Chapitre 2

Les Fondements de l'optimisation sans contraintes

Chapitre 2

Les Fondements de l'optimisation sans contraintes

2.1 Définitions

On aborde le problème d'optimisation sans contrainte qui est normalement non-linéaire. Le problème est posé dans $\mathbb{R}^n$, donc

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

où $f : \mathbb{R}^n \to \mathbb{R}$.

Définition 2.1.

- *Un vecteur x^* de $\mathbb{R}^n$ est appellé minimum local de f sur $\mathbb{R}^n$, s'il existe une boule $B(x^*, \rho)$ de centre x^* et de rayon $\rho > 0$, telle que*

$$\forall x \in B(x^*, \rho),\ f(x) \geq f(x^*).$$

- *Un minimum global de f sur $\mathbb{R}^n$ est un vecteur x^*de $\mathbb{R}^n$, tel que*

$$\forall x \in \mathbb{R}^n,\ f(x) \geq f(x^*). \tag{2.1}$$

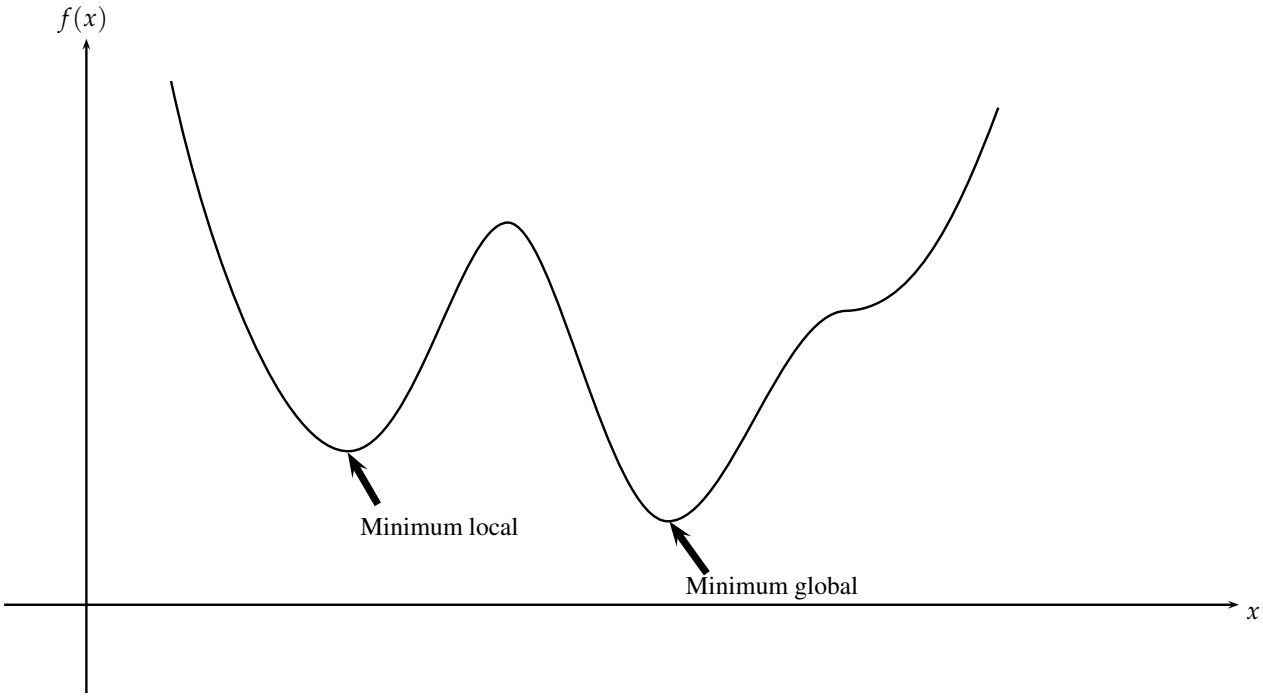

Figure 2.1 – Les différents types de minimas.

Remarque 2.1. *En dehors de cas particulièrement favorable (il n'existe qu'un seul minimum ou bien on est dans le cas de la programmation linéaire continue, ou quadratique à matrice semi-définie positive, ou en programmation convexe), on devra se contenter de découvrir des minimas locaux de f.*

2.1.1 Principe général de résolution

Pour résoudre le problème (P), on applique un algorithme qui, d'un point initial $x_0 \in \mathbb{R}^n$, construit itérativement une suite de vecteurs dans $\mathbb{R}^n$, définie par

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

où, d_k est un vecteur non nul de $\mathbb{R}^n$ appelé direction et t_k une valeur réelle non nulle qu'on appel le pas. Le choix du vecteur d_k et le scalaire t_k détermine le type d'algorithme utilisé pour résoudre (P) ainsi que la convergence et la vitesse de convergence de la suite $\{x_k\}$ engendrée par l'algorithme.

Définition 2.2. *Soit d un vecteur non nul de $\mathbb{R}^n$, s'il existe un réel $\delta > 0$ tel que pour tout t dans l'intervalle $]0, \delta[$, on a*

$$f(x + td) < f(x),$$

alors le vecteur d est dit une direction de descente en x.

Théorème 2.1. *Si f est différentiable en un point $x \in \mathbb{R}^n$ et $d \in \mathbb{R}^n$, telle que*

$$g(x)^\top d < 0. \tag{2.3}$$

Alors d est une direction de descente en x.

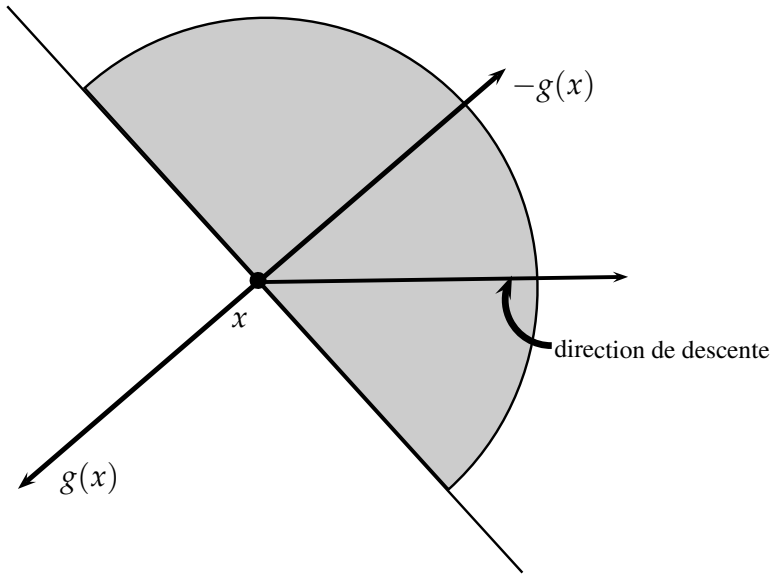

Figure 2.2 – Cône des directions de descente.

2.2 Conditions d'optimalité

Les conditions d'optimalité d'un point $x \in \mathbb{R}^n$ sont des équations, des inéquations ou des propriétés que vérifient les solutions de (P) (conditions nécessaires) ou qui assurent à un point d'être solution de (P) (conditions suffisantes). Ces conditions sont utiles pour

- Vérifier l'optimalité éventuelle d'un point $x \in \mathbb{R}^n$, voir si c'est un minimum, un maximum ou un point stationnaire.
- Calculer les solutions de (P).
- Élaboré des méthodes numériques permettant de résoudre (P).
- Définir des tests d'arrêt des itérations dans les algorithmes de résolution de (P).

2.2.1 Le cas différentiable

Conditions nécessaires d'optimalité du premier ordre

Théorème 2.2. *Soit $f : \mathbb{R}^n \to \mathbb{R}$ différentiable en $\bar{x}$. Si $\bar{x}$ est minimum local, alors*

$$g(\bar{x}) = 0. \tag{2.4}$$

Démonstration.

Par l'absurde, en supposant que $g(\bar{x}) \neq 0$ et on prend $d = -g(\bar{x})$ qui est une direction de descente (car $g(\bar{x})^\top d = -\|g(\bar{x})\|^2 < 0$) et d'après la définition 2.1, $\bar{x}$ n'est pas un optimum local.

□

Ces conditions ne sont pas suffisantes. En effet la fonction $f(x, y) = x^2 - y^2$ considérée sur $[-1, 1] \times [-1, 1]$, vérifie $f(\bar{x}) = 0$ en $\bar{x} = (0, 0)$, cependant ce point n'est pas un optimum local, mais un point de selle (voir figure 2.3).

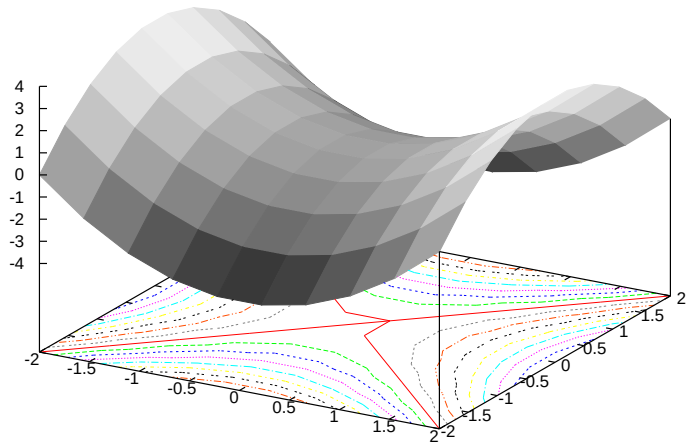

Figure 2.3 – Une fonction avec un point selle en $(x, y) = (0, 0)$.

Conditions nécessaires d'optimalité du second ordre

Théorème 2.3. *Soit $f : \mathbb{R}^n \to \mathbb{R}$ deux fois différentiable en $\bar{x}$ et supposons que $\bar{x}$ est un minimum local. Alors $g(\bar{x}) = 0$ et $H(\bar{x})$ est semi-définie positive.*

Démonstration.

Le premier point est le résultat du théorème précédent. Le deuxième se démontre par contradiction. Supposons que $H(\bar{x})$ n'est pas SDP. Alors, on peut choisir un vecteur d tel que $d^\top H(\bar{x})d < 0$. En utilisant le développement de Taylor au voisinage de $\bar{x}$, il existe un $\bar{t} > 0$ tel que pour tout $t \in]0, \bar{t}]$ on a

$$f(\bar{x} + td) = f(\bar{x}) + tg(\bar{x})^\top d + \frac{1}{2}t^2 d^\top H(\bar{x})d < f(\bar{x}).$$

□

Moyennant un léger renforcement, les conditions du théorème précédent s'avèrent suffisantes pour garantir un minimum local.

Conditions suffisantes d'optimalité du second ordre

Théorème 2.4. *Soit $f : \mathbb{R}^n \to \mathbb{R}$ deux fois différentiable en $\bar{x}$. Si $g(\bar{x}) = 0$ et $H(\bar{x})$ est définie positive, alors $\bar{x}$ est un minimum local strict.*

Démonstration.

Choisissons $\varepsilon > 0$ assez petit de sorte que pour tout $x \in B(\bar{x}, \varepsilon) = \{z \in \mathbb{R}^n : \|z - \bar{x}\| < \epsilon\}$ on a

$$f(x) = f(\bar{x}) + g(\bar{x})^\top (x - \bar{x}) + \frac{1}{2}(x - \bar{x})^\top H(\bar{x})(x - \bar{x}) = f(\bar{x}) + \frac{1}{2}(x - \bar{x})^\top H(\bar{x})(x - \bar{x}),$$

puisque $H(\bar{x})$ est défini positif, alors

$$f(x) - f(\bar{x}) = \frac{1}{2} d^\top H(\bar{x}) d > 0 \Rightarrow f(x) > f(\bar{x}).$$

□

2.2.2 Le cas convexe

La convexité est d'un grand intérêt, car elle permet une caractérisation globale des extremums et elle permet aussi d'énoncé des résultats en absence de la différentiabilité.

Définition 2.3. *Une fonction f définie sur un ensemble convexe X, est dite convexe si $\forall x_1, x_2 \in X$, et $\forall \alpha \in]0, 1[$ on a*

$$f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha) f(x_2). \tag{2.5}$$

Elle est dite strictement convexe si $\forall x_1, x_2 \in X$, $x_1 \neq x_2$ et $\forall \alpha \in]0, 1[$ on a

$$f(\alpha x_1 + (1 - \alpha)x_2) < \alpha f(x_1) + (1 - \alpha) f(x_2). \tag{2.6}$$

Lorsqu'une fonction est différentiable, il y a d'autre manières de caractériser la convexité.

Caractérisation de la convexité

Théorème 2.5. *Soit X un sous-ensemble ouvert et convexe de $\mathbb{R}^n$,*

1. *Si f est de classe C^1, alors f est convexe sur X si et seulement si*

$$f(x_2) \geq f(x_1) + g(x_1)^\top (x_2 - x_1), \forall x_1, x_2 \in X. \tag{2.7}$$

2. *Si f est de classe C^2, alors f est convexe sur X si et seulement si $H(x)$ est semi-définie positive dans tout X,*
3. *Si $H(x)$ est définie positive sur X alors f est strictement convexe sur X.*

Conditions suffisantes d'optimalité du premier ordre

La convexité permet d'énoncer des conditions nécessaires et suffisantes d'optimalité pour des minimas globaux.

Théorème 2.6. *On suppose que f est convexe sur $\mathbb{R}^n$, alors*

1. *Tout minimum local est aussi un minimum global.*
2. *Si f est différentiable en $\bar{x} \in \mathbb{R}^n$ et si*

$$g(\bar{x}) = 0. \tag{2.8}$$

alors $\bar{x}$ est un minimum global.

Démonstration.

1. Soit $\bar{x}$ un minimum local de f sur $\mathbb{R}^n$. S'il existe $x \in \mathbb{R}^n$ tel que $f(x) < f(\bar{x})$, on peut écrire par convexité de f (pour $\alpha \in]0, 1]$)

$$f(\alpha x + (1 - \alpha)\bar{x}) \leq \alpha f(x) + (1 - \alpha) f(\bar{x}) < f(\bar{x}),$$

ce qui contredit le fait que $\bar{x}$ est un minimum local.

2. Soit $\bar{x} \in \mathbb{R}^n$ tel que $g(\bar{x}) = 0$. Pour tout point $x \neq \bar{x}$ dans $\mathbb{R}^n$, la convexité de f permet d'écrire

$$f(x) \geq f(\bar{x}) + g(\bar{x})^\top (x - \bar{x}),\ \forall x \in \mathbb{R}^n,$$

Si de plus on a $g(\bar{x}) = 0$, alors pour tout $x \in \mathbb{R}^n$, $f(x) \geq f(\bar{x})$. Donc $\bar{x}$ est bien un minimum global.

□

Remarque 2.2. *Les théorèmes 2.2, 2.3, 2.4, 2.5 et 2.6 restent valables si $\mathbb{R}^n$ est remplacé par un ensemble X, ouvert (et convexe pour le théorème 2.6).*

2.3 Algorithmes

Considérons maintenant le problème non linéaire sans contraintes (P), où l'on suppose que f est continûment différentiable sur $\mathbb{R}^n$.

2.3.1 Existence

Le fait que f soit minorée n'implique pas l'existence d'une solution à (P).

Exemple 2.1. *$f(x) = e^x$ est minorée par 0 mais on ne peut jamais obtenir $x \in \mathbb{R}$, tel que $f(x) = 0$.*

Il faut donc une hypothèse de plus, le lemme suivant va dans ce sens.

Lemme 2.1. *Soit f une fonction de classe C^2, s'il existe une constant $m > 0$ vérifiant*

$$d^\top H(x) d \geq m \|d\|^2 ; \forall x, d \in \mathbb{R}^n.$$

Alors, pour tout vecteur $y \in \mathbb{R}^n$, l'ensemble

$$\mathcal{L} = \{x \in \mathbb{R}^n : \ f(x) \leq f(y)\}, \tag{2.9}$$

est convexe et compact.

Démonstration.

La fermeture de l'ensemble est une conséquence de la continuité de la la fonction f, même chose pour la convexité de l'ensemble résulte du fait que f l'est aussi. Par conséquent nous avons juste à prouver que l'ensemble est borné. Prenons d un vecteur dans $\mathbb{R}^n$ tel que $\|d\| = 1$. Soit la fonction réelle θ définie par

$$\theta(t) = f(y + td).$$

A partir des conditions sur f et la normalisation de d, on en déduit l'inégalité

$$\theta''(t) \geq m,$$

d'où il suit que la fonction définie par

$$\psi(t) = \theta(t) - \theta(0) - td^\top g(y) - \frac{1}{2}t^2 m.$$

est convexe (en effet $\psi''(t) \geq 0$). Maintenant ψ est choisie pour satisfaire $\psi(0) = \psi'(0) = 0$, donc la fonction ψ admet un minimum en $t = 0$, d'où l'inégalité

$$\theta(t) \geq \theta(0) + td^\top g(y) + \frac{1}{2}t^2 m.$$

Mais le côté droit de cette inégalité est plus grand que $\theta(0)$ si

$$|t| > 2\frac{\|g(y)\|}{m} \geq 2\frac{\left|d^\top g(y)\right|}{m}.$$

Par conséquent, et puisque le vecteur d est arbitraire, (il est possible de mettre $x = y + td$) et on a $f(x) \geq f(y)$ si $\|x - y\| > 2\frac{\|g(y)\|}{m}$. Par conséquent l'ensemble des points x satisfaisant la condition $f(x) < f(y)$ est borné.

$\square$

Remarque importante : Théoriquement, le but d'un algorithme de minimisation est l'identification d'un minimum, dans la pratique on est plus modeste, on se contente souvent de chercher un point stationnaire (c'est-à-dire un point qui vérifie l'équation (2.4)).

2.3.2 La méthode du gradient

On se propose de trouver une solution du problème (P). On remplace, $f(x+d)$ par $f(x)+g(x)^\top d$ (ce qui est possible pour $\|d\|$ assez petit), Donc

$$f(x+d) = f(x) + g(x)^\top d. \tag{2.10}$$

Soit d_k une solution du problème approché correspondant

$$\begin{cases} min\ g(x_k)^\top d, \\ \|d\| \leq 1, \end{cases} \tag{2.11}$$

ensuite x_{k+1} est rechercher le long de d_k c'est à dire

$$x_{k+1} = x_k + td_k.$$

Remarque 2.3.

- *La contrainte $\|d\| \leq 1$ est essentielle pour assurer une solution à (2.11) mais la longueur de la direction n'a aucune importance (on peu prendre $\|d\| \leq a$), car elle sera corrigée par le choix du pas t.*
- *Il est possible de de résoudre (2.11) suivant le choix de la norme $\|.\|$.*

Cas de la norme l^1

Dans ce cas, $\|x\|_1 = \sum_{i=1}^{i=n} |x_i|$. L'ensemble $S = \left\{ d \in \mathbb{R}^n : \|d\|_1 = \sum_{i=1}^{i=n} |x_i| \leq 1 \right\}$ est un polyèdre de sommées

$$E = \{e_1, e_2, ..., e_n, -e_1, -e_2, ..., -e_n\},$$

où $\{e_i\}_{1 \leq i \leq n}$ est une base de $\mathbb{R}^n$.

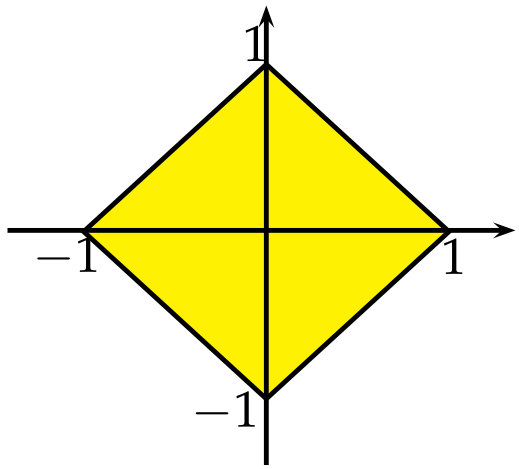

Figure 2.4 – L'ensemble S dans $\mathbb{R}^2$

Puisque la fonction : $d \to g(x_k)^\top d$ est linéaire alors le problème (2.11) est équivalent au problème d'optimisation linéaire (voir [2, 22])

$$\begin{cases} \min \ g(x_k)^\top d, \\ d \in E. \end{cases}$$

Donc

$$(2.11) \Longleftrightarrow \min_{1 \leq i \leq n} \{g_i(x_k), -g_i(x_k)\}$$

tel que, $g(x_k) = (g_1(x_k), ..., g_n(x_k))$ et

$$d_k = (0, 0, .., \varepsilon_i, 0, .., 0) = \varepsilon_i e_i,$$

où i et ε_i sont donnés par

$$\begin{cases} g_i(x_k) = \max\left\{\left|g_j(x_k)\right|; 1 \leq j \leq n\right\} \\ \varepsilon_i = -\dfrac{g_i(x_k)}{\left|g_i(x_k)\right|}. \end{cases}$$

Graphiquement la solution d_k est parallèle à un certain axe de coordonnées (correspondant à la plus grande composante du gradient). Posant $\theta(t) = f(x_k + td_k)$ et cherchons t_k une solution du problème

$$\min_{t>0} \theta(t).$$

L'idée immédiate est de chercher une solution de

$$\theta'(t) = \frac{d}{dt} f(x_k + td_k) = g(x_k + td_k)^\top d_k = 0, \ t > 0. \tag{2.12}$$

Or d_k est l'un des vecteurs de base (à un signe près), ce qui donne

$$\begin{aligned} g(x_k + td_k)^\top d_k &= g(x_k + \varepsilon_i t e_i)^\top (\varepsilon_i e_i) \\ &= \varepsilon_i g_i(x_k + \varepsilon_i t e_i) = 0, \ i = 1, 2, ..., n. \end{aligned}$$

C'est donc résoudre une seule composante du gradient, ce qui nous donne une interprétation de la méthode Gauss-Seidel. Maintenant on choisis parmi les solutions t_k possible celle qui donne

$$f(x_k + t_k d_k) < f(x_k).$$

Cas de la norme l^2

Dans ce cas $\|x\| = \|x\|_2 = \sqrt{\sum_{i=1}^{i=n} |x_i|^2}$. Abordant un Lemme essentiel pour la suite.

Lemme 2.2. *Supposant que f est différentiable en x tel que $g(x) \neq 0$. Alors le problème d'optimisation*

$$\begin{cases} \min \ g(x)^\top d \\ \quad \|d\| \leq 1 \end{cases}$$

admet une solution $d_k = -\dfrac{g(x_k)}{\|g(x_k)\|}$

- *On peut prendre $d_k = -g(x_k)$ selon la remarque précédente*
- *Comme auparavant, posons $\theta(t) = f(x_k + td_k)$ et cherchons t_k une solution de*

$$\min_{t>0} \theta(t), \tag{2.13}$$

telle que $\theta(t_k) < \theta(0)$

L'algorithme de la plus profonde descente (Steepest Descent)

(i) $d_k = -g(x_k)$

(ii) t_k solution de

$$\min_{t>0} \theta(t)$$

qui vérifie

$$\theta(t_k) < \theta(0)$$

(iii) $x_{k+1} = x_k + t_k d_k$

Remarque 2.4. *La méthode (i), (ii), (iii) donne une interprétation directe de la méthode des approximations successive vue en (1.5.2) de plus, (ii) permet de calculer avec précision le pas t_k, contrairement au théorème 1.4.*

- *La contrainte $t > 0$ est justifiée par le fait que si $t_k = 0$*

$$\theta'(0) = g(x_k)^\top d_k = g(x_k)^\top(-g(x_k)) = -\|g(x_k)\|^2 = 0 \Rightarrow g(x_k) \equiv 0$$

  et on aura $x_{k+1} = x_k$

- *Dans les deux cas (la norme l^1 ou l^2) on voit bien que le vecteur d_k est une direction de descente puisque $g(x_k)^\top d_k < 0$*

2.3.3 Étude globale de la convergence

Il est plus intéressant de démontrer la convergence d'une variante de la méthode (i),(ii),(iii) où le pas t_k est calculé de façon plus général, considérons au lieu de (ii)

(ii') t_k est tel que

$$\theta(t_k) \leq \theta(t_k^*), \tag{2.14}$$

où t_k^* est la plus petite solution de (2.13)

Pour la convergence de cette méthode, nous reprenons une proposition dù à [3].

Proposition 2.1. *On suppose le gradient g lipschitzien sur la tranche*

$$\mathcal{L} = \{x \in \mathbb{R}^n : f(x) \leq f(x_0)\}.$$

Alors la méthode définie par (i), (ii') et (iii), satisfait à

- *Soit $f(x_k)$ est non bornée inférieurement.*
- *Soit $g(x_k) \to 0$.*

Démonstration. Observons que $\theta(t)$ est décroissante sur $]0, t_k^*]$ puisque

$$\begin{cases} \theta'(0) < 0 \\ \theta'(t) \text{ continue } (g \text{ continue}) \\ t_k^* \text{ est le premier zero de } \theta'(t),\ t > 0 \end{cases} \Rightarrow \theta'(t) < 0,\ \forall t \in]0, t_k^*]$$

Donc t_k^* est le plus petit minimum de $\theta(t)$ pour $t > 0$. Si on pose $x_k^* = x_k + t_k^* d_k$, on a : $x_k^*, x_k \in \mathcal{L}$ (par construction).

(1) On va montrer que $f(x_k + td_k)$ décroit sensiblement dans le voisinage de $t = 0$: Prenons $z = x_k + td_k$, $t > 0$ assez petit de façon à ce que $z \in \mathcal{L}$, la formule de la moyenne pour z' entre x_k et z donne

$$\begin{aligned} f(z) &= f(x_k) + g(z')^\top (z - x_k) = f(x_k) + (g(z') - g(x_k))^\top (z - x_k) + \\ &\quad + g(x_k)^\top (z - x_k). \end{aligned} \tag{2.15}$$

Puisque $z' \in \mathcal{L}$, du fait que $z' = x_k + \tau d_k$, $\tau \in]0, t]$, alors

$$f(z') \leq f(z) \leq f(x_k)$$

g est lipschitzien sur $\mathcal{L}$ et grâce l'inégalité de Cauchy-Schwarz on a

$$\begin{aligned} |(g(z') - g(x_k)^\top (z - x_k)| &\leq \|g(z') - g(x_k)\| \, \|z - x_k\| \leq M \|z' - x_k\| \, \|z - x_k\| \\ &\leq M \|z - x_k\|^2, \end{aligned}$$

donc

$$f(z) \leq f(x_k) - \|g(x_k)\| \, \|z - x_k\| + M \|z - x_k\|^2 = f(x_k) - t \|g(x_k)\|^2 + Mt^2 \|g(x_k)\|^2$$

Si on prend $t \leq \dfrac{1}{2M}$, alors $M t^2 \leq t/2$ et donc pour $t \leq \min\{t_k^*, \dfrac{1}{2M}\}$, on a

$$f(z) \leq f(x_k) + (M t^2 - t) \|g(x_k)\|^2 \leq f(x_k) - \frac{t}{2} \|z - x_k\|.$$

Ce qui démontre (1).

(2) Minorer le pas t_k : Soit $x_k^* = x_k + t_k^* d_k$. On a $g(x_k^*)^\top d_k = \theta'(t_k^*) = 0$ et donc

$$0 = g(x_k^*)^\top (x_k^* - x_k) = (g(x_k^*) - g(x_k))^\top (x_k^* - x_k) + g(x_k)^\top (x_k^* - x_k).$$

En utilisant les mêmes techniques qu'en (1), on obtient

$$0 \quad \leq \quad \|x_k^* - x_k\| \left[M \|x_k^* - x_k\| - \|g(x_k)\|\right]$$

et donc

$$\|x_{k+1} - x_k\| \geq \|x_k^* - x_k\| \geq \frac{\|g(x_k)\|}{M}.$$

Puisque

$$0 < t_k^* \|g(x_k)\| = \|x_k^* - x_k\| \leq \|x_{k+1} - x_k\| = t_k \|g(x_k)\|,$$

du fait que $t_k^* \leq t_k$ (rappelons que θ est décroissante dans $]0, t_k^*]$ et de (ii')).

D'où (2) à savoir

$$\|x_{k+1} - x_k\| \geq \|x^* - x_k\| \geq \frac{\|g(x_k)\|}{M}.$$

Maintenant, prenons $t \leq \dfrac{\|g(x_k)\|}{2M} \Rightarrow z = x_k + td_k$ est entre x_k et x^* (d'après (2)) donc

$$z = x_k + td_k,\ t \in]0, t_k^*] \Rightarrow f(z) = \theta(t) \geq \theta(t_k^*) = f(x^*).$$

Et de (ii') on a

$$f(x_{k+1}) = \theta(t_k) \leq \theta(t_k^*) = f(x_k^*),$$

ce qui donne

$$f(x_{k+1}) \leq f(x_k^*) \leq f(z) \Rightarrow f(x_{k+1}) \leq f(z). \tag{2.16}$$

De l'argument (1) et (2.16)

$$f(z) \leq f(x_k) - \frac{\|g(x_k)\|^2}{4M},$$

donc $\|g(x_k)\|^2 \leq 4M[f(x_k) - f(z)] \leq 4M[f(x_k) - f(x_{k+1})]$, ce qui implique

$$\sum_{i=0}^{i=k} \|g(x_i)\| \leq 4M[f(x_0) - f(x_{k+1})].$$

Si f est bornée inférieurement on a $[f(x_0) - f(x_{k+1})] \leq C, \forall k \in \mathbb{N}$ ce qui donne

$$\sum_{i=0}^{i=\infty} \|g(x_i)\| \leq C \Rightarrow g(x_k) \to 0.$$

□

2.3.4 Etude locale de la convergence

Maintenant, on va analyser la convergence locale de cette méthode, c'est-à-dire la vitesse de convergence des suites générées par l'algorithme au voisinage du point optimum.

Théorème 2.7. *Si f est quadratique de la forme $f(x) = \frac{1}{2}x^\top Qx - b^\top x$, avec $Q \in \mathcal{M}_n$ symétrique définie positive, $b \in \mathbb{R}^n$, la méthode "steepest descent" (i),(ii),(iii), converge vers la solution du problème avec une vitesse linéaire de rapport*

$$r = \frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1},$$

où $\lambda_n = \max(\sigma(Q))$ et $\lambda_1 = \min(\sigma(Q))$.

Démonstration. La condition suffisante d'optimalité (CS2) permet d'identifier la solution de (P), il suffis de résoudre l'équation (2.12), ce qui donne

$$t_k = \frac{g(x_k)^\top g(x_k)}{g(x_k)^\top Qg(x_k)}. \tag{2.17}$$

Avec un calcul simple, on trouve

$$f(x_{k+1}) = \left[1 - \frac{(g(x_k)^\top g(x_k))^2}{(g(x_k)^\top Qg(x_k))(g(x_k)^\top Q^{-1}g(x_k))}\right] f(x_k) \tag{2.18}$$

En utilisant (iii) et grâce à l'inégalité de Kantorovich [22]

$$\frac{\|x\|^4}{\|x\|_Q^2 \, \|x\|_{Q^{-1}}^2} \geq \frac{4\lambda_1\lambda_n}{(\lambda_1+\lambda_n)^2}$$

on a

$$f(x_{k+1}) \leq \left[\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}\right]^2 f(x_k). \tag{2.19}$$

D'autre part, $f(x) = \frac{1}{2}\|x-x^*\|_Q - \frac{1}{2}\|x^*\|_Q$, le fait que Q est définie positive et grâce à (2.19), nous obtenons

$$\|x_{k+1}-x^*\|_Q \leq \left[\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}\right]^2 \|x_k-x^*\|_Q \,. \tag{2.20}$$

□

Remarque 2.5. *L'inégalité (2.9) montre que la vitesse de convergence de la suite $\{f(x_k)\}$ est linéaire, et elle dépend clairement du rapport $\frac{\lambda_n}{\lambda_1}$. Si ce dernier est très grand (dans le cas d'une matrice Q, malle conditionnée) le terme*

$$\frac{\lambda_n-\lambda_1}{\lambda_n+\lambda_1}$$

s'approchera du 1 donc, la convergence est sous-linéaire.

Dans un cadre plus général (fonction f non-linéaire), l'analyse précédente reste valable, car en s'approchant de l'optimum x^*, la fonction f peut être approximée quadratiquement par

$$f(x) \approx f(x^*) + g(x^*)^\top(x-x^*) + \frac{1}{2}(x-x^*)^\top H(x^*)(x-x^*)$$

où le hessien $H(x^*)$ est maintenant définie positive.

Algorithme 2.1.

Etape 0 (Initialisation) : *x_0 et $\varepsilon > 0$ données, $k = 0$,*

Etape 1 : *Test d'arrêt : calculer $g(x_k)$, si $\|g(x_k)\| < \varepsilon$ alors stop sinon*

Etape 2 : *Calcul de la direction* $d_k = -g(x_k)$,

Etape 3 : *Recherche linéaire : Trouver* t_k *qui vérifie "les bonnes hypothèses" en particulier* $f(x_k + t_k d_k) < f(x_k)$,

Etape 4 : *Si la recherche linéaire réussit* $x_{k+1} = x_k + t_k d_k$, *remplacer k par* $k+1$ *et aller en 1.*

Remarque 2.6. *Dans ce schéma, tout est clair sauf l'étape 3, qui ne dit ni quelle condition précise doit vérifier le pas* t_k, *ni comment le trouver. La recherche du pas s'effectue en fait par une sous-itération sur* $t > 0$ *durant laquelle s'établit un dialogue avec un sous programme qui se charge de calculer* t_k. *Le problème de la recherche linéaire est suffisamment important pour justifier le chapitre suivant à lui tout seul.*

Exercices

Exercice 2.1. *Calculer le gradient* $g(x)$ *et le hessien g(x) de la fonction Rosenbrock*

$$f(x) = 100(y - x^2)^2 + (1 - x)^2.$$

Montrer que $x^* = (1\,,\,1)^T$ *est le seul minimum local de cette fonction.*

Exercice 2.2. *Montrer que la fonction* $f(x) = 8x_1 + 12x_2 + x_1^2 - 2x_2^2$ *a un seul point stationnaire et que ce n'est ni un maximum et ni un minimum, mais un point de selle. "Croquiser" les courbes à niveaux de* f.

Exercice 2.3. *Considérons la fonction* $f : \mathbb{R}^2 \to \mathbb{R}$ *définie par* $f(x) = \|x\|^2$. *Montrer que la suite d'itérations* $\{x_k\}$ *définie par*

$$x_{k+1} = (1 + \frac{1}{2^k}) \begin{pmatrix} cos(k) \\ sin(k) \end{pmatrix},$$

satisfait $f(x_{k+1}) < f(x_k)$, $k = 1, 2, ...$*que tout point sur le cercle unité* $\{x : \|x\| = 1\}$ *est un point*

limite de $\{x_k\}$. *Astuce : Chaque valeur* $\theta \in [0\,,\,2\pi]$ *est un point limite de la sous-suite*

$$\xi_i = k\,(mod\,2\pi) = k - 2\pi\left[\frac{k}{2\pi}\right],$$

où l'opérateur $[\,.\,]$ *est la partie entière d'un réel.*

Exercice 2.4. *Montrer que tous les minimas locaux isolés sont strictes.*

Exercice 2.5. *Supposons que* f *est une fonction convexe. Montrer que l'ensemble des minimas globaux de* f *est un ensemble convexe.*

Exercice 2.6. *Considérons la fonction* $f(x_1\,,\,x_2) = (x_1 + x_2^2)^2$. *Au point* $x^T = (1\,,\,0)$, *nous considérons la direction* $d^T = (-1\,,\,1)$. *Montrer que* d *est une direction de descente et trouver toutes les minimas du problème*

$$\min_{t>0} f(x + t\,d).$$

Exercice 2.7. *Supposons que* $\tilde{f}(z) = f(x)$, *où* $x = Sz + s$ *pour un certain* $S \in \mathbb{R}^{n\times m} = \mathcal{M}^{n\times m}$ *et* $s \in \mathbb{R}^n$. *Monter que*

$$\tilde{g}(z) = \nabla\tilde{f}(z) = S^T\nabla f(x),\;\; \tilde{H}(z) = \nabla^2\tilde{f}(z) = S^T H(x) S.$$

Bibliographie

[1] L. Armijo. Minimization of functions having Lipschitz continuous first derivatives. Pacific Journal of Mathematics 16 (1966) 1-3

[2] M. S. Bazara. H.D. Sherali, C.M Shetty. Nonlinear programming. J.W&S (1993)

[3] J. F. Bonnans, J. C. Gilbert, C. Lemaréchal and C. A. Sagastizábal. Numerical optimization : theoretical and practical aspects. Springer Science & Business Media (2006).

[4] C. G. Broyden. Quasi-Newton method and their application to function minimization, Mathematics of Computation, 21,(1967).

[5] A. Bukley, A.Lenir. QN-like variable storage conjugate gradients. Math. Prog. 27,(1983)155-175.

[6] R. Byrd and J. Nocedal. A tool for the analysis of quasi-Newton methods with application to unconstrained minimization, SIAM Journal on Numerical Analysis, 26 (1989) 727-739.

[7] A. Cauchy. Méthodes générales pour la résolution des systèmes d'équations simultanées. Compte rendus Acad. Sc. Paris, Tome 25 (1847) 536-538.

[8] F. Chatelin. Eigenvalues of Matrices, J.Wiley & S (1993).

[9] W. C. Davidon. Variable metric methods for minimization, AEC Research Development, Report ANL-5990 (1959).

[10] J. E. Denis, Jr., J.J. Moré. A characterisation of superlinear convergence and its application to quasi-Newton methods, Mathematics of computation, 28 (1974) 549-560.

[11] J. E. Dennis, H. Wolkowitcz. Sizing and Least Change Secant Methods, CORR Report 93-11 (1993).

[12] R. Fletcher. A new approch to variable metric algorithms, Computer Journal, 13,(1970)317-322.

[13] R. Fletcher, M.J.D. Powell. A Rapidly convergent descent method for minimisation, Computer Journal 6,(1963)163-168.

[14] R. Fletcher. An overview of Unconstrained Optimization in Algorithms for Continuous Optimization, The State of the Art, E. Spedicato, ed., Kluwer academic Publishers, Boston, (1994)109-143.

[15] A. A. Goldestein, J.F. Price. An effective algorithm for minimization. Num. Math. 10,(1967)184-189.

[16] D. Goldfarb. A falmily of metric variablee methods derived by variation means, mathematics Of computation 24,(1970)23-26.

[17] A. Griewank, Ph.L. Toint. Local convergence analysis of partitioned quasi-Newton udaptes, Numeric Mathematics 39,(1982)429-448.

[18] T.G. Kolda. Limited-Memory matrix methods with application, Applied Mathematics program, university of Maryland (1998).

[19] A. S.Lewis. Convex analysis on the hermitian matrices, Technical report CORR 97-09, University of Waterloo (1994).

[20] D. H. Li, M. Fukushima. A global convergence of BFGS methods for nonconvex unconstrained optimization problems, SIAM Journal on Numerical Analysis.7, (1999).

[21] D. H. Li, M. Fukushima. A modified BFGS method and its global convergence in noncovex minimization, Technical Repport 98003, Département of Applied Mathematics and Physics, Kyoto University, (1998).

[22] D. G. Luenberger. Linear and Nonlinear Programming, 2nd edition, Addison-Wesley (1984).

[23] P. E. Gill, W. Murray. Conjugate gradient method for large-scale nonlinear optimization, Technical report SOL 79-15, Dept. of Operation Res. Stanford University, (1979).

[24] J. J. Moré, B.S. Garbow, K.E. Hillstrome. Testing unconstrained optimization softwear, ACM Trans. Math. Softwear, 7, (1981)17-41.

[25] L. . Conjugate gradient methods less dependent on conjugacy. SIAM Review 28,(1986)501-511.

[26] Y. E. Nestrov. On the a approach to the construction of optimal methods of minimization of smooth convex function, Economica I Matem. 24, (1988)504-517.

[27] J. Nocedal. Udapting quasi-Newton matrice with limited storage, Mathemating of computation 35, (1980)773-782.

[28] J. Nocedal. Theory of algorithms for unconstrained optimization, acta Numerica, (1999)199-242.

[29] J. Nocedal, J. W. Stephen. Numerical optimization. Springer (1999).

[30] D. P. O'Leary. A discrete Newton algorithm for minimizing a function of many variable, Math. Prog. 23, (1982)20-30.

[31] S. S. Oren, E. Spedicato, Optimal condition of self-scaling variable metric algorithms, Math. Prg., 10, (1976)70-90.

[32] J. M. Perry. A class of conjugate gradient with two-step variable-metric memory, Discussion paper 269, C.M.S.E.M.S. Northwestern University (Evanston, IL, 1977)

[33] J. D. Pearson. Variable metric methods of minimization. Computer Journal, Vol 12, (1971)21-36.

[34] M. J. D. Powell. On the convergence of the variable metric algorithm, Journal of of the Institute of Mathematics and its Applications, (1971)21-36.

[35] M. J. D. Powell. Some properties of the variable metric algorithm for minimization without exact line searches, in Nonlinear Programming, SIAM-AMS Proceeding, Vol. IX, R.W. Cottle, and C.E. Lemke, eds., SIAM, (1976)35-72.

[36] L. Schwartz. Analyse, topologie générale et analyse fonctionnelle, Hermann (1970).

[37] M. Sibony, J.C. Mardon. Analyse Numérique. Systèmes linéaires et non linéaires, HERMAN (1982).

[38] D. F. Schanno. Conditionning of quasi-Newton methods for function minimization, Mathematics of computation, 24,(1970)641-656.

[39] D. F. Schanno. On the convergence of new conjugate gradient algorithm, SIAM Journal on numerical analysis, 15,(1978a)1247-1257.

[40] T. Steihaug. The conjugate gradient method and trust regions in large-scale optimization, SIAM J. Num. Anal. 20, (1983)626-637.

[41] P. Wolfe. Convergence for ascent methods. SIAM Rev. 11,226-235, 1968.

[42] W. I. Zangwill. Nonlinear programming : a unified approach, Prentice Hall, Englewood Cliffs, RI, 1969.

[43] G. Zoutendijk. Non linear programming, computation programming. In J. Abadie (ed), integer and non linear programming. North Holland (1970)37-86.