الجمهوريـة الجزائـريـة الديمقراطيـة الشعبيـة
République Algérienne Démocratique et Populaire
وزارة التعليـم العالـي والبحـث العلمـي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

N° Réf :...............

Centre Universitaire de Mila

Institut des Sciences et de la Technologie **Département de Mathématique et informatique**

Mémoire préparé En vue de l'obtention du diplôme de licence

en filière : Mathématiques Fondamentales

METHODE DE GRADIENT AVEC PROJECTION POUR OPTIMISATION SOUS CONTRAINTE

Préparé par : Hemimes Somia
Hamideche Selma
Belmrabet Rofia
Lahmari Safia

Encadré par M^{me} :BENAOUICHA Loubna

Année universitaire : 2013/2014

Table des matières

Introduction Générale

L'optimisation joue un rôle important en recherche opérationnelle (donc en économie et microéconomie), dans les mathématiques appliquées (fondamentales pour l'industrie et l'ingénierie), en analyse et en analyse numérique, en statistique pour l'estimation du maximum de vraisemblance d'une distribution, pour la recherche de stratégies dans le cadre de la théorie des jeux, ou encore en théorie du contrôle et de la commande.

Le problème d'optimisation consiste alors à déterminer les variables de décision conduisant aux meilleures conditions de fonctionnement du système (ce qui revient à minimiser ou maximiser).

Dans ce mémoire, nous entrons dans le vif du sujet en nous intéressant à la résolution de problèmes d'optimisation sans contrainte, puis avec contraintes. Pour chacun de ces problèmes nous tenterons de répondre aux questions suivantes :

1. Existe-t-il une solution du problème considéré ? si oui, a-t-on unicité ?
2. Comment la caractériser ? (conditions d'optimalité).
3. Comment la calculer ? Quel type d'algorithme choisir ?

Dans le premiere chapitre, nous avons données une définition d'un probléme d'optimisation, rappeles et notion de calcul différentiel.

Dans le deuxième chapitre, nous avons présenté l'optimisation sans contraintes en particulier les conditions d'optimalité.

Dans le troixième chapitre, nous avons présenté l'optimisation sous contraintes, ces définition et condition d'optimalité.

Dans le dernière chapitre nous avons données une présentation génerale de la programmation quadratique, nous avons etudié la méthode de descente. Ensuit la mèthode du gradient avec projection et nous terminons par la méthode du gradient projeté.

Chapitre 1

Définition et rappels de calcul différentiel

1.1 Définition de problème d'optimisation

1.1.1 Problème d'optimisation

On définit un problème d'optimisation avec contraintes comme suit :

$$\begin{cases} min\ f(x) \\ x \in C \end{cases} \tag{p}$$

Oú $f : \mathbb{R}^n \to \mathbb{R}$ continue, $C \subseteq \mathbb{R}^n$: ensemble des contraintes.
Si $C = \mathbb{R}^n$, (p) est appelé problème d'optimisation sans contraintes.

1.1.2 Problème mathématique

Un problème mathématique (PM) est un problème d'optimisation, dans lequel l'ensemble des contraintes C est exprimé essentiellement par des inégalités, (et/ou) des égalités, c'est -à- dire C est de la forme :

$$C = \{x \ \in \ \mathbb{R}^n / f_i(x) \leq 0, i = 1, ..., n,\ g_j(x) = 0,\ j = 1...m\}$$

et

$$f_i,\ g_j : \mathbb{R}^n \ \to \mathbb{R}.$$

1.1.3 Solution optimale

Solution réalisable :

Un point $x^* \in C$ (c'est -à-dire vérifiant les contraintes de (p)) est appelé solution réalisable de (p).

Solutiolon optimal globale :

Une solution réalisable qui minimise f sur C est appelée solution optimale globale de (p), nous la noterons x^*.

$$L'ensemble\ des\ solutions\ globales\ est\ noté :\ \arg\min_{C} f(x)$$

Solution optimale local :

Un point $x^* \in C$ est une solution optimale local de (p) s'il existe un voisinage V de x^* tel que :

$$f(x^*) \leq f(x)\,;\ \forall\ x \in V$$

On note par :

$$\log\min_{C} f(x)\,(\ L'ensemble\ des\ solution\ optimales\ local\ de\ (p))\,.$$

$$Nous\ avons\ toujours\ \arg\min_{C} f(x) \subseteq \log\min_{C} f(x)\,.$$

Si le probléme (p) est convexe les deux ensenbles sont égaux.

1.2 Rappels et notion de calcul différentiel

Définition 1.2.1 *Soient E et F des espaces vectoriels normés, f une application de E dans F et $x \in E$. On dit que f est différentiable en x s'il existe $T \in L(E,F)$ (où $L(E,F)$ est l'ensemble des applications linéaires de E dans F) telle que :*

$$f(x+h) = f(x) + T(h) + \parallel h \parallel \varepsilon(h)$$

$$avec \varepsilon(h) \rightarrow 0 quand h \rightarrow 0.$$

L'application T est alors unique et on note :

$$Df(x) = T \in L(E,F)\,.$$

Remarque 1.2.2 *On peut remarquer qu'en dimention infinie. T dépend des normes associées à E et F.*

Voyons maintenant quelques cas particuliers d'espace E et F.

•Cas où $E = \mathbb{R}^n$ et $F = \mathbb{R}^p$

Soit : $f : \mathbb{R}^n \to \mathbb{R}^p$, $x \in \mathbb{R}$ et supposons que f est différentiable en x ;
alors $Df(x) \in L(E, F)$ et il existe $A(x) \in M_{p,n}(\mathbb{R})$
tel que $\underbrace{Df(x)(y)}_{\in\mathbb{R}^p} = \underbrace{Ay}_{\in\mathbb{R}^p}, \forall y \in \mathbb{R}^n$.

On confond alors l'application linéaire

$$Df(x) \in L(\mathbb{R}^n, \mathbb{R}^p)$$

et la matrice

$$A(x) \in M_{p,n}(\mathbb{R})$$

qui la représente.

On écrit donc :

$$A(x) = Df(x) = (a_{ij})_{1 \le i \le p,\ 1 \le j \le n}, \text{ où } a_{ij} = \partial_j f_i(x),$$

∂_j désignant la dérivée partielle par rapport à la j-ème variable.

Exemple 1.2.3 *Prenons $n = 3$ et $p = 2$, notons $x = (x_1, x_2, x_3)^t$ et considérons $f : \mathbb{R}^3 \to \mathbb{R}^2$ définie par :*

$$f(x) = \begin{pmatrix} x_1^2 + x_2^3 + x_3^4 \\ 2x_1 - x_2 \end{pmatrix}$$

On vérifiera par le calcul que pour $h = (h_1, h_2, h_3)^t$, on a :

$$Df(x) = \begin{pmatrix} \frac{\partial f(x_1,x_2,x_3)}{\partial x_1} & \frac{\partial f(x_1,x_2,x_3)}{\partial x_2} & \frac{\partial f(x_1,x_2,x_3)}{\partial x_3} \\ \\ \frac{\partial f(x_1,x_2,x_3)}{\partial x_1} & \frac{\partial f(x_1,x_2,x_3)}{\partial x_2} & \frac{\partial f(x_1,x_2,x_3)}{\partial x_3} \end{pmatrix}$$

donc :

$$Df(x) = \begin{pmatrix} 3x_1 & 3x_2^2 & 4x_3^3 \\ 2 & -1 & 0 \end{pmatrix}$$

$$Df(x)h = \begin{pmatrix} 2x_1h_1 + 3x_2^2h_2 + 4x_3^3h_3 \\ 2h_1 - h_2 \end{pmatrix}$$

et donc, avec les notations précédentes :

$$A(x) = \begin{pmatrix} 2x_1 & 3x_2^2 & 4x_3^3 \\ 2 & -1 & 0 \end{pmatrix}$$

•Cas où $E = \mathbb{R}^n$, $F = \mathbb{R}$ c'est un sous-cas du paragraphe précédent, puis qu'on est ici dans le cas $p = 1$.

Soit $x \in \mathbb{R}^n$ et f une fonction de E dans F différentiable en x ; on a donc :

$$Df(x) \in M_{1,n}(\mathbb{R}),$$

et on peut définir le gradient de f en x par :

$$\nabla f(x) = (Df(x))^t \in \mathbb{R}^n.$$

Pour $(x, y) \in (\mathbb{R}^n)^2$, on a donc

$$Df(x)y = \sum_{j=1}^{n} \partial_j f(x)\, y_j = \nabla f(x) . y \text{ où } \nabla f(x) = \begin{pmatrix} \partial_1 f(x) \\ \vdots \\ \vdots \\ \partial_n f(x) \end{pmatrix} \in \mathbb{R}^n$$

Définition 1.2.4 ***(matrice hessinne)***

Soit : $f : \mathbb{R}^n \to \mathbb{R}$ une fonction deux fois différentiable. La fonction notée

$$\nabla^2 f(x) : \mathbb{R}^n \to \mathbb{R}^{n*n}$$

est appelée matrice hessienne de f, définie par :

$$\nabla^2 f(x) = \begin{pmatrix} \partial^2 f(x)/\partial x_1^2 & \partial^2 f(x)/\partial x_1 \partial x_2 & \dots & \partial^2 f(x)/\partial x_1 \partial x_n \\ \partial^2 f(x)/\partial x_2 \partial x_1 & \partial^2 f(x)/\partial x^2 & \dots & \partial^2 f(x)/\partial x_2 \partial x_n \\ \vdots & \vdots & \ddots & \vdots \\ \partial^2 f(x)/\partial x_n \partial x_1 & \partial^2 f(x)/\partial x_n \partial x_2 & \dots & \partial^2 f(x)/\partial x_n^2 \end{pmatrix}$$

La matrice hessienne est toujours symétrique.

Chapitre 2

Optimisation sans contraintes

2.1 Présentation du problème

Soit $f \in C(E, \mathbb{R})$ et E un espace vecroriel normé. On cherche soit un minimum global de f, c'est - à-dire : $x^* \in E$ tel que :

$$f(x^*) \leq f(y), \ \forall y \in E \tag{1}$$

où un minimum local, c'est- à- dire : $x^* \in E$ tel que : $\exists\ \alpha > 0$,

$$f(x^*) \leq f(y), \forall y \in B(x^*, \alpha) \tag{2}$$

2.2 Condition d'optimalité

Proposition 2.2.1 *(Condition nécessaire d'optimalité)*

Soit E un espace vctoriel normé, et soient $f \in C(E, \mathbb{R})$, et $x^* \in E$ tel que f est différentiable en x^*.

Si x^* est solution de (2) alors :

$$Df(x^*) = 0.$$

Démonstration :

Supposons qu'il existe $\alpha > 0$ tel que :

$$f(x^*) \leq f(y)$$

pour tout $y \in B(x^*, \alpha)$. Soit $z \in E \setminus \{0\}$, alors si $\mid t \mid < \frac{\alpha}{\|z\|}$ on a : $\frac{-\alpha}{\|z\|} < t \leq \frac{\alpha}{\|z\|}$

$$x^* + tz \in B(x^*, \alpha)$$

(où $B(x^*, \alpha)$ désigne la boule ouverte de centre x^* et de rayon α) et on a donc

$$f(x^*) \leq f(x^* + tz)$$

comme f est diffirentiable en x^*, on a :

$$f(x^* + tz) = f(x^*) + Df(x^*)(tz) + \mid t \mid \varepsilon_z(t)$$

où $\varepsilon_z(t) \to 0$ lorsque $t \to 0$. On a donc

$$f(x^*) + Df(x^*)(z) + \mid t \mid \varepsilon_z(t) \geq f(x^*).$$

Et pour $\frac{\alpha}{\|z\|} > t > 0$, on a :

$$Df(x^*)(z) + \varepsilon_z(t) \geq 0.$$

En faisant tendre t vers 0, on obtient que

$$Df(x^*)(z) \geq 0, \forall\, z \in E.$$

On a aussi

$$Df(x^*)(-z) \geq 0, \forall\, z \in E,$$

et donc :

$$-Df(x^*)(z) \geq 0, \forall\, z \in E$$

alors :

$$Df(x^*)(z) \leq 0$$

On en conclut que :

$$Df(x^*) = 0$$

Théorème 2.2.2 *(Condition suffisantes d'optimalité)*

Soit une fonction $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ *deux fois différentiable dans un sous-ensemble ouvert* V *de* $\mathbb{R}^n$ *et soit* $x^* \in V$ *qui vérifie les conditions :*

$$\nabla f(x^*) = 0 \ et \ \nabla^2 f(x^*) \ est\ définie\ positive.$$

Alors : x^* *est un minimum local de* f.

2.3 Rappels sur la convexité

Proposition 2.3.1 *Soit* E *un espace vectoriel normé (sur* $\mathbb{R}$*) et* $f \in C^1(E, \mathbb{R})$ *alors :*

1. f *convexe si et seulement si :*

$$f(y) \geq f(x) + Df(x)(y - x), \forall x,\ y \in E.$$

2. f *est strictement convexe si et seulement si :*

$$f(y) > f(x) + Df(x)(y - x), \forall x, y \in E \ et \ x \neq y.$$

Proposition 2.3.2 *(1ére caractérisation de la convexité)*

Soit E *un espace vectoriel et* $f : E \to \mathbb{R}$*. On dit que* f *est convexe si :*

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

pour tout $(x, y) \in E^2$ *tel que.* $x \neq y$ *et* $t \in [0, 1]$.

On dit que f *est strictement convexe si :*

$$f(tx + (1-t)y) < tf(x) + (1-t)f(y)$$

pour tout $(x, y) \in E^2$ *tell que :* $x \neq y$ *et* $t \in]0, 1[$.

Proposition 2.3.3 *(2 ème caractérisation de la convexité)*

Soit $E = \mathbb{R}^n$ et $f \in C^2(E, \mathbb{R})$. Soit $H_f(x)$ la matrice hessienne de f au point x, i.e :

$$(H_f(x))_{i,j} = \partial_{i,j}^2 f(x).$$

Alors :

1. f est convexe si et seulement si $H_f(x)$ est symétrique et positive pour tout $x \in E$ (c'est-à-dire $H_f(x)^t = H_f(x)$ et $H_f(x) y.y \geq 0$ pour tout $y \in \mathbb{R}^n$)

2. f est strictement convexe si $H_f(x)$ est symétrique définie positive pour tout $x \in E$.

(Attention la réciproque est fausse .)

Proposition 2.3.4 *(Caractérisation des points tels que $f(x^*) = \min\limits_E f$)*

Soit E espace vectoriel normé et f une fonction de E dans $\mathbb{R}$.

On suppose que $f \in C^1(E, \mathbb{R})$ et que f est convexe. Soit $x^* \in E$.

Alors $(f(x^*) = \min\limits_E f) \Longleftrightarrow Df(x^*) = 0$

En particulier si $E = \mathbb{R}^n$ alors :

$$f(x^*) = \min_{x \in \mathbb{R}^n} f(x) \Longleftrightarrow \nabla f(x^*) = 0$$

2.4 Résultas d'existance et d'unicité

Théorème 2.4.1 *(Existence)*

Soit $E = \mathbb{R}^n$ et $f : E \to \mathbb{R}$ une application telle que :

(i) f est continue.

(ii) $f(x) \to +\infty$ quand $\| x \| \longrightarrow +\infty$.

Alors il existe $x^* \in \mathbb{R}^n$ tel que :

$$f(x^*) \leq f(y), \text{ pour tout } y \in \mathbb{R}^n.$$

Théorème 2.4.2 *(Condition suffisante d'unicité)*

Soit E un espace vectoriel normé et $f : E \to \mathbb{R}$ strictement convexe alors il existe au plus un $x^* \in E$ tel que :

$$f(x^*) \leq f(y), \ \forall\ y \in E.$$

Démonstration

Soit f strictement convexe, supposons qu'il existe x^* et $x^{**} \in E$ tel que :

$$f\left(x^*\right) = f\left(x^{**}\right) = \min_{\mathbb{R}^n} f.$$

Comme f est strictement convexe, si $x^* \neq x^{**}$alors :

$$f\left(\frac{1}{2}x^* + \frac{1}{2}x^{**}\right) < \frac{1}{2}f\left(x^*\right) + \frac{1}{2}f\left(x^{**}\right) = f\left(x^*\right) = \min_{R^n} f,$$

Ce qui est impossible donc :

$$x^* = x^{**}$$

Théorème 2.4.3 *(Existence et unicité)*

Soit $E = \mathbb{R}^n$, et soit $f : E \to \mathbb{R}$.

On suppose que :

(i) f continue.

(ii) $f\left(x\right) \to +\infty$ quand $\| x \| \to +\infty$.

(iii) f est strictement convexe.

Alors il existe un unique $x^* \in \mathbb{R}^n$ tel que :

$$f\left(x^*\right) = \min_{\mathbb{R}^n} f.$$

Contre-exemple

Pour montrer que la réciproque de 2 est fausse, on propose le contre-exemple suivant :

Soit $n = 1$ et $f \in C^2\left(\mathbb{R}, \mathbb{R}\right)$, on a alors :

$$H_f\left(x\right) = f''\left(x\right).$$

Si f est la fonction définie par :

$$f\left(x\right) = x^4,$$

alors f est strictement convexe car :

$$f''\left(x\right) = 12x^2 \geq 0, \forall x \in \mathbb{R},$$

mais $f''\left(0\right) = 0$.

Chapitre 3

Optimisation sous contraintes

3.1 Présentation du problème

Soit $E = \mathbb{R}^n$, soit $f \in C(E, \mathbb{R})$, et soit K un sous ensemble de E. On s'intéresse à la recherche de $x^* \in K$ tel que :

$$\begin{cases} x^* \in K \\ f(x^*) = \min_K f \end{cases} \tag{3}$$

Ce problème est un problème de minimisation sous contrainte au sens où l'on cherche x qui minimise f en astreignant x a être dans K. Voyons quelques exemples de ces contraintes (définition par l'ensemble K), qu'on va expliciter à l'aide des p fonctions continues, $g_i \in C(E, \mathbb{R})$, $i = 1, \ldots, p$.

3.2 Définitions

3.2.1 *Contraintes égalités*

On pose

$$K = \{x \in E,\ g_i(x) = 0,\ i = 1, \ldots, p\}.$$

On verra plus loin que le problème de minimisation de f peut alors être résolu grace au théorème des multiplicateurs de Lagrange (voir la page 14).

3.2.2 *Contraintes inégalités*

On pose

$$K = \{x \in E,\ g_i(x) \leq 0,\ i = 1 \ldots p\}.$$

On verra plus loin que le problème de minimisation de f peut alors être résolu grace au théorème de kuhn-Tucker (voir la page 14).

3.2.3 *Contraintes d'égalités et d'inégalité*

L'ensemble des contrainte est donné par :

$$K = \{x \in E,\ g_i(x) = 0\ ,\forall\ i = 1, \ldots, p \text{ et } g_i(x) \leq 0\ ,\forall\ i = 1, \ldots, p\}.$$

3.2.4 Programmation linéaire

Avec un tel ensemble de contraintes K, si de plus f est linéaire, c'est-à-dire qu'il existe $b \in \mathbb{R}^n$ tel que :

$$f(x) = b.x\ ,$$

et les fonctions g_i sont affines, c'est-à-dire qu'il existe $b_i \in \mathbb{R}^n$ et $c_i \in \mathbb{R}$ tel que :

$$g_i(x) = b_i.x + c_i\ ,$$

alors on dit qu'on a affaire à un problème de "programmation linéaire".

3.2.5 Programmation quadratique

Avec le même ensemble de contraintes K, si de plus f est quadratique, c'est-à-dire si f est de la forme :

$$f(x) = \frac{1}{2}Ax.x - b.x,$$

et les fonction g_i sont affines, alors on dit qu'on a affaire à un probléme de "programmation convexe".

3.2.6 *Programmation convexe*

Dans le cas où f est convexe et K est convexe, on dit qu'on a affaire un problème de "programmation convexe".

3.3 Conditions d'optimalité :

3.3.1 Contraintes égalité :

Dans tout ce paragraphe, on considérera les hypothéses et notations suivantes :
$f \in C(\mathbb{R}^n, \mathbb{R})$, $g_i \in C^1(\mathbb{R}^n, \mathbb{R})$, $i = 1...p$;
$K = \{x \in \mathbb{R}^n,\ g_i(x) = 0,\ \forall\ i = 1...p\}$;
$g = (g_1, ..., g_p)^t \in C^1\left(\mathbb{R}^n, \mathbb{R}^P\right)$.

Remarque 3.3.1 *(Quelques rappels de calcul différentiel)*
Comme : $g \in C^1(\mathbb{R}^n, \mathbb{R}^p)$, *si* $x \in \mathbb{R}^n$, *alors* $Dg(x) \in L(\mathbb{R}^n, \mathbb{R}^p)$, *ce qui revient à dire, en confondant l'aplication linéaire* $dg(x)$ *avec sa matrice, que* $dg(x) \in M_{p,n}(\mathbb{R})$.
Par définition, $\operatorname{Im}(Dg(x)) = \{Dg(x)z, z \in \mathbb{R}^n\} \subset \mathbb{R}^p$.

Théorème 3.3.2 *(Multiplicateurs de lagrange) :*
Soit $x^* \in K$ *tel que* $f(x^*) = \min_K f$. *On suppose que* f *est différentiable en* x *et* $\dim(\operatorname{Im}(Dg(x^*))) = p$ *(ou rang* $(Dg(x^*)) = p$*), alors : il existe* $(\lambda_1, \ldots, \lambda_p)^t \in \mathbb{R}^p$ *tel que :*

$$\nabla f(x^*) + \sum_{i=1}^{p} \lambda_i \nabla g_i(x^*) = 0.$$

3.3.2 Contraintes inégalités

Soit $f \in C(\mathbb{R}^n, \mathbb{R})$ et $g_i \in C^1(\mathbb{R}^n, \mathbb{R})\, i = 1, \ldots, p$. On considère maintenant un ensemble K de la forme : $K = \left\{x \in \mathbb{R}^n,\ g_{i(x)} \leq 0\ i = 1, \ldots, P\right\}$ et on cherche à résoudre le problème de minimisation qui s'écrit :

$$\begin{cases} x^* \in K \\ f(x^*) \leq f(x),\ \forall x \in K \end{cases}$$

Théorème 3.3.3 *(Kuhn- Tucker) :*
Soit $f \in C(\mathbb{R}^n, \mathbb{R})$, *soit* $g_i \in C^1(\mathbb{R}^n, \mathbb{R})$, *pour* $i = 1, \ldots, p$, *et soit :* $K = \{x \in \mathbb{R}^n,\ g_i(x) \leq 0\ \forall\ i = 1$
On suppose qu'il existe x^* *solution de (3),et on pose* $I(x^*) = \{i \in \{1, \ldots, p\}; |\ g_i(x^*) = 0\}$.
On suppose que f *est différentiable en* x^* *et que la famille (de* $\mathbb{R}^n$*)* $\{\nabla g_i(x^*),\ i \in I(x^*)\}$ *est libre. Alors il existe une famille* $(\lambda_i)_{i \in I(x^*) \subset R_+}$ *tel que :*

$$\nabla f(x^*) + \sum_{i \in I(x^*)} \lambda_i \nabla g_i(x^*) = 0.$$

3.4 Existence-Unicité-conditions d'optimalité simple

Théorème 3.4.1 *(Existence)*

Soit $E = \mathbb{R}^n$, et $f \in C(E, \mathbb{R})$.

1.Si K est un sous-ensemble fermé et borné de E, alors il existe $x^* \in K$ tel que : $f(x^*) = \min\limits_K f$.

2.Si K est un sous-ensemble fermé de E, et si f est croissante à l'infini c'est-à-dire que $f(x) \to +\infty$ quand $|x| \to +\infty$, alors $\exists\ x^* \in K$ tel que :

$$f(x^*) = \min_K f.$$

Démonstration

1. Si K est un sous-ensemble fermé bornè de E, comme f est continue, elle atteint ses bornes sur K, d'où l'existence de x^*.

2. Si f est croissant à l'infini, alors il existe $r > 0$ tel que si $\| x \| > r$ alors :

$$f(x) > f(0);$$

donc

$$\min_K f = \min_{K \cap B_R} f,$$

où B_R désigne la boule de centre 0 et de rayon R. L'ensemble $K \cap B_R$ est compact, car intersection d'un fermé et d'un compact. Donc : par ce qui précède, il existe $x^* \in K$ tel que :

$$f(x^*) = \min_{K \cap B_R} f = \min_{B_R} f.$$

Théorème 3.4.2 *(Unicité)*

Soit $E = \mathbb{R}^n$ et $f \in C(E, \mathbb{R})$. On suppose que f est strictement convexe et que K est convexe. Alors il existe au plus un élément x^*de K tel que :

$$f(x^*) = \min_K f.$$

Théorème 3.4.3 *(Existence et unicité)*

Soit $E = \mathbb{R}^n$, $f \in C(E, \mathbb{R}^n)$ une fonction strictement convexe et K un sous ensemble convexe fermé de E. Si K est borné ou si f est croissante à l'infini, c'est-à-dire si $f(x) \to +\infty$ quand $\| x \| \longrightarrow +\infty$, alors il existe un unique élément x^*de K solution du problème de minimisation (3), i.e. tel que :

$$f(x^*) = \min_K f.$$

Proposition 3.4.4 *(condition simple d'optimalité) :*

Soient $E = \mathbb{R}^n$, $f \in C(E, \mathbb{R})$, et $x^* \in K$ tel que :

$$f(x^*) = \min_K f.$$

On suppose que f est différentiable en x^* :

1. Si $x^* \in K^\circ$ alors $\nabla f(x^*) = 0$. $\left(K^\circ\right.$ l'interieur de l'ensemble $\left.K.\right)$

2. Si K est convexe, alors $\nabla f(x^*)(x - x^*) \geq 0$ pour tout $x \in K$.

Démonstration

1. Si $x^* \in K^\circ$, alors il existe $\varepsilon > 0$ tel que $B(x^*, \varepsilon) \subset K$ et $f(x^*) \leq f(x)$ $\forall\ x \in B(x^*, \varepsilon)$. Alors on a déjà vu que ceci implique $\nabla f(x^*) = 0$.

2. Soit $x \in K$. comme x^* réalisé le minimum de f sur K, on a :

$f(x^* + t(x - x^*)) = f(tx + (1-t)x^*) \geq f(x^*)$ pour tout $t \in]0, 1[$, par convexité de K. On en déduit que :

$\frac{f(x^*+t(x-x^*))-f(x^*)}{t} \geq 0$ pour tout $t \in]0, 1]$.

En passant à la limite lorsque t tend vers 0 dans cette derniére égalité, on obtient :

$$\nabla f(x^*)(x - x^*) \geq 0.$$

Chapitre 4

Méthode du gradient avec projection

4.1 Méthodes de descente :

Définition 4.1.1 *Soient* $f \in C(E, \mathbb{R})$ *et* $E = \mathbb{R}^n$.

1.Soit $x \in E$, *on dit que* $d \in E \setminus \{0\}$ *est une direction de descente en* x *s'il existe* $\rho_0 > 0$ *tel que :*

$$f(x + \rho d) < f(x), \ \forall \rho \in [0, \rho_0].$$

2.Soit $x \in E$, *on dit que* $d \in E \setminus \{0\}$ *est une direction de descente stricte en* x *s'il existe* $\rho_0 > 0$ *tel que :*

$$f(x + \rho d) < f(x), \ \forall \rho \in]0, \rho_0].$$

3.Une "méthode de descente" pour la recherche de x^* *tel que* $f(x^*) = \inf_E f$ *consiste à construire une suite* $(x_n)_n$ *de la manière suivante :*

a)Inisialisation $x_0 \in E$;

b)Itération n *: on suppose* $x_0 \ldots x_n$ *connus* $(n \geq 0)$;

i)On cherche d_n *direction de descente stricte de* x_n.

ii)On prend $x_{n+1} = x_n + \rho_n d_n$ *avec* $\rho_n > 0$ *"bien choisi".*

Proposition 4.1.2 *Soient* $E = \mathbb{R}^n$, $f \in C^1(E, \mathbb{R})$, $x \in E$ *et* $d \in E \setminus \{0\}$; *alors :*

1.Si d *direction de descente en* x *alors* $d.\nabla f(x) \leq 0$.

2.Si $\nabla f(x) \neq 0$ *alors* $d = -\nabla f(x)$ *est une direction de descente stricte en* x.

4.2 Projecton sur un convexe fermé :

Proposition 4.2.1 *Soit E un espace de Hilbert, muni d'une norme $\|.\|$ induite par un produit scalair $(.,.)$ et soit K un convexe fermé non vide de E. Alors, tout $x \in K$ il existe un unique $x_0 \in K$ tel que :*

$$\| x - x_0 \| \leq \| x - y \| \text{ pour tout } y \in K,$$

On not $x_0 = P_k(x)$. La projection orthogonale de x sur K.
On a également :

$$x_0 = p_K(x) \text{ si et seulement si } (x - x_0, x_0 - y) \geq 0. \ \forall y \in K.$$

Dans le cadre des algorithmes de minimisation avec contraintes que nous allons développer maintenant, nous considèrons $E = \mathbb{R}^n$, $f \in C^1(\mathbb{R}^n, \mathbb{R})$ une fonction convexe, et K fermé convexe non vide. On cherche à calculer une solution approchée de x^*, solution du problème (3).

4.3 Algorithme du gradient avec projection :

On considère le problème :

$$\begin{cases} \min f(x) \\ x \in K \end{cases} \tag{p}$$

avec : $\phi \neq K \subset \mathbb{R}^n$ convexe fermé et $f \in \mathbb{C}^1$sur K.
Soit $\rho > 0$ fixé. Ona la condition

$$x^* = proj_K(x^* - \rho \nabla f(x^*))$$

est une condition nécessaire pour que x^* soit une solution optimale de (p) elle est suffisante si f est convexe. (Si la condition est vérifiée, on dit que x^* est un point critique).

4.3.1 Algorithme du gradient à pas fixe avec projection sur K $(GPFK)$:

La méthode est comme suit :

1. Initialisation : on part de $x^* \in C$, on fait $k = 0$.
2. On calcule $y_k = proj_k \left(x_k - \rho \nabla f \left(x_k\right)\right)$.
3. Test d'arrêt : $y_k = x_k$.
4. Sinon on prend pour direction de descente $d_k = y_k - x_k$.
5. On fait une recherche linéaire à partire de x_k dans la direction d_k. On obtient t_k.
6. On fait $x_{k+1} = x_k + \rho t_k d_k$, $k = k + 1$ et on retourne en 2.

Si $d_k = 0$ alors x_k est un point critique et solution optimale dans le cas où f est convexe.

Vérifions dans le cas contraire que d_k est une direction de descente. Par définition de la projection on a :

$$\langle y_k - x_k + \alpha \nabla f \left(x\right), x - y_k \rangle \geq 0, \forall x \in K.$$

Prenons $x = x_k$, alors :

$$0 > -\|d_k\|^2 \geq \alpha \langle \nabla f \left(x_k\right), d_k \rangle.$$

D'autre part la convexité de K implique :

$$x_k + t \left(y_k - x_k\right) \in K \text{ pour tout } t \in \left]0, 1\right[.$$

Lemme 4.3.1 *Soit* $\left(x_n\right)_n$ *construite par l'algorithme* $(GPFK)$. *On suppose que* $x_n \to x$ *quand* $n \to \infty$. *Alors* x *est solution de* (3).

Démenstration :

Siot $P_k : \mathbb{R}^n \to K \subset \mathbb{R}^n$ la projection sur K définie par la proposition (4.2.1)

Alors P_k est continue.

Donc si $x_n \to x$ quand $n \to +\infty$ alors :

$$x = P_k\left(x - \rho\nabla f\left(x\right)\right)$$

et $x \in K$ (car $x_n \in K$ et K est fermé).

La caractérisation de $P_k\left(x - \rho\nabla f\left(x\right)\right)$ donnée dans la proposition (4.2.1) donne alors :

$$\left\langle x - \rho\nabla f\left(x\right) - x, x - y\right\rangle \geq 0 \text{ pour tout } y \in K$$

$$\left\langle -\rho\nabla f\left(x\right), x - y\right\rangle \geq 0 \text{ pour tout } y \in K$$

et comme $\rho > 0$, ce ci entraîne $\left\langle \nabla f\left(x\right), x - y\right\rangle \geq 0 \ \forall \ y \in K$.

Or f est convexe donc $f\left(y\right) \geq f\left(x\right) + \underbrace{\nabla f\left(x\right)\left(y - x\right)}_{\geq 0}$ pour tout $y \in K$,

et donc :

$$f\left(y\right) \geq f\left(x\right) \text{ pour tout } y \in K.$$

Alors : x est une solution de $\left(3\right)$.

Théorème 4.3.2 *(convergence de l'agorithme GPFK)*

Soit $f \in C^1\left(\mathbb{R}^n, \mathbb{R}\right)$, et K convexe fermé non vide. On suppose que :

1. Il existe $\alpha > 0$, tel que :

$$\left\langle \nabla f\left(x\right) - \nabla f\left(y\right), x - y\right\rangle \geq \alpha \mid x - y \mid^2, \text{ pour tout } \left(x, y\right) \in \mathbb{R}^n \times \mathbb{R}^n,$$

2. Il existe $M > 0$ tel que :

$$\mid \nabla f\left(x\right) - \nabla f\left(y\right) \mid \leq M \mid x - y \mid \text{ pour tout } \left(x, y\right) \in \mathbb{R}^n \times \mathbb{R}^n,$$

alors :

1. Il existe un unique élément $x^* \in K$ solution du probléme $\left(3\right)$.

2. Si $0 < \rho < \frac{2\alpha}{M^2}$, la suit $\left(x_n\right)$ définie par l'algorithme $\left(PGFK\right)$ converge vers x^* lorsque $n \to +\infty$.

Lemme 4.3.3 *(Propriétè de contraction de la projection orthogonale)*

Soit E un espace de Hilbert, $\| \,.\, \|$ la norme et $(.,.)$ le produit scalaire, K un convexe fermé non vide de E et P_k la projection orthgonale sur K définie par la proposition (4.3.1), *alors :*

$$\| P_k(x) - P_k(y) \| \leq \| x - y \| \;\; \textit{pour tout} \;\; (x,y) \in E^2.$$

Démonstration

Comme E est un espace de Hilbert,

$$\| P_k(x) - P_k(y) \|^2 = \langle P_k(x) - P_k(y), P_k(x) - P_k(y) \rangle.$$

.

On a donc

$\| P_k(x) - P_k(y) \|^2 = \langle P_k(x) - x + x - y + y - P_k(y), P_k(x) - P_k(y) \rangle$

$= \langle P_k(x) - x, P_k(x) - P_k(y) \rangle_E + \langle x - y, P_k(x) - P_k(y) \rangle + \langle y - P_k(y), P_k(x) - P_k(y) \rangle$.

Or :

$$\langle P_k(x) - x, P_k(x) - P_k(y) \rangle \geq 0$$

et

$$\langle y - P_k(y), P_k(x) - P_k(y) \rangle \geq 0$$

d'où :

$$\| P_k(x) - P_k(y) \| \leq \langle x - y, P_k(x) - P_k(y) \rangle$$

et donc, grace à l'inégalité de Cauchy-Schwarz,

$$\| P_k(x) - P_k(y) \| \leq \| x - y \| \| P_k(x) - P_k(y) \| \leq \| x - y \| .$$

4.3.2 Algorithme du gradient à pas optimal avec projection sur K $(GPOK)$

L'algoritme du gradient à pas optimal avec projection sur K s'écrit :

Initialisation : $x_0 \in K$

Itération : x_n connu

$d_n = -\nabla f(x_n)$; calculer ρ_n optimal dans la direction d_n tel que : ρ variable à chaque itération

$x_{n+1} = P_k(x_n + \rho_n d_n)$

La démonstration de convergence de cet algorithme se déduit de celle de l'algorithme à pas fixe.

Conclusion Générale

En conclusion, on peut dire que la méthode de gradient avec projecton est une méthode itératif pour résoudre un probléme d'optimisation avec contrainte qui parmis les plus performantes pour la resolution des problèmes d'optimisation sous contrainte.

Le but de la méthode du gradient projeté est pour optimaliser un problème sous contrainte sur un sous-ensemble convexe fermé non vide.

Bibliographie

[1] A. Keraghel, Analyse convexe : Théorie fondamentale et exercices, Editions Dar El'houda Ain Mlila , Algérie, (2001).

[2] A. Keraghel, Programmtion mathématique différentiable, Université Ferhat Abbas, Sétif Algérie.

[3] Guy Cohen, Convexité et optimisation, École Nationale des Ponts et Chaussées et INRIA, Édition 2000 (correction 2006).

[4] Mechel Bierlaire, Introduction a l'optimisation différentiable, Presses polytechniques et universitaires romandes.

[5] Beranard Marcheterre, Calcul différentiel, Gerald L.Bradley.

[6] Jean-Pol Guillement, Resume d'Optimisation, Departement de Mathématiques ,Nantes 2010/2011.