Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Badji Mokhtar
Annaba

Badji Mokhtar University
Annaba

Faculté des Sciences
Département de Mathématiques

THESE

Présentée en vue de l'obtention du diplôme de
Doctorat en Mathématiques
Option : Analyse Fonctionnelle et Optimisation

Gradient conjugué, étude unifiée

Par :
BADREDDINE Sellami

Sous la direction de
Pr. YAMINA Laskri

Devant le jury

PRESIDENT :	Benzine RACHID	Pr.	Université Badji Mokhtar d'Annaba.
EXAMINATRICE :	Rebbani FAOUZIA	Pr.	Université Badji Mokhtar d'Annaba.
EXAMINATEUR :	Benterki DJAMEL	Pr.	Université Sétif 1.
EXAMINATEUR :	Ellaggoun FATEH	M.C.A	Université 8 mai 1945 de Guelma.
EXAMINATEUR :	Abdessamad AMIR	M.C.A	Université Abdelhamid Ibn Badis de Mostaganem.

Année : 2013

Résumé

Les méthodes du gradient conjugué sont très importantes pour la résolution des problèmes d'optimisation non linéaire, en particulier pour les problèmes de grande taille. Cependant, contrairement aux méthodes quasi-Newton. Les méthodes du gradient conjugué sont généralement analysées individuellement. Dans cette thèse, nous proposons une nouvelle famille à deux paramètres des méthodes du gradient conjugué pour l'optimisation sans contraintes. La famille à deux paramètres comprend non seulement les trois méthodes du gradient conjugué non linéaire déjà existantes, mais aussi une autre famille des méthodes du gradient conjugué comme sous-famille. La famille à deux paramètres des méthodes du gradient conjugué avec la recherche linéaire de Wolfe est indiquée pour assurer la propriété de descente à chaque itération. Certains résultats généraux de convergence globale sont également établis pour la famille à deux paramètres des méthodes du gradient conjugué. Les résultats numériques montrent que cette méthode est efficace pour les problèmes d'essais donnés. En outre, les méthodes liées à cette famille dont les méthodes hybrides sont largement discutées.

Mots clés: Optimisation sans contraintes, gradient conjugué, convergence globale, recherche linéaire.

Abstract

Conjugate gradient methods are very important ones for solving nonlinear optimization problems, especially for large scale problems. However, unlike quasi-Newton methods, conjugate gradient methods were usually analyzed individually. In this thesis, we propose a new two-parameter family of conjugate gradient methods for unconstrained optimization. The two-parameter family of methods not only includes the already existing three practical nonlinear conjugate gradient methods, but has other family of conjugate gradient methods as subfamily. The two-parameter family of methods with the Wolfe line search is shown to ensure the descent property of each search direction. Some global convergence results are also established for the two-parameter family of methods. The numerical results show that this method are efficient for the given test problems. In addition, the methods related to this family are uniformly discussed.

Key words: unconstrained optimization, conjugate gradient, line search, global convergence.

TABLE DES MATIÈRES

1 Introduction

1.1 Motivation

Les problèmes d'optimisation sont incontournables et sont rencontrés dans tous les domaines des sciences de l'ingénieur. Le développement des modèles théoriques et des techniques traitant les problèmes d'optimisation ont connu une accélération spectaculaire, particulièrement après la deuxième guerre mondiale.

Durant cette période, les mathématiciens se sont trouvés face à des problèmes à la taille et à la complexité croissante, ce qui fût une motivation pour la recherche des méthodes de résolution fiables et systématiques. La plupart d'entre elles reposent sur un socle solide des résultats théoriques établissant les conditions pour leur convergence vers la solution optimale recherchée.

Parmi les plus anciennes méthodes utilisées pour résoudre les problèmes d'optimisation non linéaires, on peut citer la méthode du gradient conjugué. Cette méthode a été découverte en 1952 par *Hestenes* et *Steifel* ([14,1952]) pour la minimisation 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* ([12,1964]) (*méthode de Fletcher-Reeves*) puis en 1969 par *Polak, Ribière* ([19,1969]) et *Ployak* ([20,1969]) (*méthode de Polak-Ribière-Ployak*). Une autre variante a été étudiée en 1987 par *Fletcher* ([11,1987]) (*méthode de la descente conjuguée*).

Une dernière version a été proposée par *Dai* et *Yuan* ([3,1999]) (*méthode de Dai-Yuan*). Ces méthodes jouent un rôle très important pour résoudre les problèmes d'optimisation non linéaire.

Soit $f : IR^n \to IR$ et (P) le problème de minimisation non linéaire, sans contraintes, suivant :

$$\min \ \{f(x) : x \in IR^n\} \qquad (P) \tag{0.1}$$

Les méthodes du gradient conjugué sont une classe des méthodes importantes pour résoudre (P), notamment pour les problèmes de grande de taille. Ces méthodes sont itératives et génèrent à partir d'un point initial x_0 une suite de points $\{x_k\}_{k\in\mathbb{N}}$ de la

manière suivante :

$$x_{k+1} = x_k + \alpha_k d_k, \tag{0.2}$$

où x_k est l'itération courante, le pas $\alpha_k \in \mathbb{R}$ est déterminé par une recherche linéaire exacte ou inexacte et les directions d_k sont calculées de façon récurrente par les formules suivantes :

$$d_k = \begin{cases} -g_k & \text{si } k = 1; \\ -g_k + \beta_k d_{k-1} & \text{si } k \geq 2, \end{cases} \tag{0.3}$$

où $g_k = \nabla f(x_k)$ est le gradient de f au point 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 par $y_{k-1} = g_k - g_{k-1}$, on obtient les variantes suivantes :

$$\beta_k^{PRP} = \frac{g_k^T y_k}{\|g_{k-1}\|^2} \qquad \text{Gradient conjugué variante } \textit{Polak-Ribière-Polyak} \tag{0.4}$$

$$\beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2} \qquad \text{Gradient conjugué variante } \textit{Fletcher Reeves} \tag{0.5}$$

$$\beta_k^{CD} \frac{\|g_k\|^2}{-d_{k-1}^T g_{k-1}} \qquad \text{Gradient conjugué variante } \textit{descente conjugué} \tag{0.6}$$

$$\beta_k^{DY} = \frac{\|g_k\|^2}{d_{k-1}^T y_{k-1}} \qquad \text{Gradient conjugué variante de } \textit{Dai-Yuan.} \tag{0.7}$$

Faire de la recherche linéaire veut dire déterminer un pas α_k le long d'une direction de descente d_k, autrement dit résoudre le problème unidimensionnel suivant :

$$\min_{x \in \mathbb{R}^n} f\left(x_k + \alpha d_k\right), \quad \lambda \in \mathbb{R} \tag{0.8}$$

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

comportement local de l'objectif f sur l'itération courante x_k, la méthode choisit la "direction du mouvement" d_k (qui normalement est une direction de descente de l'objectif : $\nabla f(x)^T .d < 0$) et exécute un pas dans cette direction

$$x_k \longmapsto x_{k+1} = x_k + \alpha d_k, \tag{0.9}$$

afin de réaliser un certain progrès en valeur de l'objectif, c'est-à-dire pour assurer que

$$f(x_k + \alpha d_k) \leq f(x_k).$$

Et dans la majorité des méthodes le pas dans la direction d_k est choisi par la minimisation unidimensionnelle de la fonction

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

On bute à réduire «de façon importante »la valeur de l'objectif par un pas $x_k \longmapsto x_{k+1} = x_k + \alpha_k d_k$ de x_k dans la direction d_k, tel que $f(x_k + \alpha_k d_k) < f(x_k)$, or cette condition de décroissance stricte n'est pas suffisante pour minimiser φ au moins localement. Il existe deux grandes classes des méthodes qui s'intéressent à l'optimisation unidimensionnelle.

1- Recherches linéaires exactes :

Comme on cherche a minimiser f, il semble naturel de chercher a minimiser le critère le long de d_k et donc de déterminer le pas λ_k comme solution du problème

$$\min_{\lambda \geq 0} \varphi(\lambda).$$

C'est ce que l'on appelle la règle de *Cauchy* et le pas déterminé par cette règle est appelé pas de *Cauchy* ou pas optimal. Dans certains cas, on préférera le plus petit point stationnaire de φ qui fait décroître par cette fonction

$$\inf\{\lambda \geq 0 : \varphi'(\lambda) = 0, \varphi(\lambda) < \varphi(0)\}.$$

On parle alors de règle de *Curry* et le pas déterminé par cette règle est appelé pas

de *Curry*. De manière un peu imprécise, ces deux règles sont parfois qualifiées de recherche linéaire exacte.

Ils ne sont utilisés que dans des cas particuliers, par exemple lorsque φ est quadratique, la solution de la recherche linéaire s'obtient de façon exacte et dans un nombre fini d'itérations.

Dans le cas où f est une fonction quadratique strictement convexe avec une recherche linéaire exacte toutes ces variantes de β_k ont la même valeur.

$$\beta_k^{PRP} = \beta_k^{FR} = \beta_k^{CD} = \beta_k^{DY}.$$

2- Recherches linéaires inexactes :

Une condition naturelle est de demander que f décroisse autant qu'une paramètre $\omega_1 \in]0,1[$ parfois appelée condition d'*Armijo* ou condition de décroissance linéaire

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k \nabla f(x_k)^T d_k. \tag{0.10}$$

Dans la règle d'*Armijo* on assure la décroissance de la fonction objectif à chaque pas mais ce ne pas suffisant. Les conditions de *Goldstein* et *Price* suivantes sont suffisantes pour assurer la convergence sous certaines conditions et indépendamment de l'algorithme qui calcule le paramètre. Dans celle-ci, en ajoutant une deuxième inégalité à la règle d*'Armijo* on obtient la règle de *Goldstein*.

$$f(x_k) + \sigma_1 \alpha_k \nabla^T f(x_k) d_k \geq f(x_k + \alpha_k d_k) \geq f(x_k) + \sigma_2 \alpha_k \nabla^T f(x_k) d_k, \tag{0.11}$$

où σ_1 et σ_2 sont deux constantes vérifiant $0 < \sigma_1 < \sigma_2 < 1$.

La règle de *Wolfe* fait appel au calcul de $\varphi'(\alpha)$, 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 pour $\alpha > 0$.

$$f(x_k + \alpha d_k) \leq f(x_k) + \delta \alpha \nabla f(x_k)^T . d_k, \tag{0.12}$$

$$\nabla f(x_k + \alpha d_k)^T . d_k \geq \sigma \nabla f(x_k)^T . d_k, \tag{0.13}$$

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

On obtient des contraintes plus fortes si l'on remplace (0.13) par

$$\left|\nabla f(x_k + \alpha d_k)^T . d_k\right| \leq -\sigma \nabla f\left(x\right)^T . d_k \tag{0.14}$$

Les formules (0.12) et (0.14) sont les conditions de *Wolfe* fortes. La contrainte (0.14) entraîne que $\sigma\varphi\left(0\right) \leq \varphi'\left(\alpha\right) < -\sigma\varphi\left(0\right)$, c'est-à-dire. $\varphi'\left(\alpha\right)$ n'est pas "trop" positif.

La règle de *Wolfe* relaxée à été proposée par *Dai* et *Yuan* [1996], cette règle consiste à choisir le pas satisfaisant les conditions suivantes :

$$f(x_k + \alpha d_k) \leq f(x_k) + \delta\alpha \nabla^T f(x_k) . d_k, \tag{0.15}$$

$$\sigma_1 \nabla^T f\left(x_k\right) . d_k \leq \nabla^T f(x_k + \alpha_k d_k) . d_k \leq -\sigma_2 \nabla^T f\left(x_k\right) . d_k, \tag{0.16}$$

où $0 < \delta < \sigma_1 < 1$ et $\sigma_2 > 0$.

Pour des fonctions générales, les différents résultats pour le scalaire β_k pour les méthodes du gradient conjugués non-linéaires sont distincts [3, 10, 11, 12, 14, 16, 19, 20, 22].

En 1964, *Fletcher et Reeves* (FR) à introduit la méthode du gradient conjugué, avec :

$$\beta_k^{FR} = \frac{\left\|g_k\right\|^2}{\left\|g_{k-1}\right\|^2}, \tag{0.17}$$

où $\|.\|$ désigne la norme euclidienne.

Pour les fonctions objectives non-quadratiques, la convergence globale de la méthode de *Fletcher et Reeves* a été prouvée avec la recherche linéaire exacte ou une recherche linéaire de *Wolfe* forte avec certaines conditions. Toutefois, si les conditions de *Wolfe* forte sont satisfaits pour $\sigma < 1$, la méthode de *Fletcher et Reeves* peut assurer une direction de descente et converge globalement uniquement dans le cas où f est quadratique [5, 4]. La méthode de descente conjugué (CD) de *Fletcher* [11], où

$$\beta_k^{CD} = \frac{\left\|g_k\right\|^2}{-d_{k-1}^T g_{k-1}}, \tag{0.18}$$

assure une direction de descente pour les fonctions générales si la recherche linéaire satisfait les conditions de *Wolfe* fortes avec $\sigma < 1$. Mais la convergence globale de la méthode est prouvée [6] seulement dans le cas où la recherche linéaire satisfait (0.12) et

$$\sigma g_k^T d_k \leq g(x_k + \alpha_k d_k)^T d_k \leq 0. \tag{0.19}$$

Pour toute constante positive σ_2, un exemple est construit dans [6] montrant que la méthode de descente conjugué avec α_k satisfaisant (0.12) et

$$\sigma_1 g_k^T d_k \leq g(x_k + \alpha_k d_k)^T d_k \leq -\sigma_2 g_k^T d_k, \tag{0.20}$$

ne converge pas.

Récemment, Dai et Yuan [3] à proposé une méthode non linéaire du gradient conjugué, qui à la forme (0.2), (0.3) avec

$$\beta_k^{DY} = \frac{\|g_k\|^2}{d_{k-1}^T y_{k-1}}, \tag{0.21}$$

où $y_{k-1} = g_k - g_{k-1}$. Une propriété remarquable de la méthode de *Dai et Yuan* est qu'il fournit une direction de descente à chaque itération et converge globalement à condition que le pas α_k satisfait les conditions de *Wolfe*, à savoir, (0.12) et

$$\sigma g_k^T d_k \leq g(x_k + \alpha_k d_k)^T d_k. \tag{0.22}$$

Par des calculs directs, nous pouvons en déduire une forme équivalente pour β_k^{DY} à savoir

$$\beta_k^{DY} = \frac{g_k^T d_k}{g_{k-1}^T d_{k-1}}. \tag{0.23}$$

Bien que toutes ces méthodes ont été très importantes pour la résolution du problème (P), leur étude s'est faite de façon individuelle et chacune à part par *Flecher, Powell, Wolf, El Baali, Polak-Ribière, Polyak, Y. H. Dai et Y. Yuan,....* Il est alors légitime de se poser la question suivante :

Peut-on définir et unifier toutes ces méthodes dans une même classe et étudier par la suite ses propriétés et sa convergence comme il a été fait avec les méthodes quasi-Newtoniennes [26] **?**

Y. Dai et *Y. Yuan* ont répondu dans ([7,2003]) positivement à cette question et ont proposé une famille des méthodes du gradient conjugué globalement convergents, dans laquelle

$$\beta_k = \frac{\|g_k\|^2}{\lambda \|g_{k-1}\|^2 + (1-\lambda)(d_{k-1}^T y_{k-1})}, \tag{0.24}$$

autrement dit β_k est de la forme générale suivante :

$$\beta_k = \frac{\phi_k}{\phi_{k-1}}, \tag{0.25}$$

où ϕ_k satisfait

$$\phi_k = \lambda \|g_k\|^2 + (1-\lambda)(-g_k^T d_k), \qquad \lambda \in [0,1], \tag{0.26}$$

avec $\lambda \in [0,1]$ est un paramètre, et a prouvé que la famille des méthodes du gradient conjugué utilisant les recherches linéaire qui satisfont les conditions de *Wolfe* (0.8) et (0.13) converge globalement si les paramètres σ_1, σ_2 et λ sont telles que

$$\sigma_1 + \sigma_2 \leq \lambda^{-1}.$$

Observant que les formules (0.17), (0.18) et (0.23) partagent le même numérateur, mais possèdent trois dénominateurs différents.

Notre travail s'est inspiré et généralise les résultats obtenus par les auteurs cités ci- dessus et repose sur des travaux d'*Al-Baali* ([1,1985]), ([2,1985]) et *Dai* et *Y.Yuan* ([5,2001]) pour arriver à la direction de descente à chaque itération et à la convergence globale à condition que le pas α_k satisfait les conditions de *Wolfe*.

Le but principal de cette thèse est donc d'améliorer le travail de *Y. Dai* et *Y. Yuan* ([7, 2003]), pour cela nous proposons une nouvelle famille à deux paramètres des méthodes du gradient conjugué pour l'optimisation sans contraintes et nous étudions

ses propriétés et sa convergence avec β_k^* défini comme suit :

$$\beta_k^* = \frac{(1-\lambda_k)\left\|g_k\right\|^2 + \lambda_k(-g_k^T d_k)}{(1-\lambda_k-\mu_k)\left\|g_{k-1}\right\|^2 + (\lambda_k+\mu_k)(-g_{k-1}^T d_{k-1})}, \tag{0.27}$$

où $\lambda_k \in [0,1]$ et $\mu_k \in [0, 1-\lambda_k]$ sont des paramètres.

autrement dit β_k^* est de la forme générale suivante

$$\beta_k^* = \frac{\phi_k}{\phi'_{k-1}}, \tag{0.28}$$

où ϕ_k satisfait

$$\phi_k = (1-\lambda_k)\left\|g_k\right\|^2 + \lambda_k(-g_k^T d_k), \tag{O.29}$$

et

$$\phi'_{k-1} = (1-\lambda_k-\mu_k)\left\|g_{k-1}\right\|^2 + (\lambda_k+\mu_k)(-g_{k-1}^T d_{k-1}). \tag{0.30}$$

Il est clair que la formule (0.27) est une généralisation des trois méthodes précédentes.

1.2 Le plan de thèse

Cette thèse comporte une introduction, six chapitres et une conclusion avec des perspectives.

Le premier chapitre contient des remarques préliminaires et des notions essentielles nécessaires pour notre travail. On insiste surtout sur les notions de direction de descente, la notion de convexité ainsi que les notions de conditions d'optimalité des problèmes de minimisation sans contraintes.

Le deuxième chapitre est consacré aux recherches linéaires inexactes célèbres (*Armijo, Wolf* et *Goldstein*). On montre comment ces recherches linéaires contribuent dans la convergence des algorithmes à directions de descente.

Dans le troisième chapitre, on étudie tout d'abord la contribution de la recherche linéaire à la convergence des algorithmes à direction de descente. Ce n'est qu'une contribution, parce que la recherche linéaire seule ne pourra pas assurer la convergence des itérés. Il est clair que le choix de la direction de descente joue aussi un

rôle. Cela se traduit par une condition dite de *Zoutendijk* qui assure la convergence des méthodes du gradient conjugué.

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

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

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

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

Puis on donne les principes des méthodes de descente et du gradient ainsi qu'une description des méthodes utilisant des directions conjugués notamment la méthode du gradient conjugué.

Le chapitre 4 est une synthèse des différents résultats de convergence des différentes variantes de la méthode du gradient conjugué avec la recherche linéaire inexacte de *Wolfe* forte et de *Wolfe* faible. On s'intéressera spécialement aux méthodes de *Fletcher- Reeves, Polak-Ribière* et *Dai- Yuan.* Nous essayerons de répondre aux deux questions suivantes :

Les directions d_k définies par (0.3) sont-elles des directions de descente de f ?

Les méthodes ainsi définies sont-elles convergentes ?

Le cinquième chapitre est consacré à la classe unifiée des méthodes du gradient conjugué introduite par *Y. Dai* et *Y. Yuan* ([7,2003]). Cette classe génère des suites $\{x_k\}_{k\in\mathbb{N}}$ de la façon suivante

$$x_{k+1} = x_k + \alpha_k d_k.$$

Le pas $\alpha_k \in \mathbb{R}$ est déterminé par une optimisation unidimensionnelle ou recherche linéaire inexacte. Les directions d_k sont calculées de façon récurrente par les formules suivantes :

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

$g_k = \nabla f(x_k)$ et $\beta_k \in \mathbb{R}$. Pour la forme unifiée du gradient conjugué en question, β_k prend la forme générale suivante :

$$\beta_k = \frac{\phi_k}{\phi_{k-1}}$$

où ϕ_k satisfait

$$\phi_k = \lambda \|g_k\|^2 + (1-\lambda)(-g_k^T d_k), \qquad \lambda \in [0,1].$$

Les auteurs ont prouvé que les différentes valeurs de β_k, c'est-à-dire β_k^{FR}, β_k^{DY}, ...sont des valeurs particulières de β_k et que cette méthode peut assurer une direction de descente à chaque itération et converge globalement avec la recherche linéaire de *wolfe*.

Le dernier chapitre est la partie qui comporte le résultat original de cette thèse, une nouvelle famille à deux paramètres des méthodes du gradient conjugué est présentée pour l'optimisation sans contraintes [27], la propriété de descente est assurée à chaque itération et certains résultats généraux de convergence sont montrés pour cette famille avec la recherche linéaire de *Wolfe*. Le but principal de cette thèse est donc d'amélioré le travail de *Y. Dai et Y. Yuan.*[7]. Cette nouvelle famille à deux paramètres des méthodes du gradient conjugué pour l'optimisation sans contraintes est présentée avec toutes ses propriétés et sa convergence globale avec β_k^* de la forme générale suivante :

$$\beta_k^* = \frac{\phi_k}{\phi'_{k-1}},$$

où ϕ'_{k-1} satisfait

$$\phi'_{k-1} = (1-\lambda_k-\mu_k)\|g_{k-1}\|^2 + (\lambda_k+\mu_k)(-g_{k-1}^T d_{k-1}),$$

et

$$\phi_k = \lambda \|g_k\|^2 + (1-\lambda)(-g_k^T d_k), \qquad \lambda \in [0,1],$$

est démontrée.

Les résultats numériques montrent que cette méthode est efficace pour les pro-

blèmes d'essais donnés. En outre, les méthodes liées à cette famille dont les méthodes hybrides sont largement discutées.

Enfin, le manuscrit se termine par une conclusion générale et quelques perspectives.

CHAPITRE 1

Minimisation sans contraintes

Ce chapitre sera consacré à certaines notions, définitions de base et théorèmes concernant les méthodes itératives d'optimisation non linéaire en général et les méthodes du gradient conjugué en particulier. Il existe de nombreuses sortes des problèmes d'optimisation, certaines caractéristiques permettent de les distinguer : Comportent-ils des contraintes ? Les fonctions objectives sont-elles quadratiques ? Sont-elles convexes ? Les domaines de définition des fonctions sont-ils continus ou discrets ? Tous ces problèmes possèdent des structures différentes et ne peuvent être traités de la même façon. Le présent projet a pour vocation de se focaliser sur les problèmes d'optimisation sans contraintes et non linéaires. Ce problème sera formulé comme suit :

$$(P) \qquad \text{minimiser } f(x) \qquad x \in \mathbb{R}^n, \tag{1.1}$$

où $f : \mathbb{R}^n \longrightarrow \mathbb{R}$. La plupart des méthodes de résolution d'optimisation qu'on étudie par la suite sont de nature itérative, c'est-à-dire qu'à partir d'un point initial x_0, elles engendrent une suite infinie $x_1, x_2, ..., x_k, ..$Dont on espère qu'elle converge vers la solution optimale.

Pour construire des algorithmes de minimisation sans contraintes on fait appel à

des processus itératifs du type

$$x_{k+1} = x_k + \lambda_k d_k, \tag{1.2}$$

où d_k détermine la direction de déplacement à partir du point x_k et λ_k est solution optimale du problème d'optimisation unidimensionnel suivant :

$$\underset{\lambda \geq 0}{min} f\left(x_k + \lambda d_k\right),$$

c'est-à-dire que λ_k vérifie

$$f\left(x_k + \lambda_k d_k\right) \leq f\left(x_k + \lambda d_k\right), \forall \lambda \geq 0. \tag{1.3}$$

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

La façon avec laquelle on construit les vecteurs d_k et les scalaires λ_k détermine directement les propriétés du processus et spécialement en ce qui concerne la convergence de la suite $\{x_k\}$, la vitesse de la convergence. Pour s'approcher de la solution optimale du problème (1.1) (dans le cas général, c'est un point en lequel ont lieu peut être avec une certaine précision les conditions nécessaires d'optimalité de f) on se déplace naturellement à partir du point x_k dans la direction de la décroissance de la fonction f.

1 Rappel de quelques définitions

Nous introduisons ici les principales définitions de base qui seront utilisées par la suite.

Définition 1.1. (Convergence des algorithmes) *Un algorithme de résolution est un procédé qui permet, à partir de la donnée du point initial x_1, d'engendrer la suite $x_1, x_2 \ldots, x_k, \ldots$.*

Un algorithme est parfaitement défini par la donnée de l'application A qui as-

socie à x_k le point $x_{k+1} \in A(x_k)$. Ceci permettra de confondre un algorithme et l'application A qui lui est associée.

Définition 1.2. (**Convergence globale**) *Nous dirons qu'un algorithme décrit par une application multivoque A, est globalement convergent (ou encore : possède la propriété de convergence globale) si quel que soit le point de départ x_1 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).*

Définition 1.3. (Modes de convergence) *Soit $\{x_k\}_{k\in N}$ une suite dans $\mathbb{R}^n$ convergeant vers x^*.*

♦ *Si*

$$\limsup \frac{\|x_{k+1-}x_*\|}{\|x_k - x_*\|} = \tau < 1,$$

on dit que $\{x_k\}_{k\in N}$ converge vers x^* linéairement avec le taux τ.

♦ *Si*

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

on dit que la convergence est superlinéaire.

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

$$\limsup_{k\longrightarrow\infty} \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|^p} < +\infty,$$

on dit que la convergence est superlinéaire d'ordre p.

En particulier si

$$\limsup_{k\longrightarrow\infty} \frac{\left\|x^{k+1} - x^*\right\|}{\left\|x^k - x^*\right\|^2} < +\infty,$$

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

Définition 1.4. (Ensemble convexe).*Soit L'ensemble $C \subseteq \mathbb{R}^n$.C est convexe si seulement si*

$$\forall x, y \in C^2, \forall t \in [0,1], tx + (1-t)y \in C$$

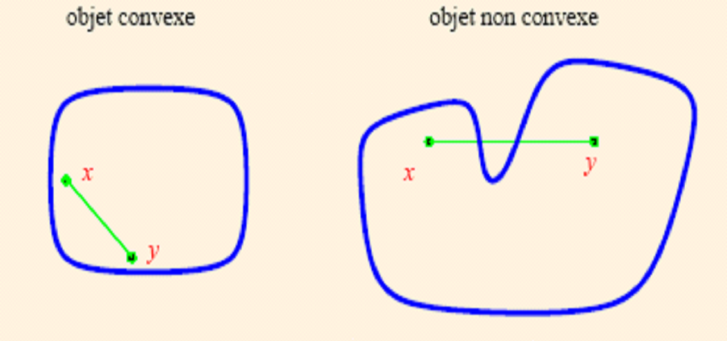

FIG. 1.1 – Objet convexe et non convexe

Remarque 1.1. *Soit* $x_1, x_2, \quad x_k \in \mathbb{R}^n$ *et* t_j *telle que* $t_j \geq 0$ *et* $\sum_{j=1}^{k} t_j = 1$. *Toute expression de la forme*

$$\sum_{j=1}^{k} t_j x_j.$$

S'appelle combinaison convexe des points x_J *ou barycentre.*

Définition 1.5. (Fonction convexe). *Soit* $C \subseteq \mathbb{R}^n$ *un ensemble convexe non vide. Une fonction* $f : C \longrightarrow \mathbb{R}$ *est convexe si seulement si*

$$\forall x, y \in C^2, \ \forall t \in [0,1], \quad f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

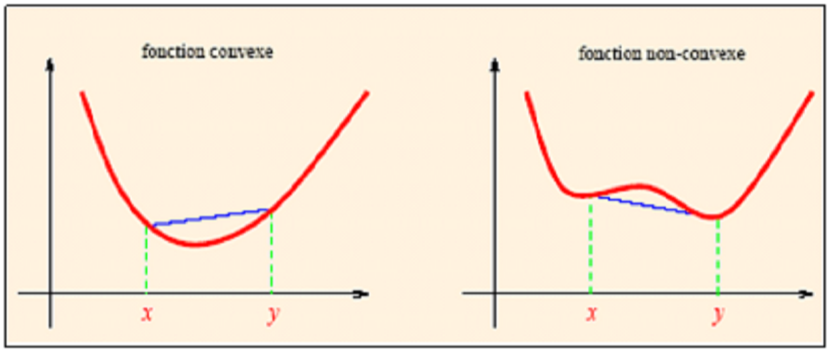

FIG. 1.2 – Fonction convexe et non convexe

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

$$\forall x, y \in C^2, x \neq y, \forall t \in [0,1], \quad f(tx + (1-t)y) < tf(x) + (1-t)f(y)$$

Définition 1.6. (Fonction fortement ou uniformément convexe de module $\upsilon > 0$).

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

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - \frac{\upsilon}{2}t(1-t)\|x-y\|^2, \quad \forall x, y \in C^2, \ \forall t \in [0,1].$$

Définition 1.7. (Fonction convexe différentiable).

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

$$f(x) = f(\widehat{x}) + A(x - \widehat{x}) + \|x - \widehat{x}\| \, \alpha(\widehat{x}, x - \widehat{x}),$$

où $\alpha(\widehat{x}, x-\widehat{x}) \underset{x\to\widehat{x}}{\longrightarrow} 0$. *On peut note le vecteur* A *comme suit :*

$$A = \nabla f(\widehat{x}) = (\frac{\partial f(\widehat{x})}{\partial x_1}, ..., \frac{\partial f(\widehat{x})}{\partial x_n}). \tag{1.4}$$

Définition 1.8. (Fonction convexe deux fois différentiable).

Soit $C \subseteq \mathbb{R}^n$ *non vide et* $f : \mathbb{R}^n \longrightarrow \mathbb{R}$. *est dite deux fois différentiables ou point* $\widehat{x} \in int(C)$ *s'il existe un vecteur* $\nabla f(\widehat{x})$ *et une matrice symétrique* $H(\widehat{x})$ *d'ordre* (n, n) *appelée matrice hessienne et une fonction* $\alpha : \mathbb{R}^n \to \mathbb{R}$ *tels que*

$$\forall x \in C : f(x) = f(\widehat{x}) + \nabla f(\widehat{x})^T (x-\widehat{x}) + \frac{1}{2}(x-\widehat{x})^T H(\widehat{x})(x-\widehat{x}) + \|x-\widehat{x}\|^2 \alpha(\widehat{x}, x-\widehat{x}) \tag{1.5}$$

où $\alpha(\widehat{x}, x-\widehat{x}) \underset{x\to\widehat{x}}{\longrightarrow} 0$. *On peut écrire :*

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

Remarque 1.2

♦ *La matrice* $H(\hat{x})$ *est dite semi-définie positive si*

$$\forall x \in \mathbb{R}^n : x^T H(\hat{x}) x \geq 0$$

♦ *La matrice* $H(\hat{x})$ *est dite définie positive si*

$$\forall x \in \mathbb{R}^n, x \neq 0 : x^T H(\hat{x}) x > 0$$

2 Conditions d'optimalité des problèmes de minimisation sans contraintes :

Considérons le problème d'optimisation sans contraintes (P)

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

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

2.1 Pourquoi avons-nous besoin de conditions d'optimalité ?

Afin d'analyser ou de résoudre de manière efficace un problème d'optimisation, il est fondamental de pouvoir disposer de conditions d'optimalité. En effet, celles-ci nous servent non seulement à vérifier la validité des solutions obtenues, mais souvent l'étude de ces conditions aboutit au développement des algorithmes de résolution eux-mêmes. Des conditions équivalentes peuvent être obtenues de diverses manières, en procédant à des analyses suivantes différentes "lignes directrices". L'approche considérée ici pour l'obtention de conditions est basée sur les notions de descente.

2.2 Minima locaux et globaux

Définition 1.9.

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

$$f(\widehat{x}) \leq f(x) : \forall x \in V_\varepsilon(\widehat{x}) \tag{1.6}$$

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

$$f(\widehat{x}) < f(x) : \forall x \in V_\varepsilon(\widehat{x}), x \neq \hat{x} \tag{1.7}$$

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

$$f(\widehat{x}) \leq f(x) : \forall x \in \mathbb{R}^n \tag{1.8}$$

Remarque 1.3. *Dans le cas d'une fonction objectif convexe, il n'y a pas de distinction entre minimum local et global c'est-à-dire tout minimum local est également global, comme l'établit le théorème suivant.*

2.3 Direction de descente

Principes généraux

Considérons le problème d'optimisation sans contraintes :

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

où $f : \mathbb{R}^n \to \mathbb{R}$ est supposé régulière.

On note aussi $\nabla f(x)$ et $\nabla^2 f(x)$ le gradient et le hessien de f en x pour ce produit scalaire.

On s'intéresse ici à une classe d'algorithmes qui sont fondés sur la notion de direction de descente.

On dit que d est une direction de descente de f en $x \in \mathbb{R}^n$ si

$$\nabla f(x)^T.d < 0. \tag{1.9}$$

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

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

L'ensemble des directions de descente de f en x,

$$\left\{d \in \mathbb{R}^n : \nabla^T f(x).d < 0\right\},$$

forme un demi-espace ouvert de $\mathbb{R}^n$

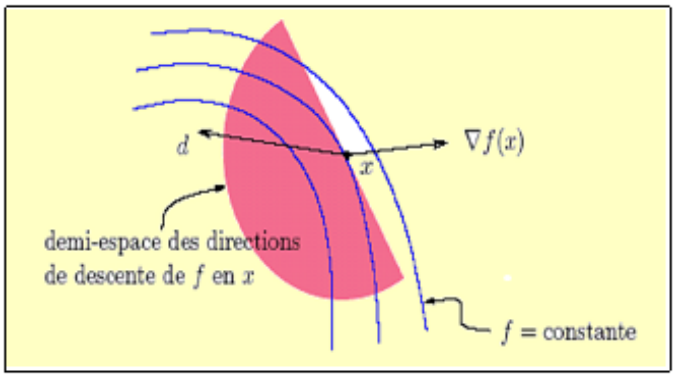

FIG. 1.3 – Demi-espace des directions de descente d de f en x.

On voit que si d est une direction de descente, $f(x+\lambda d) < f(x)$, pour tout $\lambda > 0$ suffisamment petit, et donc que f décroît strictement dans la direction d. De telles directions sont intéressantes en optimisation car pour faire décroître f, il suffit de faire un déplacement le long de d. Les méthodes à directions de descentes utilisent cette idée pour minimiser une fonction. Elles construisent la suite des itérés $\{x_k\}_{k\geq 1}$, approchant une solution x_k du problème (1.1) par la récurrence :

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

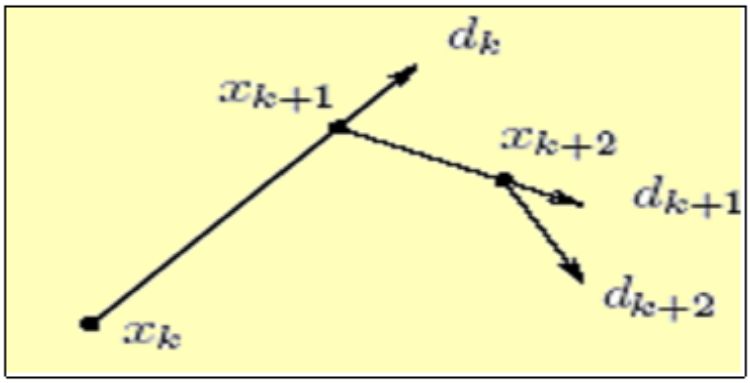

FIG. 1.4 – La direction de descente

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

Pour définir une direction 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 λ_k, c'est ce que l'on appelle la recherche linéaire.

Remarque 1.4. *Par la suite l'angle θ_k entre la direction d_k et $-g_k$ jouera un rôle important dans le processus de la convergence. Pour θ_k on a la relation classique suivante*

$$-g_k^T d_k = \|g_k\| \, \|d_k\| \cos\theta_k.$$

Définition 1.10. *Soit $f : x \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, $\hat{x} \in \mathbb{R}^n$ et $d \in \mathbb{R}^n$ est une direction de descente au point $\hat{x}$ si et seulement si il existe $\delta > 0$ tel que*

$$f(\hat{x} + \lambda d) < f(\hat{x}) : \forall \lambda \in]0, \delta[\,, \tag{1.10}$$

donnons maintenant une condition suffisante pour que soit d une direction de descente.

Théorème 1.1. *Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ différentiable au point $\hat{x} \in \mathbb{R}^n$ et $d \in \mathbb{R}^n$ une*

direction vérifiant la condition suivante :

$$f'(\hat{x}, d) = \nabla f(\hat{x})^T .d < 0.$$

Alors d est une direction de descente au point $\hat{x}$.

Preuve : f est différentiable au point $\hat{x}$. Donc

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

avec $\alpha(\hat{x}, d) \underset{\lambda \longrightarrow 0}{\longrightarrow} 0$ ceci implique que

$$f'(\hat{x}, d) = \lim_{\lambda \to 0} \frac{f(\hat{x} + \lambda d) - f(\hat{x})}{\lambda} = \nabla f(\hat{x})^T .d < 0. \tag{1.11}$$

La limite étant strictement négative, alors il existe un voisinage de zéro $V(0)$ tel que

$$\frac{f(\hat{x} + \lambda d) - f(\hat{x})}{\lambda} < 0, \qquad \forall \lambda \in V(0). \tag{1.12}$$

La relation (1.12) est particulièrement vraie pour tout $\lambda \in]-\delta, +\delta[$. On obtient le résulta cherché en multipliant la relation (1.12) par $\lambda > 0$.

□

Algorithme 2.1 (méthode à directions de descente- une itération)

Etape 0 : (initialisation)

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

Etape1 :

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

Etape2 :

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

Etape3 :

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

Etape4 :

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

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

2.4 Schémas général des algorithmes d'optimisation sans contraintes

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 :

$$x_{k+1} = x_k + \lambda_k d_k, \qquad \lambda_k \in]0, +\delta[$$

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

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

Un bon chois de d_k et λ_k de permet ainsi de construire une multitude d'algorithme d'optimisation.

Exemple de choix de direction 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. Dans le cas $d_k = -(H(x_k))^{-1}.\nabla f(x_k)$, on obtient la méthode de Newton où la matrice hessienne $H(x_k)$ est définie positive.

Exemple de choix de pas λ_k :

On choisit en général λ_k de façon optimale c'est-à-dire que λ_k doit vérifier

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

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

2.5 Conditions nécessaires d'optimalité

Etant donné un vecteur $\hat{x}$, nous souhaiterions être capables de déterminer si ce vecteur est un minimum local ou global de (P). La propriété de différentiabilitè de f

fournit une première manière de caractériser une solution optimale. Enonçons tout d'abord un théorème pour établir une première condition nécessaire d'optimalité.

Condition nécessaire d'optimalité du premier ordre

Théorème 1.2. *Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ différentiable au point $\hat{x} \in \mathbb{R}^n$.Si $\hat{x}$ est un minimum local de (P) alors $\nabla f(\hat{x}) = 0$.*

Preuve : C'est une conséquence directe du théorème (1.1).En effet, supposons que $\nabla f(\hat{x}) \neq 0$. Puisque la direction $d = -\nabla f(\hat{x})$ est une direction de descente, alors il existe $\delta > 0$ tel que

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

Ceci est contradiction avec le fait que $\hat{x}$ est une solution optimale locale de (P).

□

Remarque 1.5. *si f est convexe, la condition nécessaire du premier ordre est également suffisante pour que $\hat{x}$ soit un minimum global. Dans le cas où f est deux fois différentiable, une autre condition nécessaire est donnée par le théorème (1.3) elle est appelée condition nécessaire du second ordre car elle fait intervenir la matrice hessienne de f (que nous noterons $\nabla^2 f(x)$, dont les éléments sont ses secondes dérivées partielles).*

Condition nécessaire d'optimalité du second ordre

Théorème 1.3. *Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ deux fois différentiable au point $\hat{x} \in \mathbb{R}^n$.Si $\hat{x}$ est un minimum local de (P) alors $\nabla f(\hat{x}) = 0$ et la matrice hessiene de f au point $\hat{x}$ est semi définie positive.*

Preuve : Soit $x \in \mathbb{R}^n$ quelconque, étant f deux fois différentiable au point $\hat{x}$, on aura pour tout $\lambda \neq 0$

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

Ceci implique

$$\frac{f(\hat{x}+\lambda x)-f(\hat{x})}{\lambda^2}=\frac{1}{2}x^T H(\hat{x})x+\|x\|^2\,\alpha(\hat{x},\lambda x). \tag{1.13}$$

$\hat{x}$ est un optimum local, il existe alors $\delta>0$ tel que

$$\frac{f(\hat{x}+\lambda x)-f(\hat{x})}{\lambda^2}\geq 0, \qquad \forall\lambda\in\,]-\delta,+\delta[\,.$$

Si on prend en considération (1.13) et on passe à la limite quand $\lambda\to 0$, $\lambda\neq 0$, on obtient :

$$x^T H(\hat{x})x\geq 0, \qquad \forall x\in\mathbb{R}^n.$$

□

2.6 Conditions suffisantes d'optimalité

Les conditions données précédemment sont nécessaires (si f n'est pas convexe), c'est-à-dire qu'elles doivent être satisfaites pour tout minimum local ; cependant, tout vecteur vérifiant ces conditions n'est pas nécessairement un minimum local. Le théorème (1.5) établit une condition suffisante pour qu'un vecteur soit un minimum local, si f est deux fois différentiable.

Condition suffisante d'optimalité du second ordre

Théorème 1.4. *Soient $f:\mathbb{R}^n\to\mathbb{R}$ deux fois différentiable au point $\hat{x}\in\mathbb{R}^n$. Si $\nabla f(\hat{x})=0$ et $H(\hat{x})$ est définie positive, alors $\hat{x}$ est un minimum local strict de (P).*

Preuve : f étant deux fois différentiable au point $\hat{x}$, on aura pour tout $x\in\mathbb{R}^n$

$$f(x)=f(\widehat{x})+\frac{1}{2}(x-\widehat{x})^T H(\widehat{x})(x-\widehat{x})+\|x-\widehat{x}\|^2\,\alpha(\widehat{x},x-\widehat{x}),\quad \alpha(\widehat{x},(x-\widehat{x}))\underset{x\to\widehat{x}}{\to}0,\ (\nabla f(\hat{x})=0). \tag{1.14}$$

Supposons que $\widehat{x}$ n'est pas un optimum local strict. Alors il existe une suite $\{x_k\}_{k\in\mathbb{N}\mathbb{N}^*}$ telle que

$$x_k \neq \widehat{x} : \forall k, \qquad x_k \underset{k\to\infty}{\to} \widehat{x} \quad et \quad f(x_k) \leq f(\hat{x}). \tag{1.15}$$

Dans (1.14) prenons $x = x_k$ divisons le tout par $\|x_k - \widehat{x}\|^2$ et notons $d_k = \dfrac{(x_k - \widehat{x})}{\|x_k - \widehat{x}\|}$, on obtient

$$\frac{f(x_k) - f(\hat{x})}{\|x_k - \widehat{x}\|^2} = \frac{1}{2} d_k^T H(\hat{x}) d_k + \alpha(\widehat{x}, (x_k - \widehat{x}), \quad \alpha(\widehat{x}, (x_k - \widehat{x}) \underset{k\to\infty}{\to} 0. \tag{1.16}$$

(1.15) et (1.16) impliquent

$$\frac{1}{2} d_k^T H(\hat{x}) d_k + \alpha(\widehat{x}, (x_k - \widehat{x})) \leq 0, \qquad \forall k.$$

D'autre part la suite $\{d_k\}_{k\in\mathbb{N}^*}$ est bornée ($\|d_k\| = 1, \forall n$). Donc il existe une sous suite $\{d_k\}_{k\in\mathbb{N}_1\subset\mathbb{N}}$ telle que $d_k \underset{k\to\infty}{\to} \tilde{d}_{k\in\mathbb{N}_1}$.

Finalement lorsque $k \to \infty$ on obtient

$$\frac{1}{2} \tilde{d}^T H(\hat{x}) \tilde{d}_k \leq 0.$$

La dernière relation et le fait que $\tilde{d} \neq 0$ ($\left\|\tilde{d}\right\| = 1$) impliquent que la matrice hessienne $H(\hat{x})$ n'est pas définie positive. Ceci est en contradiction avec l'hypothèse.

□

Remarque 1.6. *Dans le cas où f est convexe, alors tout minimum local est aussi global. De plus si f est strictement convexe, alors tout minimum local devient non seulement global mais aussi unique.*

CHAPITRE 2

La recherche linéaire inexacte

1 Recherche linéaire

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

$$\min_{x \in \mathbb{R}^n} f(x_k + \lambda d_k), \quad \lambda \in \mathbb{R} \tag{2.1}$$

Notre intérêt pour la recherche linéaire ne vient pas seulement du fait que dans les applications on rencontre naturellement des problèmes unidimensionnels, mais plutôt du fait que la recherche linéaire est un composant fondamental de toutes les méthodes traditionnelles d'optimisation multidimensionnelles. En regardant le comportement local de l'objectif f sur l'itération courante x_k, la méthode choisit la "direction du mouvement" d_k (qui, normalement, est une direction de descente de l'objectif : $\nabla f(x)^T .d < 0$) et exécute un pas dans cette direction :

$$x_k \longmapsto x_{k+1} = x_k + \lambda d_k, \tag{2.2}$$

afin de réaliser un certain progrès en valeur de l'objective, c'est-à-dire, pour assurer que :

$$f(x_k + \lambda d_k) \leq f(x_k).$$

Et dans la majorité des méthodes le pas dans la direction d_k est choisi par la minimisation unidimensionnelle de la fonction :

$$\lambda \longrightarrow \varphi(\lambda) = f(x_k + \lambda d_k). \tag{2.3}$$

Ainsi, la technique de recherche linéaire est un brick de base fondamentale de toute méthode multidimensionnelle. 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 $\mathbb{R}$.

1.1 But de la recherche linéaire

On a vu que dans le cas non-quadratique les méthodes de descente :

$$x_{k+1} = x_k + \lambda d_k; \qquad \lambda > 0,$$

nécessitent la recherche d'une valeur de $\lambda_k > 0$ optimale ou non, vérifiant :

$$f(x_k + \lambda d_k) \leq f(x_k). \tag{2.4}$$

On définit comme précédemment la fonction $\varphi(\lambda) = f(x_k + \lambda d_k)$. Rappelons que si f est différentiable, le pas optimal $\hat{\lambda}$ peut être caractérisé par

$$\begin{cases} \varphi'\left(\hat{\lambda}\right) = 0, \\ \varphi\left(\hat{\lambda}\right) \leq \varphi(\lambda) \ \textit{pour} \ 0 \leq \lambda \leq \hat{\lambda}, \end{cases}$$

autrement dit, $\hat{\lambda}$ est un minimum local de φ qui assure de plus la décroissance de f.

2 Recherches linéaires exactes et inexactes

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

2.1 Recherches linéaires exactes

Comme on cherche a minimiser f, il semble naturel de chercher a minimiser le critère le long de d_k et donc de déterminer le pas λ_k comme solution du problème

$$\min_{\lambda \geq 0} \varphi(\lambda).$$

C'est ce que l'on appelle la règle de *Cauchy* et le pas déterminé par cette règle est appelé pas de *Cauchy* ou pas optimal. Dans certains cas, on préférera le plus petit point stationnaire de φ qui fait décroître par cette fonction

$$\inf \{\lambda \geq 0 : \varphi\prime(\lambda) = 0, \varphi(\lambda) < \varphi(0)\}.$$

On parle alors de règle de *Curry* et le pas déterminé par cette règle est appelé pas de *Curry*. De manière un peu imprécise, ces deux règles sont parfois qualifiées de recherche linéaire exacte.

Ils ne sont utilisés que dans des cas particuliers, par exemple lorsque φ est quadratique, la solution de la recherche linéaire s'obtient de façon exacte et dans un nombre fini d'itérations.

Remarque 2.1 *Ces deux règles ne sont utilisées que dans des cas particuliers, par exemple lorsque φ est quadratique.*

2.2 Les inconvénients de la recherche linéaire exacte

Pour une fonction non linéaire arbitraire.

♦ Il peut ne pas exister de pas de *Cauchy* ou de *Curry*.

♦ La détermination de ces pas demande en général beaucoup de temps de calcul et ne peut pas de toute façon être faite avec une précision infinie.

♦ L'efficacité supplémentaire éventuellement apportée à un algorithme par une recherche linéaire exacte ne permet pas en général de compenser le temps perdu à déterminer un tel pas.

♦ Les résultats de convergence autorisent d'autres types de règles (recherche linéaire inexacte) moins gourmandes en temps de calcul.

2.3 Intervalle de sécurité

Dans la plupart des algorithmes d'optimisation modernes on ne fait jamais de recherche linéaire exacte, car trouver λ_k signifie qu'il va falloir calculer un grand nombre de fois la fonction φ et cela peut être dissuasif du point de vue du temps de calcul. En pratique, on recherche plutôt une valeur de $\hat{\lambda}$ qui assure une décroissance suffisante de f. Cela conduit à la notion d'intervalle de sécurité.

Définition 2.1 *On dit que $[a, b]$ est un intervalle de sécurité s'il permet de classer les valeurs de λ 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. L'idée est de partir d'un intervalle suffisamment grand pour contenir $[a, b]$ et d'appliquer un bonne stratégie pour itérativement réduire cet intervalle.

Recherche d'un intervalle de départ

♦ Si λ est satisfaisant alors on s'arrête.

♦ Si λ est trop grand, alors on prend $b = \lambda$ et on s'arrête.

♦ Si λ est trop petit, on fait $c\lambda \longrightarrow \lambda$, $c > 1$ et on retourne en 1.

Cet algorithme donne un intervalle initial $[a, b]$ qu'il va falloir ensuite réduire sauf si λ est admissible auquel cas la recherche linéaire est terminée, ce peut être le cas

si la valeur initiale de λ est bien choisie.

Réduction de l'intervalle

On suppose maintenant que l'on dispose d'un intervalle $[a, b]$ mais que l'on n'a pas encore de λ satisfaisant.

Une manière simple de procéder par exemple par dichotomie, en choisissant

$$\lambda = \frac{a+b}{2},$$

puis en conservant soit $[a, \lambda]$ ou $[\lambda, b]$ suivant que λ est trop grand ou trop petit. Le problème est que cette stratégie ne réduit pas assez rapidement l'intervalle. Cependant elle n'utilise aucune informations sur φ On préfère en général procéder en construisant une approximation polynomiale $p(\lambda)$ de φ et en choisissant λ réalisant le minimum (s'il existe) de $p(\lambda)$ sur $[a, b]$. Lorsque l'on utilise la règle de *Wolfe*, on peut utiliser une approximation cubique.

Algorithme.2.1. (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 $\hat{\lambda} = \lambda_k$ 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 $b_{k+1} = \lambda_k, \quad a_{k+1} = a$, 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 va nous permettre de caractériser les valeurs de λ convenables, ainsi que les techniques utilisées pour réduire l'intervalle.

2.4 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 multidimensionnelles. 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 : $f : \mathbb{R}^n \longrightarrow \mathbb{R}$:

$$\nabla f\left(x_k\right)^T .d < 0. \tag{2.5}$$

Le but est de réduire «de façon importante» la valeur de l'objectif par un pas $x_k \longmapsto x_{k+1} = x_k + \lambda_k d_k$ de x_k dans la direction d_k. Pour cela de nombreux mathématiciens (*Armijo, Goldstein, Wolfe, Albaali, Fletcher...*) ont élaboré plusieurs règles. L'objectif de cette section consiste à présenter les principaux tests.

2.5 Schéma des recherches linéaires inexactes

Elles reviennent à déterminer, par tâtonnement un intervalle $[a, b]$ où $\hat{\lambda} \in [a, b]$ dans lequel

$$\varphi\left(\lambda_k\right) < \varphi\left(0\right)\left(f\left(x_k + \lambda_k d_k\right) < f\left(x_k\right)\right) \tag{2.6}$$

Le schéma de l'algorithme est donc :

Algorithme.2.2. (Schéma 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 ($\hat{\lambda} = \lambda_k$).

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

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

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

2.6 La règle d'Armijo

On bute à réduire « de façon importante » la valeur de l'objectif par un pas $x_k \longmapsto x_{k+1} = x_k + \lambda_k d_k$ de x_k dans la direction d_k, tel que $f(x_k + \lambda_k d_k) < f(x_k)$. Or cette condition de décroissance stricte n'est pas suffisante pour minimiser φ au moins localement. Une condition naturelle est de demander que f décroisse autant qu'une portion $\omega_1 \in]0,1[$ parfois appelée condition d'*Armijo* ou condition de décroissance linéaire

$$f(x_k + \lambda_k d_k) \leq f(x_k) + \omega_1 \lambda_k \nabla f(x_k)^T d_k. \tag{2.7}$$

Elle est de la forme $(f(x_k + \lambda_k d_k) < f(x_k))$ car ω_1 devra être choisi dans $]0,1[$. On voit bien à la figure (2.1) ce que signifie cette condition. Il faut qu'en λ_k, la fonction prenne une valeur plus petite que celle prise par la fonction affine

$$\lambda \longrightarrow f(x_k) + \omega_1 \lambda_k \nabla f(x_k)^T d_k. \tag{2.8}$$

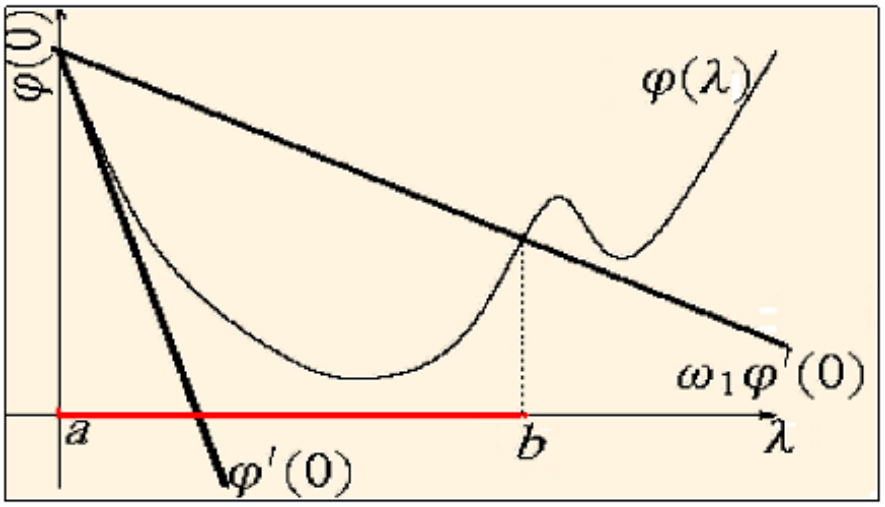

FIG. 2.1 – La règle d'Armijo

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

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

Algorithme 2.3. (Règle d'Armijo)

Etape 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.

Etape 1 :

Si $\varphi(\lambda_k) \leq \varphi(0) + \omega_1 \varphi'(0)\,\lambda_k$: STOP $\left(\hat{\lambda} = \lambda_k\right)$

Si $\varphi(\lambda_k) > \varphi(0) + \omega_1 \varphi'(0)\,\lambda_k$, alors $b_{k+1} = b$, $a_{k+1} = k$ 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 $\lambda_{k+1} \in]a_{k+1}, b_{k+1}[$.
Remplacer k par $k+1$ et aller à l'étape 1.

Remarque 2.2 ♦ *En pratique, la constante ω_1 est prise très petite, de manière a satisfaire* (2.8) *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.*

♦ *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.*

♦ *Il est clair d'après la figure* (2.1) *que l'inégalité* (2.8) *est toujours vérifiée si $\lambda_k > 0$ est suffisamment petit.*

♦ *Il était dangereux d'accepter des pas trop petits, cela pouvait conduire a une fausse convergence. Il faut donc un mécanisme supplémentaire qui empêche le pas d'être trop petit. On utilise souvent la technique de rebroussement due à Armijo en* 1966 *ou celle de Goldstein.*

2.7 La règle de Goldstein et Price

Dans la règle d'*Armijo* on assure la décroissance de la fonction objectif à chaque pas mais c'est ne pas suffisant. Les conditions de *Goldstein* et *Price* suivant sont suffisante pour assurer la convergence sous certaines conditions et indépendamment de l'algorithme qui calcule le paramètre. Dans celle-ci, en ajoutant une deuxième inégalité à la règle d*'Armijo* on obtient la règle de *Goldstein*

$$f(x_k) + \omega_1 \lambda_k \nabla^T f(x_k) d_k \geq f(x_k + \lambda_k d_k) \geq f(x_k) + \omega_2 \lambda_k \nabla^T f(x_k) d_k, \qquad (2.9)$$

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.

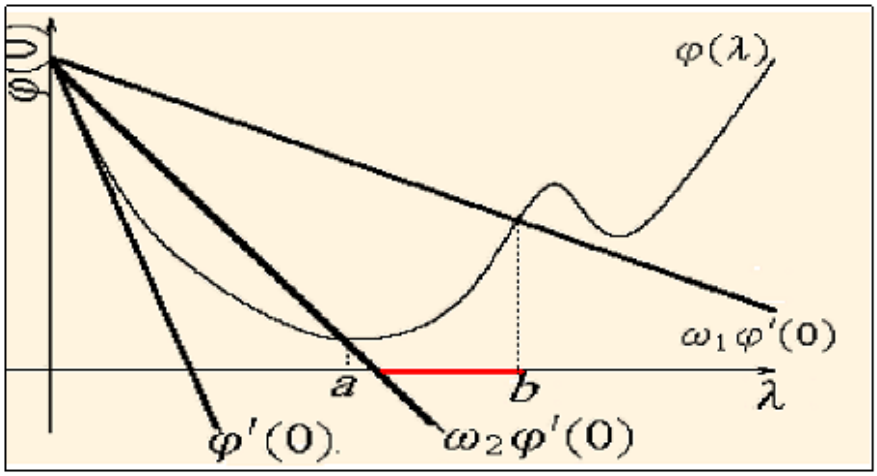

FIG. 2.2 – La règle de Goldestein

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 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 2.4. (Règle de Goldstein et Price)

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 $\varphi(0) + \omega_2\varphi'(0)\lambda \leq \varphi(\lambda_k) \leq \varphi(0) + \omega_1\varphi'(0)\lambda_k$: STOP ($\hat{\lambda} = \lambda_k$).

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.

Etape 2 :

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

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

2.8 La règle de Wolfe

Les conditions de la règle de *Goldstein et Price* peuvent exclure un minimum ce qui est peut être un inconvénient. 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 $\lambda > 0$

$$f(x_k + \lambda d_k) \leq f(x_k) + \omega_1\lambda\nabla f(x_k)^T.d_k, \qquad (E1)$$

$$\nabla f(x_k + \lambda d_k)^T.d_k \geq \omega_2\nabla f(x_k)^T.d_k, \qquad (E2)$$

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

Règle de Wolfe

♦ 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.

On voit bien à la figure (2.3) ce que signifie cette condition

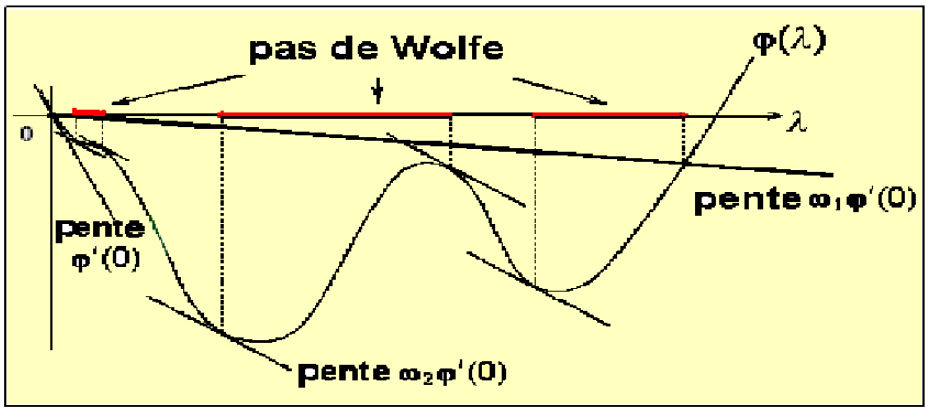

FIG. 2.3 – La règle de Wolfe

Algorithme 2..5. (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 ($\hat{\lambda} = \lambda_k$).

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 $\lambda_{k+1} \in]b_{k+1}, +\infty[$.

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

On obtient des contraintes plus fortes si l'on remplace $(E2)$ par

$$\left|\nabla f(x_k + \lambda d_k)^T . d_k\right| \leq -\omega_2 \nabla f(x)^T . d_k \qquad (E3)$$

Les $(E1)$ et $(E3)$ sont les conditions de *Wolfe* fortes. La contrainte $(E3)$ entraîne que $\omega_2 \varphi(0) \leq \varphi'(\lambda) < -\omega_2 \varphi(0)$ c'est a dire. $\varphi'(\lambda)$ n'est pas "trop" positif.

2.9 La règle de Wolfe relaxée

Proposée par *Dai* et *Yuan* [1996], cette règle consiste à choisir le pas satisfaisant les conditions

$$f(x_k + \lambda d_k) \leq f(x_k) + \omega_1 \lambda \nabla^T f(x_k).d_k, \qquad (E4)$$

$$\omega_2' \nabla^T f(x_k).d_k \leq \nabla^T f(x_k + \lambda_k d_k).d_k \leq -\omega_2'' \nabla^T f(x_k).d_k, \qquad (E5)$$

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

On voit bien que les conditions de *Wolfe* relaxée impliquent les conditions de *Wolfe* fortes. Effectivement $(E4)$ est équivalente à $(E1)$, tandis que pour le cas particulier $\omega_2' = \omega_2'' = \omega_2$, $(E5)$ est équivalente à $(E3)$

$$\begin{aligned} \omega_2' \nabla^T f(x_k).d_k \quad &\leq \quad \nabla^T f(x_k + \lambda_k d_k).d_k \leq -\omega_2'' \nabla^T f(x_k).d_k \\ &\Longrightarrow \quad \omega_2 \nabla^T f(x_k).d_k \leq \nabla^T f(x_k + \lambda d_k).d_k \leq -\omega_2 \nabla^T f(x_k).d_k \\ &\Longrightarrow \quad \left|\nabla^T f(x_k + \lambda d_k).d_k\right| \leq -\omega_2 \nabla^T f(x).d_k \qquad (E\ 3) \end{aligned}$$

CHAPITRE 3

Convergence des méthodes à directions de descente

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.

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

$$f(x_{k+1}) \leq f(x_k) - C \left\|\nabla f(x_k)\right\|^2 \cos^2 \theta_k, \tag{3.1}$$

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

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

Voici comment on se sert de la condition de *Zoutendijk*.

Proposition 3.1. *Si la suite $\{x_k\}$ générée par un algorithme d'optimisation vérifie la condition de Zoutendijk* (3.1) *et si la suite $\{f(x_k)\}$ est minorée, alors*

$$\sum_{k \geq 1} \left\|\nabla f(x_k)\right\|^2 \cos^2 \theta_k < \infty. \tag{3.2}$$

Preuve :

En sommant les inégalités (3.1), on a

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

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

Il ya des propositions précisent les circonstances dans les quelles la condition de *Zoutendijk* (3.1) est vérifie avec les règles de la recherche linéaire exacte (*Cauchy*, *Curry*) et aussi les règles de la recherche linéaire inexacte (*Armijo*, *Wolfe*).

Ce qui concernant la proposition qui est vérifie la condition de *Zoutendijk* avec la règle de *Wolfe*.

Proposition 3.2. *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction continument 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 (E1)-(E2).

Alors il existe une constante $C > 0$ telle que, pour tout $k \geq 1$, la condition de Zoutendijk (3.1) *est vérifiée.*

Preuve :

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

D'après $(E2)$

$$g_{k+1}^T d_k \geq \omega_2 g_k^T d_k$$

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

et du fait que f est continument différentiable :

$$\begin{aligned}(1-\omega_2)\left|g_k^T d_k\right| &= (1-\omega_2)\left\|g_k\right\|\left\|d_k\right\|\cos\theta_k \\ &\leq \left\|g_{k+1}-g_k\right\|\left\|d_k\right\| \\ &\Rightarrow \quad (1-\omega_2)\left\|g_k\right\|\cos\theta_k \leq L\lambda_k\left\|d_k\right\| \\ &\Rightarrow \quad \lambda_k\left\|d_k\right\| \leq \frac{(1-\omega_2)}{L}\left\|g_k\right\|\cos\theta_k,\end{aligned}$$

en utilisant $(E1)$, on aura :

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

On en déduit (3.1).

2 Méthodes itératives d'optimisation sans contraintes

A travers de ce chapitre et du suivant, on s'attachera dorénavant à la description plus spécifique des algorithmes itératifs (ou méthodes itératives) qui ont été implémentés et qui permettent la résolution des problèmes d'optimisation non linéaires. Il convient de souligner que la plupart des algorithmes d'optimisation contrainte ou non fonctionnent selon un schéma général consistant à chaque itération 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 (appelées aussi parfois méthodes d'optimisation directe), soit le problème

$$(P) \quad \text{minimiser} \quad f(x) \quad x \in \mathbb{R}^n,$$

pour lequel nous commencerons par décrire les méthodes suivantes :

- ▶ Les méthodes du gradient
- ▶ Les méthodes du Newton
- ▶ Les méthodes utilisant des directions conjuguées.

Ces méthodes utilisent des dérivées (et donc la propriété de différentiabilitè de f) à l'exception des méthodes de directions conjuguées (sauf dans le cas particulier de la méthode du gradient conjugué) basée sur des propriétés plus géométriques.

Parmi les plus anciennes méthodes utilisées pour résoudre les problèmes du type (P), on peut citer la méthode du gradient conjugué. Cette méthode est surtout utilisée pour les problèmes de grande taille.

Après les avoir décrites, nous donnons la définition d'une forme quadratique et les principes des méthodes de descente et du gradient.

Définition 3.1. *Soit H une matrice symétrique $n \times n$ et $b \in \mathbb{R}^n$ On appelle forme quadratique la fonction $f : \mathbb{R}^n \to \mathbb{R}$ définie par*

$$f(x) = \frac{1}{2} x^T H(x) x - b^T x. \tag{3.3}$$

Lorsque la matrice H est définie positive (resp. semi-définie positive), on dira que $f(x)$ est une forme quadratique définie positive (resp. semi-définie positive).

2.1 Principe des méthodes de descente

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

$$x_{k+1} = x_k + \lambda_k d_k \qquad \lambda_k > 0,$$

tout en assurant la propriété

$$f(x_{k+1}) < f(x_k). \tag{3.4}$$

Le vecteur d_k est la direction de descente en x_k. Le scalaire λ_k est appelé le pas de la méthode à l'itération k. On peut caractériser les directions de descente en x_k à l'aide du gradient.

2.2 Principe des méthodes du gradient (méthodes de la plus forte pente)

On cherche à déterminer la direction de descente qui fait décroitre $\phi(\lambda) = f(x + \lambda d)$ le plus vite possible (au moins localement). Pour cela on va essayer de minimiser la dérivée de $\phi(\lambda)$ 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 façon d'utiliser cette direction de descente. On peut par exemple utiliser un pas fixé a priori $\lambda_K = \alpha > 0, \forall k$. On obtient alors la méthode du gradient simple :

$$\begin{cases} d_k = -\nabla f(x_k), \\ x_{k+1} = x_k + \lambda d_k. \end{cases} \tag{3.5}$$

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

$$\begin{cases} d_k = -\nabla f(x_k), \\ x_{k+1} = x_k + \lambda_k d_k \end{cases} \tag{3.6}$$

Où λ_k est choisi de manière à ce que

$$f(x_k + \lambda_k d_k) \leq f(x_k + \lambda d_k), \forall \lambda > 0. \tag{3.7}$$

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

Proposition 3.3. *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction différentiable. Les directions de*

descente d_k générées par la méthode (3.6) et (3.7) vérifient

$$d_{k+1}^T d_k = 0. \tag{3.8}$$

Preuve : Si on introduit la fonction

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

on a

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

et puisque ϕ est dérivable on a nécessairement

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

donc

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

Exemple 3.1.

▶ **Calcul du pas optimal dans le cas quadratique :**

On a $f(x) = \frac{1}{2} x^T H x - b^T x$ avec $H > 0$ et on note $\phi(\lambda) = f(x_k + \lambda d_k)$. Le pas optimal λ_k est caractérisé par

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

on a donc

$$\nabla f(x_k + \lambda_k d_k)^T d_k = (H(x_k + \lambda_k d_k) - b)^T d_k = 0,$$

soit

$$(\nabla f(x_k) + \lambda_k H d_k)^T d_k = 0,$$

on obtient donc

$$\lambda_k = -\frac{\nabla f(x_k)^T d_k}{d_k^T H d_k}, \tag{3.9}$$

qui est bien positif car d_k est une direction de descente et

$$d_k^T H d_k > 0 (car\ H > 0).$$

La méthode du gradient à pas optimal peut donc s'écrire (dans le cas quadratique)

$$\begin{aligned} d_k &= b - Hx_k, \\ \lambda_k &= \frac{d_k^T d_k}{d_k^T H d_k}, \\ x_{k+1} &= x_k + \lambda_k d_k. \end{aligned}$$

Exemple 3.2.

▶ **Méthode du gradient à pas optimal dans le cas quadratique**

Dans le cas où $f(x) = \frac{1}{2}x^T H(x)x - b^T x$ la méthode du gradient simple peut s'écrire

$$\begin{aligned} d_k &= b - Hx_k, \\ x_{k+1} &= x_k + \alpha d_k, \end{aligned}$$

où $\alpha > 0$ est fixé a priori. Il existe bien sûr des conditions sur pour que la méthode converge. Nous illustrons ici le fonctionnement de la méthode dans le cas $n = 2$ sur une petite simulation.

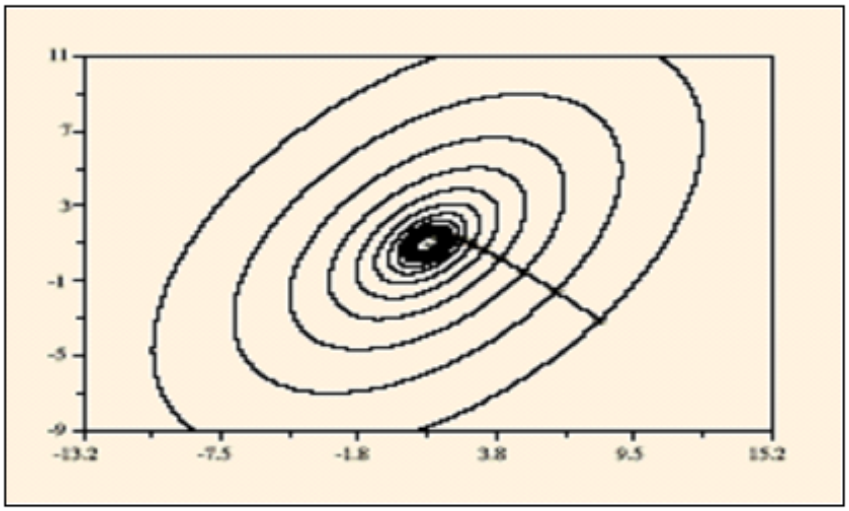

FIG. 3.1 – Illustration de la convergence plus rapide de la méthode du gradient

Algorithme de la méthode de la plus forte pente

Etape initiale :

Choisir un $\varepsilon > 0$.

Choisir un point initiale x_1.

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

Etape principale :

Si $\|\nabla f(x)\| < \varepsilon$ stop.

Sinon poser $d_k = -\nabla f(x_k)$ et soit λ_k la solution optimale de la recherche linéaire

$$\min \{f(x_k + \lambda d_k); \lambda \geq 0\}.$$

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

Remplacer k par $k + 1$et répéter l'étape principale.

Inconvénients de la méthode de la plus forte pente

▶ Lenteur de la méthode au voisinage des points stationnaires.

▶ Cette méthode travaille de façon performante dans les premières étapes de l'algorithme. Malheureusement, dès qu'on s'approche du point stationnaire, La méthode devient très lente. On peut expliquer intuitivement ce phénomène par les considérations suivantes

$$f(x_k + \lambda d_k) = f(x_k) + \lambda \nabla f(x_k)^T d + \lambda \|d\| \alpha(x_k; \lambda d),$$

où $\alpha(x_k; \lambda d) \to 0$ quand $\lambda d \to 0$.

Si $d_k = -\nabla f(x_k)$, on obtient : $x_{k+1} = x_k - \lambda \nabla f(x_k)$ et par conséquent

$$x_{k+1} - x_k = \lambda \left[- \|\nabla f(x_k)\|^2 + \|\nabla f(x_k)\| \alpha(x_k; \lambda \nabla f(x_k)) \right].$$

D'après l'expression précédente, on voit que lorsque x_k s'approche d'un point stationnaire et si f est continument différentiable alors $\|\nabla f(x_k)\|$ est proche de zéro. Donc le terme à droite s'approche de zéro, indépendamment de λ, et par conséquent $f(x_{k+1})$ ne s'éloigne pas beaucoup de $f(x_k)$ quand on passe du point x_k au point x_{k+1}.

2.3 Les méthodes utilisant des directions conjuguées

Description de la méthode

Ces méthode sont basées sur l'important concept de la conjugaison et ont été développées afin de résoudre le problème quadratique

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(x) = \frac{1}{2} x^T . Q . x + b^T . x + c.$$

Où $Q \in \mathbb{R}^{n \times n}$ est symétrique et définie positive, $b \in \mathbb{R}^n$et $c \in \mathbb{R}$. Les méthodes de direction conjuguées peuvent résoudre les problèmes de cette forme en au plus n itérations et contrairement aux méthodes présentées jusqu'à présent, elles n'utilisent pas des dérivées sauf dans le cas particulier de la méthode du gradient conjugué. Donnons la définition de la notion de "conjugaison".

Définition 3.2. *Soient Q une matrice $n \times n$ symétrique et définie positive et*

un ensemble de directions non nulles $\{d_1, d_2, ..., d_k\}$. *Ces directions sont dites Q-conjuguées si*

$$d_i^T Q d_j = 0, \qquad \forall i; j \text{ tels que } i \neq j. \tag{3.10}$$

Propriété 3.1. Si $d_1, ..., d_k$ sont Q-conjuguées, alors elles sont linéairement indépendantes.

Propriété 3.2. Comme des directions Q-conjuguées sont linéairement indépendantes, alors l'espace vectoriel engendré par un ensemble de n directions Q-conjuguées est de dimension n.

Etant donné un ensemble de n directions Q-conjuguées $d_0, d_1, ..., d_{n-1}$, la méthode de directions conjuguées correspondante est donnée par

$$x_{k+1} = x_k + \lambda_k d_k, \qquad k = 0, ..., n-1,$$

où x_0 est un vecteur de départ choisi arbitrairement et où les λ_k sont obtenus par minimisation monodimensionnelle le long de d_k.

Le principal résultat concernant les méthodes utilisant des directions conjuguées est qu'à chaque itération k, la méthode minimise f sur le sous-espace généré par les k premières directions Q-conjuguées utilisées par l'algorithme. A la nième itération au plus tard, ce sous-espace inclura alors le minimum global de f, grâce à la propriété d'indépendance linéaire des directions Q-conjuguées qui assure qu'à l'itération n, l'espace vectoriel généré par les n directions Q-conjuguées ne sera autre que $\mathbb{R}^n$.

Remarque 3.1.*la notion de conjugaison n'a pas de sens dans le cas non quadratique*

La méthode du gradient conjugué non linéaire

L'idée de la méthode est de construire itérativement des directions $d_0, ..., d_k$ 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}, les coefficients étant choisis de telle manière que d_k soit conjuguée avec toutes les directions précédentes. On s'intéresse ici à la minimisation d'une fonction $f : \mathbb{R}^n \to \mathbb{R}$, non nécessairement

quadratique :

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

et on cherche à étendre la méthode du gradient conjugué a ce problème. Il y a plusieurs manières de le faire et peu de critères permettant de dire laquelle est la meilleure. Une extension possible consiste simplement a reprendre les formules utilisées dans le cas quadratique. On se propose donc d'étudier les méthodes où la direction d_k est définie par la formule de récurrence suivante :

$$d_k = \begin{cases} -g_1 & \text{si } k = 1, \\ -g_k + \beta_k d_{k-1} \ldots\ldots\ldots & \text{si } k \geq 2, \end{cases} \tag{3.11}$$

où $g_k = \nabla f(x_k)$, $\beta_k \in \mathbb{R}$ et $\{x_k\}$ est générée par la formule :

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

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

Ces méthodes sont des extensions de la méthode du gradient conjuguée si k prend l'une des valeurs

$$\beta_k^{PRP} = \frac{g_k^T y_k}{\|g_{k-1}\|^2} \quad \text{Gradient conjugué variante Polak-Ribière-Polyak} \tag{3.12}$$

$$\beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2} \quad \text{Gradient conjugué variante Fletcher-Reeves} \tag{3.13}$$

$$\beta_k^{HS} = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}} \quad \text{Gradient conjugué variante Hestenes-Stiefel} \tag{3.14}$$

$$\beta_k^{CD} = \frac{\|g_k\|^2}{-d_{k-1}^T g_{k-1}} \quad \text{Gradient conjugué variante descente conjugué} \tag{3.15}$$

$$\beta_k^{DY} = \frac{\|g_k\|^2}{d_{k-1}^T y_{k-1}} \quad \text{Gradient conjugué variante de Dai-Yuan,} \tag{3.16}$$

où $y_{k-1} = g_k - g_{k-1}$.

Pour faciliter la présentation nous appelons les méthodes qui correspondent à (3.12)-(3.16) la méthode PRP, la méthode FR, la méthode HS, la méthode CD et la méthode DY, respectivement.

Dans le cas où f est une fonction quadratique strictement convexe avec une recherche linéaire exacte toutes ces variantes de β_k ont la même valeur.

$$\beta_k^{PRP} = \beta_k^{FR} = \beta_k^{CD} = \beta_k^{DY}.$$

Si f est quelconque, Ils n'ont pas la même valeur pour les méthodes de *Polak-Ribière-Ployak* ([19,1969])-([20,1969]) méthode de *Fletcher-Reeves* ([12,1964]), méthode de la *descente conjuguée* ([11,1987]) et la méthode de *Dai-Yuan* ([4,1999]) selon que l'on utilise $\beta_k^{PRP}, \beta_k^{FR}, \beta_k^{CD}$ ou β_k^{DY} à la place de β_k dans (3.11).

Pour que les méthodes ainsi définies soient utilisables, il faut répondre aux deux questions suivantes.

▶ Les directions d_k définies par (3.11) sont-elles des directions de descente de f ?

▶ Les méthodes ainsi définies sont-elles convergentes ?

En ce qui concerne la première question remarquons que, quel que soit $\beta_k \in \mathbb{R}$, d_k est une direction de descente si on fait de la recherche linéaire exacte, c'est a dire si le pas λ_{k-1} est un point stationnaire de $\lambda \to f(x_{k-1} + \lambda d_{k-1})$. En effet, dans ce cas $g_k^T d_{k-1} = 0$ et on trouve lorsque $g_k \neq 0$:

$$\begin{aligned} d_k^T g_k &= \left(-g_k + \beta_k d_{k-1}\right)^T g_k \\ &= -\left\|g_k\right\|^2 + \beta_k d_{k-1}^T g_k = -\left\|g_k\right\|^2 < 0. \end{aligned}$$

Cependant, il est fortement déconseillé de faire de la recherche linéaire exacte lorsque f n'est pas quadratique car le coût de détermination de k est excessif.

L'efficacité de la méthode du gradient conjugué repose essentiellement sur deux points :

▶ La recherche linéaire détermination du pas optimal doit être exacte.

▶ Les relations de conjugaison doivent être précises.

La recherche du pas optimal doit être réalisée à l'aide d'un algorithme spécifique, puisque f est quelconque. Par contre la notion de conjugaison n'a pas de sens dans le cas non quadratique.

L'étude des propriétés de convergence de quelques méthodes du gradient conjugué non linéaire est l'objectif du quatrième chapitre.

CHAPITRE 4

Synthèse sur la convergence des quelques méthodes du gradient conjugué non linéaire avec des recherches linéaires inexactes

On expose dans ce chapitre une synthèse concernant la convergence des différentes variantes du gradient conjugué avec la recherche linéaire inexacte de *Wolfe* forte et de *Wolfe* faible.

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

$$\min \{f(x) : x \in \mathbb{R}^n\}, \tag{4.1}$$

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 variantes du gradient conjugué génèrent une suite $\{x_k\}$ de la forme suivante

$$x_{k+1} = x_k + \lambda_k d_k, \tag{4.2}$$

la direction d_k est définie par la formule de récurrence suivante ($\beta_k \in \mathbb{R}$)

$$d_k = \begin{cases} -g_1 & \text{si } k = 1, \\ -g_k + \beta_k d_{k-1} \text{.....si } k \geq 2, \end{cases} \tag{4.3}$$

le pas $\lambda_k \in \mathbb{R}$ étant déterminé par une recherche linéaire. Rappelons les différentes variantes du gradient conjugué qui différent celons les valeurs que prennent les constantes β_k. On y rencontre particulièrement les variantes suivantes :

$$\beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2}, \qquad (Fletcher - Reeves) \tag{4.4}$$

$$\beta_k^{PRP} = \frac{g_k^T(g_k - g_{k-1})}{\|g_{k-1}\|^2} \qquad (Polak - Ribiére) \tag{4.5}$$

$$\beta_k^{HS} = \frac{g_k^T(g_k - g_{k-1})}{d_{k-1}^T(g_k - g_{k-1})}, \qquad (Hestenes - Stiefel) \tag{4.6}$$

$$\beta_k^{CD} = \frac{\|g_k\|^2}{-d_{k-1}^T g_{k-1}}, \qquad (méthode\ de\ la\ descente\ conjuguée) \tag{4.7}$$

où $\|.\|$ est la norme euclidienne. Récemment Dai et yuan ont aussi introduit la forme suivante :

$$\beta_k^{DY} = \frac{\|g_k\|^2}{d_{k-1}^T(g_k - g_{k-1})}. \tag{4.8}$$

La convergence globale des méthodes (4.4), (4.5), (4.6), (4.7) et (4.8) a été étudiés par beaucoup d'auteurs. Citons surtout *Al-Baali* [([1,1985]).], *Gilbert et Nocedal* ([13,1992]), *Hestenes et Stiefel* ([14,1952]), *Hu et Story* ([15,1991]) et *Zoutendijk* ([25,1970]).Un facteur clé dans l'étude de la convergence globale est comment sélectionner le paramètre λ_k. Le choix le plus naturel de λ_k est de faire une recherche linéaire exacte, cest.à dire poser $\lambda_k = \arg\min_{\lambda \geq 0} f(x_k + \lambda d_k)$. Cependant, quelque peu étonnamment ce choix naturel pourrait résulter en une séquence non-convergente dans le cas de la méthode PR et la méthode HS, comme montré par *Powell* ([37,1986]). Inspiré par le travail de *Powell*, *Gilbert et Nocedal* ([13,1992]) a mené une analyse élégante et a montré que la méthode PR est globalement convergente si β_k^{PR} est restreint pour être non-négatif et λ_k est déterminé en un pas de la

recherche linéaire satisfaisant la condition suffisante de descente

$$g_k^T d_k < -c \|g_k\|^2, \quad \text{où } c > 0, \tag{4.9}$$

en plus des conditions standard de *Wolfe* :

$$\begin{cases} f(x_k + \lambda_k d_k) \leq f(x_k) + \omega_1 \lambda_k d_k^T g_k, \\ d_k^T g_{k+1} \geq \omega_2 d_k^T g_k. \end{cases} \tag{4.10}$$

Où $0 < \omega_1 < \omega_2 < 1$. Récemment *Dai et Yuan* [4,6] ont montré que la méthode CD et la méthode FR sont globalement convergentes si les conditions de la recherche de linéaire suivantes pour λ_k sont satisfaites

$$\begin{cases} f(x_k + \lambda_k d_k) \leq f(x_k) + \omega_1 \lambda_k d_k^T g_k, \\ \omega_2' d_k^T g_k \leq d_k^T g_{k+1} \leq -\omega_2'' d_k^T g_k. \end{cases} \tag{4.11}$$

Où $0 < \omega_1 < \omega_2' < 1$ et $\omega_2'' > 0$ pour la méthode CD et de plus $\omega_2' + \omega"_2 \leq 1$ pour la méthode FR.

Par commodité, nous assumons que $g_k \neq 0$ pour tout k.

Nous adoptons la Supposition suivante sur fonction f qui est utilisée communément dans la littérature.

Supposition 4.1.

(i) L'ensemble $\pounds := \{x \in \mathbb{R}^n; f(x) \leq f(x_1)\}$ *est borné ; où* $x_0 \in \mathbb{R}^n$ *est le point initial.*

(ii) Sur un voisinage N *de* $\pounds$*, la fonction objectif* f *est continûment différentiable et son gradient est lipschitzien i.e*

$$\exists L > 0 \text{ tel que} \quad \|g(x) - g(\tilde{x})\| \leq L \|x - \tilde{x}\|, \forall x, \tilde{x} \in N. \tag{4.12}$$

Remarque 4.1 *Ces suppositions impliquent qu'il existe* $\gamma > 0$ *tel que*

$$\|g(x)\| \leq \gamma, \forall x \in \pounds. \tag{4.13}$$

Rappelons maintenant le Théorème suivant obtenu essentiellement par *Zoutendijk* [25] et *Wolfe* [24]. Ce théorème assure la satisfaction de la condition de *Zoutendijk* (voir chapitre3), pour toute méthode du type (4.2), (4.3), dans laquelle le pas λ_k est déterminé par la règle de *Wolfe* faible (4.10).

Théorème 4.1. [25]

Considérons la suite $(x_k)_k$, définie par (4.2),où d_k est une direction de descente et λ_k vérifie les conditions (4.10).Considérons aussi que la supposition (4.1) soit satisfaite, alors :

$$\sum_{k\geq 1} \cos^2\theta \, \|g_k\|^2 < \infty, \tag{4.14}$$

est vérifiée, avec

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

Remarque 4.2.

$$(4.14) \Longleftrightarrow \sum_{k\geq 1, \|d_k\|\neq 0} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty. \tag{4.15}$$

En effet :

D'après la $2^{ème}$ condition de (4.10) nous avant

$$d_k^T \left(g_{k+1} - g_k\right) \geq (\omega_2 - 1)\, d_k^T g_k.$$

D'autre part

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

d'où

$$\lambda_k \geq \left(\frac{\omega_2 - 1}{L}\right) \frac{d_k^T g_k}{\|d_k\|^2}.$$

En remplaçant ceci dans la $1^{ère}$ condition de on aura :

$$f(x_k + \lambda_k d_k) \leq f(x_k) + \mu \cos^2 \theta_k \|g_k\|^2,$$

où $\mu = \dfrac{\omega_1(\omega_2 - 1)}{L}$.

Or puisque f est bornée sur N on a :

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

Ce qui achève la démonstration. □

1 Résultats de convergence du gradient conjugué, version Fletcher-Reves

Avant d'exposer sans démonstration les différents résultats de convergence, donnons d'abord l'algorithme de la méthode de Fletcher Reeves.

1.1 Algorithme 4.1 de la méthode de Fletcher-Reeves

Etape0 : (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 : $\lambda_k = \min_{\alpha>0} f(x_k + \lambda d_k)$

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

Où

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

Le premier résultat de convergence de la méthode du gradient conjugué non linéaire (version *Flecher Reeves*) avec la recherche linéaire inexacte de *Wolfe* forte

$$\begin{cases} f(x_k + \lambda_k d_k) \leq f(x_k) + \omega_1 \lambda_k d_k^T g_k, \\ \left| d_k^T g_{k+1} \right| \leq -\omega_2 d_k^T g_k. \end{cases} \tag{4.16}$$

où $(\omega_2 < \frac{1}{2})$ a été démontré par *Al-Baali* ([3,1985]).

• *Touati Ahmed* et *Story* ([23,1990]) ont généralisé ce résultat pour

$$0 \leq \beta_k \leq \beta_k^{FR}. \tag{4.17}$$

• *Gilbert* et *Nocedal* ([13,1992]) ont généralisé ce résultat pour

$$|\beta_k| \leq \beta_k^{FR}. \tag{4.18}$$

• Récemment, *Dai et Yuan* ([4,1996]) a montré que la méthode FR es globalement convergente avec la recherche linéaire inexacte de *Wolfe relaxée.*

On donne dans ce paragraphe les principaux résultats de convergence de la méthode du gradient conjugué non linéaire avec la recherche linéaire inexacte de *Wolfe* forte et aussi avec la recherche linéaire inexacte de *Wolfe relaxée.*

Proposition 4.1. [**13**]

On considère que la supposition (4.1) est satisfaite. Considérons une méthode du type (4.2) et (4.3) avec β_k satisfaisant à (4.18) et le pas λ_k vérifiant la règle de Wolfe forte (4.16) où $\omega_2 \in \left]0, \frac{1}{2}\right[$.

Alors cette méthode génère des directions de descente. De plus on a

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

Théorème 4.2. [13] *On considère que la supposition (4.1) est satisfaite. Toute méthode du type (4.2) et (4.3) dans laquelle β_k vérifié*

$$|\beta_k| \leq \beta_k^{FR}, \quad \forall k \geq 1,$$

et le pas λ_k est déterminé par la règle de Wolfe forte (4.16) avec $0 < \omega_1 < \omega_2 < \frac{1}{2}$ est une méthode de descente ($g_k^t d_k < 0$, $\forall k \geq 1$) convergente, dans le sens où

$$\lim_{k\to\infty} \inf \|g_k\| = 0. \tag{4.19}$$

Théorème 4.3. [13] *Soit x_1 un point de départ pour lequel la supposition (4.1) soit satisfaite. Considérons une méthode du type (4.2) et (4.3) avec β_k satisfaisant à (4.4). On suppose aussi que*

♦ *Chaque direction d_k* vérifié

$$g_k^T d_k \leq 0. \tag{4.20}$$

♦ *Le pas λ_k est déterminé par la règle de Wolfe relaxée (4.11) avec $\omega"_2 < \infty$, alors :*

$$\lim_{k\to\infty} \inf \|g_k\| = 0.$$

Remarque 4.3.*D'un point de vue théorique, la méthode de Fletcher-Reeves est une bonne méthode car :*

► *Elle ne nécessite pas de stocker une matrice.*

► *Elle converge pour une classe assez large des fonctions f et sa vitesse de convergence est meilleure que celle de la méthode du gradient.*

Cependant, en pratique il est préférable d'utiliser la méthode de *Polak-Ribière* dont les performances moyennes dépassent de loin celles de la méthode de *Fletcher-Reeves.*

2 Résultats de convergence du gradient conjugué, version Polak Ribière

La variante dite de *Polak-Ribière* consiste à définir β_k par la formule (4.5). Cette méthode fut découverte par *Polak, Ribière* ([19, 1969]) et *Polyak* ([20, 1969]). L'algorithme cette méthode est le suivant :

2.1 Algorithme 4.2 de la méthode de Polak-Ribière-Polyak

Etape0 : (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 :

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

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

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.

Le résultat suivant est du à *Polak* et *Ribière* ([19,1969]). Il établit la convergence de cette méthode pour une fonction fortement convexe avec une recherche linéaire exacte.

Proposition 4.2. [**19**] *Si f est fortement convexe, de classe C^1 avec un gradient lipschitzien, alors la méthode de Polak-Ribiére avec recherche linéaire exacte génère une suite $\{x_k\}$ convergeant vers l'unique point x_* réalisant le minimum de f.*

Remarque 4.4. *Si f n'est pas convexe, la méthode de Polak-Ribière-Polyak peut ne pas converger.*

3 Résultats de convergence du gradient conjugué, version Dai et Yuan

Récemment *Dai* et *yuan* ([3,1998]) ont introduit une formule pour β_k sous la forme suivante

$$\beta_k^{DY} = \frac{\|g_k\|^2}{d_{k-1}^T y_{k-1}}, \tag{4.21}$$

où $y_{k-1} = g_k - g_{k-1}$.

Cette méthode possède plusieurs propriétés, par exemple elle possède la propriété de descente à chaque itération et la convergence si le pas est déterminé par les règles de (*Wolfe faible, Armijo et Goldstein) et* si f est strictement convexe ou régulière.

3.1 Descente de la méthode de Dai-Yuan

Théorème 3.3. ([9, 19)]) *Soit x_1 un point de départ pour lequel la supposition (4.1) est satisfaite. Considérons une méthode du type (4.2) et (4.3) avec β_k satisfaisant à (4.21).*

♦ *Si f est strictement convexe sur l'ensemble convexe £ cest à dire*

$$(g(x) - g(y))^T (x - y) > 0, \quad \textit{pour tout } x, y \in £. \tag{4.22}$$

Alors pour tout $k \geq 1$:

$$g_k^T d_k < 0. \tag{4.23}$$

En 1999 ils se sont généralisé ce résultat pour toute fonction régulière avec la recherche de *Wolfe* faible (4.10).

Théorème 4.4. ([3, 1999]) *Supposons que Supposition (4.1) soit satisfaite. Pour toute méthode du type (4.2) et (4.3) dont β_k satisfait à (4.21) et le pas λ_k satisfait les conditions de Wolfe faible (4.10), alors toutes les directions générées sont de descente, autrement dit*

$$d_k^T g_k < 0; \quad \forall k \geq 1$$

3.2 Convergence de la méthode de Dai-Yuan

Les résultats suivant sont dus à *Dai* et *Yuan* ([9,1998]). Ils assurent la convergence de cette méthode pour une fonction strictement convexe avec une recherche linéaire inexacte d'*Armijo* et *Goldstein,* (voir théorème (4.5) et (4.6)).

Théorème 4.5. [9] *Soit x_1 un point de départ pour lequel la supposition (4.1) soit satisfaite. Considérons une méthode du type (4.2) et (4.3) avec β_k satisfaisant à (4.21).*

♦ *Si f est strictement convexe sur l'ensemble convexe $\pounds$ et le pas λ_k satisfait les conditions de Goldstein*

$$\omega_2\lambda_k\nabla^T f(x_k)d_k \leq f\left(x_k+\lambda_k d_k\right)-f(x_k) \leq \omega_1\lambda_k\nabla^T f(x_k)d_k, \tag{4.24}$$

où $g_k=\nabla f(x_k)$ et ω_1 et ω_2 sont deux constantes vérifiant

$$0<\omega_1<1/2<\omega_2<1..$$

Alors

$$\lim_{k\to\infty}\inf\|g_k\|=0.$$

Théorème 4.6. [9] *Soit x_1 un point de départ pour lequel la supposition (4.1) est satisfaite. Considérons une méthode du type (4.2) et (4.3) avec β_k satisfaisant à (4.21).*

♦ *Si f est uniformément convexe sur l'ensemble convexe $\pounds$ c a d s'il existe une constante $\eta>0$ tel que*

$$(g(x)-g(y))^T(x-y)\geq\eta\left\|x-y\right\|^2, \quad \textit{pour tout } x,y\in\pounds, \tag{4.25}$$

et le pas λ_k satisfait les conditions d'Armijo

$$f\left(x_k+\lambda_k d_k\right)-f(x_k)\leq\omega\lambda_k\nabla^T f(x_k)d_k, \tag{4.26}$$

où $g_k=\nabla f(x_k)$ et $0<\omega<1$, pour lequel la condition suffisante de descente (4.9)

est satisfaite. Alors

$$\lim_{k\to\infty} \|g_k\| = 0.$$

3.3 Algorithme 4.3 de la Méthode de Dai-Yuan avec la règle de *Wolfe* faible

Etape0 : (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 vérifie les conditions (4.10) et

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

Oú

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

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

Théorème 4.7. ([9, 1999]) *Soit x_1 un point de départ pour lequel la proposition (4.1) est satisfaite.*

La suite $\{x_k\}$ générée par l'algorithme 4.3 converge dans le sens où

$$\lim_{k\to\infty} \inf \|g_k\| = 0.$$

CHAPITRE 5

La convergence globale d'une classe du gradient conjugué

1 Préliminaires, hypothèses et lemmes intermédiaires

1.1 Position du problème

Soit $f : IR^n \to IR$ et (P) le problème de minimisation non linéaire sans contraintes suivant :

$$\min \{f(x) : x \in IR^n\} \qquad (P) \tag{5.1}$$

Comme on l'a vu avant les différentes méthodes du gradient conjugué génèrent des suites $\{x_k\}_{k\in\mathbb{N}}$ de la façon suivante

$$x_{k+1} = x_k + \lambda_k d_k \tag{5.2}$$

Le pas $\lambda_k \in \mathbb{R}$ est déterminé par une optimisation unidimensionnelle ou recherche linéaire exacte ou inexacte.

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

$$d_k = \begin{cases} -g_1 & \text{si } k = 1 \\ -g_k + \beta_k d_{k-1} & \text{si } k \geq 2 \end{cases} \tag{5.3}$$

$g_k = \nabla f(x_k)$ 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

$$\beta_k^{PRP} = \frac{g_k^T y_k}{\|g_{k-1}\|^2} \qquad \text{Gradient conjugué variante Polak-Ribière-Polyak} \tag{5.4}$$

$$\beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2} \qquad \text{Gradient conjugué variante Fletcher Reeves} \tag{5.5}$$

$$\beta_k^{CD} = \frac{\|g_k\|^2}{-d_{k-1}^T g_{k-1}} \qquad \text{Gradient conjugué variante descente conjugué} \tag{5.6}$$

$$\beta_k^{DY} = \frac{\|g_k\|^2}{d_{k-1}^T y_{k-1}} \qquad \text{Gradient conjugué variante de Dai-Yuan} \tag{5.7}$$

Bien que toutes ces méthodes ont été très importantes pour la résolution du problème (P), ses étude s'est faite de façon individuelle et chacune à part, par *Flecher, Powell, Wolf, El Baali, Polak-Ribière, Polyak, Y. H. Dai et Y. Yuan*,.... Il est alors légitime de se poser le problème suivant :

Peut-on définir et unifier toutes ces méthodes dans une même classe et étudier par la suite ses propriétés et sa convergence comme il a été fait avec les méthodes quasi-Newtoniennes ?

Y. Dai et Y. Yuan ont répondu dans ([7,2003]) positivement à cette question. Le but principal de ce chapitre est d'étudier de façon approfondie le travail de *Dai et Yuan*

La forme unifiée du gradient conjugué en question est de la forme (0.2) et (0.3)

où β_k est la forme générale suivante

$$\beta_k = \frac{\phi_k}{\phi_{k-1}} \tag{5.8}$$

ϕ_k pouvant être de la forme suivante (ou d'autres)

$$\phi_k = \lambda \|g_k\|^2 + (1-\lambda)(-g_k^t d_k), \qquad \lambda \in [0,1]. \tag{5.9}$$

On montre que les différentes valeurs de β_k, c'est-à-dire $\beta_k^{FR},\ \beta_k^{DY}, ...$sont des valeurs particulières de $\beta_k = \frac{\phi_k}{\phi_{k-1}}$.

Hypothèses

Dans tout ce chapitre nous supposerons que $g_k = \nabla f(x_k) \neq 0,\ \forall k,$ et que les deux hypothèses suivantes sont satisfaites :

Hypothèse 1 :

a) *f est bornée inférieurement dans l'ensemble* $\mathcal{L}=\{x \in \mathbb{R}^n : f(x) \leq f(x_1)\}$

b) *f est différentiable dans un voisinage V de $\mathcal{L}$ et son gradient g est continument lipschitzien, c'est-à-dire qu'il existe $L > 0$ tel que :*

$$\|g(x) - g(\widetilde{x})\| \leq L \|x - (\widetilde{x})\|, \quad \textit{pour tout } x,\ \widetilde{x} \quad \in V$$

Hypothèse 2 :

l'ensemble $\mathcal{L}=\{x \in \mathbb{R}^n : f(x) \leq f(x_1)\}$ *est borné.*

Remarque 5.1 :

Si f satisfait à Hyothèse1 et Hypothèse 2, alors il existe une constante positive $\overline{\gamma}$ telle que :

$$\|g(x)\| \leq \overline{\gamma}, \quad \textit{pour tout } x \in \mathcal{L}.$$

1.2 Conditions de Wolfe

Rappelons maintenant les conditions de *Wolfe* faibles :

$$\begin{cases} f(x_k + \lambda_k d_k) \leq f(x_k) + \omega_1 \lambda_k d_k^T g_k, \\ d_k^T g_{k+1} \geq \omega_2 d_k^T g_k. \end{cases} \tag{5.10}$$

Où $0 < \omega_1 < \omega_2 < 1$.

Les conditions de *Wolfe* fortes :

$$\begin{cases} f(x_k + \lambda_k d_k) \leq f(x_k) + \omega_1 \lambda_k d_k^T g_k, \\ \left| d_k^T g_{k+1} \right| \leq -\omega_2 d_k^T g_k. \end{cases} \tag{5.11}$$

Où $0 < \omega_1 < \omega_2 < 1$.

Les conditions de *Wolfe* relaxées :

$$\begin{cases} f(x_k + \lambda_k d_k) \leq f(x_k) + \omega_1 \lambda_k d_k^T g_k, \\ \omega_2' d_k^T g_k \leq d_k^T g_{k+1} \leq -\omega_2'' d_k^T g_k. \end{cases} \tag{5.12}$$

Où $0 < \omega_1 < \omega_2' < 1$ et $\omega_2'' > 0$.

1.3 Lemmes intermédiaires

Sous l'hypothèse1 sur f, nous avons le lemme suivant :

Lemme 5.1 [7] *Supposons que x_1 soit un point initial pour lequel l'hypothèse1 soit satisfaite. Considérons une méthode itérative du type (5.2) où d_k est une direction de descente et λ_k vérifiant les conditions de Wolf faibles (5.10). Alors :*

$$\sum_{k \geq 1} \frac{(g_k^t d_k)^2}{\|d_k\|^2} < \infty.$$

Pour étudier notre problème de convergence, nous avons besoin aussi des deux lemmes suivants :

Lemme 5.2 [7] *Supposons que $\{a_i\}$ et $\{b_i\}$ soient des suites à termes positifs.*

Si

$$\sum_{k\geq 1} a_k = \infty,$$

et pour tout $k \geq 1$ *on ait*

$$b_k \leq c_1 + c_2 \sum_{i=1}^{k} a_i,$$

où c_1 *et* c_2 *sont des constantes positives, alors on a*

$$\sum_{k\geq 1} \frac{a_k}{b_k} = \infty.$$

Lemme 5.3 [7] *Considérons la fonction d'une variable réelle suivante :*

$$\rho(t) = \frac{a+bt}{c+dt}, \quad t \in \mathbb{R},$$

avec $a,\ b,\ c,\ d \neq 0$ *sont des nombres réels donnés. Si*

$$bc - ad > 0,$$

alors $\rho(t)$ *est strictement croissante pour* $t < \dfrac{-c}{d}$ *et* $t > \dfrac{-c}{d}$. *Sinon si*

$$bc - ad < 0,$$

$\rho(t)$ *est strictement décroissante pour* $t < \dfrac{-c}{d}$ *et* $t > \dfrac{-c}{d}$.

2 Convergence de la méthode générale (5.2),(5.3) et (5.8)

Dans ce chapitre, on considère la méthode définie par (5.2), (5.3) où β_k est de la forme (5.8), c'est-à-dire

$$\beta_k = \frac{\phi_k}{\phi_{k-1}}.$$

Nous donnerons un lemme utile pour la suite, puis nous présenterons deux résultats de convergence qui exigent certaines conditions sur ϕ_k. Pour simplifier l'étude, notons

$$t_k = \frac{\|d_k\|^2}{\phi_k^2}, \tag{5.13}$$

et

$$r_k = -\frac{g_k^T d_k}{\phi_k}. \tag{5.14}$$

Pour passer à les principaux résultats de convergence, nous avons besoin aussi le lemme suivant :

Lemme 5.4. *Considérons une méthode définie par (5.2), (5.3) où β_k est de la forme (5.8) et t_k définie par (5.13). Alors*

$$t_k = -2\sum_{i=1}^{k} \frac{g_i^T d_i}{\phi_i^2} - \sum_{i=1}^{k} \frac{\|g_i\|^2}{\phi_i^2}, \tag{5.15}$$

pour tout $k \geq 1$.

Preuve :

Pour $k = 1$ (5.15) est satisfaite puisque $d_1 = g_1$.

Pour $i \geq 2$, de (5.3) on a

$$d_i + g_i = \beta_i d_{k-1}. \tag{5.16}$$

Rendant carré les deux côtés d'équation (5.16), nous obtenons

$$\|d_i\|^2 = -\|g_i\|^2 - 2g_i^T d_i + \beta_i^2 \|d_{k-1}\|^2. \tag{5.17}$$

Diviser (5.17) par ϕ_i^2 et appliquer (5.8) et (5.13) on aura

$$\begin{aligned} \frac{\|d_i\|^2}{\phi_i^2} &= t_i = -\frac{\|g_i\|^2}{\phi_i^2} - 2\frac{g_i^T d_i}{\phi_i^2} + \frac{\phi_i^2}{\phi_{i-1}^2}\frac{\|d_{k-1}\|^2}{\phi_i^2} \\ &= t_{i-1} - -2\frac{g_i^T d_i}{\phi_i^2} - \frac{\|g_i\|^2}{\phi_i^2}. \end{aligned}$$

Additionner cette expression sur i, nous obtenons

$$t_k = t_1 - 2\sum_{i=2}^{k} \frac{g_i^T d_i}{\phi_i^2} - \sum_{i=2}^{k} \frac{\|g_i\|^2}{\phi_i^2}.$$

D'autre part $d_1 = -g_1$ et $t_1 = \dfrac{\|g_1\|^2}{\phi_1^2}$, la relation précitée est équivalent à (5.15).Donc elle est satisfaite pour tout $k \geq 1$. □

2.1 Principaux résultats de convergence

Nous arrivons maintenant aux deux principaux résultats de convergence. Il s'agit des théorèmes 5.1 et 5.2.

Théorème 5.1. *Soit x_1 un point de départ pour lequel l'hypothèse1 soit satisfaite. Considérons une méthode définie par (5.2),(5.3) avec β_k de la forme (5.8) et r_k définie par (5.14). Si pour tout k, d_k est une direction de descente et le pas λ_k est déterminé par la règle de Wolfe faible (5.10) et si*

$$\sum_{k\geq 1} r_k^2 = \infty, \tag{5.18}$$

alors, nous avons

$$\lim_{k\to\infty} \inf \|g_k\| = 0. \tag{5.19}$$

Preuve : D'après (5.17)

$$\|d_i\|^2 \geq -\|g_i\|^2 - 2g_i^T d_i, \tag{5.20}$$

multiplier le côtés adroite de l'équation (5.20) par $\dfrac{\|g_i\|^2}{\|g_i\|^2}$ on aura

$$-2g_i^T d_i - \|g_i\|^2 \leq \frac{(g_i^T d_i)^2}{\|g_i\|^2}. \tag{5.21}$$

Divisons (5.21) par ϕ_i^2 et appliquons (5.15) on aura

$$t_k \leq \sum_{i=1}^{k} \frac{(g_i^T d_i)^2}{\|g_i\|^2 \phi_i^2}, \tag{5.22}$$

de (5.14)

$$t_k \leq \sum_{i=1}^{k} \frac{r_i^2}{\|g_i\|^2}. \tag{5.23}$$

Supposons maintenant que

$$\lim_{k\to\infty} \inf \|g_k\| \neq 0. \tag{5.24}$$

De (5.24), il existe une constante $\delta > 0$ telle que

$$\|g_k\|^2 \geq \delta, \text{ pour tout } k.$$

Dans ce cas, de (5.23), on obtient

$$t_k \leq \frac{1}{\delta^2} \sum_{i=1}^{k} r_i^2.$$

De la relation précitée, (5.18) et lemme (5.2) on aura

$$\sum_{i\geq 1} \frac{r_i^2}{t_i^2} = \sum_{i\geq 1} \frac{(g_i^T d_i)^2 \phi_i^2}{\|d_i\|^2 \phi_i^2} = \sum_{i\geq 1} \frac{(g_i^T d_i)^2}{\|d_i\|^2} = \infty. \tag{5.25}$$

En considérant (5.25) et (5.15) on obtient une contradiction.(voir chapitre 4). Donc (5.19) est satisfaite.

Théorème 5.2. *Soit x_1 un point de départ pour lequel l'hypothèse1 soit satisfaite. Considérons une méthode définie par (5.2), (5.3) avec β_k de la forme (5.8). Si pour tout k, d_k est une direction de descente et le pas λ_k est déterminé par la règle de Wolfe faible (5.10), et si*

$$\sum_{k\geq 1} \frac{\|g_k\|^2}{\phi_k^2} = \infty, \tag{5.26}$$

alors, nous avons

$$\lim_{k\to\infty} \inf \|g_k\| = 0. \tag{5.27}$$

Preuve : D'après (5.13)

$$t_k \geq 0 \text{ pour tout } k,$$

donc, de (5.15)

$$\begin{aligned} -2\sum_{i=1}^{k} \frac{g_i^T d_i}{\phi_i^2} \quad &\geq \quad \sum_{i=1}^{k} \frac{\|g_i\|^2}{\phi_i^2} \Longleftrightarrow -4\sum_{i=1}^{k} \frac{g_i^T d_i}{\phi_i^2} \geq 2\sum_{i=1}^{k} \frac{\|g_i\|^2}{\phi_i^2} \\ &\Longleftrightarrow \quad -4\sum_{i=1}^{k} \frac{g_i^T d_i}{\phi_i^2} - \sum_{i=1}^{k} \frac{\|g_i\|^2}{\phi_i^2} \geq \sum_{i=1}^{k} \frac{\|g_i\|^2}{\phi_i^2}, \end{aligned}$$

de (5.21) on a

$$4\sum_{i=1}^{k} \frac{(g_i^T d_i)^2}{\|g_i\|^2 \phi_i^2} \geq -4\sum_{i=1}^{k} \frac{g_i^T d_i}{\phi_i^2} - \sum_{i=1}^{k} \frac{\|g_i\|^2}{\phi_i^2} \geq \sum_{i=1}^{k} \frac{\|g_i\|^2}{\phi_i^2},$$

alors, si (5.26) est satisfaite nous avons aussi

$$\sum_{k\geq 1} \frac{(g_k^T d_k)^2}{\|g_k\|^2 \phi_k^2} = \infty. \tag{5.28}$$

Puisque (5.22) est satisfaite, d'après (5.28) et le lemme (5.2) on obtient

$$\sum_{k\geq 1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} \leq \sum_{k\geq 1} \frac{(g_k^T d_k)^2}{\|g_k\|^2 \|d_k\|^2} = \infty.$$

Donc la relation (5.15) est vérifiée et par conséquent (5.27) est satisfaite. Ce qui achève la démonstration. □

2.2 Application aux méthodes de Dai-Yuan et Flecher-Reeves

Des théorèmes (5.1), (5.2), nous avons fourni deux conditions suffisantes qui assurent la convergence globale de la méthode générale définie par (5.2), (5.3) avec β_k de la forme (5.8). Pour démontrer la convergence globale de telles méthodes nous avons deux étapes à franchir.

Etape1 : Prouver la descente de la direction d_k

Etape2 : Montrer que l'une des deux conditions (5.18) ou (5.26) est vérifiée.

Application à la méthode de Dai-Yuan

Nous voyons que pour la formule de *Dai-yuan* définie par la formule

$$\beta_k^{DY} = \frac{\|g_k\|^2}{d_{k-1}^T y_{k-1}},$$

la relation (5.18) a lieu, car dans ce cas on a

$$\phi_k = -g_k^T d_k,$$

et par conséquent

$$r_k = 1.$$

Application à la méthode de Flecher-Reeves

Le cas de la méthode de Flecher-Reeves correspond à

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

le cas de la méthode de Flecher-Reeves correspond à

$$\phi_k = \|g_k\|^2,$$

Si la fonction f vérifie l'Hypothèse2, alors la relation (5.26) est vérifiée.

3 Etude de la convergence d'une famille des méthodes du gradient conjugué

Dans cette section nous allons étudier la convergence d'une une famille de gradients conjugués, formée à partir d'une combinaison convexe de la méthode de *Flecher-Reeves* et la méthode de *Dai-yuan.*

Considérons donc la famille des méthodes définie par (5.2), (5.3) avec β_k de la forme (5.8) et ϕ_k étant une combinaison convexe de β_k^{FR} et β_k^{DY} c'est-à-dire

$$\phi_k = \alpha \left\|g_k\right\|^2 + (1-\alpha)(-g_k^T d_k), \quad \alpha \in [0,1]\,. \tag{5.29}$$

D'après cette expression de ϕ_k, nous voyons que la méthode de *Flecher-Reeves* correspond à $\alpha = 1$. La méthode de *Dai-yuan* correspond à $\alpha = 0$. Le lemme suivant nous donne d'autres expressions de β_k et de ϕ_k qui nous seront utiles par la suite

Lemme 5.5.

$$\beta_k = \frac{\left\|g_k\right\|^2}{\alpha \left\|g_{k-1}\right\|^2 + (1-\alpha)(d_{k-1}^T y_{k-1})} \tag{5.30}$$

$$\phi_k = \left\|g_k\right\|^2 \frac{\alpha \left\|g_{k-1}\right\|^2 + (1-\alpha)(-g_{k-1}^T d_{k-1})}{\alpha \left\|g_{k-1}\right\|^2 + (1-\alpha) d_{k-1}^T y_{k-1}}. \tag{5.31}$$

Preuve : En effet, d'après (5.3) on a

$$g_k^T d_k = -\left\|g_k\right\|^2 + \beta_k g_k^T d_{k-1}, \tag{5.32}$$

substituons (5.32) dans (5.29) on aura

$$\begin{aligned} g_k^T d_k &= -\left\|g_k\right\|^2 + \frac{\alpha \left\|g_k\right\|^2 + (1-\alpha)(-g_k^T d_k)}{\alpha \left\|g_{k-1}\right\|^2 + (1-\alpha)(-g_{k-1}^T d_{k-1})} g_k^T d_{k-1} \\ &= \frac{-\alpha \left\|g_{k-1}\right\|^2 \left\|g_k\right\|^2 - (1-\alpha)(-g_{k-1}^T d_{k-1}) \left\|g_k\right\|^2 + \left[\alpha \left\|g_k\right\|^2 + (1-\alpha)(-g_k^T d_k)\right] g_k^T d_{k-1}}{\alpha \left\|g_{k-1}\right\|^2 + (1-\alpha)(-g_{k-1}^T d_{k-1})} \end{aligned} \tag{4.33}$$

d'autre part

$$y_{k-1} = g_k - g_{k-1} \Longleftrightarrow g_{k-1} = g_k - y_{k-1},$$

donc, de (5.33)

$$\begin{aligned} & g_k^T d_k \left[\alpha \left\|g_{k-1}\right\|^2 + (1-\alpha)(-g_k + y_{k-1})d_{k-1}\right] \\ = \; & -\alpha \left\|g_{k-1}\right\|^2 \left\|g_k\right\|^2 - (1-\alpha)(-g_{k-1}^T d_{k-1}) \left\|g_k\right\|^2 + \left[\alpha \left\|g_k\right\|^2 + (1-\alpha)(-g_k^T d_k)\right] g_k^T d_{k-1}, \end{aligned}$$

de la relation précédente, on a

$$g_k^T d_k = -\left\|g_k\right\|^2 \frac{\alpha(\left\|g_{k-1}\right\|^2 - g_k^T d_{k-1}) + (1-\alpha)(-g_{k-1}^T d_{k-1})}{\alpha \left\|g_{k-1}\right\|^2 + (1-\alpha)(d_{k-1}^T y_{k-1})} \quad (4.34)$$

$$g_k^T d_k \left[\alpha \left\|g_{k-1}\right\|^2 + (1-\alpha) d_{k-1}^T y_{k-1}\right] = -\left\|g_k\right\|^2 \left[\alpha(\left\|g_{k-1}\right\|^2 - g_k^T d_{k-1}) + (1-\alpha)(-g_{k-1}^T d_{k-1})\right] \quad (4.35)$$

alors, de (5.35) nous déduisons une forme équivalente de β_k

$$\beta_k = \frac{\left\|g_k\right\|^2}{\alpha \left\|g_{k-1}\right\|^2 + (1-\alpha) d_{k-1}^T y_{k-1}}. \quad (5.36)$$

Pour obtenir l'expression (5.31) de ϕ_k, substituons (5.34) dans (5.29), nous aurons

$$\phi_k = \left\|g_k\right\|^2 \frac{\alpha \left\|g_{k-1}\right\|^2 + (1-\alpha)(-g_{k-1}^T d_{k-1})}{\alpha \left\|g_{k-1}\right\|^2 + (1-\alpha) d_{k-1}^T y_{k-1}}. \quad \square \quad (5.37)$$

Nous pouvons maintenant énoncer les résultats de convergences. Il s'agit de trois résultats de convergence (Théorèmes 5.3, 5.4 et 5.5) correspondant chacun à l'une des trois recherches linéaires inexactes de *Wolfe* (*Wolfe* faible, *Wolfe* forte et *Wolfe* relaxée).

3.1 Résultat de convergence avec la recherche linéaire inexacte de Wolfe faible

Théorème 5.3. *Soit x_1 un point de départ pour lequel l'hypothèse1 soit satisfaite. Considérons une méthode définie par (5.2)-(5.3) avec β_k et ϕ_k de la forme (5.8) et (5.29) respectivement. Supposons que pour tout k, d_k est une direction de descente*

et que le pas λ_k est déterminé par la règle de Wolfe faible (5.10). Alors

$$\phi_k \leq \frac{1}{1-\omega_2} \|g_k\|^2 \tag{5.38}$$

et

$$\lim_{k\to\infty} \inf \|g_k\| = 0. \tag{5.39}$$

Preuve : D'après la $2^{ème}$ condition de (5.10) nous avons

$$\begin{aligned}
g_{k+1}^T d_k &\geq \omega_2 g_k^T d_k \Longleftrightarrow \\
g_k^T d_{k-1} &\geq \omega_2 g_{k-1}^T d_{k-1} \Longleftrightarrow \\
g_k^T d_{k-1} - g_{k-1}^T d_{k-1} &\geq \omega_2 g_{k-1}^T d_{k-1} - g_{k-1}^T d_{k-1} \Longleftrightarrow \\
d_{k-1}^T (g_k^T - g_{k-1}^T) &\geq (\omega_2 - 1) g_{k-1}^T d_{k-1} \Longleftrightarrow \\
d_{k-1}^T y_{k-1} &\geq (1-\omega_2)(-d_{k-1}^T g_{k-1}).
\end{aligned} \tag{5.40}$$

Donc, avec la relation (5.37), (5.38) est vrai. En effet, de (5.40)

$$\begin{aligned}
(-d_{k-1}^T g_{k-1}) &\leq \frac{d_{k-1}^T y_{k-1}}{(1-\omega_2)} \Longleftrightarrow \\
\left[\alpha \|g_{k-1}\|^2 + (1-\alpha)(-g_{k-1}^T d_{k-1})\right] \|g_k\|^2 &\leq \frac{\left[\alpha \|g_{k-1}\|^2 + (1-\alpha) d_{k-1}^T y_{k-1}\right] \|g_k\|^2}{(1-\omega_2)} \Longleftrightarrow \\
\frac{\alpha \|g_{k-1}\|^2 + (1-\alpha)(-g_{k-1}^T d_{k-1})}{\alpha \|g_{k-1}\|^2 + (1-\alpha) d_{k-1}^T y_{k-1}} \|g_k\|^2 &\leq \frac{\|g_k\|^2}{(1-\omega_2)},
\end{aligned} \tag{5.41}$$

ce qui donne

$$\phi_k \leq \frac{\|g_k\|^2}{(1-\omega_2)}. \tag{5.42}$$

D'après l'hypothèse 2 on a

$$\|g_k\|^2 \leq \gamma,$$

de (5.42)

$$\begin{aligned} \frac{\|g_k\|^2}{\phi_k} &\geq (1-\omega_2) \Longleftrightarrow \\ \frac{\|g_k\|^2}{\phi_k^2} &\geq \frac{(1-\omega_2)}{\|g_k\|^2} \geq \frac{(1-\omega_2)}{\gamma^2}. \end{aligned}$$

Alors

$$\sum_{k\geq 1} \frac{\|g_k\|^2}{\phi_k^2} = \infty.$$

Donc, (5.39) est satisfaite d'après le théorème (5.2).

3.2 Résultat de convergence avec la recherche linéaire inexacte de Wolfe relaxée

Dans cette partie on utilise la recherche linéaire inexacte de *Wolfe* relaxée (5.6) où les scalaires ω_2' et ω_2" sont soumises à certaines conditions. Contrairement à la recherche linéaire inexacte de *Wolfe* faible, l'utilisation de la recherche linéaire inexacte de *Wolfe* relaxée assure la descente de la fonction f à chaque itération. Avant de donner le principal résultat de cette section, introduisons par souci de simplicité les notations suivantes

$$\bar{r}_k = \frac{-g_k^T d_k}{\|g_k\|^2}, \tag{5.43}$$

et

$$l_k = \frac{g_{k+1}^T d_k}{g_k^T d_k}. \tag{5.44}$$

L'utilisation de (5.34) permet d'écrire

$$\bar{r}_k = \frac{\alpha + (1-\alpha+\alpha l_{k-1})\bar{r}_{k-1}}{\alpha + (1-\alpha)(1-l_{k-1})\bar{r}_{k-1}}. \tag{5.45}$$

Théorème 5.4. *Soit x_1 un point de départ pour lequel l'hypothèse1 soit satisfaite. Considérons une méthode du type (5.2), (5.3), (5.8) et (5.29) où $\alpha \in (0,1]$ et le pas λ_k est déterminé par la règle de Wolfe relaxée (5.12). Si*

$$\omega_2' + \omega_2'' \leq \frac{1}{\alpha}, \tag{5.46}$$

alors, nous avons pour tout $k \geq 1$,

$$\bar{r}_k > 0. \tag{5.47}$$

D'autre part la méthode converge globalement, c'est-à-dire qu'ona

$$\lim_{k\to\infty} \inf \|g_k\| = 0.$$

Preuve : le côté droit de (5.45) est une fonction de α, l_{k-1}, et $\bar{r}_{k-1}$. Notons cela par $\psi(\alpha, l_{k-1}, \bar{r}_{k-1})$. Tout d'abord montrons que

$$0 < \bar{r}_k < (1-\omega_2')^{-1}, \text{ pour tout } k \geq 1. \tag{5.48}$$

La démonstration se fait par récurrence

1) pour $k = 1$,

$$\bar{r}_1 = \frac{-g_1^T d_1}{\|g_1\|^2} = \frac{\|g_1\|^2}{\|g_1\|^2} = 1,$$

alors (5.48) est satisfaite pour $k = 1$.

2) Supposons que (5.48) est satisfaite pour $k-1$ et démontrons qu'elle le sera pour k.

D'après la $2^{ème}$ relation de la condition de (5.12) on a

$$-\omega_2'' \leq l_{k-1} \leq \omega_2'.$$

En utilisant lemme 5.3, nous obtenons

$$\bar{r}_k \leq \psi(\alpha, \omega_2', \bar{r}_{k-1}) < \psi(\alpha, \omega_2', (1-\omega_2')^{-1}) = (1-\omega_2')^{-1}.$$

D'autre part, en prenant en considération le lemme 5.3 et la relation (5.46), nous avons aussi

$$\bar{r}_k \geq \psi(\alpha, -\omega_2'', \bar{r}_{k-1}) > \psi(\alpha, -\omega_2'', (1-\omega_2')^{-1}) \geq 0.$$

Alors (5.48) est vraie pour tout k, donc il est vrai pour tout $k \geq 1$.Reste à montrer que

$$\lim_{k\to\infty} \inf \|g_k\| = 0.$$

En effet, en utilisant le théorème 5.1 (voir la relation (5.18)), il suffit de montrer que

$$\max\{\bar{r}_{k-1}, \bar{r}_k\} \geq c_1, \ \textit{pour tout } k \geq 2, \ \textit{et pour } c_1 > 0. \tag{5.49}$$

En effet si

$$\bar{r}_{k-1} \leq 1,$$

nous pouvons affirmer en prenant en considération le lemme 5.3

$$\bar{r}_k \geq \psi(\alpha, -\omega_2'', 1) \stackrel{\Delta}{=} c_2.$$

Puisque $c_2 \in (0,1)$, nous obtenons alors

$$\max\{\bar{r}_{k-1}, \bar{r}_k\} \geq c_2, \ \textit{pour tout } k \geq 2. \tag{5.50}$$

Si on utilise la définition (5.14) de r_k et la relation (5.29), nous obtenons

$$r_k = \frac{\bar{r}_k}{\alpha + (1-\alpha)\bar{r}_k}.$$

La relation (5.50) et le lemme 5.3 impliquent que la relation (5.49) est satisfaite avec $c_1 = c_2$. $\square$

3.3 Résultat de convergence avec la recherche linéaire inexacte de Wolfe forte

Théorème 5.5. *Soit x_1 un point de départ pour lequel l'Hypothèse1 soit satisfaite. Considérons une méthode du type (5.2), (5.3), où*

$$\beta_k = \frac{\tau_k \|g_k\|^2}{\alpha \|g_{k-1}\|^2 + (1-\alpha) d_{k-1}^T y_{k-1}}, \tag{5.51}$$

λ_k est déterminé par la règle de Wolfe forte (5.11) avec $\omega_2 \leq \frac{1}{2}$. Si pour tout $\lambda \in [0,1]$

$$\tau_k \in \left[\frac{\omega_2 - 1}{1 + (1-2\alpha)\omega_2}, 1 \right], \tag{5.52}$$

alors la méthode produit une direction de descente à chaque itération et converge globalement dans le sens (5.19), c'est àdire

$$\lim_{k \to \infty} \inf \|g_k\| = 0.$$

Preuve :

Notons

$$\bar{\beta}_k = \frac{\alpha \|g_k\|^2 + (1-\alpha)(-g_k^T d_k)}{\alpha \|g_{k-1}\|^2 + (1-\alpha)(-g_{k-1}^T d_{k-1})}, \tag{5.53}$$

et

$$\xi_k = \frac{\beta_k}{\bar{\beta}_k}. \tag{5.54}$$

Des calculs directs montrent que

$$\bar{r}_k = \frac{\alpha + [\tau_k l_{k-1} + (1-\alpha)(1-l_{k-1})]\, \bar{r}_{k-1}}{\alpha + (1-\alpha)(1-l_{k-1})\bar{r}_{k-1}}, \tag{5.55}$$

et

$$\xi_k = \frac{[\alpha + (1-\alpha)\bar{r}_{k-1}]\, \tau_k}{\alpha + (1-\alpha)(1 - l_{k-1} + \tau_k l_{k-1})\bar{r}_{k-1}}, \tag{5.56}$$

où $\bar{r}_k$ et l_k sont définis respectivement par (5.43) et (5.44).

Maintenant, la partie droite de (5.55) est une fonction de α, τ_k, l_{k-1}, et $\bar{r}_{k-1}$, qu'on pourra noter

$$\psi(\alpha, \tau_k, l_{k-1}, \bar{r}_{k-1}).$$

Montrons d'abord que

$$0 < \bar{r}_k < (1-\omega_2)^{-1}, \text{ pour tout } k \geq 1. \tag{5.57}$$

La démonstration se fait par récurrence

1) pour $k = 1$,

$$\bar{r}_1 = \frac{-g_1^T d_1}{\|g_1\|^2} = \frac{\|g_1\|^2}{\|g_1\|^2} = 1.$$

2) Supposons que (5.57) soit satisfaite pour $k-1$ et démontrons qu'elle le sera pour k.

D'après la $2^{ème}$ relation de la condition (5.12) on a

$$|l_{k-1}| \leq \omega_2. \tag{5.58}$$

Cette relation et le lemme 5.3, nous donnent

$$\begin{aligned}
\bar{r}_k &\leq \max\left\{\psi(\alpha, 1, l_{k-1}, \bar{r}_{k-1}), \psi(\alpha, \frac{\omega_2 - 1}{1 + (1-2\alpha)\omega_2}, l_{k-1}, \bar{r}_{k-1})\right\} \\
&\leq \max\left\{\psi(\alpha, 1, \omega_2, \bar{r}_{k-1}), \psi(\alpha, \frac{\omega_2 - 1}{1 + (1-2\alpha)\omega_2}, -\omega_2, \bar{r}_{k-1})\right\} \\
&< \max\left\{\psi(\alpha, 1, \omega_2, (1-\omega_2)^{-1}), \psi(\alpha, \frac{\omega_2 - 1}{1 + (1-2\alpha)\omega_2}, -\omega_2, (1-\omega_2)^{-1})\right\} \\
&= (1-\omega_2)^{-1},
\end{aligned}$$

où $\omega_2 \leq \dfrac{1}{2}$ est aussi utilisé dans l'égalité. D'autre part, nous pouvons prouver que

$$\bar{r}_k < \min\left\{\psi(\alpha, 1, -\omega_2, (1-\omega_2)^{-1}), \psi(\alpha, \frac{\omega_2 - 1}{1 + (1-2\alpha)\omega_2}, \omega_2, (1-\omega_2)^{-1})\right\} \geq 0.$$

Par conséquent (5.54) est vrai pour tout k.

Maintenant prouvons que

$$\xi_k \in [-1, 1], \ \textit{pour tout } k \geq 2. \tag{5.59}$$

Notons par D_k le dénominateur de ξ_k dans (5.56). Des calculs directs donnent

$$(1 - \xi_k)D_k = (1 - \tau_k)\left[\alpha + (1 - \alpha)(1 - l_{k-1})\bar{r}_{k-1}\right], \tag{5.60}$$

et

$$(1 - \xi_k)D_k = \left[\alpha + (1 - \alpha)(1 + l_{k-1})\bar{r}_{k-1}\right]\tau_k + \left[\alpha + (1 - \alpha)(1 - l_{k-1})\bar{r}_{k-1}\right], \tag{5.61}$$

En appliquant maintenant (5.52), (5.57) et (5.58), nous pouvons montrer que $D_k > 0$ et les termes à droite dans les relations (5.60) et (5.61) sont non-négatifs. Donc (5.59) est satisfaite. En plus, de la même façon que dans la preuve du théorème 5.4, on peut vérifier que (5.49) est aussi vraie pour $c_1 > 0$.

Finalement en prenant en considération les relations (5.49) et (5.60) et la discussion qu'on a engagé dans la section précédente, on peut conclure qu'on a bien

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

CHAPITRE 6

Nouvelle famille à deux paramètres des méthodes du gradient conjugué

1 Introduction

Consider the unconstrained optimization problem

$$\min \ f(x), \qquad x \in IR^n, \tag{6.1}$$

where f is a smooth function and its gradient is available. Conjugate gradient methods are a class of important methods for solving (6.1), especially for large scale problems, which have the following form :

$$x_{k+1} = x_k + \alpha_k d_k, \tag{6.2}$$

where x_k is the current iterate, α_k is a positive scalar and called the step length which is determined by some line search, and d_k is the search direction generated by the rule

$$d_k = \begin{cases} -g_k & \text{for } k = 1; \\ -g_k + \beta_k d_{k-1} & \text{for } k \geq 2, \end{cases} \tag{6.3}$$

where $g_k = \nabla f(x_k)$ is the gradient of f at x_k and β_k is a scalar.

The strong Wolfe conditions, namely,

$$f(x_k + \alpha_k d_k) - f(x_k) \leq \delta \alpha_k g_k^T d_k \tag{6.4}$$

$$\left| g(x_k + \alpha_k d_k)^T d_k \right| \leq -\sigma g_k^T d_k, \tag{6.5}$$

where $0 < \delta < \sigma < 1$, are often imposed on the line search (in this case, we call the line search the strong Wolfe line search). The scalar β_k is chosen so that the method (6.2), (6.3) reduces to the linear conjugate gradient method in the case when f is convex quadratic and exact line search ($g(x_k + \alpha_k d_k)^T d_k = 0$) is used.

For general functions, however, different formula for scalar β_k result in distinct nonlinear conjugate gradient methods, see [3, 10, 11, 12, 14, 16, 19, 20, 22]. Since Fletcher and Reeves (FR) introduced the linear conjugate gradient method in 1964, with

$$\beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2}, \tag{6.6}$$

where$\|.\|$ means the Euclidean norm.

For non-quadratic objective functions, the global convergence of (FR) method was proved when the exact line search or strong Wolfe line search [2, 4] was used. However, if the condition (6.5) is satisfied for $\sigma < 1$, the above method of (FR) with the strong Wolfe line search can ensure a descent search direction and converge globally provided only for the case when f is quadratic [5, 4]. The conjugate descent (CD) method of Fletcher [11], where

$$\beta_k^{CD} = \frac{\|g_k\|^2}{-d_{k-1}^T g_{k-1}}, \tag{6.7}$$

ensures a descent direction for general functions if the line search satisfies the strong Wolfe conditions $(6.4) - (6.5)$ with $\sigma < 1$. But the global convergence of the method is proved (see [6]) only for the case when the line search satisfies (6.4) and

$$\sigma g_k^T d_k \leq g(x_k + \alpha_k d_k)^T d_k \ \leq 0. \tag{6.8}$$

For any positive constant σ_2, an example is constructed in [6] showing that conjugate descent method with α_k satisfying (6.4) and

$$\sigma_1 g_k^T d_k \leq g(x_k + \alpha_k d_k)^T d_k \leq -\sigma_2 g_k^T d_k, \tag{6.9}$$

need not converge.

Recently, Dai and Yuan [3] proposed a nonlinear conjugate gradient method, which has the form (6.2) – (6.3) with

$$\beta_k^{DY} = \frac{\|g_k\|^2}{d_{k-1}^T y_{k-1}}, \tag{6.10}$$

where $y_{k-1} = g_k - g_{k-1}$. A remarkable property of the DY method is that it provides a descent search direction at every iteration and converges globally provided that the step size satisfies the Wolfe conditions, namely, (6.4) and

$$\sigma g_k^T d_k \leq g(x_k + \alpha_k d_k)^T d_k. \tag{6.11}$$

By direct calculations, we can deduce an equivalent form for β_k^{DY}, namely

$$\beta_k^{DY} = \frac{g_k^T d_k}{g_{k-1}^T d_{k-1}}. \tag{6.12}$$

In [7], Dai and Yuan proposed a family of globally convergent conjugate methods, in which

$$\beta_k = \frac{\|g_k\|^2}{\lambda \|g_{k-1}\|^2 + (1-\lambda)(d_{k-1}^T y_{k-1})}, \tag{6.13}$$

where $\lambda \in [0, 1]$ is a parameter, and proved that the family of methods using line searches that satisfy (6.4) and (6.9) converges globally if the parameters σ_1, σ_2 and λ are such that

$$\sigma_1 + \sigma_2 \leq \lambda^{-1}. \tag{6.14}$$

Observing that the formula (6.6), (6.7) and (6.10) share same numerators and three denominators, we can use combinations of these numerators and denominators to

obtain the following two-parameter family :

$$\beta_k^* = \frac{(1-\lambda_k)\|g_k\|^2 + \lambda_k(-g_k^T d_k)}{(1-\lambda_k-\mu_k)\|g_{k-1}\|^2 + (\lambda_k+\mu_k)(-g_{k-1}^T d_{k-1})}, \tag{6.15}$$

where $\lambda_k \in [0,1]$ and $\mu_k \in [0, 1-\lambda_k]$ are parameters.

We see that the above formula for β_k^* is special forms of

$$\beta_k^* = \frac{\phi_k}{\phi'_{k-1}}, \tag{6.16}$$

where ϕ_k satisfies that

$$\phi_k = (1-\lambda_k)\|g_k\|^2 + \lambda_k(-g_k^T d_k), \tag{6.17}$$

and

$$\phi'_{k-1} = (1-\lambda_k-\mu_k)\|g_{k-1}\|^2 + (\lambda_k+\mu_k)(-g_{k-1}^T d_{k-1}). \tag{6.18}$$

It is clear that the formula (6.15) is a generalization of the three previous methods.

The rest of this chapter is organized as follows. Some preliminaries are given in the next section. Section 3 provides two convergence theorems for the general method (6.2), (6.3) with β_k^* defined by (6.16). Section 4 includes the main convergence properties of the two-parameter family of conjugate gradient methods with Wolfe line search, and we study methods related to the new nonlinear conjugate gradient method (6.15). The preliminary numerical results are contained in section 5.

2 Preliminaries

For convenience, we assume that $g_k \neq 0$ for all k, for otherwise a stationary point has been found. We give the following basic assumption on the objective function.

Assumption 6.1 *(i) f is bounded below on the level set $\pounds = \{x \in \Re^n; f(x) \leq f(x_1)\}$; (ii) In some neighborhood N of $\pounds$, f is differentiable and its gradient g is Lipschitz*

continuous, namely, there exists a constant $L > 0$ *such that*

$$\|g(x) - g(\tilde{x})\| \leq L \|x - \tilde{x}\|, \qquad \text{for all } x, \tilde{x} \in N. \tag{6.19}$$

Some of the results obtained in this paper depend also on the following assumption.

Assumption 6.2 *The level set* $\mathcal{L} = \{x \in \Re^n; f(x) \leq f(x_1)\}$ *is bounded.*

If f *satisfies Assumption 6.1 and 6.2, there exists a positive constant* γ *such that*

$$\|g(x)\| \leq \gamma, \qquad \text{for all } x \in \mathcal{L}. \tag{6.20}$$

The conclusion of the following lemma, often called the Zoutendijk condition, is used to prove the global convergence of nonlinear conjugate gradient methods. It was originally given in [23, 24].

Lemma 6.3 *Suppose Assumption 6.1 holds. Let* $\{x_k\}$ *be generated by (6.2) and* d_k *satisfy* $g_k^T d_k < 0$*. If* α_k *is determined by the Wolfe line search (6.4),(6.11), then we have*

$$\sum_{k \geq 1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty. \tag{6.21}$$

In the latter section, we need the following lemma, the first of which is derived from [21], whereas the second is self-evident and will be used for many times.

Lemma 6.4 *Suppose that* $\{a_i\}$ *and* $\{b_i\}$ *are positive number sequences. If*

$$\sum_{k \geq 1} a_k = \infty, \tag{6.22}$$

and there exist two constants c_1 *and* c_2 *such that for all* $k \geq 1$*,*

$$b_k \leq c_1 + c_2 \sum_{i=1}^{k} a_i, \tag{6.23}$$

then we have that

$$\sum_{k\geq 1} \frac{a_k}{b_k} = \infty. \tag{6.24}$$

Lemma 6.5 *Consider the following 1-dimensional function,*

$$\rho(t) = \frac{a+bt}{c+dt}, \quad t \in \mathbb{R}^1, \tag{6.25}$$

where $a,\ b,\ c,$ *and* $d \neq 0$ *are given real numbers. If*

$$bc - ad > 0, \tag{6.26}$$

$\rho(t)$ *is strictly monotonically increasing for* $t < \dfrac{-c}{d}$ *and* $t > \dfrac{-c}{d}$; *.otherwise, if*

$$bc - ad < 0, \tag{6.27}$$

$\rho(t)$ *is strictly monotonically decreasing for* $t < \dfrac{-c}{d}$ *and* $t > \dfrac{-c}{d}$.

3 Algorithm and convergence analysis

Now we can present a new descent conjugate gradient method, namely NFCG method, as follows :

Algorithm 6.1

Step 0 : Given $x_1 \in \Re^n$, set $d_1 = -g_1$, $k = 1$. If $g_1 = 0$, then stop.

Step 1 : Find a $\alpha_k > 0$ satisfying the Wolfe conditions (6.4) and (6.11).

Step 2 : Let $x_{k+1} = x_k + \alpha_k d_k$ and $g_{k+1} = g(x_{k+1})$. If $g_{k+1} = 0$, then stop.

Step 3 : Compute β^*_{k+1} by the formula (6.15) and generate d_{k+1} by (6.3).

Step 4 : Set $k := k + 1$, go to Step 1.

In order to establish the global convergence result for the Algorithm 6.1, we will impose the following basic lemma.

For simplicity, we define

$$r_k = -\frac{g_k^T d_k}{\phi_k}, \tag{6.28}$$

and

$$t_k = \frac{\|d_k\|^2}{\phi_k^2}. \tag{6.29}$$

Lemma 6.6 *For the method (6.2), (6.3) with β_k^* defined by (6.16),*

$$t_k = 2\sum_{i=1}^{k} \frac{r_i}{\phi_i} - \sum_{i=1}^{k} \frac{\|g_i\|^2}{\phi_i^2}, \tag{6.30}$$

holds for all $k \geq 1$.

Proof Since $d_1 = -g_1, .$(6.30) holds for $k = 1$. For $i \geq 2$, it follows from (6.3) that

$$d_i + g_i = \beta_i^* d_{i-1}. \tag{6.31}$$

Squaring both sides of the above equation, we get that

$$\|d_i\|^2 = -\|g_i\|^2 - 2g_i^T d_i + \beta_i^{*2} \|d_{i-1}\|^2 . \tag{6.32}$$

Dividing (6.32) by ϕ_i^2 and applying (6.16) and (6.29),

$$t_i = \frac{\|d_{i-1}\|^2}{{\phi'_{i-1}}^2} + 2\frac{r_i}{\phi_i} - \frac{\|g_i\|^2}{\phi_i^2}. \tag{6.33}$$

Using (6.18), (6.17) and since, $d_1 = -g_1$ we get that

$$\frac{\|d_1\|^2}{{\phi'_1}^2} = \frac{\|g_1\|^2}{\|g_1\|^4} = \frac{\|g_1\|^2}{\phi_1^2}. \tag{6.34}$$

Summing the above expression (6.33) over i, we obtain

$$t_k = t_1 + 2\sum_{i=2}^{k} \frac{r_i}{\phi_i} - \sum_{i=2}^{k} \frac{\|g_i\|^2}{\phi_i^2}. \tag{6.35}$$

Since $d_1 = -g_1$ and $t_1 = \frac{\|g_1\|^2}{\phi_1^2}$, the above relation is equivalent to (6.30). So (6.30) holds for $k \geq 1$. □

Theorem 6.1 *Suppose that x_1 is a starting point for which Assumption 6.1 holds. Consider the method (6.2), (6.3) and (6.16), if for all k, d_k satisfy $g_k^T d_k < 0$ and α_k is determined by the Wolfe line search (6.4), (6.11), and if*

$$\sum_{k\geq 1} r_k^2 = \infty, \tag{6.36}$$

we have that

$$\lim_{k\longrightarrow\infty} \inf \|g_k\| = 0. \tag{6.37}$$

Proof (6.3) can be re-written as

$$g_i^T d_i + \|g_i\|^2 = \beta_i^* g_i^T d_{i-1}. \tag{6.38}$$

Squaring both sides of the above equation, we get that

$$-2g_i^T d_i - \|g_i\|^2 \leq \frac{(g_i^T d_i)^2}{\|g_i\|^2}, \tag{6.39}$$

dividing (6.39) by ϕ_i^2 and applying (6.30)

$$t_k \leq \sum_{i=1}^{k} \frac{r_i^2}{\|g_i\|^2}. \tag{6.40}$$

We proceed by contradiction. Assume that

$$\lim_{k\longrightarrow\infty} \inf \|g_k\| \neq 0. \tag{6.41}$$

Then there exists a positive constant γ such that

$$\|g_k\|^2 \geq \gamma, \qquad \text{for all } k. \tag{6.42}$$

We can see from (6.40) that,

$$t_k \leq \frac{1}{\gamma^2} \sum_{i=1}^{k} r_i^2. \tag{6.43}$$

The above relation, (6.36) and lemma 6.4, yield

$$\sum_{i\geq 1} \frac{r_i^2}{t_i} = \infty. \tag{6.44}$$

Thus, by the definition (6.1) and (6.29), we know that (6.44) contradicts (6.21). This concludes the proof. □

Theorem 6.2 *Suppose that x_1 is a starting point for which Assumption 6.1 holds. Consider the method (6.2), (6.3) and (6.16), if for all k, d_k satisfy $g_k^T d_k < 0$ and α_k is determined by the Wolfe line search (6.4), (6.11), and if*

$$\sum_{k\geq 1} \frac{\|g_k\|^2}{\phi_k^2} = \infty, \tag{6.45}$$

we have that

$$\lim_{k\longrightarrow\infty} \inf \|g_k\| = 0. \tag{6.46}$$

Proof Noting that

$$t_k \geq 0 \qquad \text{for all } k, \tag{6.47}$$

we can get from (6.30)

$$2\sum_{i=1}^{k} \frac{r_i}{\phi_i} \geq \sum_{i=1}^{k} \frac{\|g_i\|^2}{\phi_i^2}. \tag{6.48}$$

Direct calculation show that,

$$4\sum_{i=1}^{k} \frac{r_i^2}{\|g_i\|^2} \geq 4\sum_{i=1}^{k} \frac{r_i}{\phi_i} - 2\sum_{i=1}^{k} \frac{\|g_i\|^2}{\phi_i^2} \geq 0. \tag{6.49}$$

Or equivalently,

$$4\sum_{i=1}^{k} \frac{r_i^2}{\|g_i\|^2} \geq 4\sum_{i=1}^{k} \frac{r_i}{\phi_i} - \sum_{i=1}^{k} \frac{\|g_i\|^2}{\phi_i^2} \geq \sum_{i=1}^{k} \frac{\|g_i\|^2}{\phi_i^2}. \tag{6.50}$$

Thus if (6.45) holds, we also have that

$$\sum_{k\geq 1} \frac{(g_k^T d_k)^2}{\|g_k\|^2 \phi_k^2} = \infty. \tag{6.51}$$

Because (6.40) still holds, it follows from (6.51), the definition of r_k and lemma 6.4, that

$$\sum_{k\geq 1} \frac{(g_k^T d_k)^2}{\|g_k\|^2 \|d_k\|^2} = \infty. \tag{6.52}$$

The above relation and lemma 6.3 clearly give (6.37). This completes our proof. □

Thus we have proved two convergence theorems for the general method (6.2), (6.3) with β_k^* defined by (6.16).

It should also be noted that the sufficient descent condition, namely

$$g_k^T d_k \leq -c \|g_k\|^2, \tag{6.53}$$

where c is a positive constant, is not invoked in theorems 6.1 and 6.2. The sufficient descent condition (6.53) was often used or implied in the previous analysis of conjugate gradient methods (see [1, 13]). This condition has been relaxed to the descent condition ($g_k^T d_k < 0$) in the convergence analysis [3] of the FR method and the convergence analysis [8] of any conjugate gradient method.

4 Global convergence

In this section, we establish some global convergence of the two-parameter family of nonlinear conjugate gradient methods under certain line searches conditions and the methods related to this family are discussed.

We consider the method (6.2), (6.3) with ϕ_k satisfying

$$\phi_k = (1 - \lambda_k) \|g_k\|^2 + \lambda_k(-g_k^T d_k), \tag{6.54}$$

where $\lambda_k \in [0,1]$. (6.54) and (6.3) show that

$$\begin{aligned} g_k^T d_k &= -\|g_k\|^2 + \beta_k^* g_k^T d_{k-1} \\ &= -\|g_k\|^2 + \frac{(1-\lambda_k)\|g_k\|^2 + \lambda_k(-g_k^T d_k)}{(1-\lambda_k-\mu_k)\|g_{k-1}\|^2 + (\lambda_k+\mu_k)(-g_{k-1}^T d_{k-1})} g_k^T d_{k-1}. \end{aligned} \tag{6.55}$$

The above relation imply that

$$g_k^T d_k = -\frac{(1-\lambda_k-\mu_k)\|g_{k-1}\|^2 + \mu_k(-d_{k-1}^T g_{k-1}) + \lambda_k(y_{k-1}^T d_{k-1}) - g_k^T d_{k-1}}{(1-\lambda_k-\mu_k)\|g_{k-1}\|^2 + \mu_k(-d_{k-1}^T g_{k-1}) + \lambda_k(y_{k-1}^T d_{k-1})} \|g_k\|^2. \tag{6.56}$$

Thus by (6.55), we deduce an equivalent form of β_k^*,

$$\beta_k^* = \frac{\|g_k\|^2}{(1-\lambda_k-\mu_k)\|g_{k-1}\|^2 + \mu_k(-d_{k-1}^T g_{k-1}) + \lambda_k(y_{k-1}^T d_{k-1})}. \tag{6.57}$$

Substituting (6.56) into (6.54), we obtain that

$$\phi_k = \frac{(1-\lambda_k-\mu_k)\|g_{k-1}\|^2 + (\lambda_k+\mu_k)(-g_{k-1}^T d_{k-1})}{(1-\lambda_k-\mu_k)\|g_{k-1}\|^2 + \mu_k(-d_{k-1}^T g_{k-1}) + \lambda_k(y_{k-1}^T d_{k-1})} \|g_k\|^2. \tag{6.58}$$

By this relation, we can an important property of ϕ_k under Wolfe line searches and hence obtain the global convergence of the two-parameter family of nonlinear conjugate gradient methods (6.57) under some assumptions.

Theorem 6.3 *Suppose that x_1 is a starting point for which Assumption 6.1 and 6.2 hold. Consider the method (6.2), (6.3), (6.16) and (6.54), if $g_k^T d_k < 0$ for all k and α_k is computed by the Wolfe line search (6.4)-(6.11), then*

$$\frac{\phi_k}{\|g_k\|^2} \leq (1-\sigma)^{-1}. \tag{6.59}$$

Further, the method converges in the sense that

$$\lim_{k\longrightarrow\infty} \inf \|g_k\| = 0. \tag{6.60}$$

Proof Since (6.11), we have that

$$g(x_k + \alpha_k d_k)^T d_k \geq \sigma g_k^T d_k. \tag{6.61}$$

By direct calculations show that

$$d_{k-1}^T y_{k-1} \geq (1-\sigma)(-g_{k-1}^T d_{k-1}). \tag{6.62}$$

Dividing (6.58) by $\|g_k\|^2$ and applying (6.62) implies the truth of (6.59). Therefore, by (6.20) and (6.62) that

$$\sum_{k\geq 1} \frac{\|g_k\|^2}{\phi_k^2} = \infty. \tag{6.63}$$

Thus (6.37) follows from Theorem 6.2. □

In the following, we can show that, for any $\lambda_k \in (0, 1]$, the method (6.2), (6.3), (6.16) and (6.54) ensures the descent property of each search direction and converges globally under line search condition (6.4) and (6.9) where the scalars σ_1 and σ_2 satisfy certain condition. For this purpose, we define

$$\overline{r}_k = -\frac{g_k^T d_k}{\|g_k\|^2}, \tag{6.64}$$

and

$$l_k = \frac{g_{k+1}^T d_k}{g_k^T d_k}, \tag{6.65}$$

it is obvious that d_k is a descent direction if and only if $\overline{r}_k > 0$. For The above relation, (6.56) and (6.65), we can write

$$\overline{r}_k = \frac{(1-\lambda_k-\mu_k) + [\mu_k + \lambda_k(1-l_{k-1}) + l_{k-1}]\,\overline{r}_{k-1}}{(1-\lambda_k-\mu_k) + [\mu_k + \lambda_k(1-l_{k-1})]\,\overline{r}_{k-1}}. \tag{6.66}$$

Theorem 6.4 *Suppose that x_1 is a starting point for which Assumption 6.1 holds. Consider the method (6.2),(6.3), (6.16) and (6.54), where $\lambda_k \in [0, 1)$, $\mu_k \in [0, 1-\lambda_k]$ and α_k satisfies the line search conditions (6.4) and (6.9). If the scalars σ_1 and σ_2 in (6.9) is such that*

$$\sigma_1 + \sigma_2 \leq \frac{1 + \mu_k \sigma_1}{1 - \lambda_k}, \tag{6.67}$$

then we have for all $k \geq 1$

$$0 < \bar{r}_k < (1 - \sigma_1)^{-1}. \tag{6.68}$$

Further, the method converges in the sense that (6.37) is true.

Proof The right hand side of (6.66) is a function of λ_k, μ_k, l_{k-1} and $\overline{r}_{k-1}$, which is denoted as $h(\lambda_k, \mu_k, l_{k-1}, \bar{r}_{k-1})$. We prove (6.68) by induction. Noting that $d_1 = -g_1$ and hence $\bar{r}_1 = 1$, we see that (6.68) is true for $k = 1$. We now suppose that (6.68) holds for $k - 1$, namely,

$$0 < \bar{r}_{k-1} < (1 - \sigma_1)^{-1}. \tag{6.69}$$

It follows from (6.9)

$$-\sigma_2 \leq l_{k-1} \leq \sigma_1. \tag{6.70}$$

Then by lemma 6.5, the fact that $\lambda_k \in (0, 1]$ and the fact that $\mu_k \geq 0$, we get that

$$\begin{aligned} \bar{r}_k &\leq h(\lambda_k, \mu_k, \sigma_1, \bar{r}_{k-1}) < h(\lambda_k, \mu_k, \sigma_1, (1 - \sigma_1)^{-1}) \\ &= 1 + \frac{\sigma_1}{1 - \sigma_1 + \mu_k \sigma_1} \\ &\leq 1 + \frac{\sigma_1}{1 - \sigma_1} \\ &= (1 - \sigma_1)^{-1}. \end{aligned} \tag{6.71}$$

On the other hand, by lemma 6.5 and relation (6.67), we also have that

$$\bar{r}_k \geq h(\lambda_k, \mu_k, -\sigma_2, \bar{r}_{k-1}) > h(\lambda_k, \mu_k, -\sigma_2, (1 - \sigma_1)^{-1}) \geq 0. \tag{6.72}$$

Thus (6.68) is true for k, by induction, (6.68) holds for $k \geq 1$.

To show the truth of (6.37), by theorem 6.1, it suffices to prove that

$$\max\{\bar{r}_{k-1}, \bar{r}_k\} \geq c_1, \tag{6.73}$$

for all $k \geq 2$ and some constant $c_1 \geq 0$. In fact, if

$$\bar{r}_{k-1} \leq 1, \tag{6.74}$$

by lemma 6.5, the fact that $\lambda_k \in (0, 1]$ and the fact that $\mu_k \geq 0$, we can get that

$$\bar{r}_k \geq h(\lambda_k, \mu_k, -\sigma_2, 1) \triangleq c_2. \tag{6.75}$$

Since $c_2 \in (0, 1)$, we then obtain

$$\max\{\bar{r}_{k-1}, \bar{r}_k\} \geq c_2, \tag{6.76}$$

for all $k \geq 2$. By the definition (6.28) of r_k and relation (6.54), we have that

$$r_k = \frac{\bar{r}_k}{(1 - \lambda_k) + \lambda_k \bar{r}_k}. \tag{6.77}$$

Which, with (6.76) and lemma (6.5), implies that (6.73) holds with $c_1 = c_2$. This completes our proof. □

Thus we have some general convergence results are established for the two-parameter family of nonlinear conjugate gradient methods (6.57). It is easy to see from (6.57) that the two-parameter family of conjugate gradient methods includes the three nonlinear conjugate gradient methods mentioned above. Letting $\lambda_k \equiv 0$ and $\mu_k \equiv 0$, in theorem (6.55), we again obtain the convergence result of the FR method in [4]. For the case when $\lambda_k \equiv 1$ and $\mu_k \equiv 0$ the method is proved to generate a descent search direction at every iteration and converge globally of the DY method under the Wolfe line search conditions (6.4)-(6.11) (see [9]). If $\lambda_k \equiv 0$ and $\mu_k \equiv 1$, then the method ensures a descent direction for general functions and is proved to global convergence under strong Wolfe line search (6.4), (6.5) of the method CD (see

[6]) Further, if $\lambda_k \equiv \lambda$ and $\mu_k \equiv 0$, then the family reduces to the one-parameter family in [7]. Therefore the two-parameter family has the one-parameter family as subfamily in [7] and converge globally provided that the line search satisfies the Wolfe line search.

In addition, the methods related to the FR method and the DY method in [15, 3] can also be regarded as special cases of the two-parameter family methods (6.57). For example, to combine the nice global convergence properties of the FR method and the good numerical performances of the PRP method, namely

$$\beta_k^{PRP} = \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2}. \tag{6.78}$$

Hu and Storey [15] extended the result in [1] to any method (6.2) and (6.3) with β_k satisfying

$$\beta_k \in \left[0, \beta_k^{FR}\right]. \tag{6.79}$$

Gilbert and Nocedal [13] further extended the result to the case that

$$\beta_k \in \left[-\beta_k^{FR}, \beta_k^{FR}\right]. \tag{6.80}$$

Dai and Yuan [3] proved that the method (6.2) and (6.3) with β_k satisfying

$$\beta_k \in \left[\frac{\sigma - 1}{1 + \sigma}\overline{\beta_k}, \overline{\beta_k}\right], \tag{6.81}$$

where $\overline{\beta_k}$ stands for the formula (6.10), and with α_k chosen by the Wolfe line search give the convergence relation (6.37). If the line search conditions are (6.4) and (6.5) with $\sigma \leq \frac{1}{2}$. For methods related to the method (6.57). We have the following result, where s_k is given by

$$s_k = \frac{\beta_k}{\beta_k^*}, \tag{6.82}$$

where β_k^* stands for the formula (6.15). We prove that any method (6.20), (6.21) with the strong Wolfe line search produces a descent search direction at every iteration

and converges globally if the scalar β_k is such that

$$-c \leq s_k \leq (1-\sigma)^{-1}, \tag{6.83}$$

where $c = (1+\sigma)/(1-\sigma) > 0$.

Theorem 6.5 *Suppose that x_1 is a starting point for which Assumption 6.1 holds. Consider the method (6.2) and (6.3), where*

$$\beta_k = \frac{\tau_k \|g_k\|^2}{(1-\lambda_k-\mu_k)\|g_{k-1}\|^2 + \mu_k(-d_{k-1}^T g_{k-1}) + \lambda_k(y_{k-1}^T d_{k-1})}, \tag{6.84}$$

and where α_k is computed by the strong Wolfe line search (6.4) and (6.9) with $\sigma \leq \frac{1}{2}$. For any $\lambda_k \in [0,1]$ and $\mu_k \in [0, 1-\lambda_k]$, if

$$\tau_k \in \left[\frac{1+(2\lambda_k+\mu_k-1)\sigma}{\sigma-1}, \frac{1+(\mu_k-1)\sigma}{1-\sigma}\right], \tag{6.85}$$

and β_k is such that

$$s_k \in \left[-c, (1-\sigma)^{-1}\right], \tag{6.86}$$

then if $g_k \neq 0$ for all $k \geq 1$, we have that

$$0 < \bar{r}_k < (1-\sigma)^{-1} \qquad \text{for all } k \geq 1. \tag{6.87}$$

Further, the method converges in the sense that (6.37) is true.

Proof From relation (6.15), (6.84) and by direct calculations we can show that

$$\overline{r}_k = \frac{(1-\lambda_k-\mu_k) + [\tau_k l_{k-1} + \mu_k + \lambda_k(1-l_{k-1})]\,\overline{r}_{k-1}}{(1-\lambda_k-\mu_k) + [\mu_k + \lambda_k(1-l_{k-1})]\,\overline{r}_{k-1}}, \tag{6.88}$$

and

$$s_k = \frac{[1-\lambda_k-\mu_k + (\mu_k+\lambda_k)\overline{r}_{k-1}]\,\tau_k}{(1-\lambda_k-\mu_k) + [\mu_k + \lambda_k(1-l_{k-1}+\tau_k l_{k-1})]\,\overline{r}_{k-1}}, \tag{6.89}$$

where $\overline{r}_k$ and l_k are defined by (6.64) and(6.65). Now the right hand side of (6.88) is a function of λ_k, μ_k, τ_k, l_{k-1} and $\overline{r}_{k-1}$, which can be denoted as $h(\lambda_k, \mu_k, \tau_k, l_{k-1}, \bar{r}_{k-1})$.

We prove (6.87) by induction. Noting that $d_1 = -g_1$ and hence $\bar{r}_1 = 1$, we see that (6.87) is true for $k = 1$. We now suppose that (6.87) holds for $k-1$, namely,

$$0 < \bar{r}_{k-1} < (1-\sigma)^{-1}. \tag{6.90}$$

It follows from (6.5)

$$|l_{k-1}| \leq \sigma. \tag{6.91}$$

Then by lemma 6.5, and the fact that $\lambda_k, \mu_k \in [0,1]$, we get that

$$\begin{aligned}
\bar{r}_k &\leq \max\left\{h(\lambda_k, \mu_k, \frac{1+(\mu_k-1)\sigma}{1-\sigma}, l_{k-1}, \bar{r}_{k-1}), h(\lambda_k, \mu_k, \frac{1+(2\lambda_k+\mu_k-1)\sigma}{\sigma-1}, l_{k-1}, \bar{r}_{k-1})\right\} \\
&\leq \max\left\{h(\lambda_k, \mu_k, \frac{1+(\mu_k-1)\sigma}{1-\sigma}, \sigma, \bar{r}_{k-1}), h(\lambda_k, \mu_k, \frac{1+(2\lambda_k+\mu_k-1)\sigma}{\sigma-1}, -\sigma, \bar{r}_{k-1})\right\} \\
&< \max\left\{h(\lambda_k, \mu_k, \frac{1+(\mu_k-1)\sigma}{1-\sigma}, \sigma, (1-\sigma)^{-1}), h(\lambda_k, \mu_k, \frac{1+(2\lambda_k+\mu_k-1)\sigma}{\sigma-1}, -\sigma, (1-\sigma)^{-1})\right\} \\
&= 1 + \frac{\sigma}{1-\sigma} = (1-\sigma)^{-1},
\end{aligned}$$

where $\sigma \leq \dfrac{1}{2}$ is also used in the equality. For the opposite direction, we can prove that

$$\bar{r}_k \geq \min\left\{h(\lambda_k, \mu_k, \frac{1+(\mu_k-1)\sigma}{1-\sigma}, -\sigma, (1-\sigma)^{-1}), h(\lambda_k, \mu_k, \frac{1+(2\lambda_k+\mu_k-1)\sigma}{\sigma-1}, \sigma, (1-\sigma)^{-1})\right\} \geq 0. \tag{6.92}$$

Thus (6.87) is true for k, by induction, (6.87) holds for $k \geq 1$.

We now prove (6.37) by contradiction and assuming that

$$\|g(x)\| \geq \gamma, \qquad \text{for some } \gamma > 0 \text{ and all } k \geq 1.$$

Since $d_k + g_k = \beta_k d_{k-1}$, we have that

$$\|d_k\|^2 = \beta_k^2 \|d_{k-1}\|^2 - 2g_k^T d_k - \|g_k\|^2. \tag{6.93}$$

Dividing both sides of (6.93) by $\left(g_k^T d_k\right)^2$ and using (6.64) and (6.82), we obtain

$$\begin{aligned}\frac{\|d_k\|^2}{\left(g_k^T d_k\right)^2} &= \frac{\beta_k^2 \|d_{k-1}\|^2}{\left(g_k^T d_k\right)^2} + \frac{2}{\bar{r}_k \|g_k\|^2} - \frac{1}{\bar{r}_k^2 \|g_k\|^2} \\ &= \frac{(s_k\beta_k^*)^2 \|d_{k-1}\|^2}{\left(g_k^T d_k\right)^2} + \frac{1}{\|g_k\|^2}\left[1 - (1 - \frac{1}{\bar{r}_k})^2\right]. \end{aligned} \tag{6.94}$$

In addition, by the definition (6.64) of $\bar{r}_k$, the relation (6.3) and (6.82), we get

$$\bar{r}_k \|g_k\|^2 = -g_k^T d_k = \|g_k\|^2 - s_k\beta_k^* g_k^T d_{k-1},$$

the above relation and the definition (6.65) imply that

$$s_k\beta_k^* = \frac{(1 - \bar{r}_k)}{l_{k-1}(g_{k-1}^T d_{k-1})} \|g_k\|^2. \tag{6.95}$$

Relation (6.94) and (6.95), we obtain

$$\frac{\|d_k\|^2}{\left(g_k^T d_k\right)^2} = \frac{(1 - \bar{r}_k)^2 \|d_{k-1}\|^2}{\bar{r}_k^2 l_{k-1}^2 (g_{k-1}^T d_{k-1})^2} + \frac{1}{\|g_k\|^2}\left[1 - (1 - \frac{1}{\bar{r}_k})^2\right]. \tag{6.96}$$

Denote

$$m_k = \frac{1 - \bar{r}_k}{\bar{r}_k l_{k-1}}, \tag{6.97}$$

where $l_{k-1} \neq 0$. Now we prove that

$$|m_k| \leq 1, \qquad \text{for all } k \geq 2. \tag{6.98}$$

the right hand side of (6.97) is a function of l_{k-1} and $\bar{r}_k$, which can be denoted as $h(l_{k-1}, \bar{r}_k)$. We can get by (6.87), (6.91) and lemma 6.5 that

$$\begin{aligned} m_k &\leq \max\left\{h(\sigma, \bar{r}_k), h(-\sigma, \bar{r}_k)\right\} \\ &< \max\left\{h(\sigma, (1-\sigma)^{-1}), h(-\sigma, (1-\sigma)^{-1})\right\} = 1. \end{aligned}$$

Thus we have that

$$\begin{aligned} m_k &\geq \min\{h(-\sigma, \bar{r}_k), h(\sigma, \bar{r}_k)\} \\ &> \min\left\{h(-\sigma, (1-\sigma)^{-1}), h(\sigma, (1-\sigma)^{-1})\right\} = -1. \end{aligned}$$

Therefore (6.98) holds for all $k \geq 2$.

By (6.98) and (6.96), we obtain

$$\frac{\|d_k\|^2}{(g_k^T d_k)^2} \leq \frac{\|d_{k-1}\|^2}{(g_{k-1}^T d_{k-1})^2} + \frac{1}{\|g_k\|^2}. \tag{6.99}$$

Because $\|d_1\|^2 \diagup \left(g_1^T d_1\right)^2 = 1 \diagup \|g_1\|^2$, (6.99) shows that

$$\frac{\|d_k\|^2}{(g_k^T d_k)^2} \leq \sum_{i=1}^{k} \frac{1}{\|g_i\|^2},$$

for all k. Then we get from this that

$$\frac{\|d_k\|^2}{(g_k^T d_k)^2} \geq \frac{\gamma}{k},$$

which implies that

$$\sum_{k\geq 1} \frac{\left(g_k^T d_k\right)^2}{\|d_k\|^2} = +\infty. \tag{6.100}$$

This contradicts the Zoutendijk condition (6.21). Therefore (6.37) holds. □

5 Numerical results

In this section, we will test the following four conjugate gradient algorithms :

PRP^{SW} : the PRP method with the strong Wolfe conditions, where $\delta = 10^{-2}$ and $\sigma = 0, 1$.

PRP$_+^{SW}$: the PRP method with nonnegative values of $\beta_k = max\left\{0, \beta_k^{PRP}\right\}$ and

the strong Wolfe conditions, where $\delta = 10^{-2}$ and $\sigma = 0,1$.

NFCG^{SW} : Algorithm 6.1 with the Wolfe conditions (6.4) and (6.9), where the scalars σ_1 and σ_2 satisfy the condition (6.67), in addition, $\delta = 10^{-2}$, $\sigma_1 = \sigma_2 = \sigma = 0,1$, $\lambda_k = \lambda = 0,5$ and $\mu_k = \mu = 0,4$.

NFCG^{W} : Algorithm 6.1 with the standard Wolfe conditions, where $\delta = 10^{-2}$, $\sigma = 0,1$, $\lambda_k = \lambda = 0,5$ and $\mu_k = \mu = 0,5$.

In this paper, all codes were written in Matlab and run on PC with 3.0 GHz CPU processor and 1GB RAM memory and Linux operation system. During our experiments, the strategy for the initial step length is to assume that the first-order change in the function at iterate x_k will be the same as that obtained at the previous step [15]. In other words, we choose the initial guess α_0 satisfying :

$$\alpha_0 = \alpha_{k-1}\frac{\Psi_{k-1}}{\Psi_k} \qquad \forall k > 1,$$

where $\Psi_k = \nabla f_k^T d_k$, when $k = 1$, we choose $\alpha_0 = \dfrac{1}{\|\nabla f(x_1)\|}$. In the case when an uphill search direction does occur, we restart the algorithm by setting $d_k = -g_k$, but this case never occurs for NFCG^{SW} and NFCG^{W}. We stop the iteration if the inequality $\|g(x_k)\| < 10^{-6}$ is satisfied. The iteration is also stopped if the number of iteration exceeds 10000, but we find that this never occurs for our tested problems. The test problems we used are described in Moré et al [17]. Each problem was tested with various values of n changing from $n = 100$ to $n = 1000$.

Tables 5.1 list numerical results. The meaning of each column is as follows :

"N"	the number of the test problem.
"Problem"	the name of the test problem.
"n"	the dimension of the test problem.
"NI"	the number of iterations.
"NF"	the total number of function evaluations.
"NG"	the total number of gradient evaluations.
"CPUtime(s)"	the total CPU time in seconds which should be taken to compute all of these problems.

Table 5.1

Test results on PRP^{SW}/ PRP$^{SW}_{+}$/ NFCG^{SW}/ NFCG^{W} methods

N	Problem	n	PRP^{SW}	PRP$^{SW}_{+}$	NFCG^{SW}	NFCG^{W}
			NI/NF/NG	NI/NF/NG	NI/NF/NG	NI/NF/NG
1	Extended Rosenbrock	500	23/116/6	34/154/102	26/98/61	33/112/77
		1000	18/79/31	19/97/27	46/191/115	38/130/75
2	Extended Powell singular	500	312/568/464	75/177/140	85/198/163	80/179/136
		1000	156/324/251	48/113/102	91/229/188	87/208/161
3	Penalty 1	500	29/146/103	60/208/140	68/270/177	92/263/183
		1000	30/147/113	72/275/167	35/146/111	74/224/148
4	Penalty 2	250	855/2034/1769	873/1747/1217	118/287/156	147/325/241
		100	80/214/163	89/251/190	112/313/237	106/287/166
5	Variably dimensioned	100	18/62/51	18/62/51	18/62/51	18/62/51
		200	19/68/56	19/68/56	19/68/56	19/68/56
6	Trigonometric	100	49/138/101	46/129/100	52/146/102	52/106/67
		200	46/130/99	48/140/1107	52/156/111	59/130/72
7	Discrete boundary value	500	1246/2021/2014	1246/2021/2014	130/254/231	115/182/152
		1000	121/173/171	121/173/171	24/44/42	21/34/28
8	Discrete integral equation	500	3/11/4	3/11/4	3/11/4	3/11/4
		1000	3/11/4	3/11/4	3/11/4	3/11/4
CPU time(s)			1.3087e+2	1.3504e+2	1.1321e+2	1.1041e+2

We can see from the above table that, the average performances of the NFCG^{SW} is better than that of the PRP method. Also see that, the NFCG^{W} outperforms other three algorithms for solving these problems, especially for problems 6, 7 and 8. Table 5.1, shows the performance of these methods relative to CPU time. To solve all the 16 problems, the CPU time (in seconds) required by the PRP^{SW}, PRP$^{SW}_{+}$, NFCG^{SW} and NFCG^{W} are 130.87, 135.04, 113.21 and 110.41 respectively. These preliminary results obtained are encouraging.

Conclusion et perspectives

Dans cette thèse, nous avons proposé une famille à deux paramètres des méthodes du gradient conjugué non linéaire et étudié la convergence globale de cette méthode. Cette famille comprend non seulement les trois méthodes du gradient conjugué déjà connues, mais aussi une autre famille des méthodes du gradient conjugué comme sous-famille. Tout d'abord, nous pouvons voir que la propriété de descente de la direction joue un rôle important à établir des résultats généraux de convergence de cette méthode avec la recherche linéaire de Wolfe faible (6.4) et (6.11), même en l'absence de la condition de descente suffisante (6.53), à savoir, les théorèmes 6.1, 6.2 et 6.3. Ensuite, d'après le théorème 6.4, nous avons prouvé que la famille à deux paramètres peut assurer une direction de descente à chaque itération et converge globalement sous condition de la recherche linéaire (6.4) et (6.9) où les scalaires σ_1 and σ_2 satisfont à la condition (6.67). D'après le théorème 6.5, nous avons étudié les méthodes liées à la méthode (6.57). Notons que la taille de β_k dépend du valeur s_k puisque $\beta_k = s_k\beta_k^*$ (méthode hybride). Si τ_k et s_k appartiennent à un certain intervalle, à savoir, (6.85) et (6.86), respectivement , les méthodes correspondantes sont indiquées pour produire une direction de descente à chaque itération et assurer la convergence globale avec la recherche linéaire de Wolfe forte (6.4) et (6.9) où le scalaire $\sigma \leq \frac{1}{2}$. De plus, nos résultats numériques montrent que cette nouvelle méthode du gradient conjugué non linéaire, à savoir, la méthode NFCG^W non seulement converge globalement, mais aussi plus performante que la méthode de Polak-Ribière-Polyak (PRP). Les résultats, nous l'espérons, peuvent stimuler une étude plus approfondie sur la théorie spécialement sur les méthodes du gradient conjugué avec la recherche linéaire de Wolfe. En ce qui concerne les recherches futures, nous devons étudier la performance pratique de la nouvelle méthode. En perspectives, nous présentons plusieurs projets se situant directement dans la continuité des études effectuées dans le cadre de cette thèse. Voici quelques perspectives.

1- Dans tous les théorèmes dans le dernier chapitre, la convergence globale

$$\lim_{k\to\infty} \inf \|g_k\| = 0,$$

a été obtenue avec

$$\lambda \in (0, 1].$$

Peut-on généraliser ces résultats pour

$$\lambda < 0 \quad ou \quad \lambda > 1.$$

2- Peut-on définir et unifier plus de trois méthodes dans une même classe et étudier par la suite leurs propriétés ensemble comme il a été fait avec la famille à deux paramètres des méthodes du gradient conjugué non linéaire ?

3- Peut-on obtenu les mêmes résultats si on applique une autre recherche linéaire ?

BIBLIOGRAPHIE

[1] 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.

[2] M. Al-Baali. *New property and global convergence of the Fletcher-Reeves method with inexact line searches*. IMA J. Numer. Anal. 5(1985) 122-124.

[3] Y.H. Dai and Y. Yuan (1999). *A nonlinear conjugate gradient with a strong global convergence property*. SIAM J. Optimization, Vol. 10(1) pp.177-182.

[4] Y. H. Dai and Y. Yuan. *Convergence properties of the Fletcher-Reeves method*. IMA J. Numer. Anal. 16 (1996) 155-164.

[5] Y. H. Dai. Dai Y. Yuan Y. A three-parameter family of nonlinear conjugate gradient methods. J. Comp. Math. 2001 ;70 :1155–1167.

[6] Y. H. Dai and Y. Yuan. *Convergence properties of the conjugate descent method*. Mathematical Advances Vol. 25 No. 6 (1996), 552-562. CMP 97 :13

[7] Y. H. Dai and Y. Yuan. *A class of globally convergent conjugate gradient methods*. Sci. China Series. Math. Ser. A. 2003 ;46 :251–261.

[8] Y. H. Dai. J. Y. Han. G. H. Liu, D. F. Sun, H. X. Yin, and Y. Yuan. *Convergence properties of nonlinear conjugate gradient methods*. SIAM J. Optim. 2000 ;10 :345–358.

[9] Y. H. Dai and Y. Yuan. *Some properties of a new conjugate gradient method.* Advances in Nonlinear Programming (Kluwer, Boston, 1998), pp. 251-262.

[10] J. W. Daniel. *The conjugate gradient method for linear and nonlinear operator equations.* SIAM J. Numer. Anal., 4 (1967), 10-26.

[11] R. Fletcher. *Practical methods of optimization.* Chichester : JohnWiley ; 1987.

[12] R. Fletcher and C. Reeves (1964). *Function minimization by conjugate gradients.* Comput. J., 7, pp.149-154.

[13] 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.

[14] 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.

[15] Y. F. Hu and C. Storey. *Global convergence result for conjugate gradient methods.* J. Optim. Theory Appl. Vol. 71 No. 2 (1991) 399-405.

[16] Y. Liu and C. Storey. *Efficient Generalized Conjugate Gradient Algorithms.* Part 1 : Theory, Journal of Optimization Theory and Applications, Vol. 69 (1991), 129-137.

[17] J.J. Moreë. B.S. Garbow. K.E. Hillstrome. *Testing unconstrained optimization software.* ACM Trans. Math. Software 7 (1981) 17-41.

[18] J. Nocedal. S.J. Wright. *Numerical optimization.* Springer Series in Operations Research, Springer, New York, 1999.

[19] E. Polak and G. Ribière. *Note sur la convergence de directions conjugées.* Rev. Francaise Informat Recherche Operationelle, 3e Ann ee 16 (1969), pp. 35-43.

[20] B. T. Polyak. *The conjugate gradient method in extreme problems.* USSR Comp. Math. and Math. Phys. 9 (1969), pp. 94-112.

[21] D. Pu and W. Yu. *On the convergence properties of the DFP algorithms.* Annals of Operations Research 24 (1990) 175-184.

[22] D. F. Shanno. *Conjugate gradient methods with inexact searches.* Math. Oper. Res. 3 (1978), 244-256.

[23] D. Touati-Ahmed and C. Storey (1990). *Efficient hybrid conjugate gradient techniques.* JOTA, 64, pp. 379-397.

[24] P.Wolfe. *Convergence conditions for ascent methods.* SIAM Rev. 11 (1969) 226–235.

[25] G. Zoutendijk. *Nonlinear Programming.* Computational Methods, in : J. Abadie (Ed.), Integer and Nonlinear Programming, North-Holland, Amsterdam, 1970, pp.37–86.

[26] T. F. Coleman and P. A. Fenyes (1992). *Partitioned quasi-Newton methods for nolinear constrained optimization.* Mathematical programming Study, 53, 17-44.

[27] S. Badreddine .L. Yamina and B. Rachid. *A new two-parameter family of nonlinear conjugate gradient methods.* GOPT, 2013.