

4 Gradient à pas optimal

Leçons 162, 219, 226, 229, 233, 253(, 181, 215)

Ref : [Hiriart-Urruty] Exos I.9 & II.8

Ce développement démontre la convergence d'une méthode d'approximation du minimum d'une fonctionnelle, qui peut également servir à approcher la solution d'un système linéaire. On doit mentionner dans le plan les résultats préliminaires, et bien sûr présenter l'algorithme de gradient à pas optimal.

On fixe une matrice $A \in \mathcal{M}_n(\mathbb{R})$ symétrique définie positive, $b \in \mathbb{R}^n$ et $c \in \mathbb{R}$. On pose alors, pour $x \in \mathbb{R}^n$

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

On cherche alors à minimiser la fonctionnelle f sur \mathbb{R}^n . Comme f est strictement convexe et 1-coercive, il existe une unique solution $\bar{x} \in \mathbb{R}^n$ à ce problème, caractérisé par le fait d'être l'unique solution du système linéaire

$$Ax = b,$$

ce qui se démontre en étudiant les zéros du gradient de f .

Voici l'algorithme dont nous allons étudier la convergence.

Algorithme 1 (Gradient à pas optimal) On fixe $x^{(0)} \in \mathbb{R}^n$. Pour construire récursivement la suite $(x^{(k)})_{k \in \mathbb{N}}$, on suppose les premiers éléments construits, et on pose

$$\begin{cases} d_k = -\nabla f(x^{(k)}) = b - Ax^{(k)} \\ x^{(k+1)} = x^{(k)} + t_k d_k \end{cases}$$

où t_k est l'unique réel (positif) minimisant la fonction $t \mapsto f(x^{(k)} + td_k)$. On s'arrête quand on a atteint un seuil de tolérance $\varepsilon > 0$ fixé au préalable, c'est-à-dire quand $\|\nabla f(x^{(k)})\| < \varepsilon$.

Lemme 2 (Inégalité de Kantorovitch) Soit $A \in S_n^{++}(\mathbb{R})$, $\lambda_1 \geq \dots \geq \lambda_n$ ses valeurs propres. On a alors pour $x \in \mathbb{R}^n$

$$\|x\|^4 \leq \langle Ax, x \rangle \langle A^{-1}x, x \rangle \leq \frac{1}{4} \left(\sqrt{\frac{\lambda_1}{\lambda_n}} + \sqrt{\frac{\lambda_n}{\lambda_1}} \right)^2 \|x\|^4.$$

Démonstration. Il suffit bien sûr de montrer ce résultat pour les vecteurs de norme 1. Soit donc $x \in \mathbb{S}^{n-1}$. On se donne $P \in O_n(\mathbb{R})$ tel que $D = PA^tP$ soit diagonale, avec les λ_i sur la diagonale. Via le changement de variable $y = Px$, il faut alors montrer ¹

$$1 \leq \langle Dy, y \rangle \langle D^{-1}y, y \rangle \leq \frac{1}{4} \left(\sqrt{\frac{\lambda_1}{\lambda_n}} + \sqrt{\frac{\lambda_n}{\lambda_1}} \right)^2.$$

On suppose que A n'est pas une homothétie (sinon il n'y a rien à démontrer), et on a donc $\lambda_1 > \lambda_n$. On pose $\alpha_i = y_i^2$, de sorte que

$$\begin{cases} \langle Dy, y \rangle = \sum_{i=1}^n \alpha_i \lambda_i =: \bar{\lambda} \\ \langle D^{-1}y, y \rangle = \sum_{i=1}^n \frac{\alpha_i}{\lambda_i} \end{cases}$$

ce qui signifie que ces deux quantités sont des combinaisons convexes respectives des λ_i et des $\frac{1}{\lambda_i}$ (puisque

$\sum_{i=1}^n \alpha_i = \|y\|^2 = 1$). On considère alors le graphe de la fonction inverse sur \mathbb{R}_+^* , et on note $M_i = \left(\lambda_i, \frac{1}{\lambda_i} \right)$

les différents points du graphe correspondant aux valeurs propres de A . Le barycentre M de ces points, associés aux coefficients α_i , est alors dans le domaine \mathcal{D} délimité par le graphe et la droite $(M_1 M_n)$ (voir figure 4.1). On considère alors les points M^* et M_* , qui ont la même abscisse $\bar{\lambda}$ que M et qui sont respectivement sur la droite $(M_1 M_n)$ et sur le graphe de la fonction inverse.

1. Comme P est orthogonale, y est toujours de norme 1.

- M est d'ordonnée $\langle D^{-1}y, y \rangle$;
- M_* est d'ordonnée $\frac{1}{\bar{\lambda}}$;
- M^* est d'ordonnée $\frac{1}{\lambda_1} + \frac{1}{\lambda_n} + \frac{\bar{\lambda}}{\lambda_1 \lambda_n}$.

La convexité de la fonction inverse montre que l'ordonnée de M est plus petite que celle de M^* et plus grande que celle de M_* , soit

$$\begin{cases} \langle D^{-1}y, y \rangle \geq \frac{1}{\bar{\lambda}} & \text{i. e.} & \langle Dy, y \rangle \langle D^{-1}y, y \rangle \geq 1 \\ \langle D^{-1}y, y \rangle \leq \frac{\lambda_1 + \lambda_n - \bar{\lambda}}{\lambda_1 \lambda_n} & \text{i. e.} & \langle Dy, y \rangle \langle D^{-1}y, y \rangle \leq \frac{\bar{\lambda}(\lambda_1 + \lambda_n - \bar{\lambda})}{\lambda_1 \lambda_n} \end{cases}$$

On fait une étude de fonction pour déterminer le maximum de $u \mapsto \frac{u(\lambda_1 + \lambda_n - u)}{\lambda_1 \lambda_n}$. Celui-ci vaut

$$\frac{1}{\lambda_1 \lambda_n} \left(\frac{\lambda_1 + \lambda_n}{2} \right)^2 = \frac{1}{4} \left(\frac{\lambda_1}{\lambda_n} + \frac{\lambda_n}{\lambda_1} + 2 \right) = \frac{1}{4} \left(\sqrt{\frac{\lambda_1}{\lambda_n}} + \sqrt{\frac{\lambda_n}{\lambda_1}} \right)^2,$$

ce qui permet de conclure. □

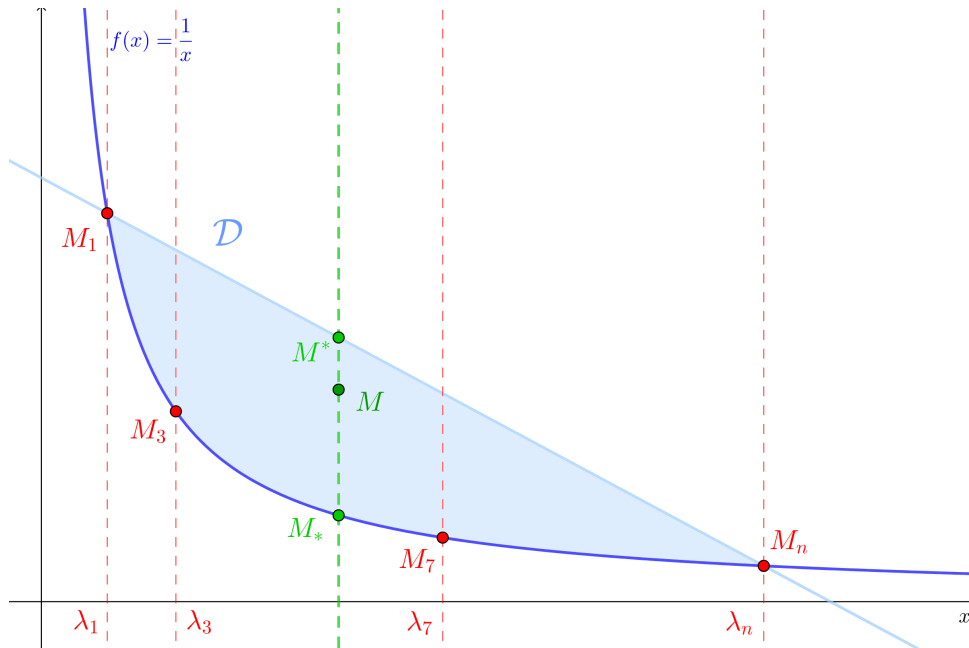


FIGURE 4.1 – Convexité dans le lemme de Kantorovitch

Théorème 3 La méthode du gradient à pas optimal converge de manière géométrique, la vitesse étant reliée au conditionnement de la matrice A , donnée par $\frac{c(A) - 1}{c(A) + 1}$.

Démonstration.

Étape 1. Expression de t_k et de $f(x^{(k+1)})$.

On étudie la fonction f_k définie sur \mathbb{R} par

$$f_k(t) := f(x^{(k)} + td_k) = f(x_k) + t \langle Ax^{(k)} - b, d_k \rangle + \frac{t^2}{2} \langle Ad_k, d_k \rangle.$$

On suppose $d_k = b - Ax^{(k)}$ non nul (sinon l'algorithme s'arrête). Comme A est définie positive et d_k non nul, f_k est un polynôme de degré 2 avec coefficient dominant strictement positif, et est donc minoré sur \mathbb{R} par sa valeur en $t_k = \frac{\|d_k\|^2}{\langle Ad_k, d_k \rangle} > 0$. On en déduit la valeur de $f(x^{(k+1)})$:

$$f(x^{(k+1)}) = f(x^{(k)}) - \frac{1}{2} \frac{\|d_k\|^4}{\langle Ad_k, d_k \rangle}. \quad (1)$$

On cherche à évaluer l'écart de $f(x^{(k)})$ au minimum \bar{f} de f , dont on sait qu'il vaut

$$\bar{f} = f(\bar{x}) = -\frac{1}{2} \langle A^{-1}b, b \rangle + c.$$

On a d'après (1)

$$f(x^{(k+1)}) - \bar{f} = \left(f(x^{(k)}) - \bar{f} \right) \left(1 - \frac{\|d_k\|^4}{2(f(x^{(k)}) - \bar{f}) \langle Ad_k, d_k \rangle} \right).$$

On observe alors que

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

$$\langle A^{-1}d_k, d_k \rangle = 2(f(x^{(k)}) - \bar{f})$$

On a donc finalement

$$f(x^{(k+1)}) - \bar{f} = \left(f(x^{(k)}) - \bar{f} \right) \left(1 - \frac{\|d_k\|^4}{\langle Ad_k, d_k \rangle \langle A^{-1}d_k, d_k \rangle} \right).$$

Étape 2. Conditionnement et vitesse de convergence.

On introduit finalement le *conditionnement* de la matrice A :

$$c(A) := \rho(A)\rho(A^{-1}) = \frac{\lambda_1}{\lambda_n}.$$

Cette valeur est celle qui dirige la convergence de la méthode du gradient. On se donne $k \in \mathbb{N}$, et on suppose que l'algorithme n'est pas encore terminé à l'étape k . On déduit de l'inégalité de Kantorovitch, toujours avec $d_k \neq 0$:

$$\frac{\|d_k\|^4}{\langle Ad_k, d_k \rangle \langle A^{-1}d_k, d_k \rangle} \geq 4 \left(\sqrt{\frac{\lambda_1}{\lambda_n}} + \sqrt{\frac{\lambda_n}{\lambda_1}} \right)^{-2} = 4 \frac{\lambda_1/\lambda_n}{(\lambda_1/\lambda_n + 1)^2} = 4 \frac{c(A)}{(c(A) + 1)^2}.$$

On déduit alors de la relation obtenue à la première étape que l'on a

$$f(x^{(k+1)}) - \bar{f} \leq \left(f(x^{(k)}) - \bar{f} \right) \left(1 - 4 \frac{c(A)}{(c(A) + 1)^2} \right) = \left(f(x^{(k)}) - \bar{f} \right) \left(\frac{c(A) - 1}{c(A) + 1} \right)^2.$$

On obtient finalement par récurrence

$$f(x^{(k)}) - \bar{f} \leq \left(f(x^{(0)}) - \bar{f} \right) \left(\frac{c(A) - 1}{c(A) + 1} \right)^{2k}. \quad (2)$$

On peut également exprimer la vitesse de convergence de la suite $(x^{(k)})$: on a

$$\begin{aligned} f(x^{(k)}) - \bar{f} &= \frac{1}{2} \langle Ax^{(k)}, x^{(k)} \rangle + \langle b, x^{(k)} \rangle + c - \bar{f} \\ &= \frac{1}{2} \langle A(x^{(k)} - \bar{x}), x^{(k)} - \bar{x} \rangle \\ f(x^{(k)}) - \bar{f} &\geq \frac{1}{2} \lambda_n \|x^{(k)} - \bar{x}\|^2 \end{aligned}$$

On déduit donc de (2)

$$\|x^{(k)} - \bar{x}\| \leq \left(\frac{2(f(x^{(0)}) - \bar{f})}{\lambda_n} \right)^{\frac{1}{2}} \left(\frac{c(A) - 1}{c(A) + 1} \right)^k \quad (3)$$

□

Remarque 4 On observe les phénomènes suivants, qui sont montrés sur la figure 4.2.

- On peut montrer que les directions successives dans lesquelles on se déplace sont orthogonales : $\langle d_k, d_{k+1} \rangle = 0$.
- On observe ainsi que plus le conditionnement est proche de 1, plus la convergence est rapide. Le cas optimal est donné par les homothéties, dont le conditionnement est 1, et pour lesquelles, en vertu de (3), le minimum est obtenu dès la première étape. On est dans le cas où les valeurs propres sont identiques. Dans le cas $n = 2$, les courbes de niveau de f sont des cercles (figure 4.2a).
- Le cadre le moins favorable est celui où les valeurs propres extrêmes sont éloignées. Les courbes de niveau sont alors des ellipses presque plates, et la convergence est beaucoup plus lente (figure 4.2b).

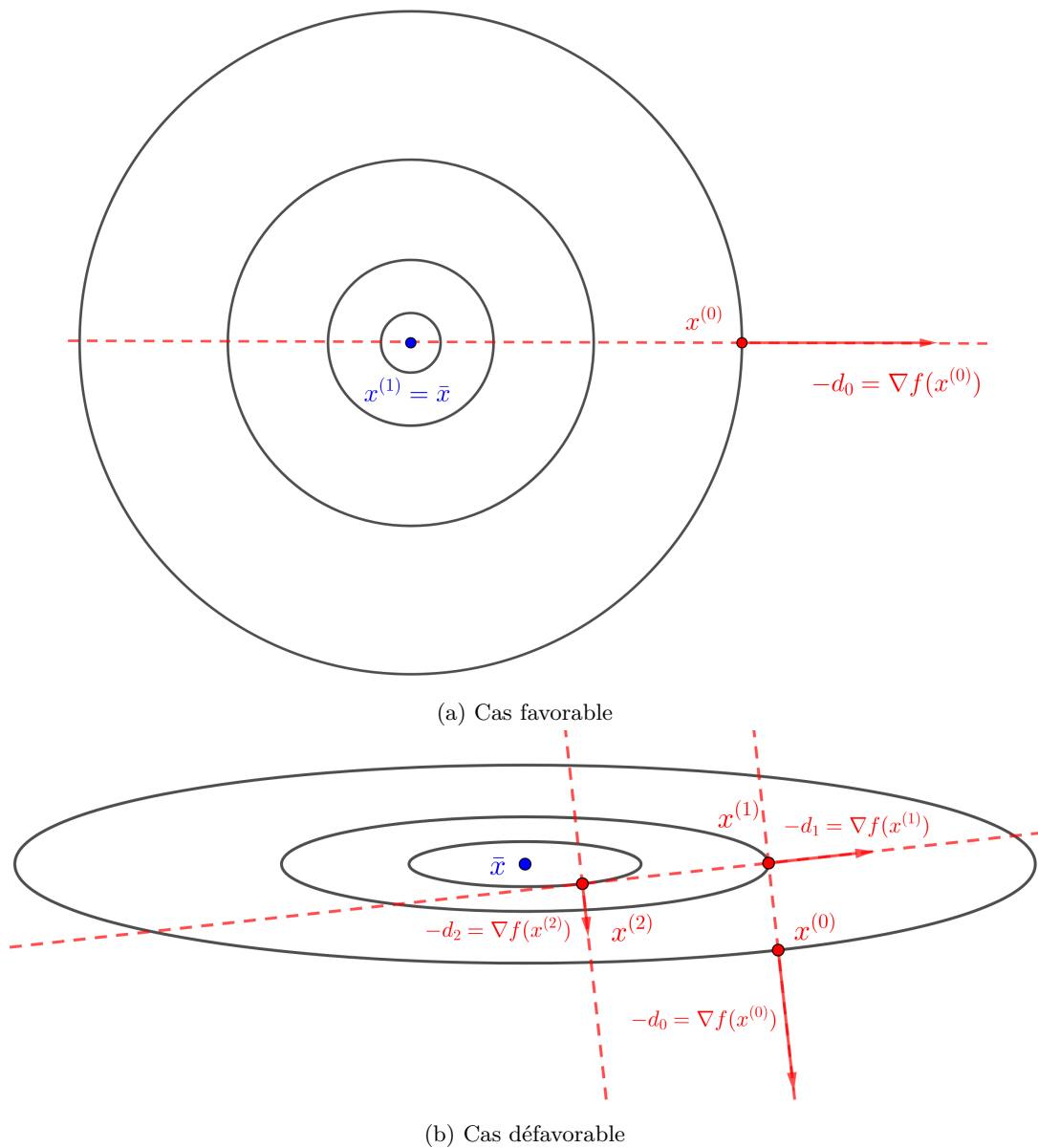


FIGURE 4.2 – Premières itérations de l'algorithme du gradient à pas optimal