilisant la recherche linéaire d'Armijo (2.3) avec $\alpha_1 > 0$. Alors il existe une constante $C > 0$ telle que, pour tout $k \geq 1$, l'une des conditions

$$f(x_{k+1}) \leq f(x_k) - C \nabla^T f(x_k) d_k,$$

ou

$$f(x_{k+1}) \leq f(x_k) - C \|\nabla f(x_k)\|^2 \cos^2 \theta_k,$$

est vérifiée.

Preuve. Si le pas $\alpha_k = \alpha_1$ est accepté, on a (2.12), car α_1 est uniformément positif. Dans le cas contraire, (2.3) n'est pas vérifiée avec un pas $\alpha'_k \leq \frac{\alpha_k}{\tau}$, c'est-à-dire $f(x_k + \alpha'_k d_k) > f(x_k) + \rho \alpha'_k \nabla^T f(x_k) d_k$. Comme f est continuellement différentiable, on a pour tout $\alpha_k > 0$:

$$\begin{aligned} f(x_k + \alpha_k d_k) &= f(x_k) + \alpha_k \nabla^T f(x_k) d_k + \int_0^1 [\nabla f(x_k + t \alpha_k d_k) - \nabla f(x_k)]^T \alpha_k d_k dt \\ &\Rightarrow f(x_k + \alpha_k d_k) \leq f(x_k) + \alpha_k \nabla^T f(x_k) d_k + C \alpha_k^2 \|d_k\|^2, \end{aligned}$$

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

$$\begin{aligned} &\begin{cases} f(x_k + \alpha'_k d_k) - f(x_k) > \rho \alpha'_k \nabla^T f(x_k) d_k \\ f(x_k + \alpha'_k d_k) - f(x_k) \leq \alpha'_k \nabla^T f(x_k) d_k + C \alpha_k'^2 \|d_k\|^2 \end{cases} \\ &\Rightarrow \rho \alpha'_k \nabla^T f(x_k) d_k \leq \alpha_k \nabla^T f(x_k) d_k + C \alpha_k'^2 \|d_k\|^2 \\ &\Rightarrow -C \alpha_k'^2 \|d_k\|^2 \leq (1 - \rho) \alpha'_k \nabla^T f(x_k) d_k, \end{aligned}$$

or

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

d'où

$$\begin{aligned} \nabla^T f(x_k) d_k &\geq \frac{-C}{1 - \rho} \alpha'_k \|d_k\|^2 \\ |\nabla^T f(x_k) d_k| &= \|\nabla f(x_k)\| \|d_k\| \cos \theta_k \leq \frac{C}{1 - \rho} \alpha'_k \|d_k\|^2, \end{aligned}$$

ce qui permet de minorer $\alpha'_k \|d_k\|$ et donc aussi $\alpha_k \|d_k\|$ par une constante fois $\|\nabla f(x_k)\| \|d_k\|$.

Cette minoration et l'expression suivante de (2.3)

$$f(x_k + \alpha_k d_k) \leq f(x_k) - \rho \alpha_k \|\nabla f(x_k)\| \|d_k\| \cos \theta_k,$$

conduit à (2.12). ■

2.3.3 Recherche linéaire de Wolfe

Proposition 2.3.3 Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction continuellement différentiable dans un voisinage de $\Gamma = \{x \in \mathbb{R}^n : f(x) \leq f(x_1)\}$.

On considère un algorithme à directions de descente d_k , qui génère une suite $\{x_k\}$ en utilisant la recherche linéaire de Wolfe (W1) – (W2). Alors il existe une constante $C > 0$ telle que, pour tout $k \geq 1$, la condition de Zoutendijk (2.12) est vérifiée.

Preuve. Noton $g_k = \nabla f(x_k)$ et $g_{k+1} = \nabla f(x_k + \lambda_k d_k)$.

D'après (W2)

$$\begin{aligned} g_{k+1}^T d_k &\geq \omega_2 g_k^T d_k \\ \Rightarrow (\omega_2 - 1) g_k^T d_k &\leq (g_{k+1} - g_k)^T d_k, \end{aligned}$$

et du fait que f est continuellement différentiable :

$$\begin{aligned} (g_{k+1} - g_k)^T d_k &\leq L \|x_{k+1} - x_k\| \|d_k\| \\ &= L \lambda_k \|d_k\| \|d_k\| \\ \Rightarrow \lambda_k &\geq \frac{(\omega_2 - 1) g_k^T d_k}{L \|d_k\|^2}, \end{aligned}$$

en utilisant (W1), on aura :

$$\begin{aligned} f(x_k + \lambda_k d_k) &\leq f(x_k) + \omega_1 \lambda_k g_k^T d_k \\ &\leq f(x_k) - \omega_1 \frac{(1 - \omega_2) g_k^T d_k}{L \|d_k\|^2} g_k^T d_k \\ &= f(x_k) - \omega_1 \frac{(1 - \omega_2)}{L} \cos^2 \theta_k \|g_k\|^2. \end{aligned}$$

On en déduit (2.12). ■

2.4 Direction du gradient

La direction du gradient d est, en réalité, l'opposé du gradient :

$$d = -\nabla f(x).$$

Il s'agit bien d'une direction de descente si $f'(x) \neq 0$, puisqu'alors

$$f'(x)^T(-\nabla f(x)) = -\|\nabla f(x)\|^2 < 0$$

2.4.1 Principe des méthodes du gradient

On cherche à déterminer la direction de descente qui fait décroître

$$\phi(\alpha) = f(x + \alpha d).$$

Le plus vite possible (au moins localement). Pour cela on va essayer de minimiser la dérivée de $\phi(\alpha)$ en 0. On a

$$\phi'(0) = \nabla f(x)^T d,$$

et on cherche d solution du problème

$$\min_{d \in \mathbb{R}^n, \|d\|=1} \phi'(0).$$

La solution est bien sûr

$$d = -\frac{\nabla f(x)}{\|\nabla f(x)\|}.$$

En vertu de l'inégalité de Schwartz.

Il y a ensuite de nombreuses façons d'utiliser cette direction de descente. On peut par exemple utiliser un pas fixé à priori $\lambda_k = \alpha > 0 ; \forall k$. On obtient alors la méthode du gradient simple :

$$\begin{cases} d_k = -\nabla f(x), \\ x_{k+1} = x_k + \lambda d_k. \end{cases} \quad (2.14)$$

Sous certaines hypothèses de régularité (f deux fois différentiable) cette méthode converge si λ est choisi assez petit. Ou bien consiste à faire les itérations suivantes

$$\begin{cases} d_k = -\nabla f(x), \\ x_{k+1} = x_k + \lambda_k d_k. \end{cases} \quad (2.15)$$

Où λ_k est choisi de manière à ce que

$$f(x_k + \lambda_k d_k) \leq f(x_k + \lambda d_k), \quad \forall \lambda > 0. \quad (2.16)$$

On obtient alors la méthode du gradient à pas optimal, cette méthode possède une propriété intéressante :

Proposition 2.4.1 *Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction différentiable. Les directions de descente d_k générées par la méthode (2.15) et (2.16) vérifient*

$$d_{k+1}^T d_k = 0. \quad (2.17)$$

Preuve. Si on introduit la fonction

$$\phi(\lambda) = f(x_k + \lambda d_k),$$

on a

$$\phi'(\lambda) = \nabla^T f(x_k + \lambda d_k) d_k,$$

et puisque ϕ est dérivable on a nécessairement

$$\phi'(\lambda) = 0,$$

donc

$$\nabla^T f(x_k + \lambda_k d_k) d_k = \nabla^T f(x_{k+1}) d_k = -d_{k+1}^T d_k = 0.$$

■

2.5 Direction du gradient conjugué

L'algorithme du gradient conjugué construit sa direction de descente en x en ajoutant à l'opposé du gradient, la direction d_{k-1} multipliée par un scalaire $\beta \in \mathbb{R}$:

$$d_k = -\nabla f(x) + \beta_k d_{k-1}.$$

Les méthodes du gradient conjugué sont utilisées pour résoudre les problèmes d'optimisation non linéaires sans contraintes spécialement les problèmes de grandes tailles. On l'utilise aussi pour résoudre les grands systèmes

linéaires. Elles reposent sur le concept des directions conjuguées parce que les gradients successifs sont orthogonaux entre eux et aux directions précédentes. L'idée initiale était de trouver une suite de directions de descente permettant de résoudre le problème

$$\min\{f(x) : x \in \mathbb{R}^n\}.$$

Où f est régulière (continûment différentiable)

2.5.1 Le principe général d'une méthode à directions conjuguées

Définition 2.5.1 Soit A une matrice symétrique $n \times n$, définie positive. On dit que deux vecteurs x et y de \mathbb{R}^n sont A -conjugués (ou conjugués par rapport à A) s'ils vérifient $x^T A y = 0$.

Description de la méthode

Soit $\{d_0, d_1, \dots, d_n\}$ une famille de vecteurs A -conjugués. On appelle alors méthode de directions conjuguées toute méthode itérative appliquée à une fonction quadratique strictement convexe de n variables : $q(x) = \frac{1}{2}x^T A x + b^T x + c$, avec $x \in \mathbb{R}^n$ et $A \in M_{n \times n}$ est symétrique et définie positive, $b \in \mathbb{R}^n$ et $c \in \mathbb{R}$ conduisent à l'optimum en n étapes au plus. Et cette méthode de la forme :

$$\begin{aligned} x_0 & \text{ donné,} \\ x_{k+1} & = x_k + \lambda_k d_k, \end{aligned} \tag{2.18}$$

où λ_k est optimal et d_0, d_1, \dots, d_n possédant la propriété d'être mutuellement conjuguées par rapport à la fonction quadratique. Si l'on note $g_k = \nabla q(x_k)$, la méthode se construit comme suit :

Calcul de λ_k

Comme λ_k minimise q dans la direction d_k , on a, $\forall k$:

$$\begin{aligned} \dot{q}(\lambda_k) & = d_k^T \nabla q(x_{k+1}) = 0 \\ d_k^T \nabla q(x_{k+1}) & = d_k^T (A x_{k+1} + b) = 0. \end{aligned}$$

Soit

$$d_k^T A(x_k + \lambda_k d_k) + d_k^T b = 0,$$

d'où l'on tire :

$$\lambda_k = \frac{-d_k^T (Ax_k + b)}{d_k^T Ad_k}. \quad (2.19)$$

Construire les directions A-conjuguées

Les directions A-conjuguées d_0, d_1, \dots, d_n peuvent être générées à partir d'un ensemble de vecteurs linéairement indépendants ξ_0, \dots, ξ_k en utilisant la procédure dite de Gram-Schmidt, de telle sorte que pour tout i entre 0 et k , le sous-espace généré par d_0, \dots, d_i soit égale au sous-espace généré par ξ_0, \dots, ξ_i . Alors d_{i+1} est construite comme suit :

$$d_{i+1} = \xi_{i+1} + \sum_{m=0}^i \varphi_{(i+1)m} d_m$$

2.5.2 La méthode du gradient conjugué dans le cas quadratique

La méthode du gradient conjugué est obtenue en appliquant la procédure de Gram-Schmidt aux gradients $\nabla q(x_0), \dots, \nabla q(x_{k-1})$, c'est-à-dire en posant $\xi_0 = -\nabla q(x_0), \dots, \xi_{n-1} = -\nabla q(x_{n-1})$. En outre, nous avons que

$$\begin{aligned} \nabla q(x) &= Ax + b, \\ \text{et } \nabla^2 q(x) &= A. \end{aligned}$$

Notons que la méthode se termine si $\nabla q(x_k) = 0$. La particularité intéressante de la méthode du gradient conjugué est que le membre de droite de l'équation donnant la valeur de d_{k+1} dans la procédure de Gram-Schmidt peut être grandement simplifié. Notons que la méthode du gradient conjugué est inspirée de celle du gradient (plus profonde pente).

Algorithme de La méthode du gradient conjugué pour les fonctions quadratiques (cas linéaire)

On suppose ici que la fonction à minimiser est quadratique sous la forme :

$$q(x) = \frac{1}{2} x^T Ax + b^T x + c.$$

Si l'on note $g_k = \nabla q(x_k)$, l'algorithme prend la forme suivante. Cet algorithme consiste à générer une suite d'itérés $\{x_k\}$ sous la forme :

$$x_{k+1} = x_k + \lambda_k d_k. \quad (2.20)$$

L'idée de la méthode est :

1-construire itérativement des directions d_0, \dots, d_k mutuellement conjuguées

A chaque étape k la direction d_k est obtenue comme combinaison linéaire du gradient en x_k et de la direction précédente d_{k-1} c'est-à-dire

$$d_{k+1} = -\nabla q(x_{k+1}) + \beta_{k+1}d_k, \quad (2.21)$$

les coefficients β_{k+1} étant choisis de telle manière que d_k soit conjuguée avec toutes les directions précédentes. Autrement dit :

$$d_{k+1}^T Ad_k = 0,$$

on en déduit :

$$\begin{aligned} d_{k+1}^T Ad_k &= 0 \Rightarrow (-\nabla q(x_{k+1}) + \beta_{k+1}d_k)^T Ad_k = 0 \\ &\Rightarrow -\nabla^T q(x_{k+1}) Ad_k + \beta_{k+1}d_k^T Ad_k = 0 \\ &\Rightarrow \beta_{k+1} = \frac{\nabla^T q(x_{k+1}) Ad_k}{d_k^T Ad_k} = \frac{g_{k+1}^T Ad_k}{d_k^T Ad_k}. \end{aligned} \quad (2.22)$$

2-déterminer le pas λ_k :

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

$$\lambda_k = \min q(x_k + \lambda d_k), \quad \lambda > 0, \quad (2.23)$$

on en déduit :

$$\lambda_k = \frac{-d_k^T g_k}{d_k^T Ad_k}. \quad (2.24)$$

Théorème 2.5.1 $\forall x_0 \in \mathbb{R}^n$, la suite $\{x_k\}$, définie par l'algorithme des directions conjuguées (définie par (2.20) et (2.24)), converge en au maximum n pas vers x^* .

Preuve. [36] Sachant que d_0, \dots, d_{n-1} engendrent \mathbb{R}^n , définissons

$$\sigma_k = \frac{d_k^T A(x^* - x_0)}{d_k^T Ad_k},$$

t.q. $x^* - x_0 = \sigma_0 d_0 + \sigma_1 d_1 + \dots + \sigma_{n-1} d_{n-1}$.

Selon l'algorithme (2.20), nous avons que $x_k = x_0 + \lambda_0 d_0 + \lambda_1 d_1 + \dots + \lambda_{n-1} d_{n-1}$.

En multipliant cette expression par $d_k^T A$, on voit que $d_k^T A(x_k - x_0) = 0$ et donc que

$$d_k^T A(x^* - x_0) = d_k^T A(x^* - x_k) = d_k^T (b - Ax_k) = -d_k^T g_k.$$

Ainsi, finalement $\sigma_k = \lambda_k$, ce qui démontre le théorème. ■

Algorithme (gradient conjugué "linéaire")

Etape 0 : (initialisation)

Soit x_0 le point de départ, $g_0 = \nabla q(x_0) = Ax_0 + b$, poser $d_0 = -g_0$.

Poser $k = 0$ et aller à l'étape 1.

Etape 1 :

Si $g_k = 0$: STOP ($x^* = x_k$). "Test d'arrêt".

Si non aller à l'étape 2.

Etape 2 :

Définir $x_{k+1} = x_k + \lambda_k d_k$ avec :

$$\begin{aligned} \lambda_k &= \frac{-d_k^T g_k}{d_k^T A d_k}, \\ d_{k+1} &= -g_{k+1} + \beta_{k+1} d_k, \\ \beta_{k+1} &= \frac{g_{k+1}^T A d_k}{d_k^T A d_k}. \end{aligned}$$

Poser $k = k + 1$ et aller à l'étape 1.

La validité de l'algorithme du gradient conjugué linéaire

On va maintenant montrer que l'algorithme ci-dessus définit bien une méthode de directions conjuguées.

Théorème 2.5.2 *A une itération k quelconque de l'algorithme où l'optimum de $q(x)$ n'est pas encore atteint (c'est-à-dire $g_i \neq 0, i = 0, 1, \dots, k$) on a*

a)

$$\lambda_k = \frac{g_k^T g_k}{d_k^T A d_k} \neq 0 \tag{2.25}$$

b)

$$\beta_{k+1} = \frac{g_{k+1}^T [g_{k+1} - g_k]}{g_k^T g_k} \tag{2.26}$$

$$= \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k} \tag{2.27}$$

c) Les directions d_0, d_1, \dots, d_{k+1} engendrées par l'algorithme sont mutuellement conjuguées.

Preuve. [39]

On raisonne par récurrence sur k en supposant que d_0, d_1, \dots, d_k sont mutuellement conjuguées

a) Montrons d'abord l'équivalence de (2.24) et de (2.25).

On a : $d_k = -g_k + \beta_k d_{k-1}$.

Donc (2.24) s'écrit :

$$\begin{aligned} \lambda_k &= \frac{-d_k^T g_k}{d_k^T Ad_k} \\ &= \frac{-[-g_k + \beta_k d_{k-1}]^T g_k}{d_k^T Ad_k} \\ &= \frac{g_k^T g_k}{d_k^T Ad_k} - \beta_k \frac{d_{k-1}^T g_k}{d_k^T Ad_k}. \end{aligned}$$

Comme $(d_0, d_1, \dots, d_{k-1})$ sont mutuellement conjuguées, x_k est l'optimum de $q(x)$ sur la variété v^k passant par x_0 et engendrée par $(d_0, d_1, \dots, d_{k-1})$.

Donc $d_{k-1}^T g_k = 0$ d'où l'on déduit (2.25).

b) Pour démontrer (2.26) remarquons que :

$$\begin{aligned} g_{k+1} - g_k &= A(x_{k+1} - x_k) = \lambda_k Ad_k \\ \Rightarrow Ad_k &= \frac{1}{\lambda_k} [g_{k+1} - g_k]. \end{aligned}$$

On a alors :

$$g_{k+1}^T Ad_k = \frac{1}{\lambda_k} g_{k+1}^T [g_{k+1} - g_k],$$

et en utilisant (2.25)

$$\lambda_k = \frac{g_k^T g_k}{d_k^T Ad_k},$$

il vient

$$\begin{aligned} g_{k+1}^T Ad_k &= \frac{d_k^T Ad_k}{g_k^T g_k} g_{k+1}^T [g_{k+1} - g_k] \\ \Rightarrow \frac{g_{k+1}^T Ad_k}{d_k^T Ad_k} &= \frac{g_{k+1}^T [g_{k+1} - g_k]}{g_k^T g_k}. \end{aligned}$$

Or de (2.22) on aura :

$$\beta_{k+1} = \frac{g_{k+1}^T Ad_k}{d_k^T Ad_k} = \frac{g_{k+1}^T [g_{k+1} - g_k]}{g_k^T g_k},$$

ce qui démontre (2.26).

(2.27) découle alors du fait que :

$$g_{k+1}^T g_k = 0,$$

car

$$g_k = d_k - \beta_k d_{k-1}.$$

Appartient au sous-espace engendré par (d_0, d_1, \dots, d_k) et que g_{k+1} est orthogonal à ce sous-espace.

c) Montrons enfin que d_{k+1} est conjuguée par rapport à (d_0, d_1, \dots, d_k) . On a bien $d_{k+1}^T Ad_k$ car, en utilisant $d_{k+1} = -g_{k+1} + \beta_{k+1} d_k$ on aura :

$$\begin{aligned} d_{k+1}^T Ad_k &= (-g_{k+1} + \beta_{k+1} d_k)^T Ad_k \\ &= -g_{k+1}^T Ad_k + \beta_{k+1} d_k^T Ad_k \\ &= -g_{k+1}^T Ad_k + \frac{g_{k+1}^T Ad_k}{d_k^T Ad_k} d_k^T Ad_k \\ &= 0. \end{aligned}$$

Vérifions maintenant que :

$$d_{k+1}^T Ad_i = 0 \text{ pour } i = 0, 1, \dots, k-1.$$

On a :

$$d_{k+1}^T Ad_i = -g_{k+1}^T Ad_i + \beta_{k+1} d_k^T Ad_i.$$

Le seconde terme est nul par l'hypothèse de récurrence $((d_0, d_1, \dots, d_k)$ sont mutuellement conjuguées).

Montrons qu'il en est de même du premier terme. Puisque $x_{i+1} = x_i + \lambda_i d_i$ et que $\lambda_i \neq 0$ on a :

$$\begin{aligned} Ad_i &= \frac{1}{\lambda_i} (Ax_{i+1} - Ax_i) \\ &= \frac{1}{\lambda_i} (g_{i+1} - g_i). \end{aligned}$$

En écrivant :

$$\begin{aligned} g_{i+1} &= d_{i+1} - \beta_i d_i \\ g_i &= d_i - \beta_{i-1} d_{i-1}, \end{aligned}$$

on voit que Ad_i est combinaison linéaire de d_{i+1}, d_i et de d_{i-1} seulement. Mais puisque (d_0, d_1, \dots, d_k) sont mutuellement conjuguées, on sait que le point x_{k+1} est l'optimum de $q(x)$ sur la variété v^{k+1} , engendrée par (d_0, d_1, \dots, d_k) . Donc g_{k+1} est orthogonal au sous-espace engendré par (d_0, d_1, \dots, d_k) et comme Ad_i appartient à ce sous-espace pour $i = 0, 1, \dots, k-1$, on en déduit $g_{k+1}^T Ad_i = 0$ ce qui achève la preuve. ■

Remarque 2.5.1 Dans ce cas d_k est une direction de descente puisque

$$\begin{aligned} d_k^T \nabla q(x_k) &= (-\nabla q(x_k) + \beta_k d_{k-1})^T \nabla q(x_k) \\ &= -\nabla q(x_k)^T \nabla q(x_k) + \beta_k d_{k-1}^T \nabla q(x_k) \\ &= -\|\nabla q(x_k)\|^2 \quad (\text{car } d_{k-1}^T \nabla q(x_k) = 0) \\ &< 0. \end{aligned}$$

Algorithme des différentes méthodes du gradient conjugué non linéaire

Etape 0 : (initialisation)

Soit x_0 le point de départ, $g_0 = \nabla f(x_0)$, poser $d_0 = -g_0$.

Poser $k = 0$ et aller à l'étape 1.

Etape 1 :

Si $g_k = 0$: STOP ($x^* = x_k$). "Test d'arrêt".

Si non aller à l'étape 2.

Etape 2 :

Définir $x_{k+1} = x_k + \lambda_k d_k$ avec :

λ_k : calculer par la recherche linéaire

$d_{k+1} = -g_{k+1} + \beta_{k+1} d_k$,

où β_{k+1} : définir selon la méthode.

Poser $k = k + 1$ et aller à l'étape 1.

2.6 Direction de Newton

La direction de Newton est définie en un point x en lequel le hessien de f est inversible par

$$d = -[H(x_k)]^{-1} \nabla f(x_k).$$

Considérons le problème d'optimisation sans contraintes (P)

$$(P) : \min_{x \in \mathbb{R}^n} f(x) \quad \text{où } f : \mathbb{R}^n \rightarrow \mathbb{R}.$$

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

$$f(x) = f(x_k) + \nabla^T f(x_k)(x - x_k) + \frac{1}{2}(x - x_k)^T H(x_k)(x - x_k) + o(\|x - x_k\|^2),$$

posons

$$q(x) = f(x_k) + \nabla^T f(x_k)(x - x_k) + \frac{1}{2}(x - x_k)^T H(x_k)(x - x_k).$$

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

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

donc

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

Avantages

Si le point x_1 est assez proche de la solution optimale locale x_* telle que $H(x_*)$ soit définie positive, alors l'algorithme de Newton converge de façon quadratique vers la solution x_* , c'est à dire que l'on a,

$$\|x_{k+1} - x_*\| \leq \gamma \|x_k - x_*\|^2, \quad \gamma \geq 0$$

Inconvénients

1-Cette méthode fonctionne très bien pour les problèmes de petite dimension ($1 \leq n \leq 10$), lorsque on peut calculer facilement la matrice Hessienne H et sont inverse. Ce calcul nécessite des itérations plus nombreuses et coûteuses dans les problèmes de grandes tailles.

2-Comme $x_{k+1} = x_k - [H(x_k)]^{-1} \nabla f(x_k)$. On voit bien que le successeur x_{k+1} de x_k n'est pas toujours bien défini.

3-Même si $[H(x_k)]^{-1}$ existe la direction $d_k = - [H(x_k)]^{-1} \nabla f(x_k)$ n'est pas toujours une direction de descente.

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 à été découverte par Davidon en 1959 et a été développée et popularisée par Fletcher et Powell en 1963. 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).

2.7 Direction de quasi Newton

Une méthode de quasi Newton est une méthode de type :

$$\begin{cases} x_{k+1} = x_k + \lambda_k d_k, \\ d_k = -B_k g_k. \end{cases} \quad (2.28)$$

Où B_k est une matrice destinée à approcher l'inverse du Hessien de f (respectivement le Hessien de f) en x_k . Le problème posé est : quelle stratégie à adopter pour faire cette approximation ? On peut par exemple poser $B_0 = I$, mais comment ensuite mettre à jour l'approximation B_k au cours des itérations ?

L'idée est la suivante : on sait que au point x_k , le gradient et le hessien de f vérifient la relation

$$g_{k+1} = g_k + H(x_k)(x_{k+1} - x_k) + \epsilon(x_{k+1} - x_k).$$

Si on suppose que l'approximation quadratique est bonne, on peut alors négliger le reste et considérer que l'on a

$$g_{k+1} \simeq g_k + H(x_k)(x_{k+1} - x_k),$$

cela conduit à la notion de relation de quasi-Newton

Définition 2.7.1 *On dit que les matrice B_{k+1} et H_{k+1} vérifient une relation de quasi-Newton si on a*

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

ou

$$x_{k+1} - x_k = B_{x+1} [\nabla f(x_{k+1}) - \nabla f(x_k)].$$

Il reste un problème à résoudre : comment mettre à jour B_k tout en assurant $B_k > 0$?

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

$$\begin{cases} x_{k+1} = x_k + \lambda_k d_k \\ d_k = -B_k g_k, \end{cases}$$

à appliquer une formule du type

$$B_{k+1} = B_k + \Delta_k, \tag{2.29}$$

avec Δ_k symétrique, assurant la relation de quasi-Newton

$$x_{k+1} - x_k = B_{k+1} [g_{k+1} - g_k],$$

ainsi que $B_{k+1} > 0$, sous l'hypothèse que $B_k > 0$.

La formule (2.29) permet d'utiliser les nouvelles informations obtenues lors de l'étape k de l'algorithme, c'est à dire essentiellement le gradient $g_{k+1} = \nabla f(x_{k+1})$ au point x_{k+1} , obtenu par recherche linéaire (exacte ou approchée) dans la direction d_k . Il existe différentes formules du type (2.29). Suivant que Δ_k est de rang un ou deux, on parlera de correction de rang un ou de rang deux .

2.7.1 Méthode de correction de rang un

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

$$B_{k+1} = B_k + a_k u_k u_k^T, \quad u_k \in \mathbb{R}^n$$

donc la condition de quasi -Newton s'écrit comme suit

$$s_k = (B_k + a_k u_k u_k^T) y_k,$$

ou encore

$$s_k - B_k y_k = a_k u_k u_k^T y_k.$$

D’où l’on déduit que u_k est proportionnel à $s_k - B_k y_k$, avec un facteur qui peut être pris en compte dans a_k . Un choix évident pour vérifier cette dernière équation est de prendre $u_k = s_k - B_k y_k$ et a_k tel que $a_k (u_k^T y_k) = 1$, on obtient :

$$B_{k+1} = B_k + \frac{(s_k - B_k y_k) (s_k - B_k y_k)^T}{(s_k - B_k y_k)^T y_k}. \quad (2.30)$$

Théorème 2.7.1 *Si f est quadratique, de matrice Hessienne H définie positive et si s_1, s_2, \dots, s_n sont des vecteurs indépendants, alors la méthode de correction de rang un converge au plus dans $(n + 1)$ itérations et $(B_{n+1})^{-1} = H$.*

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 linéaires exactes.

Inconvénients :

Même si la fonction est quadratique, et même si son Hessien est défini positif, il se peut que la matrice B_k ne soit pas définie positive. Le dénominateur $(s_k - B_k y_k)^T y_k$ peut devenir nul ou très petit, ce qui rend le procédé instable.

2.7.2 Méthode de Davidon Fletcher Powell (DFP)

Cette méthode a été proposée par Davidon en 1959 et développé plus tard en 1963 par Fletcher. La formule de mise à jour de **DFP** est une formule de correction de rang deux. De façon plus précise construisons B_{k+1} en fonction de B_k de la forme :

$$B_{k+1} = B_k + A_k + \Delta_k, \quad (2.31)$$

avec Δ_k et A_k deux matrices de rang un tel que

$$A_k = a_k u_k u_k^T, \quad \Delta_k = b_k v_k v_k^T,$$

a_k, b_k sont des constantes, u_k, v_k sont deux vecteurs de \mathbb{R}^n . B_{k+1} doit satisfaire la condition quasi -Newton c’est à dire

$$x_{k+1} - x_k = B_{k+1} [g_{k+1} - g_k].$$

Si on pose par suite

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

donc

$$\begin{aligned} s_k &= B_{k+1}y_k \\ &= (B_k + a_k u_k u_k^T + b_k v_k v_k^T) y_k, \end{aligned} \quad (2.32)$$

par suite

$$a_k u_k u_k^T y_k + b_k v_k v_k^T y_k = s_k - B_k y_k.$$

Un choix évident pour satisfaire cette équation est de prendre

$$u_k = s_k, \quad v_k = B_k y_k, \quad a_k u_k^T y_k = 1, \quad b_k v_k^T y_k = -1,$$

d'où

$$B_{k+1} = B_k + \frac{s_k s_k^T}{s_k^T y_k} - \frac{B_k y_k y_k^T B_k}{y_k^T B_k y_k}. \quad (2.33)$$

Remarque 2.7.1 *Le résultat suivant montre que sous certaines conditions, la formule (2.31) conserve la définie positivité des matrices B_k .*

Théorème 2.7.2 *On considère la méthode définie par*

$$\begin{cases} x_{k+1} = x_k + \lambda_k d_k, \\ d_k = -B_k g_k, \end{cases}$$

où λ_k optimal, $B_0 > 0$ est donnée ainsi que x_0 , alors les matrices B_k sont définies positives.

Théorème 2.7.3 *Appliqué à une forme quadratique f , l'algorithme DFP décrit par la relation*

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

engendre des directions conjuguées d_1, d_2, \dots, d_k vérifiant

$$d_i^T H d_j = 0, \quad 1 \leq i \leq j \leq k, \quad (2.34)$$

$$B_{k+1} H d_i = d_i, \quad 1 \leq i \leq k. \quad (2.35)$$

La méthode **DFP** se comporte donc dans le cas quadratique, comme une méthode de directions conjuguées. On peut aussi remarquer qu'on a pour $k = n - 1$ la relation

$$B_{n+1}Hd_i = d_i, \quad i = 1, \dots, n - 1.$$

Et comme les d_i sont linéairement indépendants (car mutuellement conjugués) on en déduit que $B_{n+1} = H^{-1}$.

Avantages :

1-Pour des fonctions quadratiques (avec une recherche linéaire exacte) :

- L'algorithme converge dans au plus n étapes avec $B_{n+1} = H^{-1}$.
- Elles engendrent des directions conjuguées.

2-Pour les fonctions quelconques :

-La matrice B_k reste définie positive, ce qui est nécessaire pour que la direction soit une direction de descente.

Inconvénients :

-La méthode DFP est sensible à la précision de la recherche linéaire.

2.7.3 Méthode de Broyden, Fletcher, Goldfarb et Shanno (BFGS)

La formule de mise à jour de Broyden, Fletcher, Goldfarb et Shanno(**BFGS**) est une formule de correction de rang deux, qui s'obtient à partir de la formule DFP en intervertissant les rôles de s_k et y_k . La formule obtenue permet de mettre à jour une approximation B_k de Hessien lui même et non de son inverse comme dans le cas de la méthode DFP. On exigera que posée dans les mêmes propriétés, à savoir B_{k+1} reste définie positive si B_k et vérifiant la relation de quasi-Newton

$$B_{k+1}s_k = y_k.$$

On obtient donc

$$B_{k+1} = B_k + \frac{y_k y_k^T}{y_k^T s_k} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}. \quad (2.36)$$

1-Notons que la direction d_k est obtenue par une résolution d'un système linéaire. En particulier la mise à jour de B_k est faite directement sur le facteur de Cholesky C_k où $B_k = C_k C_k^T$ ce qui ramène le calcul de d_k au même coût que pour la formule de DFP .

2-La méthode BFGS possède les mêmes propriétés que la méthode DFP dans le cas quadratique. Les directions engendrées sont conjuguées. Cette

méthode est reconnue comme étant beaucoup moins sensible que la méthode DFP aux imprécisions dans la recherche linéaire, du point de vue de vitesse de convergence. Elle est donc tout à fait adaptée quand la recherche linéaire est faite de façon économique, avec par exemple la règle de Goldstein ou la règle de wolfe et Powell.

3-La relation (2.36) permet de construire une approximation de la matrice Hessienne elle même (et non pas son inverse).

En effet : Posons

$$C_k = \frac{y_k y_k^T}{y_k^T s_k} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}. \quad (2.37)$$

Nous avons

$$H_{k+1} = [B_{k+1}]^{-1} = [B_k + C_k]^{-1}.$$

Par application de la formule de Sherman-Morrison-Woodbury suivante

$$(A + ab^T)^{-1} = A^{-1} - \frac{A^{-1} a b^T A^{-1}}{1 + b^T A^{-1} a}. \quad (2.38)$$

Où A est une matrice inversible, et b est un vecteur de \mathbb{R}^n , et en supposant que $b^T A^{-1} a \neq -1$, alors on a

$$H_{k+1} = [B_{k+1}]^{-1} = [B_k + C_k]^{-1} = \left[B_k + \frac{y_k y_k^T}{y_k^T s_k} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} \right]^{-1}.$$

Posons

$$A = B_k + \frac{y_k y_k^T}{y_k^T s_k}, \quad a = -\frac{B_k s_k}{s_k^T B_k s_k}, \quad b^T = s_k^T B_k,$$

donc

$$H_{k+1} = \left[B_k + \frac{y_k y_k^T}{y_k^T s_k} \right]^{-1} + \frac{\left[B_k + \frac{y_k y_k^T}{y_k^T s_k} \right]^{-1} \frac{B_k s_k}{s_k^T B_k s_k} s_k^T B_k \left[B_k + \frac{y_k y_k^T}{y_k^T s_k} \right]^{-1}}{1 - s_k^T B_k \left[B_k + \frac{y_k y_k^T}{y_k^T s_k} \right]^{-1} \frac{B_k s_k}{s_k^T B_k s_k}}, \quad (2.39)$$

on doit calculer $\left[B_k + \frac{y_k y_k^T}{y_k^T s_k} \right]^{-1}$ pour cela on applique la formule de Sherman-Morrison-Woodbury une deuxième fois, on pose :

$$A = B_k, \quad a = \frac{y_k}{y_k^T s_k}, \quad b^T = y_k^T$$

$$\begin{aligned} \left[B_k + \frac{y_k y_k^T}{y_k^T s_k} \right]^{-1} &= [B_k]^{-1} - \frac{[B_k]^{-1} \frac{y_k y_k^T}{y_k^T s_k} [B_k]^{-1}}{1 + y_k^T [B_k]^{-1} \frac{y_k}{y_k^T s_k}} \\ &= [B_k]^{-1} - \frac{[B_k]^{-1} y_k y_k^T [B_k]^{-1}}{y_k^T s_k + y_k^T [B_k]^{-1} y_k}. \end{aligned}$$

Remplaçons cette dernière dans la formule (2.39) et d'après un calcul on obtient

$$\begin{aligned} H_{k+1} &= [B_{k+1}]^{-1} & (2.40) \\ &= [B_k]^{-1} + \frac{y_k y_k^T}{y_k^T s_k} - \frac{[B_k]^{-1} s_k s_k^T [B_k]^{-1}}{s_k^T [B_k]^{-1} s_k} \\ &= H_k + \frac{y_k y_k^T}{y_k^T s_k} - \frac{H_k s_k s_k^T H_k}{s_k^T H_k s_k}. \end{aligned}$$

Algorithme de Broyden, Fletcher, Goldfarb et Shanno

1. Choisir x_0 et H_0 définie positive quelconque (par exemple $H_0 = I$)
2. A l'itération k , calculer la direction de déplacement

$$d_k = -H_k^{-1} \nabla f(x_k),$$

déterminer le pas optimal λ_k et poser

$$x_{k+1} = x_k + \lambda_k d_k$$

3. Poser $s_k = \lambda_k d_k$ et $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$ puis calculer

$$H_{k+1} = H_k + \frac{y_k y_k^T}{y_k^T s_k} - \frac{H_k s_k s_k^T H_k}{s_k^T H_k s_k}$$

4. Faire $k \leftarrow k + 1$. Retourner en 2 sauf si le critère d'arrêt est vérifié. Comme critère d'arrêt on retiendra par exemple $\|g_{k+1}\| < \epsilon$.

2.8 Direction de Gauss-Newton

2.8.1 Problèmes de moindres carrés

Le cas particulier des problèmes de moindres carrés linéaires et non linéaires. Un problème de moindres carrés est un problème d'optimisation de

la forme :

$$\min f(x) = \frac{1}{2} \|F(x)\|^2 : x \in \mathbb{R}^n, \quad (2.41)$$

où F désigne une application de \mathbb{R}^n dans \mathbb{R}^m , en supposant $m \geq n$. Remarquons que si le système $F(x) = 0$ a des solutions, alors ce sont également les solutions du problème (2.41).

Remarquons tout d'abord que l'on a des expressions explicites des dérivées de f en fonction de F et de sa jacobienne J_F :

$$\begin{aligned} \nabla f(x) &= J_F(x)^T F(x) = \sum_{i=1}^m F_i(x) \nabla F_i(x), \\ H_f(x) &= J_F(x)^T J_F(x) + \sum_{i=1}^m F_i(x) H_{F_i}(x). \end{aligned}$$

Une caractéristique des algorithmes de moindres carrés est qu'à condition de connaître la jacobienne de F en tout point, on peut calculer une partie du hessien H_f sans calcul supplémentaire. De plus ce terme $J_F(x)^T J_F(x)$ est souvent prépondérant par rapport au second terme $\sum_{i=1}^m F_i(x) H_{F_i}(x)$ pour deux raisons : tout d'abord à cause de la presque linéarité du modèle au voisinage de la solution (i.e. $H_{F_i}(x)$ petit), ou bien à cause du résidu F_j petit au voisinage de la solution.

La plupart des algorithmes de moindres carrés non linéaires existants exploitent cette structure particulière du Hessien de f .

Problèmes de moindres carrés linéaires

On se place dans le cas où la fonction F est linéaire,

$$F(x) = Ax - b, \quad \text{avec} \quad A \in M_{n,p}(\mathbb{R}),$$

et on résout donc le problème de moindres carrés linéaire suivant :

$$\min f(x) = \frac{1}{2} \|Ax - b\|^2 : x \in \mathbb{R}^n. \quad (2.42)$$

On calcule sans difficulté la jacobienne de F en tout point : $J_F(x) = A$, puis :

$$\nabla f(x) = A^T(Ax - b) \quad \text{et} \quad H_f(x) = J_F(x)^T J_F(x) = A^T A.$$

La hessienne de f est donc positive en tout point, le problème de moindres carrés linéaire est donc convexe. Cherchons les points critiques du problème (2.42). i.e. vérifiant : $\nabla f(x) = 0$, d'où

le système suivant à résoudre :

$$A^T Ax = A^T b,$$

appelé système d'équations normales du problème de moindres carrés linéaire (2.42). Ainsi, x^* est solution du problème (2.42) si et seulement si x^* vérifie les équations normales. De plus si A est de rang plein, alors la fonction f est strictement convexe et x^* est l'unique solution de (2.42).

Appliquons maintenant la méthode de Newton au problème (2.42) : la direction de Newton $d_k = x_{k+1} - x_k$ est solution du système :

$$A^T A d_k = -A^T (Ax_k - b), \quad \text{d'où : } A^T Ax_{k+1} = A^T b,$$

et ceci quel que soit x_k . On reconnaît ici le système d'équations normales du problème (2.42). Ceci signifie donc que la méthode de Gauss-Newton identifie la solution en une seule itération lorsque la fonction F est linéaire.

2.8.2 Algorithme de Gauss-Newton

Revenons au problème de moindres carrés non linéaire général (2.41), l'idée est de remplacer à chaque itération le problème de moindres carrés non linéaire par un problème approché de moindres carrés linéaire.

Soit $x_k \in \mathbb{R}^n$ le point courant. Remplaçons au voisinage de x_k le problème (2.41) par :

$$\min \tilde{f}(y) = \frac{1}{2} \|F(x_k) + J_F(x_k)(y - x_k)\|^2 : y \in \mathbb{R}^n, \quad (2.43)$$

où l'on a remplacé F par une approximation du premier ordre en x_k . Le problème (2.43) est un problème de moindres carrés linéaire dont les solutions x_{k+1} vérifient les équations normales associées, à savoir :

$$J_F(x_k)^T J_F(x_k)(x_{k+1} - x_k) = -J_F(x_k)^T F(x_k).$$

La direction $d_k = x_{k+1} - x_k$ ainsi définie, est appelée direction de Gauss-Newton.

ALGORITHME DE GAUSS-NEWTON.

Données : F fonction différentiable, x_0 point initial, $\epsilon > 0$ précision demandée.

Sortie : une approximation de la solution du problème de moindres carrés :

$$\min f(x) = \frac{1}{2}F(x)^T F(x), x \in \mathbb{R}^n.$$

1. $k := 0$;
2. Tant que critère d'arrêt à définir,
 - (a) Calcul d'une direction de recherche : calculer d_k solution de :

$$J_F(x_k)^T J_F(x_k)d = -J_F(x_k)^T F(x_k).$$

- (b) $x_{k+1} = x_k + d_k$;
 - (c) $k := k + 1$;
3. Retourner x_k .

Les avantages de cette méthode par rapport à la méthode de Newton sont les suivants :

1. L'approximation faite ne nécessite pas le calcul des dérivées secondes.
2. Dans de nombreuses applications, cette approximation est de bonne qualité.
3. Si $J_F(x_k)$ est de rang plein et si $\nabla f(x_k)$ est non nul, alors la direction de Gauss-Newton est une direction de descente de f en x_k et donc adaptée pour l'utilisation d'un algorithme de recherche linéaire.

La vitesse de convergence de l'algorithme de Gauss-Newton dépend de la qualité de l'approximation faite du problème initial. Dans le cas où le résidu est effectivement petit au voisinage de la solution, la convergence est rapide. Si la hessienne de F est nulle à la solution, alors la convergence est quadratique.

Sous des hypothèses supplémentaires (par exemple dans [45], valeurs singulières de $J_F(x_k)$ uniformément bornées hors d'un voisinage de l'origine et pas de Wolfe), on peut démontrer la convergence globale des itérés vers un point stationnaire de f . Un désavantage de la méthode de Gauss-Newton est que l'on n'a pas de résultat de convergence lorsque la jacobienne de F en x_k n'est pas de rang plein. Pour pallier cet inconvénient, on a recours à la méthode de Levenberg-Marquardt.

2.8.3 Algorithme de Levenberg-Marquardt

L'algorithme de Levenberg-Marquardt peut être vu comme une régularisation de l'algorithme de Gauss-Newton. Toutefois il existe un lien étroit avec les méthodes dites de région de confiance et c'est par cet aspect là que nous abordons ici la description de cet algorithme.

Soit x_k l'itéré courant. Comme pour la méthode de Gauss-Newton, on remplace, au voisinage de x_k , le problème (2.41) de moindres carrés non linéaires par le problème suivant :

$$\min \frac{1}{2} \|F(x_k) + J_F(x_k)(y - x_k)\|^2 : y \in \mathbb{R}^n,$$

mais en ajoutant cette fois-ci la contrainte : $\|y - x_k\| \leq \Delta_k$. Le paramètre $\Delta_k > 0$ est appelé rayon de la région de confiance. A chaque itération, le problème initial est remplacé par le problème quadratique suivant :

$$\begin{aligned} (P_k) & : \min \frac{1}{2} \|F(x_k) + J_F(x_k)(y - x_k)\|^2 : y \in \mathbb{R}^n, \\ \text{sous} & : \|y - x_k\| \leq \Delta_k. \end{aligned}$$

En introduisant le Lagrangien associé au problème (P_k) :

$$L_k(y; \lambda) = \frac{1}{2} \|F(x_k) + J_F(x_k)(y - x_k)\|^2 + \frac{1}{2} \lambda (\|y - x_k\|^2 - \Delta_k^2), \lambda \in \mathbb{R}^n,$$

on peut écrire les conditions nécessaires d'optimalité du premier ordre, à savoir :

$$\begin{aligned} (J_F(x_k)^T J_F(x_k) + \lambda I)d & = -J_F(x_k)^T F(x_k) \\ \lambda & = 0 \text{ ou } \|y - x_k\| = \Delta_k \\ \lambda & \geq 0. \end{aligned}$$

Dans le cas où la contrainte de région de confiance est inactive (i.e. $\|y - x_k\| < \Delta_k$ et $\lambda = 0$), on retrouve une itération de l'algorithme de Gauss-Newton. Dans le cas contraire ($\|y - x_k\| = \Delta_k$), la solution est sur le bord de la région de confiance. Le pas effectué peut être vu comme une régularisation du pas de Gauss-Newton dans le sens où la matrice $(J_F(x_k)^T J_F(x_k) + \lambda I)$ est maintenant définie positive : le pas effectué est alors un pas de descente de la fonction f en x_k .

CHAPITRE 2. MÉTHODES À DIRECTIONS DE DESCENTE

ALGORITHME DE LEVENBERG-MARQUARDT.

Données : F fonction différentiable, x_0 point initial, $\epsilon > 0$ précision demandée.

Sortie : une approximation de la solution du problème de moindres carrés :

$$\min f(x) = \frac{1}{2}F(x)^T F(x), x \in \mathbb{R}^n.$$

1. $k := 0$;
2. Tant que critère d'arrêt à définir,
 - (a) Calcul d'une direction de recherche : calculer d_k solution de :

$$(J_F(x_k)^T J_F(x_k) + \lambda I)d = -J_F(x_k)^T F(x_k).$$

- (b) $x_{k+1} = x_k + d_k$;
 - (c) Mise à jour du paramètre λ .
 - (d) $k := k + 1$;
3. Retourner x_k .

Chapitre 3

Classes itératives à directions de descente

Il convient de souligner que la plupart des algorithmes d'optimisation, contrainte ou non, fonctionnent selon un schéma général consistant, à chaque itération, à se rapprocher du minimum par la résolution d'un sous-problème de minimisation. Nous considérons ici les méthodes permettant de résoudre un problème d'optimisation sans contraintes, soit le problème :

$$\min \{f(x) : x \in \mathbb{R}^n\}.$$

3.1 Méthodes du gradient conjugué non linéaire

Notre problème consiste de minimiser une fonction f de n variables de valeurs réelles

$$(P) : \min \{f(x) : x \in \mathbb{R}^n\}, \quad (3.1)$$

où f est régulière (continûment différentiable) et g est son gradient. Notons par g_k le gradient de f au point x_k .

Rappelons que les différentes méthodes du gradient conjugué génèrent des suites $\{x_k\}$ de la forme suivante :

$$x_{k+1} = x_k + \lambda_k d_k, \quad \lambda_k > 0, \quad (3.2)$$

où la direction recherchée est définie par la formule de récurrence suivante :

$$d_k = \begin{cases} -g_1 & \text{si } k = 1, \\ -g_k + \beta_k d_{k-1} & \text{si } k \geq 2, \end{cases} \quad (3.3)$$

Le coefficient β_k déterminent la méthode du gradient conjugué en question (Fletcher Reeves, Polak Ribière Polyak, . . .).

Le pas $\lambda_k \in \mathbb{R}^+$ étant déterminé par une recherche linéaire.

Proposition 3.1.1 (i) *L'ensemble $\mathcal{L} := \{x \in \mathbb{R}^n; f(x) \leq f(x_1)\}$ est borné; où $x_1 \in \mathbb{R}^n$ est le point initial.*

(ii) *Sur un voisinage \mathcal{N} de \mathcal{L} , la fonction objectif f est continûment différentiable et son gradient est lipschitzien*

$$\exists L > 0 \text{ tel que } \|g(x) - g(\tilde{x})\| \leq L \|x - \tilde{x}\|, \quad \forall x, \tilde{x} \in \mathcal{N} \quad (3.4)$$

Remarque 3.1.1 *Ces suppositions impliquent qu'il existe $\gamma > 0$ tel que*

$$\|g(x)\| \leq \gamma, \quad \forall x \in \mathcal{L} \quad (3.5)$$

Définition 3.1.1 [37]

On dit que d_k est une direction de descente suffisante si

$$g_k^T d_k < -c \|g_k\|^2. \quad (3.6)$$

Les conditions de Wolfe faibles :

$$f(x_k + \lambda d_k) \leq f(x_k) + \omega_1 \lambda \nabla f(x_k)^T d_k. \quad (3.7)$$

$$\nabla f(x_k + \lambda d_k)^T d_k \geq \omega_2 \nabla f(x_k)^T d_k, \quad (3.8)$$

avec $0 < \omega_1 < \omega_2 < 1$.

Les conditions de Wolfe fortes :

$$f(x_k + \lambda d_k) \leq f(x_k) + \omega_1 \lambda \nabla f(x_k)^T d_k. \quad (3.9)$$

$$|\nabla f(x_k + \lambda d_k)^T d_k| \leq -\omega_2 \nabla f(x_k)^T d_k. \quad (3.10)$$

Les conditions de Wolfe relaxée :

$$f(x_k + \lambda d_k) \leq f(x_k) + \omega_1 \lambda \nabla^T f(x_k) d_k \quad (3.11)$$

$$\omega_2 \nabla^T f(x_k) d_k \leq \nabla^T f(x_k + \alpha_k d_k) d_k \leq -\omega_2 \nabla^T f(x_k) d_k, \quad (3.12)$$

où $0 < \omega_1 < \omega_2 < 1$ et $\omega_2 > 0$.

Définition 3.1.2 *Présentons maintenant un théorème fondamental qui assure la satisfaction de la condition de Zoutendijk, pour toute méthode du type (3.2), (3.3) dans laquelle le pas λ est déterminé par la règle de Wolfe faible (3.7), (3.8). Ce théorème a été démontré par Zoutendijk [65] et Powell [48].*

Théorème 3.1.1 *Considérons une méthode du type (3.2) et (3.3) dans laquelle d_k est une direction de descente et le pas λ est déterminé par la règle de Wolfe faible (3.7), (3.8) avec $0 < \omega_1 < 1/2$. Considérons aussi que la supposition (3.1.1) soit satisfaite. Alors pour une telle méthode la condition de Zoutendijk suivante :*

$$\sum_{k \geq 1} \cos^2 \theta_k \|g_k\|^2 \leq \infty. \quad (3.13)$$

3.2 Méthode de Fletcher-Reeves

La méthode de Fletcher-Reeves [30] est une extension directe de la méthode du gradient conjugué linéaire au cas des fonctions quelconques. Appliquée à une fonction quadratique, elle est identique au gradient conjugué linéaire. La méthode de Fletcher-Reeves est :

$$\beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2} \text{ Gradient conjugué variante Fletcher-Reeves} \quad (3.14)$$

3.2.1 Algorithme de la méthode de FR

Cette méthode est très intéressante, d'une part parce qu'elle nécessite le stockage de très peu d'informations, d'autre part, par sa vitesse de convergence très supérieure à celle des algorithmes du gradient classique.

Algorithme de la méthode de Fletcher-Reeves

Etape 0 : (initialisation)

Soit x_0 le point de départ, $g_0 = \nabla f(x_0)$, poser $d_0 = -g_0$.

Poser $k = 0$ et aller à l'étape 1.

Etape 1 :

Si $g_k = 0$: STOP ($x^* = x_k$). "Test d'arrêt"

Sinon aller à l'étape 2.

Etape 2 :

Définir $x_{k+1} = x_k + \lambda_k d_k$ avec : $\lambda_k = \arg \min_{\lambda > 0} f(x_k + \lambda d_k)$
 $d_{k+1} = -g_{k+1} + \beta_{k+1}^{FR} d_k$,

où

$$\beta_{k+1}^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2}.$$

3.2.2 La propriété de descente de la méthode de FR

Powell [50] a démontré la satisfaction de la propriété de descente de la fonction objectif pour la méthode de Fletcher-Reeves avec recherche linéaire exacte. Al-Baali [2] a démontré la satisfaisse de la propriété de descente de la fonction objectif pour la méthode de Fletcher-Reeves avec la recherche linéaire inexacte de Wolfe forte. Gilbert et Nocedal [32] ont généralisé ce résultat pour toute méthode du type (3.2)-(3.3) dont

$$|\beta_k| \leq \beta_k^{FR}. \quad (3.15)$$

Théorème 3.2.1 *On considère que la supposition (3.1.1) est satisfaite. Considérons une méthode du type (3.2) et (3.3) avec β_k satisfaisant à (3.14) et le pas λ_k vérifiant la règle de Wolfe forte (3.8) où $\omega_2 \in]0, \frac{1}{2}[$. Alors cette méthode génère des directions de descente. De plus on a :*

$$\frac{-1}{1 - \omega_2} \leq \frac{d_k^T g_k}{\|g_k\|^2} \leq \frac{2\omega_2 - 1}{1 - \omega_2}; \quad k = 1, \dots \quad (3.16)$$

Preuve. La démonstration se fait par récurrence.

1) Pour $k = 1$:

$$\frac{d_1^T g_1}{\|g_1\|^2} = \frac{-\|g_1\|^2}{\|g_1\|^2} = -1.$$

D'autre part :

$$0 < \omega_2 < \frac{1}{2} \Rightarrow \begin{cases} \frac{-1}{1-\omega_2} \leq -1 \\ \frac{2\omega_2-1}{1-\omega_2} \geq -1 \end{cases}$$

2) Supposons que (3.16) est satisfaite pour k et démontrons qu'elle le sera pour $k + 1$:

Supposons que :

$$\frac{-1}{1 - \omega_2} \leq \frac{d_k^T g_k}{\|g_k\|^2} \leq \frac{2\omega_2 - 1}{1 - \omega_2}; \quad k = 1, \dots \quad (3.17)$$

On a :

$$\frac{d_{k+1}^T g_{k+1}}{\|g_{k+1}\|^2} = \frac{(-g_{k+1} + \beta_{k+1} d_k)^T g_{k+1}}{\|g_{k+1}\|^2} = -1 + \beta_{k+1} \frac{d_k^T g_{k+1}}{\|g_k\|^2}.$$

D'autre part de (3.14) on aura :

$$\beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2} \implies \frac{1}{\beta_{k+1}^{FR}} = \frac{\|g_k\|^2}{\|g_{k+1}\|^2},$$

d'où :

$$\frac{d_{k+1}^T g_{k+1}}{\|g_{k+1}\|^2} = -1 + \frac{\beta_{k+1}}{\beta_k^{FR}} \frac{d_k^T g_{k+1}}{\|g_k\|^2}. \quad (3.18)$$

En utilisant la condition de recherche linéaire (3.8) on obtient :

$$|d_k^T g_{k+1}| \leq -\omega_2 d_k^T g_k \implies \omega_2 |\beta_{k+1}| d_k^T g_k \leq \beta_{k+1} d_{k+1}^T g_{k+1} \leq -\omega_2 |\beta_{k+1}| d_k^T g_k.$$

Remplaçons ceci dans (3.18)

$$-1 + \omega_2 \frac{|\beta_{k+1}|}{\beta_{k+1}^{FR}} \frac{d_k^T g_k}{\|g_k\|^2} \leq \frac{d_{k+1}^T g_{k+1}}{\|g_{k+1}\|^2} \leq -1 - \omega_2 \frac{|\beta_{k+1}|}{\beta_{k+1}^{FR}} \frac{d_k^T g_k}{\|g_k\|^2}.$$

De (3.17) on aura

$$-1 - \frac{|\beta_{k+1}| \omega_2}{\beta_{k+1}^{FR} (1 - \omega_2)} \leq \frac{d_{k+1}^T g_{k+1}}{\|g_{k+1}\|^2} \leq -1 + \frac{|\beta_{k+1}| \omega_2}{\beta_{k+1}^{FR} (1 - \omega_2)},$$

et de (3.15)

$$\frac{-\omega_2}{1 - \omega_2} \leq \frac{|\beta_{k+1}|}{\beta_{k+1}^{FR}} \leq 1,$$

d'où :

$$\frac{-1}{1 - \omega_2} \leq \frac{d_{k+1}^T g_{k+1}}{\|g_{k+1}\|^2} \leq \frac{2\omega_2 - 1}{1 - \omega_2}.$$

Ce qui découle la démonstration. ■

Corollaire 3.2.1 *Sous les mêmes hypothèses et conditions du Théorème précédent, on déduit que pour tout $k \geq 1$, d_k est une directions de descente suffisante i.e., il existe une constante $C > 0$ telle que*

$$g_k^T d_k \leq -C \|g_k\|^2.$$

Preuve.

$$\begin{aligned} \frac{-1}{1-\omega_2} &\leq \frac{d_k^T g_k}{\|g_k\|^2} \leq \frac{2\omega_2 - 1}{1-\omega_2} \\ &\Rightarrow d_k^T g_k \leq -C \|g_k\|^2, \end{aligned}$$

ou $C = \frac{1-2\omega_2}{1-\omega_2}$. ■

Remarque 3.2.1 *La méthode de Fletcher-Reeves avec une recherche linéaire exacte génère des directions de descente. En effet, à chaque itération $k \geq 1$ on a :*

$$\begin{aligned} d_{k+1}^T g_{k+1} &= (-g_{k+1} + \beta_{k+1}^{FR} d_k)^T g_{k+1} \\ &= -g_{k+1}^T g_{k+1} + \beta_{k+1}^{FR} d_k^T g_{k+1} \\ &= -\|g_{k+1}\|^2. \end{aligned}$$

Puisque

$$\lambda_k = \arg \min_{\lambda > 0} f(x_k + \lambda d_k) = \arg \min_{\lambda > 0} \varphi_k(\lambda)$$

Donc λ_k vérifie la condition nécessaire d'optimalité

$$\dot{\varphi}_k(\lambda_k) = \nabla f(x_k + \lambda_k d_k)^T d_k = g_{k+1}^T d_k = 0, \forall k \geq 1.$$

3.2.3 Convergence de la méthode de FR

Le premier résultat de convergence de la méthode du gradient conjugué non linéaire (version Flecher Reeves) avec des recherches linéaires inexactes (recherche linéaire inexacte de Wolfe forte (3.9), (3.10) où $\omega_2 < \frac{1}{2}$) était démontré par [2]. Touati Ahmed et Story [56] ont généralisé ce résultat pour $0 \leq \beta_k \leq \beta_k^{FR}$. Gilbert et Nocedal [32]. ont généralisé ce résultat pour $|\beta_k| \leq \beta_k^{FR}$.

Théorème 3.2.2 *Supposons que l'hypothèse (3.1.1) soit satisfaite. Considérons une méthode du type (3.2) et (3.3) dont β_k satisfait (3.14) et le pas λ_k satisfait aux conditions de Wolfe fortes (3.9), (3.10) où $\omega_2 \in]0, \frac{1}{2}[$. Alors cette méthode est globalement convergente, dans le sens suivant :*

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0. \quad (3.19)$$

Preuve. Puisque les conditions du théorème (3.2.1) sont satisfaites alors on a :

$$\frac{-1}{1 - \omega_2} \leq \frac{d_k^T g_k}{\|g_k\|^2} \leq \frac{2\omega_2 - 1}{1 - \omega_2} \Rightarrow -\omega_2 d_{k-1}^T g_{k-1} \leq \frac{\omega_2}{1 - \omega_2} \|g_{k-1}\|^2.$$

D'autre part de (3.10)

$$|d_k^T g_{k+1}| \leq -\omega_2 d_k^T g_k \Rightarrow |d_{k-1}^T g_k| \leq -\omega_2 d_{k-1}^T g_{k-1},$$

d'où

$$|d_{k-1}^T g_k| \leq -\omega_2 d_{k-1}^T g_{k-1} \leq \frac{\omega_2}{1 - \omega_2} \|g_{k-1}\|^2. \quad (3.20)$$

De (3.2), (3.14) et (3.20) :

$$\begin{aligned} \|d_k\|^2 &= \left| \|g_k\|^2 - 2\beta_k d_{k-1}^T g_k + \beta_k^2 \|d_{k-1}\|^2 \right| \\ &\leq \|g_k\|^2 + |2\beta_k d_{k-1}^T g_k| + \beta_k^2 \|d_{k-1}\|^2 \\ &\leq \frac{1 + \omega_2}{1 - \omega_2} \|g_k\|^2 + (\beta_k^{FR})^2 \|d_{k-1}\|^2. \end{aligned}$$

Posons $\hat{\omega} = \frac{1 + \omega_2}{1 - \omega_2}$ on aura :

$$\begin{aligned} \|d_k\|^2 &\leq \hat{\omega} \|g_k\|^2 + (\beta_k^{FR})^2 \|d_{k-1}\|^2 \\ &\leq \hat{\omega} \|g_k\|^2 + (\beta_k^{FR})^2 \left[\hat{\omega} \|g_{k-1}\|^2 + (\beta_k^{FR})^2 \|d_{k-2}\|^2 \right] \\ &\leq \hat{\omega} \|g_k\|^4 \sum_{j=2}^k \|g_j\|^{-2} + \hat{\omega} \|g_k\|^4 \|g_1\|^{-2} \\ &= \hat{\omega} \|g_k\|^4 \sum_{j=1}^k \|g_j\|^{-2}. \end{aligned} \quad (3.21)$$

Supposons que g_k est borné en dehors du zéro ($\lim_{k \rightarrow \infty} \inf \|g_k\| \neq 0$), c'est-à-dire :

$$\|g_k\| \geq \omega > 0; \forall k \Rightarrow \|g_k\|^{-2} \leq \omega^{-2},$$

de (3.21) on a :

$$\begin{aligned} \|d_k\|^2 &\leq \hat{\omega} \|g_k\|^4 \sum_{j=1}^k \|g_j\|^{-2} \leq \hat{\omega} \frac{\lambda^4}{\omega^2} \sum_{j=1}^k 1 \\ &\Rightarrow \|d_k\|^2 \leq \hat{\omega} \frac{\lambda^4}{\omega^2} k, \end{aligned}$$

d'où

$$\sum_{k \geq 1} \frac{1}{\|d_k\|^2} \geq \frac{\omega^2}{\hat{\omega} \lambda^4} \sum_{k \geq 1} \frac{1}{k} > \infty. \quad (3.22)$$

Ce qui veut dire que $\sum_{k \geq 1} \frac{1}{\|d_k\|^2}$ est divergente. D'autre part, puisque les conditions du Théorème (3.2.1), et du théorème (3.2.2) sont satisfaites on a :

$$\sum_{k \geq 1} \cos^2 \theta_k \|g_k\|^2 < \infty,$$

et

$$c_1 \frac{\|g_k\|}{\|d_k\|} \leq \cos \theta_k \leq c_2 \frac{\|g_k\|}{\|d_k\|},$$

d'où

$$\begin{aligned} \sum_{k \geq 1} c_1^2 \frac{\|g_k\|^2}{\|d_k\|^2} \|g_k\|^2 &\leq \sum_{k \geq 1} \cos^2 \theta_k \|g_k\|^2 < \infty \\ &\Rightarrow \sum_{k \geq 1} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty \\ &\Rightarrow \sum_{k \geq 1} \frac{\omega^4}{\|d_k\|^2} < \infty \\ &\Rightarrow \sum_{k \geq 1} \frac{1}{\|d_k\|^2} < \infty. \end{aligned}$$

Ce qui contredit (3.22), d'où le résultat. Donc : $\lim_{k \rightarrow \infty} \inf \|g_k\| = 0$. ■

3.3 Méthode de Polak-Ribière-Polyak

Cette méthode fut découverte par Polak, Ribière [46] et Polyak [47].
La méthode de Polak-Ribière-Polyak la variante β_k est :

$$\beta_k^{PRP} = \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2} \quad (3.23)$$

Algorithme de la méthode de Polak, Ribière-Polyak

Etape 0 : (initialisation)

Soit x_0 le point de départ, $g_0 = \nabla f(x_0)$, poser $d_0 = -g_0$,
Poser $k = 0$ et aller à l'étape 1.

Etape 1 :

Si $g_k = 0$: STOP ($x^* = x_k$). "Test d'arrêt"
Si non aller à l'étape 2.

Etape 2 :

Définir $x_{k+1} = x_k + \lambda_k d_k$ avec :

$$\lambda_k = \min_{\lambda > 0} f(x_k + \lambda d_k)$$

$$d_{k+1} = -g_{k+1} + \beta_{k+1}^{PRP} d_k,$$

où

$$\beta_{k+1}^{PRP} = \frac{g_{k+1}^T [g_{k+1} - g_k]}{\|g_k\|^2} = \frac{g_{k+1}^T y_k}{\|g_k\|^2}.$$

Poser $k = k + 1$ et aller à l'étape 1.

3.3.1 Convergence de la méthode de PRP

La convergence de cette méthode est assurée pour une fonction fortement convexe avec recherche linéaire, mais si f n'est pas convexe elle ne converge pas. Pour remédier a cet inconvénient Powell [49] a modifié le choix de β_k^{PRP} , afin d'assurer la convergence avec une recherche linéaire exacte.

Théorème 3.3.1 *Si f est fortement convexe, continûment différentiable avec un gradient lipschitzien, alors la méthode de Polak Ribière avec recherche linéaire exacte génère une suite $\{x_k\}$ converge vers l'unique point x^* réalisant le minimum de f .*

Preuve. Montrons dans un premier temps que

$$\cos \theta_k = \frac{-d_k^T g_k}{\|g_k\| \|d_k\|},$$

est uniformément positif. Grâce à la recherche linéaire exacte, on a

$$\begin{aligned} d_{k-1}^T y_{k-1} &= d_{k-1}^T (g_k - g_{k-1}) \\ &= -d_{k-1}^T g_{k-1} = -(-g_{k-1} + \beta_{k-1}^{PRP} d_{k-2})^T g_{k-1} \\ &= \|g_{k-1}\|^2. \end{aligned}$$

La forte convexité de f implique que

$$d_{k-1}^T y_{k-1} = \frac{1}{\alpha_{k-1}} (x_k - x_{k-1})^T y_{k-1} \geq \frac{\eta}{\alpha_{k-1}} \|x_k - x_{k-1}\|^2,$$

où $\eta > 0$ est le module de forte convexité de f .

On en déduit, en utilisant la constante de lipschitz L de g :

$$\begin{aligned} |\beta_k^{PRP}| &= \frac{|g_k^T y_{k-1}|}{\|g_{k-1}\|^2} = \frac{|g_k^T y_{k-1}|}{d_{k-1}^T y_{k-1}} \\ &\leq \frac{\alpha_{k-1} L \|g_k\| \|x_k - x_{k-1}\|}{\eta \|x_k - x_{k-1}\|^2} = \frac{L}{\eta} \frac{\|g_k\|}{\|d_{k-1}\|}. \end{aligned}$$

On peut alors borner $\|d_k\|$ par :

$$\begin{aligned} \|d_k\| &\leq \|g_k\| + |\beta_k^{PRP}| \|d_{k-1}\| \\ &\leq \|g_k\| + \frac{L}{\eta} \frac{\|g_k\|}{\|d_{k-1}\|} \|d_{k-1}\| \\ &\leq \left(1 + \frac{L}{\eta}\right) \|g_k\|. \end{aligned}$$

En suite

$$\begin{aligned} d_k^T g_k &= -\|g_k\|^2 \\ &\leq -\left(1 + \frac{L}{\eta}\right) \|g_k\| \|d_k\|, \end{aligned}$$

ou encore

$$\cos \theta_k = \frac{-d_k^T g_k}{\|g_k\| \|d_k\|} \geq \left(1 + \frac{L}{\eta}\right)^{-1}.$$

D'après la proposition (3.1.1) et la recherche linéaire exacte, la condition de Zoutendijk est vérifiée. Mais f et donc $\{f(x_k)\}$ est bornée inférieurement (car f est fortement convexe). On en déduit que $g_k \rightarrow 0$. D'autre part, $\{x_k\}$ est bornée et possède donc des sous suites convergentes.

La limite de celles-ci ne peut être que l'unique minimum x^* de f (car $g_k \rightarrow 0$). Donc la suite $\{x_k\}$ converge vers x^* . ■

Remarque 3.3.1 *Si f n'est pas convexe, la méthode de Polak-Ribière-Polyak peut ne pas converger. Powell [50] a donné un exemple de fonction (de 3 variables, deux fois continûment différentiable) pour laquelle l'algorithme génère une suite $\{x_k\}$ dont aucun des points d'adhérence n'est stationnaire. En 1986 Powell [49], il a modifié la variante β_k^{PRP} en évitant les valeurs négatifs, autrement dit si à l'itération k on a : si $\beta_k^{PRP} < 0$, on redémarre en posant $\beta_k^{PRP} = 0$ (prendre la direction de la plus profonde pente)*

$$\beta_k = \max \{0, \beta_k^{PRP}\}.$$

Ce choix assure la convergence si le pas λ_k est déterminé par la règle de Wolfe forte.

Remarque 3.3.2 *Gilbert et Nocedal [32] ont assuré la convergence avec une recherche linéaire exacte ou inexacte, en hybridant les méthodes de FR et de PRP.*

Remarque 3.3.3 *Donc cette nouvelle méthode consiste à prendre la va-*

riante β_k comme suit :

$$\beta_k = \begin{cases} -\beta_k^{FR} & \text{si } \beta_k^{PRP} < -\beta_k^{FR} \\ \beta_k^{PRP} & \text{si } |\beta_k^{PRP}| \leq \beta_k^{FR} \\ \beta_k^{FR} & \text{si } \beta_k^{PRP} > \beta_k^{FR} \end{cases}$$

3.4 Méthode de la descente conjuguée

Cette méthode fut proposée en 1987 par Fletcher et Reeves [29],

La méthode de la descente conjuguée la variante β_k est :

$$\beta_k^{CD} = \frac{\|g_k\|^2}{-d_{k-1}^T g_{k-1}} \quad (3.24)$$

Algorithme de la méthode de descente conjuguée

Etape 0 : (initialisation)

Soit x_0 le point de départ, $g_0 = \nabla f(x_0)$, poser $d_0 = -g_0$.

Poser $k = 0$ et aller à l'étape 1.

Etape 1 :

Si $g_k = 0$: STOP ($x^* = x_k$). "Test d'arrêt"

Si non aller à l'étape 2.

Etape 2 :

Définir $x_{k+1} = x_k + \lambda_k d_k$ avec :

$$\lambda_k = \arg \min_{\lambda > 0} f(x_k + \lambda d_k)$$

$$d_{k+1} = -g_{k+1} + \beta_{k+1}^{CD} d_k,$$

où

$$\beta_{k+1}^{CD} = \frac{\|g_{k+1}\|^2}{-d_k^T g_k}.$$

Poser $k = k + 1$ et aller à l'étape 1.

3.4.1 La propriété de descente de la méthode de la DC

Fletcher [29]) a démontré que la méthode de la descente conjuguée est une méthode de descente si le pas λ_k est déterminé par la règle forte de Wolfe avec $\omega_2 \leq \frac{1}{2}$.

Dai et Yuan [16], ont démontré que cette méthode avec la règle de Wolfe relaxée où $0 < \omega_1 < \omega_2 < 1$ et $0 \leq \hat{\omega}_2 \leq 1$ génère des directions de descente à chaque itération $k \geq 1$.

Théorème 3.4.1 *Supposons que l'hypothèse (3.1.1) est satisfaite.*

Pour toute méthode du type (3.2) et (3.3) dont β_k satisfait à (3.24) et le pas λ_k satisfait aux conditions de Wolfe relaxées suivantes : (3.11) et (3.12)

$$\begin{aligned} f(x_k + \lambda_k d_k) &\leq f(x_k) + \omega_1 \lambda_k d_k^T g_k \\ \text{et } \omega_2 d_k^T g_k &\leq d_k^T g_{k+1} \leq -\hat{\omega}_2 d_k^T g_k, \end{aligned}$$

où $0 < \omega_1 < \omega_2 < 1$ et $0 \leq \hat{\omega}_2 \leq 1$.

Alors la méthode génère des directions de descente suffisante à chaque itération $k \geq 1$.

Preuve. On a

$$\begin{aligned}
 -d_k^T g_k &= -(-g_k + \beta_k^{CD} d_{k-1})^T g_k \\
 &= \|g_k\|^2 \left[1 + \frac{d_{k-1}^T g_k}{d_{k-1}^T g_{k-1}} \right] \\
 \Rightarrow \frac{-d_k^T g_k}{\|g_k\|^2} &= 1 + \frac{d_{k-1}^T g_k}{d_{k-1}^T g_{k-1}}.
 \end{aligned}$$

D'autre part de (3.12)

$$\begin{aligned}
 \omega_2 d_k^T g_k &\leq d_k^T g_{k+1} \leq -\acute{\omega}_2 d_k^T g_k \\
 \Rightarrow 1 - \acute{\omega}_2 &\leq 1 + \frac{d_{k-1}^T g_k}{d_{k-1}^T g_{k-1}} \leq 1 + \omega_2,
 \end{aligned}$$

d'où

$$1 - \acute{\omega}_2 \leq \frac{-d_k^T g_k}{\|g_k\|^2} \leq 1 + \omega_2.$$

Donc si $\|g_k\| \neq 0$, on a :

$$d_k^T g_k \leq -C \|g_k\|^2 \text{ où } C = 1 - \acute{\omega}_2 > 0,$$

et donc d_k est une direction de descente suffisante. ■

3.4.2 Convergence de la méthode de la DC

Yuan [?] a démontré la convergence au sens (3.19) de cette méthode avec un pas satisfaisant aux conditions (3.11) et (3.12) si $\omega_2 < \frac{1}{2}$ et $\acute{\omega}_2 = 0$.

Dai et Yuan [16] ont démontré ce résultat pour $\omega_2 < 1$ et $\acute{\omega}_2 = 0$.

Théorème 3.4.2 *Supposons que l'hypothèse (3.1.1) est satisfaite.*

Toute méthode du type (3.2) et (3.3) dont β_k vérifie (3.24) et le pas λ_k est déterminé par la règle de Wolfe relaxée où $0 < \omega_1 < \omega_2 < 1$ et $\acute{\omega}_2 = 0$ est de descente convergente dans le sens où

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0.$$

Preuve. Du théorème (3.4.1) on a :

$$\begin{aligned}
 1 - \omega_2 &\leq \frac{-d_k^T g_k}{\|g_k\|^2} \leq 1 + \omega_2 \\
 \Rightarrow 1 &\leq \frac{-d_k^T g_k}{\|g_k\|^2} \leq 1 + \omega_2 \\
 \Rightarrow (1 + \omega_2)^{-1} &\leq \frac{\|g_k\|^2}{-d_k^T g_k} \leq 1 \\
 \Rightarrow (1 + \omega_2)^{-1} &\leq \frac{\|g_{k+1}\|^2 \|g_k\|^2}{-d_k^T g_k \|g_{k+1}\|^2} \leq 1 \\
 \Rightarrow (1 + \omega_2)^{-1} &\leq \frac{\beta_{k+1}^{CD}}{\beta_{k+1}^{FR}} \leq 1 \\
 \Rightarrow \beta_{k+1}^{CD} &\leq \beta_{k+1}^{FR}.
 \end{aligned}$$

Donc β_{k+1}^{CD} vérifie l'inégalité (3.15). Donc : $\lim_{k \rightarrow \infty} \inf \|g_k\| = 0$. ■

3.5 Méthode de Dai-Yuan

Cette méthode été découverte par Y. H. Dai et Y. Yuan [13] en 1999, β_k est égale à :

$$\beta_{k+1}^{DY} = \frac{\|g_{k+1}\|^2}{d_k^T y_k}; y_k = g_{k+1} - g_k. \quad (3.25)$$

Remarquons que pour cette variante on a le même numérateur et la même dénominateur que pour les variantes de Fletcher-Reeves et Hestenes-Stiefel respectivement. Cette méthode possède plusieurs propriétés, par exemple elle possède la propriété de descente à chaque itération, la convergence au sens (3.19) si le pas est déterminé par la règle de Wolfe faible.

Algorithme de la Méthode de Dai-Yuan avec la règle de Wolfe faible

Etape 0 : (initialisation)

Soit x_0 le point de départ, $g_0 = \nabla f(x_0)$, poser $d_0 = -g_0$

Poser $k = 0$ et aller à l'étape 1.

Etape 1 :

Si $g_k = 0$: STOP ($x^* = x_k$). "Test d'arrêt"

Si non aller à l'étape 2.

Etape 2 :

Définir $x_{k+1} = x_k + \lambda_k d_k$ avec :

$$d_{k+1} = -g_{k+1} + \beta_{k+1}^{DY} d_k,$$

où

$$\beta_{k+1}^{DY} = \frac{\|g_{k+1}\|^2}{d_k^T [g_{k+1} - g_k]} = \frac{\|g_{k+1}\|^2}{d_k^T y_k}.$$

Poser $k = k + 1$ et aller à l'étape 1.

3.5.1 La propriété de descente de la méthode de DY

Dai et Yuan [26] ont démontré que à chaque itération $k \geq 1$, la direction recherchée par cette méthode avec la recherche de Wolfe faible, est de descente si la fonction objectif f est strictement convexe. En 1999, dans [13], ils s'ont généralisé ce résultat pour toute fonction régulière.

Théorème 3.5.1 *Supposons que L'hypothèse (3.1.1) soit satisfaite. Pour toute méthode du type (3.1) et (3.3) dont β_k satisfait à (3.25) et le pas λ_k satisfait aux conditions de Wolfe faibles suivantes :*

$$\begin{aligned} f(x_k + \lambda d_k) &\leq f(x_k) + \omega_1 \lambda \nabla f(x_k)^T \cdot d_k \\ \text{et } d_k^T g_{k+1} &\geq \omega_2 d_k^T g_k, \end{aligned}$$

où $0 < \omega_1 < \omega_2 < 1$, alors toutes les directions générées sont de descente, autrement dit :

$$d_k^T g_k < 0; \quad \forall k \geq 1. \quad (3.26)$$

Preuve. La démonstration se fait par récurrence.

1) Pour $k = 1$:

$$d_1^T g_1 = -\|g_1\|^2 < 0$$

2) Supposons que (3.26) est satisfaite pour $k > 1$ et démontrons qu'elle le sera pour $k + 1$: Supposons que :

$$d_k^T g_k < 0; \quad k > 1.$$

En utilisant (3.7), on aura :

$$\begin{aligned} d_k^T y_k &= d_k^T (g_{k+1} - g_k) > d_k^T (g_{k+1} - g_k) \\ &= (\omega_2 - 1) d_k^T g_k > 0. \end{aligned}$$

D'autre part :

$$\begin{aligned} d_{k+1}^T g_{k+1} &= (-g_{k+1} + \beta_{k+1}^{DY} d_k)^T g_{k+1} \\ &= -\|g_{k+1}\|^2 + \beta_{k+1}^{DY} d_k^T g_{k+1} \\ &= -\|g_{k+1}\|^2 + \frac{\|g_{k+1}\|^2}{d_k^T y_k} d_k^T g_{k+1} \\ &= -\|g_{k+1}\|^2 + \frac{\|g_{k+1}\|^2}{d_k^T y_k} d_k^T (y_k + g_k) \\ &= -\|g_{k+1}\|^2 + \|g_{k+1}\|^2 + \frac{\|g_{k+1}\|^2}{d_k^T y_k} d_k^T g_k \\ &= \frac{\|g_{k+1}\|^2}{d_k^T y_k} d_k^T g_k, \end{aligned}$$

or puisque : $d_k^T g_k < 0$, $d_k^T y_k > 0$, il en résulte

$$d_{k+1}^T g_{k+1} < 0.$$

Ce qui découle la démonstration. ■

3.5.2 Convergence de la méthode de DY

Dai-Yuan ont démontré la convergence au sens (3.19) si le pas λ_k est déterminé par la règle de Wolfe faible. On a alors le théorème suivant.

Théorème 3.5.2 *Supposons que la proposition (3.1.1) est satisfaite. Pour toute méthode du type (3.1) et (3.3) dont β_k satisfait à (3.25) et le pas λ_k satisfait aux conditions de Wolfe faibles. La suite $\{x_k\}$ générée par l'algorithme de Dai-Yuan converge dans le sens*

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0$$

Preuve. En utilisant le théorème (3.1.1), on aura :

$$\sum_{k \geq 1} \cos^2 \theta \|g_k\|^2 < \infty,$$

d'autre part on a

$$\begin{aligned} \|d_{k+1} + g_{k+1}\|^2 &= \|\beta_{k+1}^{DY} d_k\|^2 \\ \Rightarrow \|d_{k+1}\|^2 &= (\beta_{k+1}^{DY})^2 \|d_k\|^2 - 2d_{k+1}^T g_{k+1} - \|g_{k+1}\|^2. \end{aligned} \quad (3.27)$$

De (3.25)

$$\begin{aligned} d_{k+1}^T g_{k+1} &= (-g_{k+1} + \beta_{k+1}^{DY} d_k)^T g_{k+1} \\ &= \frac{\|g_{k+1}\|^2}{d_k^T g_k} d_k^T g_k = \beta_{k+1}^{DY} d_k^T g_k. \end{aligned}$$

Donc

$$\beta_{k+1}^{DY} = \frac{d_{k+1}^T g_{k+1}}{d_k^T g_k},$$

remplaçons ceci dans (3.26), on aura

$$\begin{aligned} \frac{\|d_{k+1}\|^2}{(d_{k+1}^T g_{k+1})^2} &= \frac{(\beta_{k+1}^{DY})^2 \|d_k\|^2}{(d_{k+1}^T g_{k+1})^2} - \frac{2d_{k+1}^T g_{k+1}}{(d_{k+1}^T g_{k+1})^2} - \frac{\|g_{k+1}\|^2}{(d_{k+1}^T g_{k+1})^2} \\ &= \frac{\|d_k\|^2}{(d_k^T g_k)} - \left[\frac{1}{\|g_{k+1}\|^2} + 2 \frac{1}{d_{k+1}^T g_{k+1}} - \frac{\|g_{k+1}\|^2}{(d_{k+1}^T g_{k+1})^2} \right] + \frac{1}{\|g_{k+1}\|^2} \\ &= \frac{\|d_k\|^2}{(d_k^T g_k)} - \left[\frac{1}{\|g_{k+1}\|} + \frac{\|g_{k+1}\|}{d_{k+1}^T g_{k+1}} \right]^2 + \frac{1}{\|g_{k+1}\|^2} \\ &\leq \frac{\|d_k\|^2}{(d_k^T g_k)} + \frac{1}{\|g_{k+1}\|^2}, \end{aligned}$$

d'où

$$\begin{aligned} \frac{\|d_k\|^2}{(d_k^T g_k)^2} &\leq \frac{1}{\|g_k\|^2} + \frac{\|d_{k-1}\|^2}{(d_{k-1}^T g_{k-1})} \\ &\leq \frac{1}{\|g_k\|^2} + \frac{1}{\|g_{k-1}\|^2} + \frac{\|d_{k-2}\|^2}{(d_{k-2}^T g_{k-2})} \\ &\leq \sum_{i=1}^k \frac{1}{\|g_i\|^2}. \end{aligned}$$

Supposons maintenant :

$$\exists \omega > 0 \text{ tel que } \|g_k\| > \omega; \forall k.$$

On aura :

$$\begin{aligned} \frac{\|d_k\|^2}{(d_k^T g_k)^2} &\leq \frac{1}{\|g_i\|^2} \leq \frac{1}{\omega^2} \sum_{i=1}^k 1 = \frac{1}{\omega^2} k \\ &\Rightarrow \sum_{k \geq 1} \frac{(d_k^T g_k)^2}{\|d_k\|^2} \geq \omega^2 \sum_{k \geq 1} \frac{1}{k}, \end{aligned}$$

d'où

$$\sum_{k \geq 1} \frac{(d_k^T g_k)^2}{\|d_k\|^2} = \infty,$$

ce qui contredit (3.18). Ce qui achève la démonstration. ■

3.6 Méthode de Hestenes et Stiefel HS

Cette méthode a été proposée en 1952, dans sa version linéaire par Hestenes et Steifel [34]. C'est d'ailleurs le premier article publié qui parle de la méthode du gradient conjugué. Les auteurs utilisent cette méthode itérative pour résoudre des systèmes linéaires, β_k est égale à :

$$\beta_k^{HS} = \frac{g_k^T (g_k - g_{k-1})}{d_{k-1}^T (g_k - g_{k-1})}.$$

La méthode HS possède la propriété de conjugaison suivante

$$d_{k-1}^T (g_k - g_{k-1}) = 0,$$

indépendamment de la recherche linéaire.

Modification de la méthode PRP en agissant sur le coefficient β_k

Sur la base des connaissances tirées de l'exemple de Powell, Dai a suggéré [50] la modification suivante dans le paramètre de mise à jour pour la méthode PRP :

$$\beta_k^{PRP+} = \max \{0, \beta_k^{PRP}\}.$$

Dans [32] Gilbert et Nocedal ont prouvé la convergence de la méthode PRP+. L'analyse de Gilbert et Nocedal s'applique à une classe d'algorithmes du gradient conjugué qui possèdent la propriété suivante :

Considérons une méthode itérative de la forme (3.2) et (3.3), et supposons la suite des gradients $\{g_k\}$ vérifie la condition suivante :

il existe deux constantes positives γ et $\bar{\gamma}$ telles que :

$$\forall k \in \mathbb{N} : 0 < \gamma \leq \|g_k\| \leq \bar{\gamma}.$$

Théorème 3.6.1 [32] *Considérons une méthode itérative du gradient conjugué de la forme (3.2) et (3.3), qui satisfait aux conditions suivantes :*

- (a) $\beta_k \geq 0$
- (b) *Les directions de recherche sont des directions de descente suffisante i.e., il existe une constante $C > 0$ telle que : $g_k^T d_k \leq -C \|g_k\|^2 ; k = 1, 2, \dots$*
- (c) *La condition de Zoutendijk ($\sum_{k=0}^{\infty} \frac{(g_k^T d_k)^2}{\|g_k\|^2} < \infty$) est vérifiée*
- (d) *La propriété : 'il existe des constantes $b > 1$ et $\lambda > 0$ telle que : $\forall k \in \mathbb{N} : |\beta_k| \leq b$ et $\|x_{k+1} - x_k\| \leq \lambda$ implique $|\beta_k| \leq \frac{1}{2b}$ est vérifiée.*
- (e) *La condition de Lipschitz et la condition de bornitude sont assurées. Alors la suite générée par cet algorithme est globalement convergente i.e.*

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0.$$

Comme corollaire de ce résultat, Si on associe à la méthode PRP+, des directions de recherche satisfaisant la condition de descente suffisante DS et des recherches linéaires de Wolfe faible, alors la méthode PRP+ est globalement convergente.

3.6.1 Convergence de la méthode HS

Pour une recherche linéaire exacte, $\beta_k^{HS} = \beta_k^{PRP}$. Par conséquent, les propriétés de convergence de la méthode de HS doivent être similaires aux propriétés de convergence de la méthode de PRP. En particulier, si on prend en considération l'exemple de Powell [50], la méthode HS associée à une recherche linéaire exacte, peut ne pas converger pour une fonction non linéaire. Il est facile de vérifier que si les directions de recherche satisfont à la condition de descente suffisante et si on utilise une recherche linéaire de Wolfe, alors la méthode de HS satisfait la propriété.

Comme la méthode PRP +, si nous notons,

$$\beta_k^{HS+} = \max \{ \beta_k^{HS}, 0 \},$$

il résulte du théorème (3.6.1), que la méthode HS+ est globalement convergente.

3.7 Méthode de Liu et Storey LS

La méthode Liu et Storey LS est également identique à la méthode de PRP pour une recherche linéaire exacte. Bien que peu de recherches ont été faites sur ce choix pour le paramètre de mise à jour, à l'exception de l'article [38], nous nous attendons à ce que les techniques développées pour l'analyse de la méthode PRP devraient s'appliquer à la méthode LS.

3.8 Combinaison unifiée

Les méthodes quasi-Newtonniennes ont été regroupées en familles dites de Broyden [6] et ont été étudiées ensemble. De la même manière, Dai et Yuan [11, 10], ont unifié plusieurs méthodes du gradient conjugué et ont proposé une famille à un paramètre des méthodes du gradient conjugué en posant

$$\beta_k^{DY-unif} = \frac{\|g_{k+1}\|^2}{\lambda \|g_k\|^2 + (1 - \lambda) d_k^T y_k},$$

où $\lambda_k \in [0; 1]$ est un paramètre. Les méthodes unifiées sont en fait une combinaison convexe des méthodes FR et DY. Le procédé de FR correspond à $\lambda_k = 1$, tandis que la méthode de DY

correspond à $\lambda_k = 0$. Dans [22], cette famille est étudiée en considérant $\lambda_k \in]-\infty, +\infty[$. Si la condition de Lipschitz est vérifiée et si on utilise une recherche linéaire de Wolfe avec

$$\sigma_2 - 1 \leq (\sigma_2 + \acute{\sigma}_2) \lambda_k \leq 1, \tag{3.28}$$

alors on a la convergence globale pour chaque membre de la famille. En considérant des combinaisons convexes des numérateurs et des dénominateurs

de β_k^{FR} et β_k^{HS} , Nazareth [44] propose et de façon indépendante, une famille à deux paramètres des méthodes du gradient conjugué avec :

$$\beta_k = \frac{\lambda_k \|g_{k+1}\|^2 + (1 - \lambda_k) g_{k+1}^T y_k}{\mu_k \|g_k\|^2 + (1 - \mu_k) d_k^T y_k}.$$

Constatant que les six méthodes du gradient conjugué qu'on a vu précédemment ont deux numérateurs et trois dénominateurs, Dai et Yuan [24] ont introduit une nouvelle famille de gradients conjugués encore plus large à trois paramètres, ils ont choisi β_k comme suit :

$$\beta_k = \frac{(1 - \lambda_k) \|g_{k+1}\|^2 + \lambda_k g_k^T y_k}{(1 - \mu_k - \omega_k) \|g_k\|^2 + \mu_k d_k^T y_k - \omega_k d_k^T g_k},$$

où $\lambda_k \in [0, 1]$, $\mu_k \in [0, 1]$, $\omega_k \in [0, 1 - \mu_k]$ sont des paramètres.

Cette famille à trois paramètres comprend les six méthodes standard du gradient conjugué, les familles à 1 ou 2 paramètres vues ci dessus et de nombreuses méthodes hybrides comme cas particuliers. Afin de veiller à ce que les directions de recherche générés par cette famille soient des directions de descente, le critère de redémarrage de Powell [51] est utilisé. Posons $d_k = -g_k$, si

$$|g_k^T g_{k-1}| > \xi \|g_k\|^2,$$

où $\xi > 0$ est une certaine constante fixe. Si on utilise une recherche linéaire inexacte de Wolfe forte (3.9)-(3.10) avec

$$(1 + \xi) \sigma_2 \leq \frac{1}{2}.$$

Dai et Yuan [24] ont montré que les directions de recherche sont des directions de descente. Les résultats de convergence globale sont également établis. Dans [23] Dai et Liao modifient le numérateur du paramètre de mise à jour pour obtenir HS

$$\begin{aligned} \beta_k^{DL} &= \frac{g_k^T (g_{k+1} - g_k - t s_k)}{d_k^T (g_{k+1} - g_k)} \\ &= \beta_k^{HS} - t \frac{g_{k+1}^T s_k}{d_k^T (g_{k+1} - g_k)}, s_k = x_{k+1} - x_k = \alpha_k d_k, \end{aligned}$$

où $t > 0$ est une constante. Pour une recherche linéaire exacte, g_{k+1} est orthogonal à s_k . Ainsi, pour une recherche linéaire exacte, $\beta_k^{DL} = \beta_k^{HS}$.

Par conséquent la méthode du gradient conjugué DL se réduit à la méthode HS et la méthode PRP. Encore une fois, en raison de l'exemple de Powell, la méthode DL peut ne pas converger pour une recherche linéaire exacte. Semblable à la méthode PRP+, Dai et Liao ont également modifié leur formule de la façon suivante pour assurer la convergence :

$$\beta_k^{DL+} = \max \left\{ \frac{g_{k+1}^T (g_{k+1} - g_k)}{d_k^T (g_{k+1} - g_k)}, 0 \right\} - t \frac{g_{k+1}^T s_k}{d_k^T (g_{k+1} - g_k)}.$$

Si les hypothèses de Lipschitz et de bornétude sont vérifiées et si d_k satisfait à la condition de descente suffisante (2.5), il est montré dans [23] que la méthode DL+, mise en œuvre une recherche linéaire inexacte de Wolfe forte (3.9)-(3.10), est globalement convergente. Très récemment, dans un nouveau développement de cette stratégie de mise à jour, Yabe et Takano [60] proposent le choix suivant pour le paramètre de mise à jour, basée sur une condition sécante modifiée donnée par Zhang et al. [63, 64] :

$$\beta_k^{YT} = \frac{g_{k+1}^T (z_k - t s_k)}{d_k^T z_k},$$

où

$$\begin{aligned} z_k &= (g_{k+1} - g_k) + \left(\frac{\rho \theta_k}{s_k^T \eta_k} \right) \eta_k \\ \theta_k &= 6(f_k - f_{k+1}) + 3(g_k + g_{k+1})^T s_k, \end{aligned}$$

$\rho \geq 0$ est une constante et $\eta_k \in \mathbb{R}^n$ satisfait $s_k^T \eta_k \neq 0$. Encore une fois, et suivant la démarche appliquée à la méthode PRP+, Yabe et Takano ont modifié leur

formule de la façon suivante pour assurer la convergence :

$$\beta_k^{YT+} = \max \left\{ \frac{g_{k+1}^T z_k}{d_k^T z_k}, 0 \right\} - t \frac{g_{k+1}^T s_k}{d_k^T z_k}.$$

Ils montrent que la méthode YT+ est globalement convergente si les hypothèses de Lipschitz et de bornétude sont vérifiées et d_k satisfait à la condition de descente suffisante, et si une recherche linéaire inexacte de Wolfe forte (3.9)-(3.10) est employée avec

$$0 \leq \rho < \frac{1 - \sigma}{3(1 + \sigma - 2\delta)}.$$

Sellami, Laskri et Benzine [53] ont proposé une nouvelle famille à deux paramètres des méthodes du gradient conjugué pour l'optimisation sans contraintes et ont étudié ses propriétés et sa convergence. β_k^* est de la forme générale suivante :

$$\beta_k^* = \frac{\varphi_k}{\dot{\varphi}_{k-1}},$$

où $\dot{\varphi}_{k-1}$ satisfait

$$\dot{\varphi}_{k-1} = (1 - \mu_k - \omega_k) \|g_{k-1}\|^2 - (\mu_k + \omega_k) g_{k-1}^T d_{k-1}$$

et

$$\varphi_k = \mu \|g_k\|^2 + (1 - \mu)(-g_k^T d_k), \quad \mu \in [0, 1].$$

Avec ces données, nous obtenons

$$\beta_k^* = \frac{\|g_k\|^2}{(1 - \mu_k - \omega_k) \|g_{k-1}\|^2 - \omega_k d_{k-1}^T g_{k-1} + \mu_k g_{k-1}^T d_{k-1}}$$

Chapitre 4

La convergence de la méthode de descente d'une famille à deux paramètres des méthodes du gradient conjugué avec une formule fixe du pas

Nous présentons dans ce chapitre la convergence globale d'une famille à deux paramètres des méthodes du gradient conjugué [5], en utilisant une nouvelle formule de pas qui a été suggérée par Wu [59], en plus des conclusions numériques confirmant l'efficacité de la formule proposée et sa comparaison graphiquement avec le pas de Sun et Chen [8, 55].

4.1 Introduction

Considérons le problème de minimisation sans contrainte suivante :

$$\min f(x), x \in \mathbb{R}^n, \quad (4.1)$$

où f est une fonction différentiable, et

$$x_{k+1} = x_k + \alpha_k d_k, \quad (4.2)$$

où

CHAPITRE 4. LA CONVERGENCE DE LA MÉTHODE DE
DESCENTE D'UNE FAMILLE À DEUX PARAMÈTRES DES
MÉTHODES DU GRADIENT CONJUGUÉ AVEC UNE FORMULE
FIXE DU PAS

$$d_k = \begin{cases} -g_k & \text{for } k = 1, \\ -g_k + \beta_k d_{k-1} & \text{for } k \geq 2, \end{cases} \quad (4.3)$$

où $g_k = \nabla f(x_k)$, est le gradient de f à x_k , avec les formules suivantes

$$D_k = (1 - \mu_k - \omega_k) \|g_{k-1}\|^2 + \mu_k d_{k-1}^T y_{k-1} - \omega_k d_{k-1}^T g_{k-1}, \quad (4.4)$$

$$\beta_k^{\mu_k, \omega_k} = \frac{g_k^T y_{k-1}}{D_k}, \quad (4.5)$$

où $\|\cdot\|$ signifie la norme euclidien et $y_{k-1} = g_k - g_{k-1}$, et D_k dépend des paramètres $\mu_k \in [0, 1[$ et $\omega_k \in [0, 1 - \mu_k[$. Remarquons que la direction de descente d_k est définie telle que

$$g_k^T d_k = -c \|g_k\|^2, \quad (4.6)$$

où $0 < c < 1 - \mu_k - \omega_k$.

L'expression paramétrisé (4.5) est prise [8]. Elle couvre uniquement un sous-ensemble d'une grande famille introduite par Dai et Yan [24]. Trois versions classiques du GC non linéaire sont des cas particuliers de formule (4.5) :

$$\beta_k^{HS} = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}, \text{ Hestenes-Stiefel [34]} \quad (4.7)$$

$$\beta_k^{PRP} = \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2}, \text{ Polak-Ribière-Polyak [47]} \quad (4.8)$$

$$\beta_k^{LS} = \frac{g_k^T y_{k-1}}{d_{k-1}^T g_{k-1}}. \text{ Liu-Storey [38]} \quad (4.9)$$

D'autres cas sont très importants, la méthode de Fletcher-Reeves [30], la méthode de descente conjugué [29], et la méthode de Dai-Yuan [21]. Dans la mise en œuvre de toute méthode du gradient conjugué (CG), la taille de pas est souvent déterminée par certaines conditions de recherche linéaire comme les conditions de Wolfe [58]. Ces types de recherche linéaire impliquent beaucoup de calculs des valeurs de la fonction et des gradients, qui devient souvent une charge importante pour les problèmes à grande échelle, qui à

CHAPITRE 4. LA CONVERGENCE DE LA MÉTHODE DE
DESCENTE D'UNE FAMILLE À DEUX PARAMÈTRES DES
MÉTHODES DU GRADIENT CONJUGUÉ AVEC UNE FORMULE
FIXE DU PAS

poussé Sun et Zhang [55] pour suivre la méthode du gradient conjugué où ils ont calculé le pas au lieu de la recherche linéaire selon la formule suivante :

$$\alpha_k = -\delta g_k^T d_k / \|d_k\|_{Q_k}^2, \quad (4.10)$$

où $\|d_k\|_{Q_k} = \sqrt{d_k^T Q_k d_k}$, $\delta \in (0, \nu_{\min}/\tau)$, τ est une constante de Lipschitz de f et $\{Q_k\}$ est une séquence de matrices définies positives satisfaisant pour des constantes positives ν_{\min} et ν_{\max}

$$\nu_{\min} d^T d \leq d^T Q_k d \leq \nu_{\max} d^T d, \quad \forall k, \quad \forall d \in \mathbb{R}^n.$$

Dernièrement, Wu [59] a réussi à obtenir une formule de pas. Pour simplifier, nous introduisons quelques notations comme suit :

f_k, g_k, \bar{f}_{k+1} et \bar{g}_{k+1} désignent $f(x_k), \nabla f(x_k), f(x_k + d_k)$, et $\nabla f(x_k + d_k)$, respectivement.

Motivant par les méthodes de quasi-Newton dans [63] et [64], nous dérivons une formule de mise à jour pour $d_k^T Q_k d_k$ comme suit.

Par la formule de Taylor aux termes de troisième ordre, f et $g^T d$ peuvent être écrit comme suit respectivement

$$f_k = \bar{f}_{k+1} - \bar{g}_{k+1}^T d_k + \frac{1}{2!} d_k^T \bar{G}_{k+1} d_k - \frac{1}{3!} d_k^T (T_{k+1} d_k) d_k + O(\|d_k\|^4), \quad (4.11)$$

$$g_k^T d_k = \bar{g}_{k+1}^T d_k - d_k^T \bar{G}_{k+1} d_k + \frac{1}{2!} d_k^T (T_{k+1} d_k) d_k + O(\|d_k\|^4), \quad (4.12)$$

où $T_{k+1} \in \mathbb{R}^{n \times n \times n}$ est le tenseur de f en $x_k + d_k$. Annuler les termes qui incluent le tenseur dans (4.11) et (4.12), alors :

$$d_k \bar{G}_{k+1} d_k = (\bar{g}_{k+1} - g_k)^T d_k + 6(f_k - \bar{f}_{k+1}) + 3(g_k + \bar{g}_{k+1})^T d_k + O(\|d_k\|^4). \quad (4.13)$$

Ensuite, après la substitution de $d_k \bar{G}_{k+1} d_k$ comme une approximation de $d_k^T Q_k d_k$, nous obtenons une nouvelle formule du α_k :

$$\alpha_k = \frac{-\delta g_k^T d_k}{(\bar{g}_{k+1} - g_k)^T d_k + \gamma \theta_k}, \quad (4.14)$$

où

$$\theta_k = 6(f_k - \bar{f}_{k+1}) + 3(g_k + \bar{g}_{k+1})^T d_k, \quad (4.15)$$

CHAPITRE 4. LA CONVERGENCE DE LA MÉTHODE DE
DESCENTE D'UNE FAMILLE À DEUX PARAMÈTRES DES
MÉTHODES DU GRADIENT CONJUGUÉ AVEC UNE FORMULE
FIXE DU PAS

δ et γ sont des paramètres satisfaisant

$$\delta \in (0, \kappa/\tau), \quad (4.16)$$

et

$$\gamma \geq 0 \text{ si } \tau = \kappa, \text{ ou } \gamma \in \left(0, \frac{\kappa - \delta\tau}{3(\tau - \kappa)}\right) \text{ si } \tau > \kappa, \quad (4.17)$$

κ et τ sont définis dans l'hypothèse (4.2.1) ci-dessous.

Il a prouvé que la formule ci-dessus pour α_k peut assurer la convergence globale des méthodes de CD, FR et PR.

Dans ce chapitre, notre objectif est l'utilisation de la formule (4.14) pour prouver la convergence de la famille à deux paramètres de la méthode du gradient conjugué.

4.2 Préliminaires

Nous adoptons l'hypothèse suivante sur la fonction f qui est couramment utilisé dans la littérature.

Proposition 4.2.1 *la fonction f est LC^1 et fortement convexe dans \mathbb{R}^n , i.e, il existe des constantes $\tau > 0$ et $\kappa \geq 0$ tels que*

$$\|\nabla f(x_{k+1}) - \nabla f(x_k)\| \leq \tau \|x_{k+1} - x_k\|, \quad \forall x_{k+1}, x_k \in \mathbb{R}^n, \quad (4.18)$$

et

$$[\nabla f(x_{k+1}) - \nabla f(x_k)]^T (x_{k+1} - x_k) \geq \kappa \|x_{k+1} - x_k\|^2, \quad \forall x_{k+1}, x_k \in \mathbb{R}^n, \quad (4.19)$$

ou de manière équivalente,

$$f(x_{k+1}) - f(x_k) \geq \nabla f(x_k)^T (x_{k+1} - x_k) + \frac{\kappa}{2} \|x_{k+1} - x_k\|^2, \quad \forall x_{k+1}, x_k \in \mathbb{R}^n. \quad (4.20)$$

Notez que l'hypothèse (4.2.1) implique que l'ensemble

$$L = \{x \in \mathbb{R}^n / f(x) \leq f(x_1)\} \text{ est bornée.}$$

Si le pas est donnée par la formule (4.14), le titulaire récursion suivant.

CHAPITRE 4. LA CONVERGENCE DE LA MÉTHODE DE
DESCENTE D'UNE FAMILLE À DEUX PARAMÈTRES DES
MÉTHODES DU GRADIENT CONJUGUÉ AVEC UNE FORMULE
FIXE DU PAS

Lemme 4.2.1 *Supposons que x_k est donné par (4.2), (4.3) et (4.14). alors*

$$g_{k+1}^T d_k = \rho_k g_k^T d_k, \quad (4.21)$$

pour tous k , où

$$0 < \rho_k = 1 - \delta \phi_k \|d_k\|^2 / \left[(\bar{g}_{k+1} - g_k)^T d_k + \gamma \theta_k \right], \quad (4.22)$$

et

$$\phi_k = \begin{cases} 0 & \text{pour } \alpha_k = 0, \\ (g_{k+1} - g_k)^T (x_{k+1} - x_k) / \|x_{k+1} - x_k\|^2 & \text{pour } \alpha_k \neq 0. \end{cases} \quad (4.23)$$

Preuve. Si $\alpha_k = 0$, alors $\rho_k = 1$ et $x_{k+1} = x_k$. Ainsi, (4.22) est vrai.

Maintenant, nous supposons que $\alpha_k \neq 0$. D'après (4.22) et (4.23), nous avons

$$\begin{aligned} g_{k+1}^T d_k &= g_k^T d_k + (g_{k+1} - g_k)^T d_k \\ &= g_k^T d_k + \alpha_k^{-1} (g_{k+1} - g_k)^T (x_{k+1} - x_k) \\ &= g_k^T d_k + \alpha_k^{-1} \phi_k \|x_{k+1} - x_k\|^2 \\ &= g_k^T d_k + \alpha_k \phi_k \|d_k\|^2 \\ &= g_k^T d_k - \left\{ \delta g_k^T d_k / \left[(\bar{g}_{k+1} - g_k)^T d_k + \gamma \theta_k \right] \right\} \phi_k \|d_k\|^2 \\ &= \left\{ 1 - \delta \phi_k \|d_k\|^2 / \left[(\bar{g}_{k+1} - g_k)^T d_k + \gamma \theta_k \right] \right\} g_k^T d_k \\ &= \rho_k g_k^T d_k. \end{aligned}$$

La preuve est complète. ■

Lemme 4.2.2 *Supposons que l'hypothèse (4.2.1) est satisfaite. On a les inégalités suivantes*

$$\kappa \|x_{k+1} - x_k\|^2 \leq (g_{k+1} - g_k)^T (x_{k+1} - x_k) \leq \tau \|x_{k+1} - x_k\|^2, \quad (4.24)$$

et

$$[\kappa + 3\gamma(\kappa - \tau)] \|d_k\|^2 \leq (\bar{g}_{k+1} - g_k)^T d_k + \gamma \theta_k \leq (1 + 3\gamma) \tau \|d_k\|^2, \quad (4.25)$$

pour tout k .

CHAPITRE 4. LA CONVERGENCE DE LA MÉTHODE DE
DESCENTE D'UNE FAMILLE À DEUX PARAMÈTRES DES
MÉTHODES DU GRADIENT CONJUGUÉ AVEC UNE FORMULE
FIXE DU PAS

Preuve. voir [59], lemma 1 ■

Corollaire 4.2.1 *Supposons que l'hypothèse (4.2.1) est satisfaite. On a*

$$1 - \frac{\delta\tau}{\kappa + 3\gamma(\kappa - \tau)} \leq \rho_k \leq 1 - \frac{\delta\kappa}{(1 + 3\gamma)\tau} \quad (4.26)$$

pour tout k .

Preuve. D'après (4.22) et (4.23) on a

$$1 - \rho_k = \delta \frac{(g_{k+1} - g_k)^T (x_{k+1} - x_k) \|d_k\|^2}{\|x_{k+1} - x_k\|^2 \left[(\bar{g}_{k+1} - g_k)^T d_k + \gamma\theta_k \right]}. \quad (4.27)$$

(4.16) et (4.17) impliquent que $\kappa + 3\gamma(\kappa - \tau) > 0$.

(4.24) et (4.25), il s'ensuit que

$$\frac{\delta\kappa}{(1 + 3\gamma)\tau} \leq 1 - \rho_k \leq \frac{\delta\tau}{\kappa + 3\gamma(\kappa - \tau)}. \quad (4.28)$$

Cette inégalité implique que (4.26) est vrai. ■

Lemme 4.2.3 *Supposons que l'hypothèse (4.2.1) est satisfaite et que x_k est donnée par (4.2), (4.3) et (4.14). alors*

$$\sum_{d_k \neq 0} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty. \quad (4.29)$$

Preuve. voir [59], lemma 3 ■

Lemme 4.2.4 *Supposons que l'hypothèse (4.2.1) est satisfaite, alors*

$$\sum_k \alpha_k^2 \|d_k\|^2 < \infty \quad (4.30)$$

Preuve. D'après (4.14), (4.25) et (4.29) on a

$$\begin{aligned} \sum_k \alpha_k^2 \|d_k\|^2 &= \sum_k \frac{(\delta g_k^T d_k)^2}{\left[(\bar{g}_{k+1} - g_k)^T d_k + \gamma\theta_k \right]^2} \|d_k\|^2 \\ &= \left[\frac{\delta}{\kappa + 3\gamma(\kappa - \tau)} \right]^2 \sum_{d_k \neq 0} \frac{(g_k^T d_k)^2}{\|d_k\|^2} \\ &< \infty \end{aligned} \quad (4.31)$$

■

CHAPITRE 4. LA CONVERGENCE DE LA MÉTHODE DE
DESCENTE D'UNE FAMILLE À DEUX PARAMÈTRES DES
MÉTHODES DU GRADIENT CONJUGUÉ AVEC UNE FORMULE
FIXE DU PAS

Lemme 4.2.5 *Pour $k \in \mathbb{N}$. Sous l'hypothèse (4.2.1), on a*

$$|g_{k+1}^T d_k| \leq -g_k^T d_k \left(1 + \frac{\tau \delta}{\kappa + 3\gamma(\kappa - \tau)} \right), \quad (4.32)$$

et

$$-g_{k+1}^T d_k \leq -g_k^T d_k \left(1 - \frac{\kappa \delta}{(1 + 3\gamma)\tau} \right) \quad (4.33)$$

Preuve. D'après (4.18), on a $|\phi_k| \leq \tau$, selon l'inégalité de Cauchy-Schwartz et (4.23), on a

$$g_{k+1}^T d_k = g_k^T d_k + y_k^T d_k = g_k^T d_k + \alpha_k \phi_k \|d_k\|^2. \quad (4.34)$$

Selon (4.25), (4.34), et $g_k^T d_k \leq 0$, on en déduit que

$$|g_{k+1}^T d_k| \leq -g_k^T d_k - \frac{\tau \delta g_k^T d_k}{\kappa + 3\gamma(\kappa - \tau)}.$$

D'après (4.19) on a $\phi_k \geq \kappa$. Selon (4.25), (4.34), et $g_k^T d_k \leq 0$, alors

$$g_{k+1}^T d_k \geq g_k^T d_k - \frac{\kappa \delta g_k^T d_k}{(1 + 3\gamma)\tau}.$$

Cela se termine notre preuve. ■

Lemme 4.2.6 *Supposons que l'hypothèse (4.2.1) est satisfaite. Les méthodes de CD, DY et FR satisfaites*

$$\alpha_k g_k^T d_k \leq -\|g_k\|^2, \quad \forall k. \quad (4.35)$$

Preuve. D'après (4.3) nous avons que

$$g_1^T d_1 = -\|g_k\|^2.$$

Pour la méthode de CD, si $k > 1$, d'après (4.3), (4.21) et (4.26), on a

$$\begin{aligned} g_k^T d_k &= -\|g_k\|^2 + \beta_k^{CD} g_k^T d_{k-1} \\ &= (-1 - \rho_{k-1}) \|g_k\|^2 \\ &< -\|g_k\|^2. \end{aligned}$$

CHAPITRE 4. LA CONVERGENCE DE LA MÉTHODE DE
DESCENTE D'UNE FAMILLE À DEUX PARAMÈTRES DES
MÉTHODES DU GRADIENT CONJUGUÉ AVEC UNE FORMULE
FIXE DU PAS

Alors (4.35) satisfaite pour la méthode de CD.

Pour la méthode de DY, D'après (4.3), (4.21) et (4.26), on a

$$\begin{aligned}
 g_k^T d_k &= -\|g_k\|^2 + \beta_k^{DY} g_k^T d_{k-1} \\
 &= \frac{g_{k-1}^T d_{k-1}}{(g_k - g_{k-1})^T d_{k-1}} \|g_k\|^2 \\
 &= -\frac{1}{1 - \rho_{k-1}} \|g_k\|^2 \\
 &< -\|g_k\|^2.
 \end{aligned}$$

La dernière inégalité découle $\rho_k \in (0, 1)$. Donc (4.35) vérifie pour DY.

Ensuite, nous considérons la méthode de FR. Supposons que (4.35) satisfaite pour $k - 1$, i.e.,

$$g_{k-1}^T d_{k-1} \leq -\|g_{k-1}\|^2. \quad (4.36)$$

D'après (4.3), (4.21) et (4.36), on a

$$\begin{aligned}
 g_k^T d_k &= -\|g_k\|^2 + \beta_k^{DY} g_k^T d_{k-1} \\
 &= \left(-1 + \rho_{k-1} \frac{g_{k-1}^T d_{k-1}}{\|g_{k-1}\|^2} \right) \|g_k\|^2 \\
 &= (-1 - \rho_{k-1}) \|g_k\|^2 \\
 &< -\|g_k\|^2.
 \end{aligned}$$

Donc (4.35) est satisfaite pour tout k . Le lemme est prouvé. ■

Théorème 4.2.1 *Supposons que l'hypothèse (4.2.1) est satisfaite. La méthode de CD engendre une suite $\{x_k\}$ tel que*

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0. \quad (4.37)$$

La même conclusion vaut pour la méthode de DY.

Preuve. Nous procédons par la contradiction, supposons que

$$\liminf_{k \rightarrow \infty} \|g_k\| \neq 0.$$

*CHAPITRE 4. LA CONVERGENCE DE LA MÉTHODE DE
DESCENTE D'UNE FAMILLE À DEUX PARAMÈTRES DES
MÉTHODES DU GRADIENT CONJUGUÉ AVEC UNE FORMULE
FIXE DU PAS*

Alors il existe une constante $\varepsilon > 0$ tel que

$$\|g_k\| > \varepsilon, \text{ pour tout } k \geq 1. \quad (4.38)$$

D'après (4.35), on a $-g_k^T d_k \geq \|g_k\|^2 > 0$. Ainsi,

$$-\frac{2}{g_k^T d_k} \leq \frac{2}{\|g_k\|^2} \leq \frac{2}{\varepsilon^2}. \quad (4.39)$$

Depuis $d_k + g_k = \beta_k^{CD} d_{k-1}$, nous avons que

$$\|d_k\|^2 = (\beta_k^{CD})^2 \|d_{k-1}\|^2 - 2g_k^T d_k - \|g_k\|^2.$$

En divisant les deux côtés de l'égalité ci-dessus par $(g_k^T d_k)^2$, en utilisant (4.35) et (4.39), nous obtenons

$$\begin{aligned} \frac{\|d_k\|^2}{(g_k^T d_k)^2} &= (\beta_k^{CD})^2 \frac{\|d_{k-1}\|^2}{(g_k^T d_k)^2} - \frac{2}{g_k^T d_k} - \frac{\|g_k\|^2}{(g_k^T d_k)^2} \\ &\leq \frac{\|d_{k-1}\|^2}{(g_{k-1}^T d_{k-1})^2} + \frac{2}{\varepsilon^2}. \end{aligned} \quad (4.40)$$

A l'aide de manière récursive (4.40) et notant que $d_1 = -g_1$, nous avons

$$\frac{\|d_k\|^2}{(g_k^T d_k)^2} < A(k-1) + 1,$$

c'est-à-dire

$$\frac{(g_k^T d_k)^2}{\|d_k\|^2} > \frac{1}{A(k-1) + 1}, \quad (4.41)$$

où $A = 2/\varepsilon^2$. Donc

$$\sum_{d_k \neq 0} (g_k^T d_k)^2 / \|d_k\|^2 = \infty.$$

Cela contredit le lemme (4.2.3). D'où, (4.37) satisfaite pour la méthode de CD.

Ensuite, nous considérons la méthode DY. Nous supposons que $\|g_k\| > \varepsilon$ pour tout k .

CHAPITRE 4. LA CONVERGENCE DE LA MÉTHODE DE
DESCENTE D'UNE FAMILLE À DEUX PARAMÈTRES DES
MÉTHODES DU GRADIENT CONJUGUÉ AVEC UNE FORMULE
FIXE DU PAS

D'après (4.3) et β_k^{DY} , nous obtenons

$$g_k^T d_k = \frac{\|g_k\|^2}{(g_k - g_{k-1})^T d_{k-1}} g_{k-1}^T d_{k-1}.$$

Ainsi $\beta_k^{DY} = g_k^T d_k / g_{k-1}^T d_{k-1}$. Cela donne

$$\begin{aligned} \frac{\|d_k\|^2}{(g_k^T d_k)^2} &= (\beta_k^{DY})^2 \frac{\|d_{k-1}\|^2}{(g_{k-1}^T d_{k-1})^2} - \frac{2}{g_k^T d_k} - \frac{\|g_k\|^2}{(g_k^T d_k)^2} \\ &\leq \frac{\|d_{k-1}\|^2}{(g_{k-1}^T d_{k-1})^2} + \frac{2}{\varepsilon^2}. \end{aligned} \quad (4.42)$$

Le reste de la preuve est la même que la preuve de la méthode CD. ■

Théorème 4.2.2 *Supposons que l'hypothèse (4.2.1) est satisfaite. La méthode de FR engendre une suite $\{x_k\}$ tel que*

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0. \quad (4.43)$$

Preuve. D'après (4.3) et β_k^{FR} , on a

$$\begin{aligned} -g_k^T d_k &= -g_k^T (-g_k + \beta_k^{FR} d_{k-1}) \\ &= \|g_k\|^2 + \frac{\|g_k\|^2}{\|g_{k-1}\|^2} \rho_{k-1} g_{k-1}^T d_{k-1} \\ &\leq \|g_k\|^2 + \bar{\rho} \frac{\|g_k\|^2}{\|g_{k-1}\|^2} (-g_{k-1}^T d_{k-1}), \end{aligned} \quad (4.44)$$

où $\bar{\rho} = 1 - \delta\kappa / [(1 + 3\gamma)\tau] > 0$ et la dernière inégalité découle (4.21) et $\rho_k \leq \bar{\rho}$. A l'aide de manière récursive (4.44) et notant que $d_1 = -g_1$, nous avons

$$-g_k^T d_k \leq \|g_k\|^2 \sum_{i=1}^{k-1} \rho^{-i} \leq \frac{1}{1 - \bar{\rho}} \|g_k\|^2. \quad (4.45)$$

Supposons que (4.43) n'est pas vrai.

Alors il existe un $\varepsilon > 0$ tel que $\|g_k\| > \varepsilon$ pour tout k . D'après (4.3), nous obtenons

$$\|d_k\|^2 = (\beta_k^{FR})^2 \|d_{k-1}\|^2 - 2g_k^T d_k - \|g_k\|^2.$$

CHAPITRE 4. LA CONVERGENCE DE LA MÉTHODE DE
DESCENTE D'UNE FAMILLE À DEUX PARAMÈTRES DES
MÉTHODES DU GRADIENT CONJUGUÉ AVEC UNE FORMULE
FIXE DU PAS

En divisant les deux côtés de l'égalité ci-dessus par $\|g_k\|^4$, d'après β_k^{FR} et (4.45), nous avons

$$\begin{aligned} \frac{\|d_k\|^2}{\|g_k\|^4} &= (\beta_k^{FR})^2 \frac{\|d_{k-1}\|^2}{\|g_k\|^4} - \frac{2g_k^T d_k}{\|g_k\|^4} - \frac{1}{\|g_k\|^2} \\ &\leq \frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^4} + \left(\frac{2}{1-\bar{\rho}} - 1 \right) \frac{1}{\|g_k\|^2} \\ &\leq \frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^4} + B, \end{aligned} \tag{4.46}$$

où $B = (1 + \bar{\rho}) / [(1 - \bar{\rho}) \varepsilon^2] > 0$. A l'aide de manière récursive (4.46) et notant que $d_1 = -g_1$, nous avons

$$\frac{\|d_k\|^2}{\|g_k\|^4} \leq B(k-1) + 1.$$

D'après (4.35), nous obtenons $(g_k^T d_k)^2 \geq \|g_k\|^4$. Ainsi,

$$\frac{(g_k^T d_k)^2}{\|d_k\|^2} \geq \frac{\|g_k\|^4}{\|d_k\|^2} \geq \frac{1}{B(k-1) + 1},$$

ce qui implique que

$$\sum_{d_k \neq 0} (g_k^T d_k)^2 / \|d_k\|^2 = \infty.$$

Il s'agit d'une contradiction à le lemme (4.2.3). Donc, (4.43) est vrai. ■

Théorème 4.2.3 *Supposons que l'hypothèse (4.2.1) est satisfaite. La méthode de PR engendre une suite $\{x_k\}$ tel que $\liminf_{k \rightarrow \infty} \|g_k\| = 0$.*

Preuve. Le théorème peut être prouvé par la même manière de la preuve du théorème 10 in Ref [55], avec un peu de modification. ■

4.3 La convergence globale de la famille à deux paramètres

Dans cette section, nous discutons des propriétés de convergence d'une famille à deux paramètres des méthodes de gradient conjugué, dans lequel β_k est donnée par (4.4) et (4.5).

Tout d'abord, nous donnons l'algorithme suivant

Etape 0 : On se donne $x_1 \in R^n$, poser $d_1 = -g_1, k = 1$.

Etape 1 : Si $\|g_k\| = 0$ alors stop. sinon aller à l'étape 2.

Etape 2 : Poser $x_{k+1} = x_k + \alpha_k d_k$ où d_k définie par (4.3), et α_k définie par (4.14).

Etape 3 : Calculer $\beta_{k+1}^{\mu_{k+1}, \omega_{k+1}}$ utiliser la formule (4.5).

Etape 4 : Poser $k := k + 1$, aller à l'étape 1.

Lemme 4.3.1 *Sous l'hypothèse (4.2.1), la méthode définie par (4.2), (4.3), (4.14) et (4.5) engendre une suite $\{x_k\}$ tel que $f(x_{k+1}) \leq f(x_k)$.*

Preuve. D'après le théorème de la valeur moyenne, on a

$$f(x_{k+1}) - f(x_k) = \bar{g}^T (x_{k+1} - x_k), \quad (4.47)$$

où $\bar{g} = \nabla f(\bar{x})$ pour certains $\bar{x} \in [x_k, x_{k+1}]$. D'après l'inégalité de Cauchy-Schwartz, (4.14), et l'hypothèse (4.2.1), nous obtenons

$$\begin{aligned} \bar{g}^T (x_{k+1} - x_k) &= g_k^T (x_{k+1} - x_k) + (\bar{g} - g_k)^T (x_{k+1} - x_k) \\ &\leq g_k^T (x_{k+1} - x_k) + \|\bar{g} - g_k\| \|x_{k+1} - x_k\| \\ &\leq g_k^T (x_{k+1} - x_k) + \tau \|x_{k+1} - x_k\|^2 \\ &= \alpha_k g_k^T d_k + \tau \alpha_k^2 \|d_k\|^2 \\ &= \alpha_k g_k^T d_k - \tau \alpha_k \delta g_k^T d_k \|d_k\|^2 / [(\bar{g}_{k+1} - g_k)^T d_k + \gamma \theta_k] \\ &= \alpha_k g_k^T d_k \left(1 - \frac{\tau \delta \|d_k\|^2}{(\bar{g}_{k+1} - g_k)^T d_k + \gamma \theta_k} \right). \end{aligned} \quad (4.48)$$

D'après (4.16), (4.17) et (4.25), nous obtenons

$$\begin{aligned} \alpha_k g_k^T d_k &= -\frac{\delta}{(\bar{g}_{k+1} - g_k)^T d_k + \gamma \theta_k} (g_k^T d_k)^2 \\ &\leq -\frac{\delta}{(1 + 3\gamma)\tau} \frac{(g_k^T d_k)^2}{\|d_k\|^2}, \end{aligned} \quad (4.49)$$

CHAPITRE 4. LA CONVERGENCE DE LA MÉTHODE DE
DESCENTE D'UNE FAMILLE À DEUX PARAMÈTRES DES
MÉTHODES DU GRADIENT CONJUGUÉ AVEC UNE FORMULE
FIXE DU PAS

et

$$1 - \frac{\tau\delta \|d_k\|^2}{(\bar{g}_{k+1} - g_k)^T d_k + \gamma\theta_k} \geq 1 - \frac{\delta\tau}{\kappa + 3\gamma(\kappa - \tau)} \quad (4.50)$$

$$> 0. \quad (4.51)$$

Il s'ensuit que

$$\Omega = \frac{\delta}{(1 + 3\gamma)\tau} \left(1 - \frac{\tau\delta}{\kappa + 3\gamma(\kappa - \tau)} \right) > 0.$$

Ainsi d'après (4.47) nous obtenons

$$f(x_{k+1}) - f(x_k) \leq -\Omega \frac{(g_k^T d_k)^2}{\|d_k\|^2} \leq 0, \quad (4.52)$$

ce qui implique $f(x_{k+1}) \leq f(x_k)$. Ceci termine la preuve. ■

Lemme 4.3.2 *Supposons que l'hypothèse (4.2.1) est satisfaite. On a*

$$D_k \geq (1 - \mu_k - \omega_k) \|g_{k-1}\|^2 - \left(\omega_k + \frac{\delta\kappa\mu_k}{(1 + 3\gamma)\tau} \right) d_{k-1}^T g_{k-1} \geq 0 \quad (4.53)$$

Preuve. Depuis $y_{k-1} = g_k - g_{k-1}$, et (4.21)

$$\begin{aligned} D_k &= (1 - \mu_k - \omega_k) \|g_{k-1}\|^2 + (\mu_k \rho_{k-1} - \mu_k - \omega_k) d_{k-1}^T g_{k-1} \\ &= (1 - \mu_k - \omega_k) \|g_{k-1}\|^2 - (\mu_k (1 - \rho_{k-1}) + \omega_k) d_{k-1}^T g_{k-1}. \end{aligned}$$

Selon (4.28) et $d_{k-1}^T g_{k-1} \leq 0$ la conclusion est immédiate. ■

Remarque 4.3.1 *Supposons que l'hypothèse (4.2.1) est satisfaite. Si $D_k = 0$, (4.53) implique*

$$(1 - \mu_k - \omega_k) \|g_{k-1}\|^2 - \left(\omega_k + \frac{\delta\kappa\mu_k}{(1 + 3\gamma)\tau} \right) d_{k-1}^T g_{k-1} = 0.$$

Puisque le côté gauche est la somme des deux termes non négatifs, nous obtenons

$$d_k = \begin{cases} (1 - \mu_k - \omega_k) \|g_{k-1}\|^2 = 0 & (a), \\ \left(\omega_k + \frac{\delta\kappa\mu_k}{(1 + 3\gamma)\tau} \right) d_{k-1}^T g_{k-1} = 0 & (b), \end{cases}$$

CHAPITRE 4. LA CONVERGENCE DE LA MÉTHODE DE
DESCENTE D'UNE FAMILLE À DEUX PARAMÈTRES DES
MÉTHODES DU GRADIENT CONJUGUÉ AVEC UNE FORMULE
FIXE DU PAS

1 : Si $\mu_k + \omega_k < 1$, (a) se résume à $\|g_{k-1}\|^2 = 0$, ce qui signifie que la convergence est atteinte à l'itération $k - 1$.

2 : Si $\mu_k = 0$, donc le numérateur de (4.5) est nul aussi. Dans ce cas, on a

$$\begin{cases} (1 - \omega_k) \|g_{k-1}\|^2 = 0, \\ \omega_k d_{k-1}^T g_{k-1} = 0. \end{cases}$$

Si $\omega_k < 1$, la conclusion est la même que dans le 1 cas .

Lemme 4.3.3 Sous l'hypothèse (4.2.1) on a

$$\liminf_{k \rightarrow \infty} \|g_k\| > 0 \Rightarrow \lim_{k \rightarrow \infty} \beta_k^{\mu_k, \omega_k} = 0,$$

où $\beta_k^{\mu_k, \omega_k}$ est définie par (4.5).

Preuve. D'après (4.29), nous concluons que

$$\lim_{k \rightarrow \infty} \|x_{k+1} - x_k\|^2 = \lim_{k \rightarrow \infty} \|\alpha_k d_k\|^2 = 0.$$

Parce que f est continûment différentiable, $\|g_k\|$ est bornée, d'après l'hypothèse (4.2.1) et la bornitude de L , nous avons aussi $\lim_{k \rightarrow \infty} y_{k-1} = 0$ et

$$\lim_{k \rightarrow \infty} g_k^T y_{k-1} = 0. \quad (4.54)$$

Si $\liminf_{k \rightarrow \infty} \|g_k\| > 0$, il existe une constante positive $\psi > 0$ telle que

$$\|g_k\| > \psi \quad \text{for all } k. \quad (4.55)$$

D'après (4.5), nous obtenons

$$|g_k^T y_{k-1}| = |\beta_k^{\mu_k, \omega_k}| |D_k|. \quad (4.56)$$

Considérons l' indice d'itération k tel que $\mu_k + \omega_k \in [0, 1/2]$. On utilise (4.53), (4.56) et $d_{k-1}^T g_{k-1} \leq 0$, on obtient

$$|g_k^T y_{k-1}| \geq |\beta_k^{\mu_k, \omega_k}| (1 - \mu - \omega) \|g_{k-1}\|^2,$$

ce qui conduit à

$$|g_k^T y_{k-1}| \geq |\beta_k^{\mu_k, \omega_k}| \psi^2 / 2, \quad (4.57)$$

d'où (4.55).

CHAPITRE 4. LA CONVERGENCE DE LA MÉTHODE DE
DESCENTE D'UNE FAMILLE À DEUX PARAMÈTRES DES
MÉTHODES DU GRADIENT CONJUGUÉ AVEC UNE FORMULE
FIXE DU PAS

Nous créer un résultat semblable dans le cas plus complexe $\mu_k + \omega_k \in (1/2, 1[$. Dans un premier temps, nous devons montrer que

$$g_k^T d_k \leq -\psi^2/2, \quad (4.58)$$

pour toutes les valeurs de k suffisamment grandes. D'après la Remark (4.3.1), dans le cas $g_{k-1}^T d_{k-1} = 0$, nous avons $\beta_k^{\mu_k, \omega_k} = 0$, alors $d_k = -g_k$ et (4.58) est valide selon la (4.55).

Maintenant, considérons le cas où $g_{k-1}^T d_{k-1} < 0$. D'après (4.3), (4.5) et (4.21) on obtient

$$\begin{aligned} g_k^T d_k &= g_k^T (-g_k + \beta_k^{\mu_k, \omega_k} d_{k-1}) \\ &= -\|g_k\|^2 + \frac{g_k^T y_{k-1}}{(1 - \mu_k - \omega_k) \|g_{k-1}\|^2 + \mu_k d_{k-1}^T y_{k-1} - \omega_k d_{k-1}^T g_{k-1}} g_k^T d_{k-1} \\ &= -\|g_k\|^2 + \frac{\rho_{k-1} g_k^T y_{k-1}}{(1 - \mu_k - \omega_k) \|g_{k-1}\|^2 + (\mu_k \rho_{k-1} - \mu_k - \omega_k) d_{k-1}^T g_{k-1}} g_{k-1}^T d_{k-1} \\ &\leq -\psi^2 + \frac{|\rho_{k-1} g_k^T y_{k-1}|}{|\mu_k \rho_{k-1} - \mu_k - \omega_k|}, \end{aligned} \quad (4.59)$$

pour k est suffisamment grand, donc (4.58) est vrai.

D'après (4.5), (4.58) et (4.26) on obtient

$$\begin{aligned} |g_k^T y_{k-1}| &\geq |\beta_k^{\mu_k, \omega_k}| \left| (1 - \mu_k - \omega_k) \psi^2 - (\mu_k (1 - \rho_{k-1}) + \omega_k) \psi^2/2 \right| \\ &\geq |\beta_k^{\mu_k, \omega_k}| \left\{ (1 - \mu_k - \omega_k) \psi^2 + \left[\frac{\mu_k \delta \kappa}{(1 + 3\gamma) \tau} + \omega_k \right] \psi^2/2 \right\} \\ &= |\beta_k^{\mu_k, \omega_k}| \left[1 - \omega_k/2 - \left(1 - \frac{\delta \kappa}{2(1 + 3\gamma) \tau} \right) \mu_k \right] \psi^2, \end{aligned} \quad (4.60)$$

pour toutes les valeurs de k suffisamment grandes. Si $\mu_k + \omega_k \in (1/2, 1[$, donc

$$|g_k^T y_{k-1}| \geq |\beta_k^{\mu_k, \omega_k}| S \psi^2. \quad (4.61)$$

Pour $S \leq 1/2$: (4.61) est impliquée par (4.57), pour que (4.61) satisfaite dans l'ensemble du domaine $\mu_k \in [0, 1[$, $\omega_k \in [0, 1 - \mu_k[$.

Finalement, (4.54) et (4.61) implique $\lim_{k \rightarrow \infty} |\beta_k^{\mu_k, \omega_k}| = 0$.

D'autre part, considérer le cas où l'hypothèse (4.2.1) n'est pas nécessairement valide.

CHAPITRE 4. LA CONVERGENCE DE LA MÉTHODE DE
DESCENTE D'UNE FAMILLE À DEUX PARAMÈTRES DES
MÉTHODES DU GRADIENT CONJUGUÉ AVEC UNE FORMULE
FIXE DU PAS

Si $\mu_k = 0$, ensuite, nous avons

$$|g_k^T y_{k-1}| \geq |\beta_k^{0, \omega_k}| \psi^2 / 2.$$

La preuve est semblable à celle de (4.61), où sont les deux cas à examiner $\omega_k \in [0, 1/2]$ et $\omega_k \in (1/2, 1[$.

Finalement, d'après (4.54) on obtient $\lim_{k \rightarrow \infty} |\beta_k^{0, \omega_k}| = 0$. ■

Théorème 4.3.1 *Sous l'hypothèse (4.2.1), la méthode définie par (4.2), (4.3), (4.14) et (4.5) engendre une suite $\{x_k\}$ tel que*

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0.$$

Preuve. Supposons au contraire que $\|g_k\| \geq \psi$ pour tout k .

Puisque $\liminf_{k \rightarrow \infty} \|g_k\| \neq 0$, d'après le lemme (4.3.3) nous avons $\beta_k^{\mu_k, \omega_k} \rightarrow 0$, comme $k \rightarrow 0$. Puisque L est bornée, les deux $\{x_k\}$ et $\{g_k\}$ sont bornées. En utilisant

$$\|d_k\| \leq \|g_k\| + |\beta_k^{\mu_k, \omega_k}| \|d_{k-1}\|, \quad (4.62)$$

nous concluons que le $\{\|d_k\|\}$ est uniformément bornée donc

$$\begin{aligned} |g_k^T d_k| &= |g_k^T (-g_k + \beta_k^{\mu_k, \omega_k} d_{k-1})| \\ &\geq \|g_k\|^2 - |\beta_k^{\mu_k, \omega_k}| \|g_k\| \|d_{k-1}\|. \end{aligned} \quad (4.63)$$

D'après (4.5), (4.21) et l'inégalité de Cauchy-Schwarz, nous avons

$$\begin{aligned} |\beta_k^{\mu_k, \omega_k}| &= \left| \frac{g_k^T (g_k - g_{k-1})}{(1 - \mu_k - \omega_k) \|g_{k-1}\|^2 + \mu_k d_{k-1}^T y_{k-1} - \omega_k d_{k-1}^T g_{k-1}} \right| \\ &\leq \frac{\|g_k\| \|g_k - g_{k-1}\|}{|(1 - \mu_k - \omega_k) \|g_{k-1}\|^2 - [(1 - \rho_{k-1})\mu_k + \omega_k] d_{k-1}^T g_{k-1}|}. \end{aligned} \quad (4.64)$$

D'après (4.18) et (4.31) on obtient

$$\begin{aligned} \|g_k - g_{k-1}\| &\leq \tau \alpha_{k-1} \|d_{k-1}\| \\ &\leq \frac{\tau \delta}{\kappa + 3\gamma(\kappa - \tau)} \frac{g_{k-1}^T d_{k-1}}{\|d_{k-1}\|}. \end{aligned} \quad (4.65)$$

CHAPITRE 4. LA CONVERGENCE DE LA MÉTHODE DE
DESCENTE D'UNE FAMILLE À DEUX PARAMÈTRES DES
MÉTHODES DU GRADIENT CONJUGUÉ AVEC UNE FORMULE
FIXE DU PAS

D'après (4.16) et (4.17) on en déduit que

$$\|g_k - g_{k-1}\| \leq \frac{g_{k-1}^T d_{k-1}}{\|d_{k-1}\|}, \quad (4.66)$$

par (4.6) nous avons

$$\begin{aligned} |(1 - \mu_k - \omega_k) \|g_{k-1}\|^2 - [(1 - \rho_{k-1})\mu_k + \omega_k] d_{k-1}^T g_{k-1}| &\geq \frac{|-1 + \mu_k + \omega_k|}{c} |g_{k-1}^T d_{k-1}| \\ &= m |g_{k-1}^T d_{k-1}|, \quad (m > 1) \end{aligned} \quad (4.67)$$

par (4.65), (4.66), et (4.67) on obtient

$$|\beta_k^{\mu_k, \omega_k}| \|d_{k-1}\| \leq \frac{\|g_k\|}{m}. \quad (4.68)$$

Par conséquent, en substituant (4.68) dans (4.63), nous avons

$$g_k^T d_k \geq A \|g_k\|^2, \quad A = \frac{m-1}{m}, \quad (4.69)$$

pour k est suffisamment grand. Ainsi, nous avons

$$\frac{(g_k^T d_k)^2}{\|d_k\|^2 \|g_k\|^2} \geq A^2 \frac{\|g_k\|^2}{\|d_k\|^2}. \quad (4.70)$$

Puisque $\|g_k\| \geq \psi$ et $\|d_k\|$ est bornée, alors il existe $\varepsilon > 0$ pour que

$$\frac{(g_k^T d_k)^2}{\|d_k\|^2 \|g_k\|^2} \geq \varepsilon,$$

ce qui implique

$$\sum_{d_k \neq 0} \frac{(g_k^T d_k)^2}{\|d_k\|^2} = \infty. \quad (4.71)$$

Ceci contredit le lemme (4.2.3). ■

4.4 Tests numériques et discussions

Dans cette partie, nous présentons les expériences numériques de la nouvelle formule (4.14) avec l'expression (4.5), ordinateur (Processeur : Intel(R)core(TM)i3-3110M cpu@2.40GHZ, Ram 4.00 GB) grâce au programme Matlab.

20 tests problèmes sont tirées [1]. Cela nous conduit à tester les propriétés globales de convergence de notre méthode. Critères d'arrêt est réglé par $\|g_k\| \leq \varepsilon$ où $\varepsilon = 10^{-6}$. Prenant en considération les paramètres suivants : $\delta = 0.75$, $\gamma = 0.01$, $\mu_k = \mu = 0.5$ et $\omega_k = \omega = 0.4$.

Tableau 1 liste des résultats numériques. La signification de chaque colonne est la suivante :

“Problème” le nom du problème de test.

“N” la dimension du problème de test.

“k” le nombre d'itérations.

“ $\|g_k\|$ ” la norme du gradient.

*CHAPITRE 4. LA CONVERGENCE DE LA MÉTHODE DE
DESCENTE D'UNE FAMILLE À DEUX PARAMÈTRES DES
MÉTHODES DU GRADIENT CONJUGUÉ AVEC UNE FORMULE
FIXE DU PAS*

Tableau 1

	Problème	N	k	$\ g_k\ $
1	Raydan 1	10	19	1.120790293916461e-07
		50	51	1.389923874174295e-07
2	Diagonal 5	500	342	1.026183259422179e-07
		1000	631	1.003760103090495e-07
3	Perturbed quadratic diagonal	100	23	2.170295100224989e-07
		500	29	1.155266677665186e-07
4	Extended quadratic penalty 1	100	18	1.884874483160104e-07
		500	19	1.016400612612853e-07
5	QUARTC	100	15	1.643581275426157e-07
		500	16	1.550430128841257e-07
6	White and Holst	100	24	2.070125493427497e-07
		500	27	1.810192036538758e-07
7	Beale	100	145	1.244578200376349e-07
		500	221	1.067324585387475e-07
8	Diagonal 7	100	15	1.279002965759446e-07
		500	16	1.099362564544676e-07
9	Extended Maratos	100	173	1.594462458701618e-07
		500	286	1.024825161025436e-07
10	Diagonal 8	50	20	3.247344971185945e-07
		200	21	1.653875889473304e-07
11	Freudenstein and Roth	100	21	2.457891545157802e-07
		500	24	1.623516986314701e-07
12	Rosenbrock	100	19	1.002347534497171e-07
		500	21	1.001365348958891e-07

*CHAPITRE 4. LA CONVERGENCE DE LA MÉTHODE DE
DESCENTE D'UNE FAMILLE À DEUX PARAMÈTRES DES
MÉTHODES DU GRADIENT CONJUGUÉ AVEC UNE FORMULE
FIXE DU PAS*

13	Penalty	100	71	1.532487962140 017e-07
		500	89	1.360078512369044e-07
14	Cliff	100	34	2.0056653220821 97e-07
		500	39	1.897012587633221e-07
15	LIARWHD	100	248	1.185263304998455e-07
		500	409	1.058770638716538e-07
16	Almost Perturbed Quadratic	2	21	1.148131830118244e-07
		500	29	1.017343210015877e-07
17	Diagonal 4	2	30	1.319452728458 628e-07
		500	31	1.001055919601577e-07
18	Staircase 1	2	21	2.292660111210135e-07
		500	25	1.025016344917332e-07
19	Power	2	21	2.251830705880973e-07
		500	27	1.095143872286877e-07
20	Full Hessian 2	2	27	1.516469135063661e-07
		500	29	1.100214389422758e-07

Les résultats du tableau 1 indiquent que l'expression (4.14) fournit le temps et la mémoire mieux que l'expression (4.10) qui utilise les matrices $\{Q_k\}$ qui peuvent devenir un fardeau à certains moments pour montrer la convergence.

La valeur de δ est trop grande dans test, la méthode peut parfois engendrer un x_{k+1} tel que $f(x_k) < f(x_{k+1})$. En théorie, cela ne peut pas se produire si δ est suffisamment petit, (4.3.1). Plus la valeur de δ est petit, ce fait permet la convergence se produire à la fin.

Si la valeur γ est petit, le nombre d'itérations diminue et $\|g_k\|$ converge rapidement parce qu'elle est liée aux paramètres τ et κ , alors que les applications des deux valeurs μ et ω affecte plus la valeur $\|g_k\|$ par rapport aux nombres d'itérations.

Il y a un certain nombre de 20 à grande échelle les problèmes de test d'optimisation sans contrainte généralisée ou étendue de collection de [1]. Pour chaque fonction test, nous avons pris six expériences numériques avec le nombre de variables augmentant comme $n = 1000, 2000, 4000, 6000, 8000, 10000$.

Nous adoptons les profils de performance par Delan and Moré [42] pour comparer les performances entre les algorithmes suivants

T-PF1 : famille à deux paramètres des méthodes de gradient conjugué avec la formule (4.10).

T-PF2 : famille à deux paramètres des méthodes de gradient conjugué

*CHAPITRE 4. LA CONVERGENCE DE LA MÉTHODE DE
DESCENTE D'UNE FAMILLE À DEUX PARAMÈTRES DES
MÉTHODES DU GRADIENT CONJUGUÉ AVEC UNE FORMULE
FIXE DU PAS*

avec la formule (4.14).

Figure 1,2,3 et 4 donnent des profils de performances pour le nombre d'itérations, le temps de CPU, les évaluations de fonction, et les évaluations de gradient respectivement.

A partir des quatre Figures, nous pouvons voir que toutes les méthodes sont efficaces. La nouvelle méthode T-PF2 fonctionne mieux que la méthode T-PF1, pour les problèmes des tests donnés. Ces résultats préliminaires obtenus sont en effet encourageants.

Dans le trace de profil de performance, la courbe supérieure correspond à la méthode qui a résolu le plus de problèmes dans le temps qui se trouve à un facteur t du meilleur temps. Le pourcentage des problèmes de test pour lesquels une méthode est signalée comme la plus rapide est donnée sur l'axe gauche de la parcelle. Le côté droit de la parcelle donne le pourcentage des problèmes de test qui ont été résolus avec succès par chacune des méthodes. En substance, le côté droit est une mesure de la robustesse de l'algorithme.

Dans les figures 1 et 2 nous utilisons les itérations et le temps CPU pour comparer les performances des codes de gradient conjugué T-PF1 et T-PF2. Les figures 1 et 2 montrent que, T-PF2 est la plus rapide (environ 83%) et (plus rapide pour 68%) par rapport T-PF1, pour des valeurs de $t \geq 6$, 5 et $t \geq 3$ respectivement

Dans la figure 3, nous comparons la performance en fonction du nombre d'évaluations de la fonction, étant donné que la courbe supérieure de la figure 3 correspond à T-PF2, cet algorithme est clairement le plus rapide pour cet ensemble de 20 problèmes de test. Aussi, il est intéressant d'observer sur la figure 3 que le code T-PF2 sont les plus performants, par rapport au nombre d'évaluations de fonction, pour des valeurs de $t \geq 6$.

Dans la figure 4, nous comparons la performance en fonction du nombre d'évaluations de gradient. T-PF2 est la plus rapide (plus rapide pour 72%), alors T-PF1. Par conséquent, le code T-PF2 sont les plus performants, par rapport au nombre d'évaluations de gradient, pour des valeurs de $t \geq 4.5$.

En conclusion, la fig. 1-4 suggèrent que notre méthode proposée T-PF2 présente la meilleure performance car il représente la plus grande probabilité d'être le solveur optimal, suivi par la méthode du gradient conjugué T-PF1 par rapport à tous les indicateurs de performance.

CHAPITRE 4. LA CONVERGENCE DE LA MÉTHODE DE DESCENTE D'UNE FAMILLE À DEUX PARAMÈTRES DES MÉTHODES DU GRADIENT CONJUGUÉ AVEC UNE FORMULE FIXE DU PAS

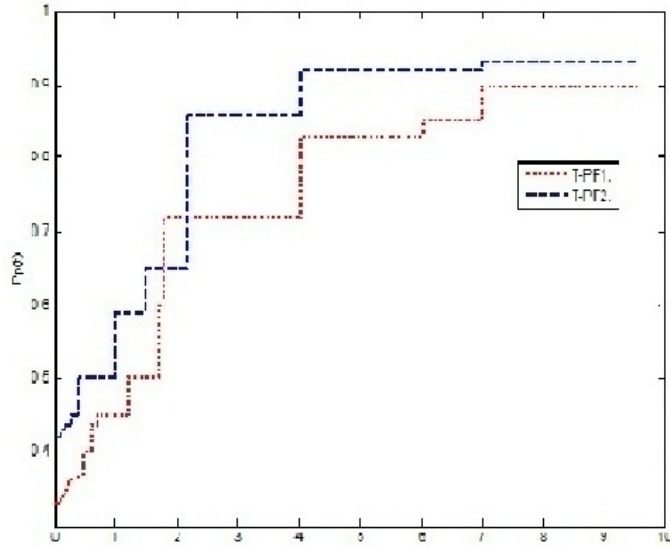


Figure 1. Performance basée sur le nombre d'itérations

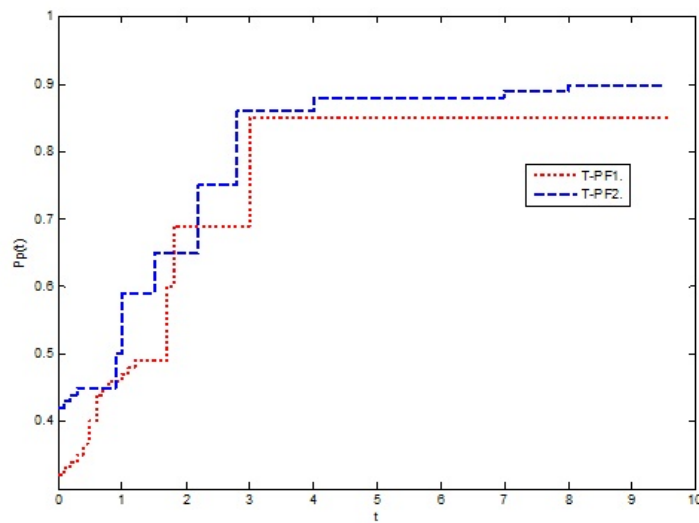


Figure 2. Performance basée sur le temps CPU

CHAPITRE 4. LA CONVERGENCE DE LA MÉTHODE DE
DESCENTE D'UNE FAMILLE À DEUX PARAMÈTRES DES
MÉTHODES DU GRADIENT CONJUGUÉ AVEC UNE FORMULE
FIXE DU PAS

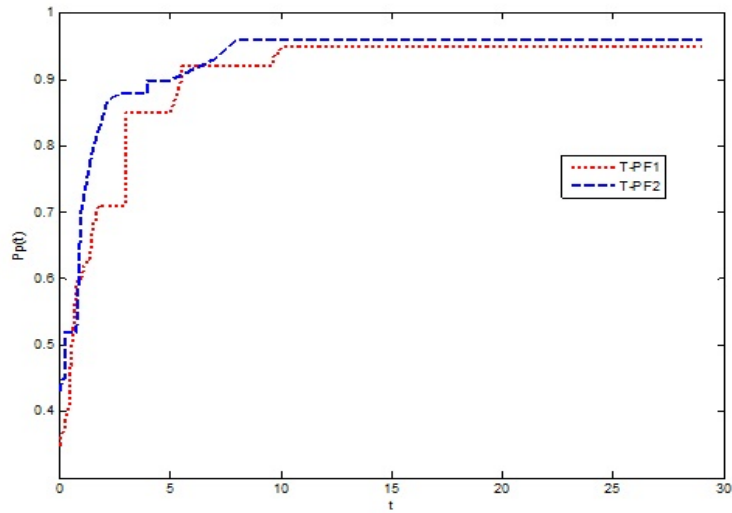


Figure 3. Performance basée sur d'évaluations de fonction

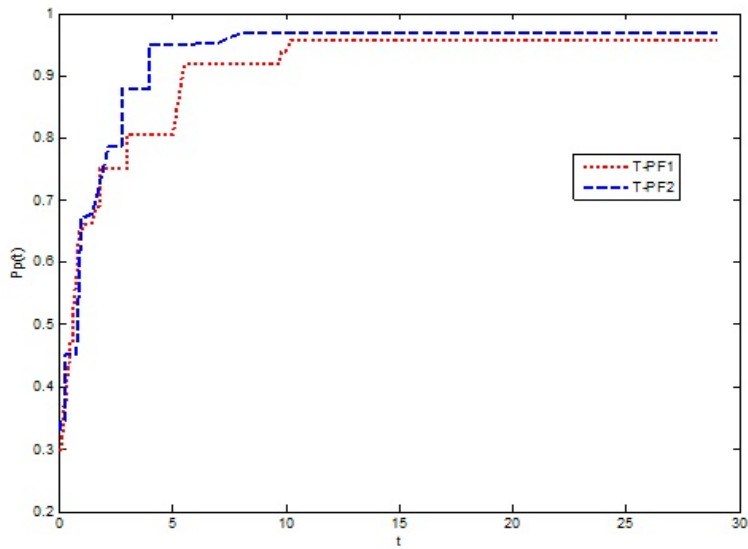


Figure 4. Performances basée sur d'évaluations de gradient

Bibliographie

- [1] N. Andrei. An unconstrained optimization test functions collection. *Adv. Model. Optim.*, 10(1) :147–161, 2008.
- [2] M. Al-Baali, Descent property and global convergence of the Fletcher-Reeves method with inexact line search, *IMA J. Numer. Anal.* 5 (1985), pp. 121-124.
- [3] M.S. Bazaraa, H.D. Sherali, et C.M. Shett, *Nonlinear Programming, Theory and Algorithms*, Wiley-Interscience,(1993).
- [4] A. Buckley, Extending the relationship between the conjugate gradient and BFGS algorithms. *Math.Program.* 15(1978), 343–348.
- [5] K. Bouaziz and Y. Laskri, Convergence of a Two-parameter Family of Conjugate Gradient Methods with a Fixed Formula of Stepsize, *Boletim da Sociedade Paranaense de Matemática*, doi :10.5269/bspm.v38i6.35641
- [6] C. G. Broyden, The convergence of a class of double-rank minimization algorithms 1. general considerations, *J. Inst. Math. Appl.*, 6 (1970), pp. 76–90.
- [7] A. Cauchy, *Analyse mathématique, Méthode générale pour la résolution des systèmes d'équations simultanées*, *Comptes Rendus de l'Académie des Sciences de Paris*, (1847) t-25, pp. 536-538.
- [8] X. Chen, J. Sun, Global convergence of a two-parametre family of conjugate gradient methods without line search, *Journal of Computational and Applied Mathematics* 146 (2002) 37-45.
- [9] Y. H. Dai, Analysis of nonlinear conjugate gradient method, Ph.D. thesis, Institute of Computational Mathematics and Scientific/ Engineering Computing, Chinese Academy of Sciences, 1997.
- [10] Y. H. Dai and Y. Yuan, A class of globally convergent conjugate gradient methods, *Sci. China Ser. A*, 46 (2003), pp. 251–261.

-
- [11] Y. H. Dai and Y. Yuan, A class of globally convergent conjugate gradient methods, Research report ICM-98-030, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, 1998. (submitted to IMA J. Numer. Anal.)
- [12] Y. H. Dai and Y. Yuan, A nonlinear conjugate gradient method with nice global convergence properties, Research report ICM-95-038, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, 1995 (accepted by SIAM J. Optimization).
- [13] Y.H. Dai, Y. Yuan, A nonlinear conjugate gradient method with a strong global convergence property, SIAM J. Optim. 10, No 1 (1999) 177–182.
- [14] Y.H. Dai, Convergence of Polak–Ribière–Polyak conjugate gradient method with constant stepsizes, Manuscript, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, 2001.
- [15] Y.H. Dai, Convergence properties of the BFGS algorithm, SIAM J. Optim. 13, No. 3 (2002), pp. 693–701.
- [16] Y.H. Dai, Y. Yuan, Convergence of the FR method under a generalized Wolfe line search, Numer. Math. J. Chinese Univ. 6 (1996) 552–562.
- [17] Y. Dai, J. Han, G. Liu, D. Sun, Y.X. Yuan, and H. Yin, Convergence properties of nonlinear conjugate gradient methods, Research report ICM-98-024, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, 1998 (accepted by SIAM J. Optimization).
- [18] Y. H. Dai and Y. Yuan, Convergence properties of the Fletcher-Reeves method, IMA J. Numer. Anal. Vol. 16 No. 2 (1996), 155–164.
- [19] Y. H. Dai and Y. Yuan, Convergence properties of the conjugate descent method, Advances in Mathematics, 6 (1996), pp. 552–562.
- [20] Y. H. Dai and Y. Yuan, Convergence properties of Beale-Powell restart method, Sciences in China (series A), Vol. 28, No. 5, pp. 424–432.
- [21] Y.H. Dai, J.Y. Han, G.H. Liu, D.F. Sun, H.X. Yin and Y.X. Yuan, Convergence properties of nonlinear conjugate gradient methods, SIAM J. Optim., 10 (1999) 345–358.
- [22] Y. H. Dai and Y. Yuan, Extension of a class of conjugate gradient methods, Research report ICM-98-049, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, 1998, (submitted to BIT).

- [23] Y. H. Dai and L. Z. Liao, New conjugacy conditions and related nonlinear conjugate gradient methods, *Appl. Math. Optim.*, 43(2001), pp. 87–101.
- [24] Y. Dai and Y. Yuan. A three-parameter family of nonlinear conjugate gradient methods. *Mathematics of Computation*, 70(235) :1155–1167, 2001.
- [25] Y. H. Dai and Y. Yuan, Some properties of a new conjugate gradient method, in : *Advances in Nonlinear Programming*, ed. Kluwer, Boston, (1998), pp. 251-262.
- [26] Y. H. Dai, Some new properties of a nonlinear conjugate gradient method, Research report ICM-98-010, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, 1998.
- [27] J. W. Daniel, The conjugate gradient method for linear and nonlinear operator equations, *SIAM J. Numer. Anal.*, 4 (1967), 10-26.
- [28] T. Diallo, Travail pratique de master, Sujet : Etude et illustration de méthodes itératives d’optimisation non linéaire, Responsable : Benjamin Leroy-Beaubier, Ecole polytechnique Fédérale de Lausanne.(2006)
- [29] R. Fletcher, *Practical Methods of Optimization vol. 1- Unconstrained optimization*, John Wiley & Sons (New York), 1987.
- [30] R. Fletcher and C. Reeves, Function minimization by conjugate gradients, *Comput. J.* 7 (1964), pp. 149-154.
- [31] J. C. Gilbert, *Eléments d’optimisation différentiable : Théorie et algorithmes*, notes de cours, École Nationale Supérieure de Techniques Avancées, Paris (2007).
- [32] J. C. Gilbert and J. Nocedal, Global convergence properties of conjugate gradient methods for optimization, *SIAM. J. Optimization*. Vol. 2 No. 1 (1992), pp. 21-42.
- [33] L. Grippo and S. Lucidi, A globally convergent version of the Polak-Ribière conjugate gradient method, *Math. Prog.* 78 (1997), pp. 375-391.
- [34] M. R. Hestenes and E. L. Stiefel, Methods of conjugate gradients for solving linear systems, *J. Res. Nat. Bur. Standards Sect. 5*, 49 (1952), 409-436.
- [35] Y. F. Hu and C. Storey, Global convergence result for conjugate gradient methods, *JOTA* : Vol. 71 No. 2 (1991) 399-405.

-
- [36] D. Kauth, Optimisation numerique méthodes du gradient conjugue linéaire (Chapitre 5.1), Université de Fribourg, le 5 novembre (2009)
- [37] K. M. Khoda, Y. Liu, and C. Storey, Generalized Polak-RiWe Algorithm, *Journal of Optimization Theory and Applications*, Vol. 75, No. 2 (1992), 345-354.
- [38] Y. Liu and C. Storey, Efficient generalized conjugate gradient algorithms, Part 1 - Theory, *Journal of Optimization Theory and Applications*, Vol. 69 (1991), 129-137.
- [39] G. Liu, J. Han and H. Yin, Global convergence of the Fletcher-Reeves algorithm with inexact line search, *Appl. Math. J. Chinese Univ. Ser. B.* 10 (1995) 75-82.
- [40] M. Minoux , *Programmation mathématique, théorie et algorithmes*, tome 1 , Dunod, (1983)
- [41] J.J. Moré, B.S. Garbow and K.E. Hillstrom, Testing unconstrained optimization software, *ACM Transactions on Mathematical Software* 7 (1981) 17–41.
- [42] J. More and E. D. Dolan, Benchmarking optimization software with performance files, *Math. Program*, 91(2), 2001-2013, (2002).
- [43] S. Mottelet, *Optimisation non-linéaire*, Université de Technologie de Compiègne, RO04/TI07 Printemps (2003).
- [44] L. Nazareth, Conjugate-gradient methods, to appear in : *Encyclopedia of Optimization* (C.Floudas and P. Pardalos, eds.), Kluwer Academic Publishers, Boston, USA and Dordrecht, The Netherlands (1999).
- [45] J. Nocedal, S.J. Wright. *Numerical optimization*. Springer Series in Operations Research and Financial Engineering. Springer, New York, second edition, 2006.
- [46] B. Polak and G. Ribière, Note sur la convergence de directions conjuguées, *Rev. ev. Francaise Informat Recherche Operationelle*, 3e Année 16 (1969), pp. 35-43.
- [47] B.T.Polyak, The conjugate gradient method in extreme problems, *U.S.S.R. Comput. Math. and Math. Phys.* 9 (1969) 94–112.
- [48] M.J.D. Powell, A new algorithm for unconstrained optimization. In : Rosen, J. B. , Mangasarian, O. L. , Ritter, K. (eds.) *Nonlinear Programming*, pp. 31-66. Academic Press, New York (1970).

-
- [49] M.J.D. Powell, Convergence properties of algorithms for nonlinear optimization, *SIAM rev*, 28 (1986), pp. 487-500.
- [50] M.J.D. Powell, Nonconvex minimization calculations and the conjugate gradient method, in : *Lecture Notes in Mathematics* vol. 1066, Springer-Verlag (Berlin) (1984), pp. 121-141.
- [51] M.J.D. Powell, Some convergence properties of the conjugate gradient method, *Math. Prog*, 11 (1976), pp. 42–49.
- [52] M. Raydan, The Barzilai and Borwein gradient method for the large scale unconstrained minimization problem, *SIAM J, Optim* 7, 26–33 (1997).
- [53] B. Sellami, Y. Laskri, and R. Benzine. A new two-parameter family of nonlinear conjugate gradient methods. *Optimization*, (ahead-of-print) :1–17, 2013.
- [54] D.F.Shanno, Conjugate gradient methods with inexact searches. *Math, Oper, Res* 3(1978), 244–256
- [55] J. Sun, J. Zhang, Global convergence of conjugate gradient methods without line search, *Ann. Oper. Res.* 103 (2001) 161–173.
- [56] D. Touati-Ahmed and C. Storey , Efficient hybrid conjugate gradient techniques, *JOTA*, 64 (1990), pp. 379-397.
- [57] P. Wolfe, Convergence conditions for ascent methods, *SIAM Review*, 11 (1969) 226-235.
- [58] P. Wolfe, Convergence conditions for ascent methods. ii : Some corrections. *SIAM review*, 13(2) :185–188, 1971.
- [59] Q.-j. Wu, A nonlinear conjugate gradient method without line search and its global convergence. In *Computational and Information Sciences (IC-CIS)*, 2011 International Conference on, pages 1148–1152. IEEE, 2011.
- [60] H. Yabe and M. Takano, Global convergence properties of nonlinear conjugate gradient methods with modified secant condition, *Comput. Optim. Appl.*, 28, (2004), pp. 203–225.
- [61] Y. Yuan, Analysis on the conjugate gradient method, *Optim. Methods Softw*, 2 (1993), pp.19–29.
- [62] Y. Yuan, *Numerical Methods for Nonlinear Programming*, Shanghai Scientific & Technical Publishers, 1993.

- [63] J. Z. Zhang, N. Y. Deng, and L. H. Chen, New quasi-Newton equation and related methods for unconstrained optimization, *J. Optim. Theory Appl.*, 102 (1999), pp. 147–167.
- [64] J. Z. Zhang and C. X. Xu, Properties and numerical performance of quasi-Newton methods with modified quasi-Newton equations, *J. Comput. Appl. Math.*, 137 (2001), pp. 269–278.
- [65] G. Zoutendijk, Nonlinear programming, computational methods, in : *Integer and Nonlinear Programming*, J. Abadie, ed, North-Holland (1970) 37-86.