UNIVERSITÉ DE MONCTON
EDMUNDSTON MONCTON SHIPPAGAN

Optimisation

MATH 3163

Ibrahima Dione

Hiver 2024

Optimisation sans contraintes

Sommaire

Le problème étudié ici est celui de la recherche du minimum (ou du maximum) d'une fonction réelle $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, de n variables réelles $x_1, x_2, \cdots, x_n$, chacune de ces variables pouvant prendre n'importe quelle valeur de $-\infty$ à $+\infty$

$$\min_{x \in \mathbb{R}^n} f(x) \tag{1}$$

Le problème (1) s'agit donc de déterminer un point $x^* = (x_1^*, \ldots, x_n^*)^\top \in \mathbb{R}^n$ tel que :

$$\forall\, x \in \mathbb{R}^n,\ f(x^*) \leqslant f(x),$$

c'est-à-dire un minimum global de f sur $\mathbb{R}^n$ [2]. Lorsque l'inégalité stricte $f(x^*) < f(x)$ est vérifiée pour tout $x \in \mathbb{R}^n$ avec $x \neq x^*$, le minimum global x^* est alors unique.

Pour beaucoup de problèmes d'optimisation sans contraintes, les principales méthodes de résolution connues ne permettent pas la détermination d'un minimum global : il faut alors se contenter d'optimums locaux. Nous allons voir maintenant comment de tels points peuvent être caractérisés.

1 Conditions d'optimalité

1.1 Conditions nécessaires d'optimalité

Théorème 1.1

Condition nécessaire du premier ordre

Soit f une fonction de classe C^1. Une condition nécessaire pour que x^* soit un minimum local de f est :

$$\nabla f(x^*) = 0, \quad \text{(stationnarité)}. \tag{2}$$

Preuve :

- x^* est un minimum local de f, alors il existe un voisinage de x^* noté $V_\varepsilon(x^*)$ t.q.

$$f(x^*) \leq f(x), \quad \text{pour tout} \quad x \in V_\varepsilon(x^*). \tag{3}$$

- Soit $\{\varepsilon_k\}_{k\in\mathbb{N}}$ d'autre part une suite de nombres réels strictement positifs, qui tend vers 0. Construisons ainsi la suite $\{x^k\}_{k\in\mathbb{N}}$ comme suit

$$x^k = x^* - \varepsilon_k \nabla f(x^*). \tag{4}$$

Ainsi, en remarquant que pour tout $k \in \mathbb{N}$ nous avons l'égalité

$$\begin{aligned} \|x^k - x^*\| &= \|-\varepsilon_k \nabla f(x^*)\| \\ &= \varepsilon_k \|\nabla f(x^*)\| \end{aligned}$$

alors à partir d'un certain k suffisamment grand, on a $x^k \in V_\varepsilon(x^*)$. Et d'après (3), on a alors $f(x^*) \leq f(x^k)$ pour $k \in \mathbb{N}$ suffisamment grand.

- Finalement, f étant de classe C^1, d'après la formule de la moyenne d'ordre 1 on a :

$$0 \leqslant f\left(x^k\right) - f(x^*) = \nabla f\left(\xi^k\right)^\top \cdot \left(x^k - x^*\right), \quad \text{où} \quad \xi^k \in \left[x^k, x^*\right]. \tag{5}$$

Ainsi tenant compte de (4), on obtient à partir de (5)

$$0 \leqslant f\left(x^k\right) - f(x^*) = -\varepsilon_k \nabla f\left(\xi^k\right)^\top \nabla f(x^*).$$

Ce qui entraine $\nabla f\left(\xi^k\right)^\top \nabla f(x^*) \leqslant 0$, pour $k \in \mathbb{N}$ suffisamment grand. Or

$$\lim_{k\to+\infty} \nabla f\left(\xi^k\right) = \nabla f(x^*),$$

car ∇f est continue du fait que f est de classe C^1. On obtient finalement

$$\begin{aligned} \lim_{k\to+\infty} \nabla f\left(\xi^k\right)^\top \nabla f(x^*) &= \nabla f(x^*)^\top \nabla f(x^*) \\ &= \|\nabla f(x^*)\|^2 \leqslant 0, \ \text{i.e.}\ , \nabla f(x^*) = 0. \end{aligned}$$

□

Exemple 1.1 : Considérons le problème de minimisation suivant

$$\min_{x_1, x_2 \geq 0} f(x_1, x_2), \quad \text{où} \quad f(x_1, x_2) = x_1^2 + 0.5x_2^2 + 3x_2 + 4.5.$$

Est-ce que les points suivants sont des minimums de ce problème ?

$$x = (1, 3)^\top, \quad x = (0, 0)^\top$$

Réponse :

- Le point $(x_1, x_2) = (1, 3)^\top$ est un point intérieur de

$$S = \{(x_1, x_2) : x_1 \geq 0, x_2 \geq 0\}.$$

En ce point $(x_1, x_2) = (1, 3)^\top$, nous avons

$$\nabla f(x_1, x_2) = (2x_1, x_2 + 3)^\top = (2, 6)^\top \neq (0, 0)^\top.$$

Par conséquent, la condition de stationnarité $\nabla f(x_1, x_2) = 0$ n'est pas vérifiée. Le point $(x_1, x_2) = (1, 3)^\top$ n'est pas un minimum local.

Si de plus f est de classe C^2 alors on a le théorème :

Théorème 1.2

Condition nécessaire du $2^{\text{ième}}$ ordre

On suppose que f est une fonction de classe C^2. Une condition nécessaire pour que x^* soit un minimum local de f est que le Hessien $H(x^*)$ est une matrice semi-définie positive, c'est-à-dire

$$y^\top H(x^*)\, y \geqslant 0, \; \forall\, y \in \mathbb{R}^n.$$

Preuve :

- Soit $z \in \mathbb{R}^n$ et $\{\varepsilon_k\}_{k \in \mathbb{N}}$ une suite de nombres réels qui tend vers 0. On définit la suite $\{x^k\}_{k \in \mathbb{N}}$ comme suit

$$x^k = x^* + \varepsilon_k z. \tag{6}$$

- En appliquant la formule de la moyenne d'ordre 2, on obtient

$$\begin{aligned} 0 &\leq f\left(x^k\right) - f\left(x^*\right) \\ &= \nabla f\left(x^*\right)^\top \left(x^k - x^*\right) + \frac{1}{2}\left(x^k - x^*\right)^\top H\left(\xi^k\right)\left(x^k - x^*\right) \\ &= \frac{1}{2}\varepsilon_k^2 z^\top H\left(\xi^k\right) z, \quad \text{où} \;\; \xi^k \in \left[x^k, x^*\right]. \end{aligned} \tag{7}$$

car $\nabla f(x^*) = 0$ (d'après le théorème 1). Alors, on obtient à partir de (7) que

$$z^\top \cdot H\left(\xi^k\right) \cdot z \geqslant 0,\ \forall\, k \ \Rightarrow \lim_{k\to+\infty} z^\top \cdot H\left(\xi^k\right) \cdot z \geqslant 0,$$
$$z^\top \cdot H(x^*) \cdot z \geqslant 0,\ \forall\, z \in \mathbb{R}^n. \qquad \square$$

Note :

- Les théorèmes 1.1 et 1.2 permettent de caractériser les minimums locaux et non de les calculer, donc d'exclure les points qui ne sont pas solutions.
- Un point x^* qui vérifie la condition (2), c'est-à-dire $\frac{\partial f}{\partial x_i} = 0,\ i = 1, \cdots, n$, est appelé un point stationnaire.

Exemple 1.2 : Considérons le problème de minimisation suivant

$$\min_{(x,y)\in\mathbb{R}^2} f(x,y), \quad \text{où} \quad f(x,y) = x^3 - x^2 y + 2y^2.$$

- Déterminons les points qui annulent le gradient. On a,

$$\nabla f(x,y) = \begin{pmatrix} 3x^2 - 2xy \\ -x^2 + 4y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \begin{cases} 3x^2 - 2xy = 0 \\ -x^2 + 4y = 0. \end{cases}$$

  $(x,y) = (0,0)$ est une solution évidente du système. Si $x \neq 0$, alors on obtient à partir de la première équation du système $y = (3/2)x$. Ainsi, la deuxième équation devient $-x^2 + 6x = 0$, d'où $x = 6$ et $y = 9$. Il y a alors deux points stationnaires :

$$(0,0) \ \text{ et } \ (6,9).$$

- Calculons le Hessien de f :

$$H(x,y) = \begin{bmatrix} 6x - 2y & -2x \\ -2x & 4 \end{bmatrix}$$

  On a alors,

$$H(6,9) = \begin{bmatrix} 18 & -12 \\ -12 & 4 \end{bmatrix}$$

  qui n'est pas semi-définie positive, par conséquent $(6,9)$ n'est pas un minimum.

1.2 Conditions suffisantes d'optimalité

Théorème 1.3

Condition suffisante du 2ième ordre

Soit f de classe C^2. Une condition suffisante pour que x^* soit un minimum local de f est :

$$\nabla f\left(x^*\right)=0 \text{ et le Hessien } H\left(x^*\right) \text{ est une matrice définie positive.}$$

Preuve :

- Considérons un point x^* vérifiant les conditions suivantes :

$$\nabla f\left(x^*\right)=0 \text{ et } H\left(x^*\right) \text{ définie positive.}$$

- Le développement de Taylor de f au voisinage de x^* s'écrit :

$$f(x)=f\left(x^*\right)+\frac{1}{2}\left(x-x^*\right)^{\top} \cdot H\left(x^*\right) \cdot\left(x-x^*\right)+o\left(\left\|x-x^*\right\|^2\right), \tag{8}$$

où $o\left(\left\|x-x^*\right\|^2\right)=\left\|x-x^*\right\|^2 \varepsilon\left(x-x^*\right)$ avec $\varepsilon\left(x-x^*\right) \rightarrow 0$ lorsque $\left\|x-x^*\right\| \rightarrow 0$. Pour toute direction de déplacement $d \in \mathbb{R}^n$ ($\|d\|=1$), en prenant $x=x^*+\theta d$, on obtient à partir de (8) la nouvelle formule

$$f\left(x^*+\theta d\right)=f\left(x^*\right)+\frac{1}{2} \theta^2 d^{\top} \cdot H\left(x^*\right) \cdot d+o\left(\theta^2\right).$$

- Comme $H\left(x^*\right)$ est définie positive, alors pour θ suffisamment petit, on aura :

$$f\left(x^*+\theta d\right) \geqslant f\left(x^*\right).$$

Donc, x^* est bien un minimum local. □

Note : Dans le cas d'une fonction f définie sur $\mathbb{R}^2$ de classe C^2, on peut montrer le critère suivant : Pour tout point $\left(x^*, y^*\right)$ de $\mathbb{R}^2$ tel que $\nabla f\left(x^*, y^*\right)=0$, posons

$$\Delta=\frac{\partial^2 f}{\partial x^2}\left(x^*, y^*\right) \times \frac{\partial^2 f}{\partial y^2}\left(x^*, y^*\right)-\left(\frac{\partial^2 f}{\partial x \partial y}\left(x^*, y^*\right)\right)^2.$$

Alors :

1. $\left(x^*, y^*\right)$ est un minimum local, si $\Delta>0$ et $\frac{\partial^2 f}{\partial x^2}\left(x^*, y^*\right)>0$.
2. $\left(x^*, y^*\right)$ est un maximum local, si $\Delta>0$ et $\frac{\partial^2 f}{\partial x^2}\left(x^*, y^*\right)<0$.
3. Si $\Delta<0$, alors $\left(x^*, y^*\right)$ est un point-selle.

4 Si $\Delta = 0$, on ne peut rien conclure.

Exemple 1.3 : Revenons à l'exemple de minimisation précédant $\min x^3 - x^2y + 2y^2$.

- On peut voir que $(6, 9)$ est un point stationnaine et qu'il n'est pas un minimum.
- En réalité c'est un point-selle, car on observe de plus que

$$\Delta = 18 \times 4 - 12 \times 12 = 8(9 - 18) < 0$$

- Par contre pour le point stationncine $(0, 0)$, on ne peut rien conclure car $\Delta = O$ du fait que

$$H(0,0) = \begin{pmatrix} 0 & 0 \\ 0 & 4 \end{pmatrix}.$$

1.3 Cas des fonctions convexes et formes quadratiques

1.3.1 - Cas des fonctions convexes

Pour une fonction convexe, on a le théorème suivant.

Théorème 1.4

Soit f une fonction convexe de classe C^1. Une condition nécessaire et suffisante pour que x^* soit un optimum global de f est que $\nabla f(x^*) = 0$.

Preuve :

- **Condition nécessaire :** Si x^* optimum global, alors x^* est un optimum local et donc

$$\nabla f(x^*) = 0, \quad \text{(d'après le théorème 2.1)}$$

- **Condition suffisante :** f étant convexe, alors d'après Théorème 4.1 - chap. I, on a

$$f(x) \geqslant f(x^*) + \nabla f(x^*)^\top \cdot (x - x^*), \; \forall\, x \in \mathbb{R}^n$$

Et comme $\nabla f(x^*) = 0$, alors on obtient finalement

$$f(x) \geqslant f(x^*), \; \forall\, x \in \mathbb{R}^n, \quad \text{i.e.,}$$

x^* est un minimum global. $\square$

Théorème 1.5

Un minimum d'une fonction strictement convexe est unique.

Preuve :

- Supposons qu'il existe deux minimums x^* et ξ^* avec $x^* \neq \xi^*$. Alors,

$$f(x^*) \leqslant f(x), \ \forall\, x \in \mathbb{R}^n$$

et

$$f(\xi^*) \leqslant f(x), \ \forall\, x \in \mathbb{R}^n$$

Ce qui implique, en particulier, que $f(x^*) = f(\xi^*)$.

- Or f est strictement convexe, donc

$$\begin{aligned} & f(\lambda x^* + (1-\lambda)\xi^*) < \lambda f(x^*) + (1-\lambda) f(\xi^*), \ \forall\, \lambda \in]0,1[, \\ \text{i.e.}\quad & f(\lambda x^* + (1-\lambda)\xi^*) < f(x^*). \end{aligned}$$

D'où la contradiction ! Par suite, on a $x^* = \xi^*$. □

• 1.3.2 - Cas des formes quadratiques

Soit A une matrice symétrique d'ordre n. Soit la fonctionnelle quadratique

$$\begin{aligned} J : \mathbb{R}^n & \longmapsto \mathbb{R} \\ x & \longrightarrow J(x) = \frac{1}{2}(Ax,\, x) + (b,\, x) + c \\ & \qquad\quad = \frac{1}{2}\sum_{i,j=1}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n} b_i x_i + c, \end{aligned}$$

où b est un vecteur de $\mathbb{R}^n$ et $c \in \mathbb{R}$ une constante.

Note :

- Un point critique de la fonctionnelle quadratique $J(x) = \frac{1}{2}(Ax,\, x) + (b,\, x) + c$, s'obtient en résolvant le système linéaire

$$Ax + b = 0.$$

- Si A est semi-définie positive alors J est convexe et dans ce cas x^* est un minimum si et seulement si $Ax^* + b = 0$.
- Si de plus A est définie positive, alors J admet un minimum unique donné par

$$x^* = -A^{-1} b.$$

2 Méthodes numériques : cas des fonctions différentiables

Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonction de classe C^1. Comme, dans tous les cas, la stationnarité de f est une condition nécessaire d'optimalité, pratiquement toutes les méthodes d'optimisation sans contraintes consistent à rechercher un point x^* stationnaire, i.e.,

$$\nabla f\left(x^*\right) = 0.$$

Ce problème est équivalent à la résolution du système d'équations non linéaires :

$$\frac{\partial f}{\partial x_i}\left(x^*\right) = 0, \;\; i = 1, \ldots, n.$$

On peut chercher à résoudre directement ce système, ce qui conduit à la méthode de Newton. Cette méthode suppose la fonction f de classe C^2 et nécessite le calcul des dérivées secondes en chaque point.

C'est pourquoi les méthodes les plus couramment utilisées procèdent différemment : il s'agit de procédures itératives où l'on engendre une suite de points $x^0, x^1, \ldots, x^k$ convergeant vers un optimum local de f.

A chaque étape k, x^{k+1} est défini par $x^{k+1} = x^k + \alpha_k d^k$, où d^k est une direction de déplacement qui peut être ($\alpha_k > 0$) :

1. soit l'opposé du gradient de f en x_k : $d^k = -\nabla f\left(x^k\right)$;
2. soit une direction calculée à partir du gradient $\nabla f\left(x^k\right)$;
3. soit choisie de façon à constituer une direction de descente, c'est-à-dire telle que

$$\nabla f\left(x_k\right)^\top \cdot d^k < 0.$$

Note : La condition sur la direction d^k

$$\nabla f\left(x^k\right)^\top \cdot d^k < 0, \quad \text{(i.e. direction de descente)}$$

peut être expliquée de la manière suivante :

- A la $k^{\text{ième}}$ itération, on choisit une direction d^k et un scalaire $\alpha_k > 0$ tel que

$$f\left(x^{k+1}\right) < f\left(x^k\right) \tag{9}$$

- La direction d^k et le scalaire α_k permettent de calculer x^{k+1} par

$$x^{k+1} = x^k + \alpha_k d^k \tag{10}$$

- On peut caractériser la direction d^k en considérant les développements en série de Taylor de $f\left(x^{k+1}\right)$ en fonction de $f\left(x^k\right)$ et $\nabla f\left(x^k\right)$ quand α_k tend vers 0,

$$f\left(x^{k+1}\right) = f\left(x^k + \alpha_k d^k\right) \simeq f\left(x^k\right) + \alpha_k \nabla f\left(x_k\right)^\top \cdot d^k$$

Pour satisfaire (9), le choix de d^k doit être tel que $\nabla f\left(x^k\right)^T \cdot d^k < 0$.

2.1 Méthode de gradient

Il s'agit ici de se déplacer dans la direction opposée au gradient, c'est-à-dire

$$d^k = -\frac{\nabla f\left(x^k\right)}{\|\nabla f\left(x^k\right)\|}.$$

Algorithme : L'algorithme est le suivant :

- On part d'un point x^0, on calcule le gradient $\nabla f\left(x^0\right)$ et définit le point x^1 par

$$x^1 = x^0 - \alpha_0 \frac{\nabla f\left(x^0\right)}{\|\nabla f\left(x^0\right)\|}, \quad \text{où} \quad \alpha_0 > 0.$$

- La procédure est répétée et engendre les points $x^0, x^1, \ldots, x^k$ suivant la relation

$$x^{k+1} = x^k - \alpha_k \frac{\nabla f\left(x^k\right)}{\|\nabla f\left(x^k\right)\|}, \quad \text{où} \quad \alpha_k > 0, \ \forall\, k.$$

2.2 Méthode du gradient à pas déterminé

C'est la méthode de gradient dans laquelle on choisit a priori les valeurs de déplacements α_k. L'inconvenient de cette méthode est que la convergence peut être très lente. Le principal intérêt des méthodes de gradient à pas déterminé est de se généraliser au cas des fonctions non partout différentiables.

2.3 Méthode de la plus forte pente (steepest descent)

• 2.3.1 - Fonction quelconque de classe C^1

Dans cette méthode d'utilisation très courante, on choisit $d^k = -\nabla f\left(x^k\right)$ et α_k est choisi de façon à minimiser la fonction d'une variable α :

$$g(\alpha) = f\left(x^k - \alpha \nabla f\left(x^k\right)\right)$$

sur l'ensemble des $\alpha \geqslant 0$ (minimisation unidimensionnelle). On est alors conduit à la méthode itérative suivante :

Algorithme :

1. Choisir un point de départ x^0 (correspondant à $k = 0$).
2. À l'itération k :
   - Calculer : $d^k = -\nabla f\left(x^k\right)$.
   - Rechercher α_k tel que :

$$f\left(x^k + \alpha_k d^k\right) = \min_{\alpha \geqslant 0} f\left(x^k + \alpha d^k\right) \iff g(\alpha_k) = \min_{\alpha \geqslant 0} g(\alpha).$$

- Faire $x^{k+1} = x^k + \alpha_k d^k$.

3 Test d'arrêt (ou critère d'arrêt) :

- Si vérifié : FIN.
- Sinon, faire $k \longleftarrow k+1$ et retourner en 2.

Comme la convergence n'est pas, en général, finie on doit définir un test d'arrêt. Donnons, à titre indicatif, quelques uns des critères d'arrêt les plus utilisés :

- Critère 1 :

$$\max_{i=1,\dots,n} \left| \frac{\partial f}{\partial x_i} \right| < \varepsilon$$

- Critère 2 :

$$\|\nabla f\|^2 = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i} \right)^2 < \varepsilon$$

- Critère 3 :

$$\left| f\left(x^{k+1}\right) - f\left(x^k\right) \right| < \varepsilon$$

où $\varepsilon > 0$, est donné.

Concernant la convergence globale de la plus forte pente, on peut énoncer

Proposition 2.1

Si f est une fonction de classe C^1 et si $f(x) \longrightarrow +\infty$ pour $\|x\| \longrightarrow +\infty$, alors pour tout point de départ x^0, la méthode de la plus forte pente (avec optimisation unidimensionnelle exacte ou approchée) converge vers un point stationnaire de f.

Note :

- Rappelons ici que la convergence globale ne signifie pas que l'on obtient nécessairement un optimum de f.
- Si f est de classe C^1, tout ce que l'on peut dire c'est que l'on obtient un point stationnaire x^* de f.
- Si f est de classe C^2 et que $H\left(x^*\right)$ est définie positive, alors x^* est un minimum local de f.

C'est seulement dans des cas bien particuliers, (f convexe différentiable par exemple) que l'on peut être sûr d'obtenir un minimum global de f.

Note : Le principal défaut de la méthode de la plus forte pente est que la convergence peut être très lente pour certains types de fonctions. En effet ;

- α_k doit minimiser la fonction d'une variable

$$g(\alpha) = f\left(x^k + \alpha d^k\right), \quad \text{pour} \quad \alpha \geqslant 0.$$

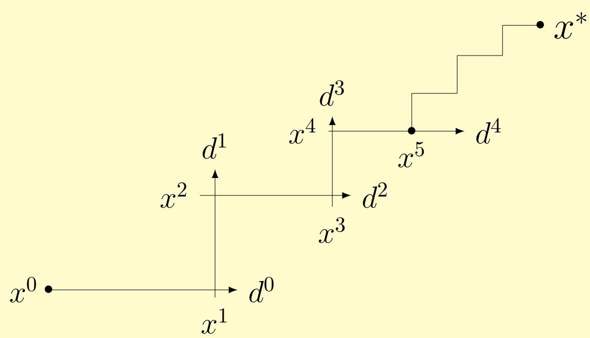
- Alors, on doit avoir

$$\begin{aligned}\frac{dg}{d\alpha}(\alpha_k) &= \left(d^k\right)^\top \cdot \nabla f\left(x^k + \alpha_k d^k\right)\\ &= \left(d^k\right)^\top \cdot \nabla f\left(x^{k+1}\right) = 0.\end{aligned}$$

- D'où l'on obtient :

$$\left(d^k\right)^\top \cdot d^{k+1} = 0.$$

Ce qui prouve que les directions de déplacement successives sont orthogonales.



Exemple 2.1 : Nous utilisons la méthode de la plus forte pente pour minimiser

$$f(x_1, x_2, x_3) = (x_1 - 4)^4 + (x_2 - 3)^2 + 4(x_3 + 5)^4.$$

Nous allons effectuer trois itérations à partir du point initial $\boldsymbol{x}^{(0)} = (4,\ 2,\ -1)^\top$.

- Pour la première itération, Nous trouvons d'abord le gradient de la fonction f

$$\nabla f(x_1, x_2, x_3) = \left(4(x_1 - 4)^3,\ 2(x_2 - 3),\ 16(x_3 + 5)^3\right)^\top.$$

Ainsi,

$$\nabla f\left(\boldsymbol{x}^{(0)}\right) = (0,\ -2,\ 1024)^\top.$$

Pour calculer $\boldsymbol{x}^{(1)}$, nous avons besoin

$$\begin{aligned}\alpha_0 &= \arg\min_{\alpha\geq 0} f\left(\boldsymbol{x}^{(0)} - \alpha\nabla f\left(\boldsymbol{x}^{(0)}\right)\right)\\ &= \arg\min_{\alpha\geq 0}\left(0 + (2+2\alpha-3)^2 + 4(-1-1024\alpha+5)^4\right)\\ &= \arg\min_{\alpha\geq 0}\phi_0(\alpha).\end{aligned}$$

En utilisant la méthode sécante, nous obtenons

$$\alpha_0 = 3.967\times 10^{-3}.$$

Ainsi

$$\begin{aligned}\boldsymbol{x}^{(1)} &= \boldsymbol{x}^{(0)} - \alpha_0\nabla f\left(\boldsymbol{x}^{(0)}\right)\\ &= (4.000,\ 2.008,\ -5.062)^\top.\end{aligned}$$

- Pour trouver $\boldsymbol{x}^{(2)}$, nous déterminons d'abord

$$\nabla f\left(\boldsymbol{x}^{(1)}\right) = (0.000, -1.984, -0.003875)^\top.$$

Ensuite, nous trouvons α_1, où

$$\begin{aligned}\alpha_1 &= \arg\min_{\alpha\geq 0}\left(0 + (2.008+1.984\alpha-3)^2 + 4(-5.062+0.003875\alpha+5)^4\right)\\ &= \arg\min_{\alpha\geq 0}\phi_1(\alpha).\end{aligned}$$

En utilisant à nouveau la méthode sécante, on obtient $\alpha_1 = 0.5000$. Ainsi

$$\boldsymbol{x}^{(2)} = \boldsymbol{x}^{(1)} - \alpha_1\nabla f\left(\boldsymbol{x}^{(1)}\right) = (4.000, 3.000, -5.060)^\top.$$

- Pour trouver $\boldsymbol{x}^{(3)}$, nous déterminons d'abord

$$\nabla f\left(\boldsymbol{x}^{(2)}\right) = (0.000,\ 0.000,\ -0.003525)^\top$$

et

$$\begin{aligned}\alpha_2 &= \arg\min_{\alpha\geq 0}\left(0.000 + 0.000 + 4(-5.060+0.003525\alpha+5)^4\right)\\ &= \arg\min_{\alpha\geq 0}\phi_2(\alpha).\end{aligned}$$

On procède comme dans les itérations précédentes pour obtenir $\alpha_2 = 16.29$. La valeur de $x^{(3)}$ est

$$\boldsymbol{x}^{(3)} = (4.000,\ 3.000,\ -5.002)^\top.$$

Notez que le minimiseur de f est $(4,\ 3,\ -5)^\top$, et il semble donc que nous soyons arrivés au minimiseur en seulement trois itérations.

- ### 2.3.2 - Cas d'une fonction quadratique

Soit $f(x)=\frac{1}{2}(Ax,x)+(b,x)+c$, où A est une matrice symétrique d'ordre n, b un vecteur de $\mathbb{R}^n$ et c un scalaire. On a $d^k=-\nabla f\left(x^k\right)=-\left(Ax^k+b\right)$, et donc

$$\begin{aligned} g(\alpha)=f\left(x^k+\alpha d^k\right) &= \frac{1}{2}\left(A\left(x^k+\alpha d^k\right), x^k+\alpha d^k\right)+\left(b,\, x^k+\alpha d^k\right)+c \\ &= \frac{1}{2}\left(Ax^k,\, x^k\right)+\left(b,\, x^k\right)+c \\ &+ \frac{\alpha^2}{2}\left(Ad^k, d^k\right)+\alpha\left(Ax^k,\, d^k\right)+\alpha\left(b,\, d^k\right). \end{aligned}$$

Alors, la dérivée de g par rapport à α donne :

$$\left.\frac{dg(\alpha)}{d\alpha}\right|_{\alpha=\alpha_k}=0 \Rightarrow \alpha_k\left(Ad^k,\, d^k\right)+\left(Ax^k+b,\, d^k\right)=0.$$

Or $\nabla f\left(x^k\right)=Ax^k+b$ et $d^k=-\nabla f\left(x^k\right)$, donc $\alpha_k\left(Ad^k,\, d^k\right)=\left(d^k,\, d^k\right)=\left\|d^k\right\|^2$. D'où

$$\alpha_k=\frac{\left\|d^k\right\|^2}{\left(Ad^k,\, d^k\right)}.$$

Cette dernière formule a un sens dans le cas (par exemple) où A est définie positive.

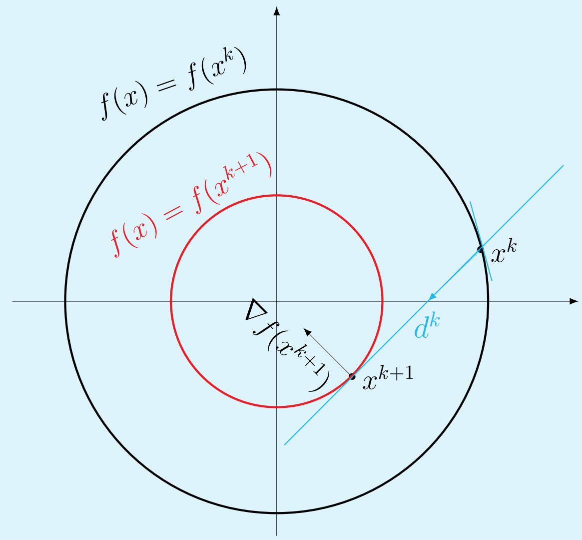
Une itération de la méthode devient (steepest descent dans le cas quadratique) :

- $d^k=-\left(Ax^k+b\right)$
- $\alpha_k=\frac{\left\|d^k\right\|^2}{\left(Ad^k,d^k\right)}$
- $x^{k+1}=x^k+\alpha_k d^k$

Avant de donner un exemple, donnons une interprétation géométrique dans $\mathbb{R}^2$ de la méthode de la plus forte pente.

Soit f une fonction définie sur $\mathbb{R}^2$, on notera par x le point de composante (x_1,x_2).

- L'ensemble des points x pour lesquels $f(x)=$ cte forme une courbe du plan $\mathbb{R}^2$.
- Pour les différentes valeurs de x^k, on obtient une famille de courbes $f(x)=f\left(x^k\right)$ et qui représentent les courbes de niveau.
- Le vecteur $\nabla f\left(x^{k+1}\right)$ est orthogonal à la courbe de niveau $f(x)=f\left(x^{k+1}\right)$.
- Comme le vecteur $d^{k+1}=-\nabla f\left(x^{k+1}\right)$ est orthogonal au vecteur d^k, alors d^k est tangent à la courbe de niveau $f(x)=f\left(x^{k+1}\right)$.



Exemple 2.2 : Soit à résoudre le problème de minimisation suivant :

$$\min_{(x,y)\in\mathbb{R}^2} f(x,y) = \min_{(x,y)\in\mathbb{R}^2} \frac{1}{2}\left(x^2 + ay^2\right), \quad \text{où } \ a \geq 1 \ \text{ est donné.}$$

- Ce problème rentre dans le cadre de formes quadratiques. En effet, on a :

$$f(x,y) = \frac{1}{2}\left(A\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix}\right), \quad \text{où } \ A = \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix}$$

- Remarquons que A est définie positive, et donc f est strictement convexe. Alors, f admet un minimum global unique qui est caractérisé par $\nabla f(x^*, y^*) = (0,0)$:

$$A\begin{pmatrix} x^* \\ y^* \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{i.e.} \quad \begin{cases} x^* = 0 \\ ay^* = 0 \end{cases} \quad \text{c'est-à-dire} \quad (x^*, y^*) = (0,0),$$

est le minimmm (ce qui est évident à voir à partir de la définition de f).

- Maintenant, appliquons la méthode de la plus forte pente à ce poblème. On a

$$\nabla f(x,y) = \begin{pmatrix} x \\ ay \end{pmatrix} \quad \text{donc} \quad \begin{cases} d^k = -\nabla f\left(x_k, y_k\right) = -\begin{pmatrix} x_k \\ ay_k \end{pmatrix} \\ \text{et} \\ \left\|d^k\right\|^2 = x_k^2 + a^2 y_k^2 \end{cases}$$

Sachant que $\alpha_k = \frac{\|d^k\|^2}{(Ad^k, d^k)}$, alors on effectue d'abors les calculs

$$Ad^k = \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix}\begin{pmatrix} -x_k \\ -ay_k \end{pmatrix} = \begin{pmatrix} -x_k \\ -a^2 y_k \end{pmatrix}$$

$$\left(Ad^k, d^k\right) = x_k^2 + a^3 y_k^2$$

Donc le calcul de α_k donne $\alpha_k = \frac{x_k^2+a^2y_k^2}{x_k^2+a^3y_k^2}$, et on a l'algorithme suivant :

$$\begin{cases} x_{k+1} = x_k + \alpha_k\left(-x_k\right) \\ y_{k+1} = y_k + \alpha_k\left(-ay_k\right) \end{cases} \iff \begin{cases} x_{k+1} = x_k - \frac{x_k^2+a^2y_k^2}{x_k^2+a^3y_k^2} \times x_k \\ \\ y_{k+1} = y_k - a \times \frac{x_k^2+a^2y_k^2}{x_k^2+a^3y_k^2} \times y_k \end{cases}$$

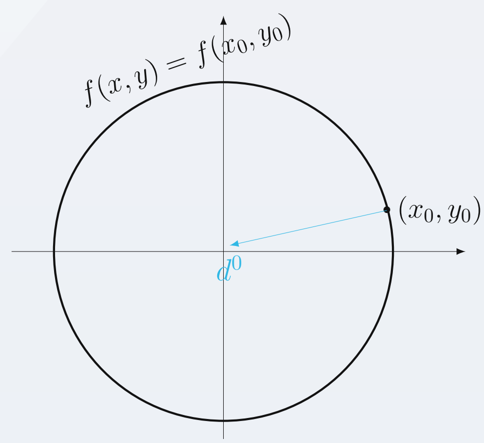
$$\text{donc} \quad \begin{cases} x_{k+1} = \frac{a^2(a-1)y_k^2}{x_k^2+a^3y_k^2} \times x_k \\ \\ y_{k+1} = \frac{(1-a)x_k^2}{x_k^2+a^3y_k^2} \times y_k \end{cases}$$

- Il y a deux cas à distinguer pour les valeurs du paramètre a :
  - ✓ 1^er^ cas, $a = 1$

    Alors on a, pour tout point de départ (x_0, y_0) : $\begin{cases} x_1 = 0 \\ y_1 = 0 \end{cases}$; c'est-à-dire que le minimum est atteint en une seule itération. Du point de vu géométrique, l'équation suivante

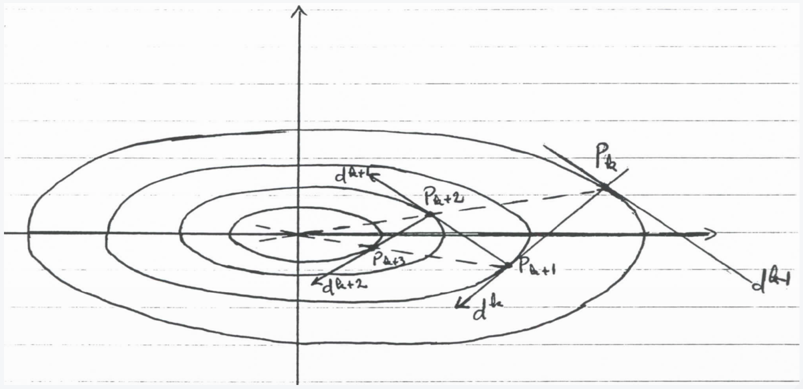
    $$\frac{1}{2}\left(x^2+y^2\right) = \text{ cte},$$

    est celle d'un cercle centré à l'origine. D'où, on a le schéma suivant

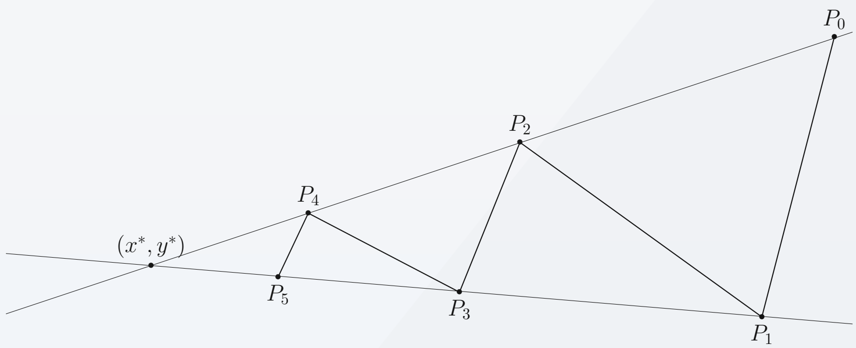
    

    de n'importe quel point (x_0, y_0), le gradient pointe vers le centre (qui est l'origine).
  - ✓ 2^e^ cas, $a > 1$

    Dans ce cas, les courbes de niveau $\frac{1}{2}\left(x^2+ay^2\right) = \text{cte}$ sont des éllipses



Les itérés successifs sont situés sur deux droites passant par l'origine.



Ce dernier résultat, peut être montré analytiquement. En effet, on a :

$$\begin{cases} x_{k+1} = \frac{a^2(a-1)y_k^2}{x_k^2+a^3y_k^2} \times x_k \\ \\ y_{k+1} = \frac{(1-a)x_k^2}{x_k^2+a^3y_k^2} \times y_k \end{cases}$$

donc en divisant ces deux équations terme à terme, on obtient l'égalité

$$\frac{y_{k+1}}{x_{k+1}} = -\frac{(1-a)x_k^2 y_k}{a^2(1-a)y_k^2 x_k} = -\frac{1}{a^2}\frac{x_k}{y_k}$$

Ainsi, on n'en tire le rapport $\frac{y_k}{x_k} = -\frac{1}{a^2}\frac{x_{k-1}}{y_{k-1}}$, et donc $\frac{y_{k+1}}{x_{k+1}} = \frac{y_{k-1}}{x_{k-1}}$. Par suite, en prenant $k = 1, 3, 5, \cdots, 2p-1$, on a

$$\frac{y_{2p}}{x_{2p}} = \frac{y_{2p-2}}{x_{2p-2}} = \cdots = \frac{y_0}{x_0}$$

Et en prenant $k = 2, 4, 6, \cdots, 2p$, on obtient

$$\frac{y_{2p+1}}{x_{2p+1}} = \frac{y_{2p-1}}{x_{2p-1}} = \cdots = \frac{y_1}{x_1}$$

On en tire les équations suivantes

$$x_0 y_{2k} - y_0 x_{2k} = 0$$

et

$$x_1 y_{2k+1} - y_1 x_{2k+1} = 0$$

d'où, on obtient les équations de droite suivantes

$$\begin{aligned} P_{2k} \in D_0 : \; x_0 y - y_0 x = 0, \\ P_{2k+1} \in D_1 : \; x_1 y - y_1 x = 0. \end{aligned}$$

2.4 Méthode de directions conjuguées : principe général

Pour améliorer la convergence des méthodes de gradient, d'autres méthodes plus élaborées ont été proposées. Elles sont basées sur la remarque suivante :
Puisque, en un minimum local x^* on a $\nabla f(x^*) = 0$, le développement de f en série de Taylor s'écrit alors au voisinage de x^* comme suit :

$$f(x) \simeq f(x^*) + \frac{1}{2}(x - x^*)^\top \cdot H(x^*) \cdot (x - x^*).$$

La fonction f se comporte donc, au voisinage de x^* comme une fonction quadratique. Il en résulte qu'une méthode itérative générale de minimisation, pour être efficace, doit au moins converger rapidement sur des fonctions quadratiques.

Dans cette optique, les méthodes de directions conjuguées sont des méthodes itératives qui, appliquées à une fonction quadratiques de n variables conduisent à l'optimum en n itérations au plus. Considérons une fonction quadratique quelconque :

$$\begin{aligned} J : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ x &\longrightarrow J(x) = \frac{1}{2}(Ax, x) + (b, x) + c, \end{aligned}$$

où A est une matrice symétrique, définie positive et d'ordre n, $b \in \mathbb{R}^n$ et $c \in \mathbb{R}$.
Le principe des méthodes de directions conjuguée consiste, à partir d'un point x^0, à minimiser $J(x)$ successivement suivant n directions linéairement indépendantes $d^0, d^1, \ldots, d^{n-1}$ possédant la propriété d'être mutuellement A - conjuguées, i.e.,

$$\left(Ad^i, d^j\right) = 0, \; \forall \, i \neq j \quad (i, j = 0, \cdots, n-1) \tag{11}$$

Supposons donc que, $\forall \, k = 0, 1, \ldots, n-1$, x^{k+1} soit déterminé à partir de x^k par :

$$x^{k+1} = x^k + \lambda_k d^k,$$

où λ_k est la valeur de λ qui minimise $J\left(x^k + \lambda d^k\right)$:

$$J\left(x^k + \lambda_k d^k\right) = \min_\lambda J\left(x^k + \lambda d^k\right), \; \forall \, k = 0, 1, \ldots, n-1.$$

Nous allons voir que, dans ces conditions le point obtenu à la $n^{\text{ième}}$ itération, c'est-à-dire

$$x^n = x^0 + \sum_{j=0}^{n-1} \lambda_j d^j,$$

est nécessairement l'optimum du problème, autrement dit, vérifie

$$Ax^n + b = \nabla J\left(x^n\right) = 0.$$

Comme λ_k minimise J dans la direction d^k, on a pour tout $k = 0, 1, \ldots, n-1$:

$$\left(\nabla J\left(x^k + \lambda_k d^k\right), d^k\right) = \left(A\left(x^k + \lambda_k d^k\right) + b, d^k\right) = 0.$$

d'où l'on tire

$$\lambda_k = -\frac{\left(Ax^k + b, d^k\right)}{\left(Ad^k, d^k\right)}.$$

On peut alors démontrer la propriété suivante, caractéristique de toutes les méthodes de directions conjuguées !

Proposition 2.2

$x^n = x^0 + \sum_{j=0}^{n-1} \lambda_j d^j$ est l'optimum de la fonction quadratique $J(x)$ sur $\mathbb{R}^n$.

Preuve :

- Observons que pour tout $i = 0, \cdots, n-1$, nous avons :

$$\begin{aligned} \left(Ax^n, d^i\right) &= \left(Ax^0, d^i\right) + \sum_{j=0}^{n-1} \lambda_j \left(Ad^j, d^i\right) \\ &= \left(Ax^0, d^i\right) + \lambda_i \left(Ad^i, d^i\right), \quad \text{grâce à } (11) \end{aligned}$$

- Comme $\lambda_i = -\frac{\left(Ax^i + b, d^i\right)}{\left(Ad^i, d^i\right)}$ et que

$$\begin{aligned} \left(Ax^i, d^i\right) &= \left(Ax^0, d^i\right) + \sum_{j=0}^{i-1} \lambda_j \left(Ad^j, d^i\right) \\ &= \left(Ax^0, d^i\right) \end{aligned}$$

alors $\lambda_i = -\frac{\left(Ax^0 + b, d^i\right)}{\left(Ad^i, d^i\right)}$ donc,

$$\begin{aligned} \left(Ax^n, d^i\right) &= \left(Ax^0, d^i\right) - \left(Ax^0 + b, d^i\right) \\ &= -\left(b, d^i\right) \end{aligned}$$

et par suite pour tout $i = 0, 1, \ldots, n-1$, on a : $\left(Ax^n + b,\, d^i\right) = 0$.

- Ceci entraine que $Ax^n + b$ est orthogonal à $\mathbb{R}^n$ (car $d^0, d^1, \ldots, d^{n-1}$ sont linéairement indépendantes et donc engendrent $\mathbb{R}^n$), et par conséquent $Ax^n + b = 0$. □

Note :

- Une méthode de directions conjuguée converge donc de façon finie en au plus n itérations dans le cas d'une fonction quadratique dans $\mathbb{R}^n$.
- Comme il existe un certain nombre de degrés de liberté pour le choix des directions $(d^0, d^1, \ldots, d^{n-1})$, on peut imaginer de nombreux algorithmes basés sur le principe général des directions conjuguées.
- Nous étudierons plus particulièrement la méthode dite du gradient conjugué pour les fonctions quadratiques et la méthode de Fletcher et Reeves pour les fonctions quelconques.

2.5 Méthode de gradient conjugué pour les fonctions quadratiques

• 2.5.1 - Analyse de la méthode

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

$$J(x) = \frac{1}{2}(Ax,\, x) + (b,\, x) + c.$$

L'idée de la méthode de gradient conjugué est de construire progressivement des directions $d^0, d^1, \ldots, d^k$ mutuellement A - conjuguées.
A l'étape k, la direction d^k est obtenue par combinaison linéaire

- ✓ du gradient $-\nabla J\left(x^k\right)$ en x^k,
- ✓ et des directions précédentes $d^0, d^1, \ldots, d^{k-1}$,
- ✓ les coefficients de la combinaison linéaire étant choisis de telle sorte que d^k soit conjuguée par rapport à toutes les directions précédentes.

A partir de x^0 donné, x^{k+1} est choisi sous la forme :

$$\begin{aligned} x^{k+1} &= x^k + \alpha_k d^k \\ \text{où} \qquad \alpha_k &= -\frac{\left(Ax^k + b,\, d^k\right)}{\left(Ad^k,\, d^k\right)} \qquad \left(\begin{array}{l} \alpha_k \text{ minimisant} \\ J\left(x^k + \alpha d^k\right) \end{array} \right) \\ &= \frac{\left(r^k,\, d^k\right)}{\left(Ad^k,\, d^k\right)} \end{aligned}$$

avec, par difinition, le résidu r^k est

$$r^k = -\left(Ax^k + b\right)\left(= -\nabla J\left(x^k\right)\right),$$

et où la direction d^k est choisie par la formule suivante

$$d^k = r^k + \beta_k d^{k-1}, \ \text{ pour } \ k \geq 1,$$

où $d^0 = r^0$. Alors, le coefficient β_k est choisi à travers la formule suivante

$$\beta_k = -\frac{\left(Ad^{k-1}, r^k\right)}{\left(Ad^{k-1}, d^{k-1}\right)} \qquad \left(\begin{array}{l}\beta_k \text{ est obtenu à travers} \\ \text{la condition } \left(Ad^k, d^{k-1}\right) = 0\end{array}\right)$$

On a finalement le résultat suivant :

Théorème 2.1

1. $\left(r^{k+1},\, d^k\right) = 0$
2. $\left(r^{k+1},\, r^k\right) = 0$
3. $\alpha_k = \frac{\left(r^k, r^k\right)}{\left(Ad^k, d^k\right)}$
4. $\beta_k = \frac{\left(r^k, r^k\right)}{\left(r^{k-1}, r^{k-1}\right)}$
5. $r^n = -\left(Ax^n + b\right) = 0.$

Preuve :

1

$$\begin{aligned}
\left(r^{k+1},\, d^k\right) &= -\left(Ax^{k+1} + b,\, d^k\right) \\
&= -\left(Ax^k + \alpha_k Ad^k + b,\, d^k\right) \\
&= -\left(Ax^k + b,\, d^k\right) - \alpha_k \left(Ad^k,\, d^k\right) \\
&= \left(r^k,\, d^k\right) - \frac{\left(r^k,\, d^k\right)}{\left(Ad^k,\, d^k\right)} \left(Ad^k,\, d^k\right) = 0.
\end{aligned}$$

2

$$\begin{aligned}
\left(r^{k+1},\, r^k\right) &= \left(r^{k+1},\, d^k - \beta_k d^{k-1}\right) \\
&= -\beta_k \left(r^{k+1},\, d^{k-1}\right) \\
&= \beta_k \left(Ax^{k+1} + b,\, d^{k-1}\right) \\
&= \beta_k \left(Ax^k + \alpha_k Ad^k + b,\, d^{k-1}\right) \\
&= \beta_k \left(Ax^k + b, d^{k-1}\right) + \beta_k \alpha_k \left(Ad^k,\, d^{k-1}\right) \\
&= -\beta_k \left(r^k,\, d^{k-1}\right) = 0. \quad (\text{pour } k \geqslant 1)
\end{aligned}$$

Pour $k = 0$, on a $\left(r^1,\, r^0\right) = \left(r^1, d^0\right) = 0$, d'après le point 1.

3

$$\begin{aligned}
\alpha_k = \frac{\left(r^k, d^k\right)}{\left(Ad^k, d^k\right)} &= \frac{\left(r^k, r^k + \beta_k d^{k-1}\right)}{\left(Ad^k, d^k\right)} \\
&= \frac{\left(r^k, r^k\right) + \beta_k \left(r^k, d^{k-1}\right)}{\left(Ad^k, d^k\right)} = \frac{\left(r^k, r^k\right)}{\left(Ad^k, d^k\right)}.
\end{aligned}$$

4

$$\text{On a}\quad \beta_k = -\frac{\left(Ad^{k-1},\, r^k\right)}{\left(Ad^{k-1},\, d^{k-1}\right)},\quad \text{or}\quad d^{k-1} = \frac{1}{\alpha_{k-1}}\left(x^k - x^{k-1}\right)$$

$$\text{donc}\quad Ad^{k-1} = \frac{1}{\alpha_{k-1}}\left(Ax^k - Ax^{k-1}\right)$$

$$= \frac{1}{\alpha_{k-1}}\left(\left(Ax^k + b\right) - \left(Ax^{k-1} + b\right)\right)$$

$$\text{i.e.,}\quad Ad^{k-1} = \frac{1}{\alpha_{k-1}}\left(r^{k-1} - r^k\right)$$

$$\text{d'où}\quad \beta_k = \frac{1}{\alpha_{k-1}}\frac{\left(r^k - r^{k-1},\, r^k\right)}{\left(Ad^{k-1},\, d^{k-1}\right)}$$

$$= \frac{1}{\alpha_{k-1}}\frac{\left(r^k,\, r^k\right)}{\left(Ad^{k-1},\, d^{k-1}\right)}$$

$$= \frac{\left(Ad^{k-1},\, d^{k-1}\right)}{\left(r^{k-1},\, r^{k-1}\right)}\frac{\left(r^k,\, r^k\right)}{\left(Ad^{k-1},\, d^{k-1}\right)}$$

$$\text{i.e.,}\quad \beta_k = \frac{\left(r^k,\, r^k\right)}{\left(r^{k-1}, r^{k-1}\right)}$$

5 D'après la démonstration de la proposition pécédente, il suffit de montrer que $(d^0, d^1, \ldots, d^{n-1})$ sont linéairement indépendantes et qu'elles sont mutuellement A - conjuguées.

- Commençons par montrer que $(d^0, d^1, \ldots, d^{n-1})$ sont mutuellement A - conjuguées. Nous allons raisonner par récurrence :
  - ✓ Tout d'abord pour (d^0, d^1), on a $(Ad^1, d^0) = 0$ par définition de β_1.
  - ✓ Supposons que $\left(d^0, d^1, \ldots, d^k\right)$ sont mutuellement A - conjuguées et montrons que c'est vrai pour $\left(d^0, d^1, \ldots, d^{k+1}\right)$.
  - ✓ Tout d'abord on a $\left(Ad^{k+1},\, d^k\right) = 0$ par définition de β_{k+1}. Alors, il reste à montrer que

$$\left(Ad^{k+1}, d^i\right) = 0,\ \forall\ i \in \{0, 1, \ldots, k-1\}.$$

Pour cela, vérifions tout d'abord que r^{k+1} est orthogonal à d^i, $\forall\ i\ \in$

$\{0,1,\ldots,k\}$. In effet, on a

$$\begin{aligned}
\left(Ax^{k+1}, d^i\right) &= \left(Ax^0, d^i\right) + \sum_{j=0}^{k} \alpha_j \left(Ad^j,\, d^i\right) \\
&= \left(Ax^0,\, d^i\right) + \alpha_i \left(Ad^i,\, d^i\right) \quad \text{(hypothèse de récurrence)} \\
&= \left(Ax^0,\, d^i\right) - \frac{\left(Ax^0 + b,\, d^i\right)}{\left(Ad^i,\, d^i\right)} \left(Ad^i,\, d^i\right) \\
&= -\left(b,\, d^i\right)
\end{aligned}$$

i.e., $\left(Ax^{k+1} + b,\, d^i\right) = 0,\ \forall\, i \in \{0,1,\ldots,k\}$

c.a.d. $\left(r^{k+1},\, d^i\right) = 0,\ \forall\, i \in \{0,1,\ldots,k\}$

Maintenant, soit $i \in \{0,1,\ldots,k-1\}$, on a $d^{k+1} = r^{k+1} + \beta_{k+1} d^k$ donc

$$\begin{aligned}
\left(Ad^{k+1}, d^i\right) &= \left(Ar^{k+1},\, d^i\right) + \beta_{k+1} \left(Ad^k,\, d^i\right) \\
&= \left(r^{k+1},\, Ad^i\right)
\end{aligned}$$

or $d^i = \dfrac{1}{\alpha_i}\left(x^{i+1} - x^i\right)$ donc $Ad^i = -\dfrac{1}{\alpha_i}\left(r^{i+1} - r^i\right)$

$$= -\frac{1}{\alpha_i}\left(d^{i+1} - \beta_{i+1} d^i - d^i + \beta_i d^{i-1}\right)$$

c.à.d. Ad_i est combinaison linéaire de d^{i-1}, d^i et d^{i+1}, qu'on écrit sous la forme $Ad^i = ad^{i-1} + bd^i + cd^{i+1}$. Donc

$$\begin{aligned}
\left(Ad^{k+1},\, d^i\right) &= \left(r^{k+1},\, Ad^i\right) \\
&= a\left(r^{k+1},\, d^{i-1}\right) + b\left(r^{k+1},\, d^i\right) + c\left(r^{k+1},\, d^{i+1}\right)
\end{aligned}$$

avec $i \leqslant k-1 \Rightarrow i+1 \leqslant k$, d'où $\left(Ad^{k+1},\, d^i\right) = 0, \quad \forall\, i \leq k-1$, car $\left(r^{k+1}, d^j\right) = 0,\ \forall\, j \leqslant k$.

- Il reste à montrer que $(d^0, d^1, \ldots, d^{n-1})$ sont linéairement indépendantes. Soit

$$\begin{aligned}
\sum_{i=0}^{n-1} a_i d^i = 0 \ &\Rightarrow\ \sum_{i=0}^{n-1} a_i \left(Ad^i, d^j\right) = 0 \\
&\Rightarrow\ a_i \left(Ad^i, d^i\right) = 0 \\
&\Rightarrow\ \quad a_i = 0, \ \text{car } A \text{ est définie positive,}
\end{aligned}$$

(à condition que $d^i \neq 0$ mais si $d^i = 0$ alors l'optimum est atteint). Ce qui termine la démonstration. $\square$

Algorithme : L'algorithme du gradient conjugué pour minimiser la quadratique

$$J(x) = \frac{1}{2}(Ax, x) + (b, x) + c$$

s'écrit alors :

i x^0 donné, $d^0 = r^0 = -(Ax^0 + b)$.

ii Etape k :

1. $\alpha_k = \frac{\|r^k\|^2}{(Ad^k, d^k)}$
2. $x^{k+1} = x^k + \alpha_k d^k$
3. $r^{k+1} = r^k - \alpha_k A d^k$
4. $\beta_{k+1} = \frac{\|r^{k+1}\|^2}{\|r^k\|^2}$
5. $d^{k+1} = r^{k+1} + \beta_{k+1} d^k$
6. faire $k \longleftarrow k+1$ et retourner en ii.

Exemple 2.3 : Soit à résoudre le problème

$$\min_{(x,y)\in\mathbb{R}^2} f(x,y), \quad \text{où} \quad f(x,y) = \frac{1}{2}\left(x^2 + ay^2\right) \text{ avec } a > 0.$$

- $\nabla f(x,y) = \begin{pmatrix} x \\ ay \end{pmatrix}$ et $A = H(x,y) = \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix}$ qui est difinie positive pour $a > 0$.
- Nous allons appliquer l'algorithme du gradient conjugué à ce poblème en partant du point $(x_0, y_0) = (1, 1)$.

$$d^0 = r^0 = -Ax^0 = -\begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -a \end{pmatrix}$$

$$\alpha_0 = \frac{\|r^0\|^2}{(Ad^0, d^0)}, \quad \|r^0\|^2 = 1 + a^2$$

$$Ad^0 = \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix}\begin{pmatrix} -1 \\ -a \end{pmatrix} = \begin{pmatrix} -1 \\ -a^2 \end{pmatrix}$$

$$\left(Ad^0, d^0\right) = 1 + a^3$$

Donc $\alpha_0 = \frac{1+a^2}{1+a^3}$

$$\begin{aligned} x^1 &= x^0 + \alpha_0 d^0 \\ &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{1+a^2}{1+a^3} \begin{pmatrix} -1 \\ -a \end{pmatrix} = \begin{pmatrix} 1 - \frac{1+a^2}{1+a^3} \\ 1 - a\frac{1+a^2}{1+a^3} \end{pmatrix} \\ &= \frac{1}{1+a^3} \begin{pmatrix} a^3 - a^2 \\ 1 - a \end{pmatrix} = \frac{a-1}{1+a^3} \begin{pmatrix} a^2 \\ -1 \end{pmatrix} \end{aligned}$$

e.i. $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \frac{a-1}{1+a^3} \begin{pmatrix} a^2 \\ -1 \end{pmatrix}$

$$\begin{aligned} r^1 &= r^0 - \alpha_0 A d^0 \\ &= \begin{pmatrix} -1 \\ -a \end{pmatrix} - \frac{1+a^2}{1+a^3} \begin{pmatrix} -1 \\ -a^2 \end{pmatrix} = \begin{pmatrix} -1 + \frac{1+a^2}{1+a^3} \\ -a + a^2\frac{1+a^2}{1+a^3} \end{pmatrix} \end{aligned}$$

alors

$$\begin{aligned} r^1 &= \frac{1}{1+a^3} \begin{pmatrix} -a^3 + a^2 \\ -a + a^2 \end{pmatrix} = \frac{a(a-1)}{1+a^3} \begin{pmatrix} -a \\ 1 \end{pmatrix} \\ \beta_1 &= \frac{\left\|r^1\right\|^2}{\left\|r^0\right\|^2} \\ \left\|r^1\right\|^2 &= \frac{a^2(a-1)^2}{\left(1+a^3\right)^2}\left(a^2+1\right), \ \left\|r^0\right\|^2 = 1 + a^2 \end{aligned}$$

donc $\beta_1 = \frac{a^2(a-1)^2}{\left(1+a^3\right)^2}$

$$\begin{aligned} d^1 &= r^1 + \beta_1 d^0 \\ &= \frac{a(a-1)}{1+a^3} \begin{pmatrix} -a \\ 1 \end{pmatrix} + \frac{a^2(a-1)^2}{\left(1+a^3\right)^2} \begin{pmatrix} -1 \\ -a \end{pmatrix} \\ &= \frac{a(a-1)}{1+a^3} \begin{pmatrix} -a - \frac{a(a-1)}{1+a^3} \\ 1 - \frac{a^2(a-1)}{1+a^3} \end{pmatrix} \\ &= \frac{a(a-1)}{\left(1+a^3\right)^2} \begin{pmatrix} -a^4 - a^2 \\ 1 + a^2 \end{pmatrix} = \frac{a(a-1)\left(1+a^2\right)}{\left(1+a^3\right)^2} \begin{pmatrix} -a^2 \\ 1 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \alpha_1 &= \frac{\left\|r^1\right\|^2}{\left(Ad^1, d^1\right)} \\ \left\|r^1\right\|^2 &= \frac{a^2(a-1)^2}{\left(1+a^3\right)^2}\left(a^2+1\right) \end{aligned}$$

$$
\begin{aligned}
Ad^1 &= \frac{a(a-1)\left(1+a^2\right)}{\left(1+a^3\right)^2}\begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix}\begin{pmatrix} -a^2 \\ 1 \end{pmatrix} \\
&= \frac{a(a-1)\left(1+a^2\right)}{\left(1+a^3\right)^2}\begin{pmatrix} -a^2 \\ a \end{pmatrix} \\
&= \frac{a^2(a-1)\left(1+a^2\right)}{\left(1+a^3\right)^2}\begin{pmatrix} -a \\ 1 \end{pmatrix} \\
\left(Ad^1, d^1\right) &= \frac{a^3(a-1)^2\left(1+a^2\right)^2}{\left(1+a^3\right)^4}\left(a^3+1\right) \\
&= \frac{a^3(a-1)^2\left(1+a^2\right)^2}{\left(1+a^3\right)^3}, \quad \text{d'où} \quad \alpha_1 = \frac{1+a^3}{a\left(1+a^2\right)}
\end{aligned}
$$

$$
\begin{aligned}
x^2 &= x^1 + \alpha_1 d^1 \\
&= \frac{a-1}{1+a^3}\begin{pmatrix} a^2 \\ -1 \end{pmatrix} + \frac{1+a^3}{a\left(1+a^2\right)}\left(\frac{a(a-1)\left(1+a^2\right)}{\left(1+a^3\right)^2}\right)\begin{pmatrix} -a^2 \\ 1 \end{pmatrix} \\
&= \frac{a-1}{1+a^3}\begin{pmatrix} a^2 \\ -1 \end{pmatrix} + \frac{a-1}{1+a^3}\begin{pmatrix} -a^2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}
\end{aligned}
$$

i.e. $x^2 = (0, 0)$.

- Donc, le minimum est atteint en deux itérations en partant du point $(1, 1)$.
- En partant du point $(1, 0)$ (par exemple), on atteint le minimum $(0, 0)$ en une itération. Dans tous les cas, on atteint le minimum en au plus 2 itérations.

• 2.5.2 - Convergence de la méthode de la plus forte pente

Rappelons qu'une itération de la méthode de steepest descent pour une forme quadratique $f(x) = \frac{1}{2}(Ax, x) + (b, x) + c$ est

$$
\begin{cases}
d^k = -\left(Ax^k + b\right) = -\nabla f\left(x^k\right) \\
\alpha_k = \frac{\left\|d^k\right\|^2}{\left(Ad^k, d^k\right)} \\
x^{k+1} = x^k + \alpha_k d^k
\end{cases}
$$

On suppose que A est définie positive. Notons par $\lambda_{\min}$ (resp. $\lambda_{\max}$) la plus petite (resp. la plus grande) valeur propre de A. On définit le conditionnement de la matrice A par

$$
\kappa(A) = \frac{\lambda_{\max}}{\lambda_{\min}}
$$

Notons aussi par x^* l'optimum de $f(x)$. Alors, on a le résultat suivant :

Théorème 2.2

$$\left\|x^k - x^*\right\| \leq C \left(\frac{\kappa(A)-1}{\kappa(A)+1}\right)^k, \text{ avec } C = \left(\frac{2\left(f\left(x^0\right) - f\left(x^*\right)\right)}{\lambda_{\min}}\right)^{1/2}.$$

Preuve :

- On a $x^{k+1} = x^k + \alpha_k d^k,\ d^k = -\left(Ax^k + b\right)$

$$\begin{aligned} f\left(x^{k+1}\right) &= \frac{1}{2}\left(Ax^k + \alpha_k Ad^k,\ x^k + \alpha_k d^k\right) + \left(b,\ x^k + \alpha_k d^k\right) + c \\ &= f\left(x^k\right) + \frac{1}{2}\alpha_k^2\left(Ad^k,\ d^k\right) + \alpha_k\left(Ax^k + b,\ d^k\right) \\ &= f\left(x^k\right) - \alpha_k\left(d^k,\ d^k\right) + \frac{\alpha_k^2}{2}\left(Ad^k,\ d^k\right) \end{aligned}$$

or $\alpha_k = \frac{\left\|d^k\right\|^2}{\left(Ad^k, d^k\right)}$, donc $f\left(x^{k+1}\right) = f\left(x^k\right) - (1/2)\frac{\left\|d^k\right\|^4}{\left(Ad^k, d^k\right)}$. D'autre part ;

$$\begin{aligned} f\left(x^k\right) - f\left(x^*\right) =& \frac{1}{2}\left(Ax^k,\ x^k\right) + \left(b,\ x^k\right) - \frac{1}{2}\left(Ax^*,\ x^*\right) - \left(b,\ x^*\right) \\ =& \frac{1}{2}\left(A\left(x^k - x^*\right),\ x^k - x^*\right) + \left(Ax^*,\ x^k\right) \\ & - \left(Ax^*, x^*\right) + \left(b, x^k\right) - \left(b, x^*\right) \\ =& \frac{1}{2}\left(A\left(x^k - x^*\right),\ x^k - x^*\right) \end{aligned}$$

car $Ax^* + b = 0$. Et $d^k = -\left(Ax^k + b\right) = -\left(Ax^k - Ax^*\right) = -A\left(x^k - x^*\right)$, donc $\left(x^k - x^*\right) = -A^{-1}d^k$. D'où

$$f\left(x^k\right) - f\left(x^*\right) = \frac{1}{2}\left(A^{-1}d^k, d^k\right).$$

Alors

$$\begin{aligned} f\left(x^{k+1}\right) &= f\left(x^k\right) - (1/2)\frac{\left\|d^k\right\|^4}{\left(Ad^k, d^k\right)} \\ &= f\left(x^*\right) + \frac{1}{2}\left(A^{-1}d^k, d^k\right) - \frac{1}{2}\frac{\left\|d^k\right\|^4}{\left(Ad^k, d^k\right)} \end{aligned}$$

d'où

$$\frac{f\left(x^{k+1}\right) - f\left(x^*\right)}{f\left(x^k\right) - f\left(x^*\right)} = 1 - \frac{\left\|d^k\right\|^4}{\left(A^{-1}d^k, d^k\right)\left(Ad^k, d^k\right)}$$

- Maintenant, on utilise l'inégalité de Kantorovich qui dit :

$$(Ax,\ x)\left(A^{-1}x,\ x\right) \leqslant \frac{\left(\lambda_{\max} + \lambda_{\min}\right)^2}{4\lambda_{\max}\lambda_{\min}}\|x\|^4,\ \forall\ x \in \mathbb{R}^n$$

d'où

$$\frac{f\left(x^{k+1}\right)-f\left(x^{*}\right)}{f\left(x^{k}\right)-f\left(x^{*}\right)} \leqslant 1-\frac{4\lambda_{\max}\lambda_{\min}}{\left(\lambda_{\max}+\lambda_{\min}\right)^{2}}$$

$$\frac{f\left(x^{k}+1\right)-f\left(x^{*}\right)}{f\left(x^{k}\right)-f\left(x^{*}\right)} \leqslant \left(\frac{\lambda_{\max}-\lambda_{\min}}{\lambda_{\max}+\lambda_{\min}}\right)^{2}$$

- D'autre part, on a

$$\begin{aligned} f\left(x^{k}\right)-f\left(x^{*}\right) &= \frac{1}{2}\left(A\left(x^{k}-x^{*}\right), x^{k}-x^{*}\right) \\ &\geqslant \frac{\lambda_{\min}}{2}\left\|x^{k}-x^{*}\right\|^{2} \end{aligned}$$

alors

$$\left\|x^{k}-x^{*}\right\|^{2} \leqslant \frac{2}{\lambda_{\min}}\left(f\left(x^{k}\right)-f\left(x^{*}\right)\right)$$

Comme

$$f\left(x^{k}\right)-f\left(x^{*}\right) \leq \left(\frac{\lambda_{\max}-\lambda_{\min}}{\lambda_{\max}+\lambda_{\min}}\right)^{2}\left(f\left(x^{k-1}\right)-f\left(x^{*}\right)\right),$$

on en déduit que

$$f\left(x^{k}\right)-f\left(x^{*}\right) \leqslant \left(\frac{\lambda_{\max}-\lambda_{\min}}{\lambda_{\max}+\lambda_{\min}}\right)^{2k}\left(f\left(x^{0}\right)-f\left(x^{*}\right)\right).$$

On en conclut que

$$\left\|x^{k}-x^{*}\right\|^{2} \leq \frac{2}{\lambda_{\min}}\left(\frac{\lambda_{\max}-\lambda_{\min}}{\lambda_{\max}+\lambda_{\min}}\right)^{2k}\left(f\left(x^{0}\right)-f\left(x^{*}\right)\right)$$

i.e.

$$\left\|x^{k}-x^{*}\right\| \leqslant \sqrt{\frac{2\left(f\left(x^{0}\right)-f\left(x^{*}\right)\right)}{\lambda_{\min}}}\left(\frac{\kappa(A)-1}{\kappa(A)+1}\right)^{k}.$$

Ce qui achève la démonstration. □

La méthode du gradient à paramètre local optimal est convergente. Sa rapidité de convergence dépend de $\frac{\kappa(A)-1}{\kappa(A)+1}$.

- Plus que $\kappa(A)$ est proche de 1, plus la méthode converge vite.
- Lorsque $\kappa(A)=1$, la méthode converge en une itération.
- Lorsque $\kappa(A)$ est trop grand, la convergence est lente.

Exemple 2.4 : Reprenons l'exemple (déjà traité) où on avait la fonction suivante

$$f(x,y) = \frac{1}{2}\left(x^2 + ay^2\right), \quad \text{avec } a \geqslant 1.$$

- Le gradient et le Hessien sont donnés par

$$\nabla f(x,y) = \begin{pmatrix} x \\ ay \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix}$$

- Ainsi, on peut caculer le conditionnement et obtenir

$$\kappa(A) = a, \quad \text{si } a > 1$$
$$\kappa(A) = 1, \quad \text{si } a = 1$$

- Si a est très grand, la convergence est lente.
- Si $a = 1$, l'optimum est atteint en une itération. Ceci confirme les résultats déjà obtenus géométriquement.

2.5.3 - Convergence de la méthode de Richardson (pas constant)

La méthode de Richardson est une méthode du gradient à paramètre α_k constant. On prend toujours une direction de descente (celle du gradient) et on choisit α indépendant de k de façon que la suite des points $\left\{x^k\right\}$ converge vers la solution de $Ax^* + b = 0$:

$$\begin{aligned} x^{k+1} &= x^k + \alpha d^k, \ \alpha > 0 \\ d^k &= -\left(Ax^k + b\right) = A\left(x^* - x^k\right) \end{aligned}$$

L'erreur à la $(k+1)^{\text{ième}}$ itération $e^{k+1} = x^{k+1} - x^*$ peut s'exprimer en fonction de l'erreur à la $k^{\text{ième}}$ itération e^k :

$$e^{k+1} = (I - \alpha A)e^k.$$

En effet ;

$$\begin{aligned} e^{k+1} &= x^{k+1} - x^* \\ &= x^k + \alpha d^k - x^* \\ &= x^k - x^* + \alpha A\left(x^* - x^k\right) \\ &= (I - \alpha A)\left(x^k - x^*\right) \\ &= (I - \alpha A)e^k. \end{aligned}$$

D'où l'on obtient

$$e^k = (I - \alpha A)^k e^0, \ \text{ où } \ e^0 = x^0 - x^*.$$

Donc, en prenant la norme de cette équation on obtient $\|e^k\| \leq \|I - \alpha A\|^k \|e^0\|$. Or

$$\begin{aligned} e^k \longrightarrow 0 &\iff \|I - \alpha A\| < 1 \\ &\iff \max_{i=1,\dots,n} |1 - \alpha\lambda_i| < 1, \end{aligned}$$

où les λ_i $(i = 1, \dots, n)$ sont les valeurs propres de la matrice A. On a alors

$$\begin{aligned} e^k \longrightarrow 0 &\iff 0 < \alpha < \frac{2}{\lambda_i},\ \forall\, i = 1, \dots, n, \\ &\iff 0 < \alpha < \frac{2}{\lambda_{\max}}. \end{aligned}$$

En effet,

$$\max_{i=1,\dots,n} |1 - \alpha\lambda_i| < 1 \iff |1 - \alpha\lambda_i| < 1,\ \forall\, i = 1, \dots, n,$$

Il y a deux cas à distinguer :

- ⋆ Soit que $1 - \alpha\lambda_i \geqslant 0$ et donc $|1 - \alpha\lambda_i| < 1 \ \Rightarrow\ 1 - \alpha\lambda_i < 1 \Rightarrow \alpha > 0$.
- ⋆ Soit que $1 - \alpha\lambda_i \leq 0$ et donc $|1 - \alpha\lambda_i| < 1 \ \Rightarrow\ \alpha\lambda_i - 1 < 1 \Rightarrow \alpha < \frac{2}{\lambda_i}$.

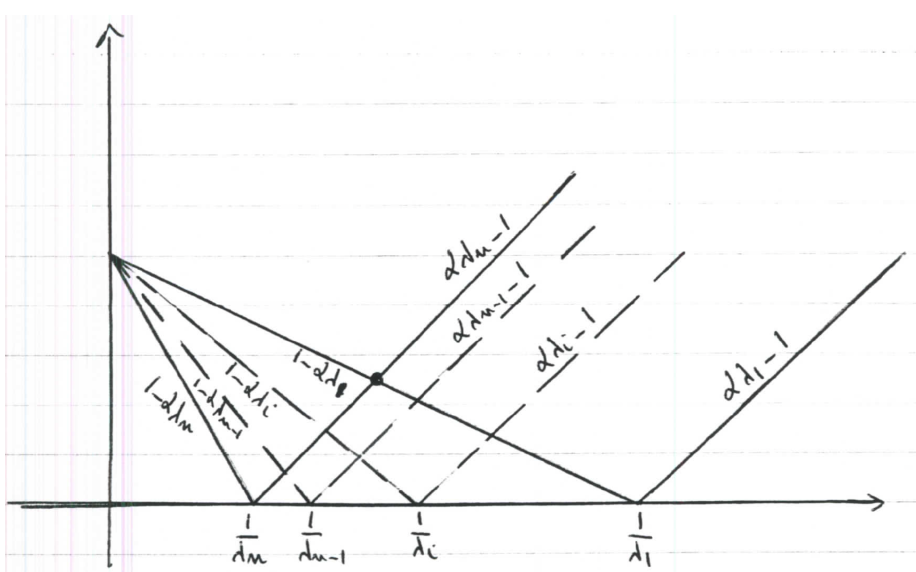
Le meilleur choix de α est celui qui minimise $\|I - \alpha A\|$. Or

$$\|I - \alpha A\| = \max_{i=1,\dots,n} |1 - \alpha\lambda_i|\,.$$

En classant les valeurs propres comme suit

$$0 < \lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_i \leqslant \cdots \leqslant \lambda_n,$$

où on note $\lambda_1 = \lambda_{\min}$ et $\lambda_n = \lambda_{\max}$, on obtient le graphique suivant



alors α_{opt} est celui qui vérifie $1-\alpha\lambda_1 = \alpha\lambda_n - 1$ i.e., $\alpha = \frac{2}{\lambda_1+\lambda_n}$, c'est-à-dire

$$\alpha_{\text{opt}} = \frac{2}{\lambda_{\text{min}} + \lambda_{\text{max}}}$$

On obtient alors

$$\begin{aligned}
\|I-\alpha_{\text{opt}}A\| &= \max\left\{|1-\alpha_{\text{opt}}\lambda_{\text{min}}|, |1-\alpha_{\text{opt}}\ \lambda_{\text{max}}|\right\} \\
&= \max\left\{\left|\frac{\lambda_{\text{max}}-\lambda_{\text{min}}}{\lambda_{\text{max}}+\lambda_{\text{min}}}\right|, \left|\frac{\lambda_{\text{min}}-\lambda_{\text{max}}}{\lambda_{\text{max}}+\lambda_{\text{min}}}\right|\right\} \\
&= \frac{\lambda_{\text{max}}-\lambda_{\text{min}}}{\lambda_{\text{max}}+\lambda_{\text{min}}} = \frac{\kappa(A)-1}{\kappa(A)+1}
\end{aligned}$$

et dans ce cas, on a

$$\begin{aligned}
\left\|e^k\right\| &\leq \|I-\alpha_{\text{opt}}A\|^k \left\|e^0\right\| \\
\text{i.e.} \qquad \left\|x^k - x^*\right\| &\leq \left(\frac{\kappa(A)-1}{\kappa(A)+1}\right)^k \left\|x^0 - x^*\right\|.
\end{aligned}$$

Le facteur de réduction de l'erreur (taux de convergence) est de l'ordre de $\frac{\kappa(A)-1}{\kappa(A)+1}$.

Exemple 2.5 :

$$\min_{\mathbb{R}^2} f(x,y), \quad \text{où} \quad f(x,y) = x^2 + y^2 + xy + 4x + 7y$$

- À partir de la matrice A, on calcule les valeurs propres

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

$$\begin{aligned}
\begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} &= (2-\lambda)^2 - 1 = (2-\lambda-1)(2-\lambda+1) \\
&= (1-\lambda)(3-\lambda)
\end{aligned}$$

- Donc, les valeurs propres de A sont $\lambda = 1$ et $\lambda = 3$, d'où $\kappa(A) = 3$ et le taux de convergence est de l'ordre

$$\frac{\kappa(A)-1}{\kappa(A)+1} = \frac{2}{4} = 1/2$$

Note : Pour la méthode du gradient conjugué (dans le cas quadratique), on a la majoration d'erreur suivante :

$$\left\|x^k - x^*\right\|_A \leqslant 2\left(\frac{\sqrt{\kappa(A)}-1}{\sqrt{\kappa(A)}+1}\right)^k \left\|x^0 - x^*\right\|_A, \text{ où } \|x\|_A^2 = (Ax, x).$$

2.6 Cas des fonctions quelconques

On suppose $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ de classe C^1. Nous allons décrire certaines méthodes numériques pour la résolution du problème

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

- **2.6.1 - La méthode du gradient à pas constant:méthode de Richardson**

Cette méthode a été déjà étudiée dans le cadre des fonctions quadratiques et consiste à utiliser la méthode du gradient avec le paramètre α_k constant pour toutes les itérations, i.e.,

$$\begin{cases} d^k & = -\nabla f\left(x^k\right) \\ x^{k+1} & = x^k + \alpha d^k, \ \alpha > 0. \end{cases}$$

- **2.6.2 - Les méthodes de Fletcher-Reeves et de Polak-Ribière.**

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

Algorithme :

i x^0 donné, $d^0 = -\nabla f\left(x^0\right)$.

ii Etape k :

- choisir α_k minimisant $g(\alpha) = f\left(x^k + \alpha d^k\right)$
- poser $x^{k+1} = x^k + \alpha_k d^k$
- et $d^{k+1} = -\nabla f\left(x^{k+1}\right) + \beta_{k+1} d^k$, avec $\beta_{k+1} = \frac{\left\|\nabla f\left(x^{k+1}\right)\right\|^2}{\left\|\nabla f\left(x^k\right)\right\|^2}$

iii Test d'arrêt. Si vérifié : FIN, sinon faire $k \longleftarrow k+1$ et retouner en ii.

La méthode de Polak-Ribière est une variante de la méthode de Fletcher-Reeves. Elle consiste à définir β_{k+1} par la formule :

$$\beta_{k+1} = \frac{\left(\nabla f\left(x^{k+1}\right) - \nabla f\left(x^k\right), \nabla f\left(x^{k+1}\right)\right)}{\left\|\nabla f\left(x^k\right)\right\|^2}.$$

2.7 La méthode de Newton

On suppose $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ de classe C^2. L'idée consiste à remplacer, au voisinage du point courant x^k, la fonction f par son approximation quadratique :

$$J(x) = f\left(x^k\right) + \nabla f\left(x^k\right)^\top \cdot \left(x - x^k\right) + \frac{1}{2}\left(x - x^k\right)^\top \cdot H\left(x^k\right)\left(x - x^k\right)$$

où $H\left(x^k\right) = \nabla^2 f\left(x^k\right)$. On prend alors comme point x^{k+1} le minimum de $J(x)$ lorsqu'il existe. Ceci ne peut être le cas que si $H\left(x^k\right)$ est une matrice définie positive.

La fonction $J(x)$ a alors un minimum unique x^{k+1} définie par :

$$\nabla J\left(x^{k+1}\right) = 0, \quad \text{c'est-à-dire} \;\; H\left(x^k\right) \cdot \left(x^{k+1} - x^k\right) + \nabla f\left(x^k\right) = 0.$$

D'où la formule itérative :

$$x^{k+1} = x^k - H\left(x^k\right)^{-1} \nabla f\left(x^k\right).$$

Cette formule n'est autre que la méthode de Newton appliquée à la résolution du système d'équations non linéaires :

$$\frac{\partial f}{\partial x_i}(x) = 0, \; i = 1, \ldots, n.$$

Note :

- On a $x^{k+1} = x^k + \alpha_k d^k$ avec $\alpha_k = 1$ et $d^k = -H\left(x^k\right)^{-1} \nabla f\left(x^k\right)$. Donc, dans ce cas la direction et le pas de déplacement sont fixés.
- Une propriété intéressante de la méthode de Newton est qu'elle converge en une seule itération, lorsqu'elle est appliquée à une fonction quadratiqne strictement convexe.

En effet, soit

$$f(x) = \frac{1}{2}(Ax, x) + (b, x) + c$$

avec A symétrique définie positive. On a :

$$\nabla f(x) = Ax + b \;\; \text{et} \;\; H(x) = A.$$

Donc, étant donné x^0 (quelconque) on a

$$\begin{aligned} x^1 &= x^0 - H\left(x^0\right)^{-1} \nabla f\left(x^0\right) \\ &= x^0 - A^{-1}\left(Ax^0 + b\right) = -A^{-1}b \end{aligned}$$

i.e., $\quad x^1 = -A^{-1}b$ donc x^1 est l'optimum.

- Au voisinage de x^*, la convergence de la méthode de Newton est quadratique. Cependant, elle ne possède pas la propriété de la convergence globale : si le point de départ x^0 est trop éloigné de x^* la méthode de Newton peut ne pas converger.

Exemple 2.6 : Considérons le problème suivant

$$\min_{\mathbb{R}^3} f(x,y,z), \text{ où la fonction est définie par}$$
$$f(x,y,z) = (2x-1)^2 + (y-2)^2 + (3z-3)^2 + xz - 4$$

- Nous avons le gradient et le Hessien donnés par

$$\nabla f(x,y,z) = \begin{pmatrix} 4(2x-1)+z \\ 2(y-2) \\ 6(3z-3)+x \end{pmatrix}, \quad H(x,y,z) = \begin{pmatrix} 8 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 18 \end{pmatrix}$$

Ainsi, on a $\nabla f(x,y,z) = 0 \iff x = \frac{54}{143},\ y = 2,\ z = \frac{140}{143}$.

- Il est facile de vérifier que $H(x,y,z)$ est définie positive. Alors, le minimum est atteint au point $(x^*, y^*, z^*) = \left(\frac{54}{143}, 2, \frac{140}{143}\right)$.
- Appliquons la méthode de Newton à ce problème en partant du point $(x_0, y_0, z_0) = (1,1,1)$. La matrice H^{-1} est

$$H^{-1} = \begin{pmatrix} \frac{18}{143} & 0 & -\frac{1}{143} \\ 0 & \frac{1}{2} & 0 \\ \frac{-1}{143} & 0 & \frac{8}{143} \end{pmatrix}$$

Alors $x^{(1)} = x^{(0)} - H\left(x^0\right)^{-1} \nabla f\left(x^0\right)$ avec

$$\nabla f\left(x^0\right) = \begin{pmatrix} 5 \\ -2 \\ 1 \end{pmatrix}$$
$$H\left(x^0\right)^{-1} \nabla f\left(x^0\right) = \frac{1}{143} \begin{pmatrix} 18 & 0 & -1 \\ 0 & \frac{143}{2} & 0 \\ -1 & 0 & 8 \end{pmatrix} \begin{pmatrix} 5 \\ -2 \\ 1 \end{pmatrix}$$
$$= \frac{1}{143} \begin{pmatrix} 89 \\ -143 \\ 3 \end{pmatrix}$$

$$\text{d'où} \quad x^{(1)} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{143} \begin{pmatrix} 89 \\ -143 \\ 3 \end{pmatrix} = \frac{1}{143} \begin{pmatrix} 54 \\ 2 \times 143 \\ 140 \end{pmatrix}$$

$$\text{i.e.} \quad x^{(1)} := \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = \begin{pmatrix} \frac{54}{143} \\ 2 \\ \frac{140}{143} \end{pmatrix} = \begin{pmatrix} x^* \\ y^* \\ z^* \end{pmatrix}$$

Afin de rendre la méthode de Newton "globalement" convergente, certaines modifications ont été apportées à cette méthode. Nous allons en citer deux.

- ⋆ La première modification est d'introduire le pas de déplacement α_k dépendant de k, i.e.,
$$x^{k+1} = x^k - \alpha_k H\left(x^k\right)^{-1} \nabla f\left(x^k\right),$$
où α_k est choisi de telle sorte qu'il minimise la fonction d'une variable
$$g(\alpha) = f\left(x^k - \alpha H\left(x^k\right)^{-1} \nabla f\left(x^k\right)\right),$$
sur l'ensemble des $\alpha \geqslant 0$.
- ⋆ La deuxième modification est de changer le Hessien $H\left(x^k\right)$ par $H\left(x^k\right) + \alpha_k I$, i.e.,
$$x^{k+1} = x^k - \left[H\left(x^k\right) + \alpha_k I\right]^{-1} \nabla f\left(x^k\right)$$
où $\alpha_k > 0$. (Le scalaire α_k peut être choisi minimal avec la contrainte que toutes les valeurs propres de $H\left(x^k\right) + \alpha_k I$ soient supérieures ou égales à une constante $\delta > 0$ donnée). Cette méthode est connue sous le nom de la méthode de Levenberg-Marquardt.
Lorsque $\alpha_k = 0$, on retrouve la méthode de Newton. L'intérêt de cette méthode est que dans le cas où $H\left(x^k\right)$ n'est pas définie positive on peut trouver un $\alpha_k > 0$ tel que $H\left(x^k\right) + \alpha_k I$ soit définie positive.
- ⋆ Une autre variante de la méthode de Newton est proposée dans le but de diminuer la quantité de calcul (notamment le calcul du Hessien $H\left(x^k\right)$ et son inverse à chaque itération). Cette méthode consiste à partir d'un point x^0 donné, de calculer le Hessien. $H = H\left(x^0\right)$ et d'utiliser le même Hessien H dans la méthode de Newton pendant p itérations, puis changer H par $H\left(x^p\right)$ et répéter la procédure.

Algorithme : L'algorithme peut s'écrire pour $m = 0, 1, \ldots,$ par

i $H = H\left(x^{mp}\right)$.

ii $x^{k+1} = x^k - H^{-1} \nabla f\left(x^k\right)$, pour $mp \leq k \leqslant (m+1)p - 1$.

p varie entre 2 et 4 itérations. Si $p = 1$, on retrouve la méthode de Newton.

Bibliographie :

[1] **V. Karmanov. Programmation mathématique. Editions Mir, Moscou, 1977.**

[2] **M. Minoux. Programmation mathématique. Théorie et algorithmes. Lavoisier, Paris, 2008.**

[3] **B. T. Polyak. Introduction to Optimization. Optimization Software, New York, 1987.**