P. Maurer

ENS Rennes

Recasages : 162, 215, 219, 226, 229, 233.

Référence : Hiriat-Urruty, Optimisation et analyse convexe

Très fortement inspiré du travail de Florent Lemonnier.

Algorithme de la descente de gradient à pas optimal

On se place sur l'espace euclidien $\mathbb{R}^n$, muni de son produit scalaire $\langle .,. \rangle$ classique, et la norme $\|.\|$ associée.

Lemme 1. *(Inégalité de Kantorovitch)*

Soit $A \in \mathcal{S}_n^{++}(\mathbb{R})$. On note λ_1 et λ_n la plus petite (respectivement la plus grande) valeur propre de A. Alors :

$$\begin{aligned}\forall x \in \mathbb{R}^n \quad \langle Ax, x \rangle \langle A^{-1}x, x \rangle &\leq \frac{1}{4}\left(\sqrt{\frac{\lambda_1}{\lambda_n}} + \sqrt{\frac{\lambda_n}{\lambda_1}}\right)^2 \|x\|^4 \\ &= \frac{1}{4}\left(\sqrt{c(A)} + \sqrt{\frac{1}{c(A)}}\right)^2 \|x\|^4,\end{aligned}$$

où $c(A) \stackrel{\text{def}}{=} \dfrac{\lambda_1}{\lambda_n}$ désigne le contenu de A.

Démonstration. A étant symétrique définie positive, elle est diagonalisable dans une base orthonormée et ses valeurs propres sont toutes strictement positives. Par ailleurs, on montre que les valeurs propres de A^{-1} sont les inverses des valeurs propres de A.

En notant $(x_1, \ldots, x_n)$ les coordonnées de x dans une telle base, l'inégalité de Cauchy-Schwarz donne alors

$$\begin{aligned}\langle Ax, x \rangle \langle A^{-1}x, x \rangle &= \left(\sum_{i=1}^n \lambda_i x_i^2\right)\left(\sum_{i=1}^n \frac{1}{\lambda_i} x_i^2\right) \\ &\geq \left(\sum_{i=1}^n \frac{\sqrt{\lambda_i}}{\sqrt{\lambda_i}} x_i \cdot x_i\right)^2 \\ &= \left(\sum_{i=1}^n x_i^2\right)^2 \\ &\geq 0\end{aligned}$$

On peut donc prendre la racine de cette quantité, et en utilisant l'inégalité $\sqrt{ab} \leq \frac{1}{2}(a+b)$ pour tout $a, b \in \mathbb{R}_+$ (qui vient du développement de $(a+b)^2$), on obtient

$$\begin{aligned}\sqrt{\langle Ax, x \rangle \langle A^{-1}x, x \rangle} &= \sqrt{\frac{\lambda_1}{\lambda_n}} \sqrt{\left(\sum_{i=1}^n \frac{\lambda_i}{\lambda_1} x_i^2\right)\left(\sum_{i=1}^n \frac{\lambda_n}{\lambda_i} x_i^2\right)} \\ &\leq \frac{1}{2}\sqrt{\frac{\lambda_1}{\lambda_n}} \sum_{i=1}^n \left(\frac{\lambda_i}{\lambda_1} + \frac{\lambda_n}{\lambda_i}\right) x_i^2.\end{aligned}$$

La fonction $t \mapsto \dfrac{t}{\lambda_1} + \dfrac{\lambda_n}{t}$ est décroissante sur $[\lambda_1, \sqrt{\lambda_1 \lambda_n}]$ et croissante sur $[\sqrt{\lambda_1 \lambda_n}, \lambda_n]$. Elle admet un maximum en λ_1 ou en λ_n, mais on a $\dfrac{\lambda_1}{\lambda_1} + \dfrac{\lambda_n}{\lambda_1} = \dfrac{\lambda_n}{\lambda_1} + \dfrac{\lambda_n}{\lambda_n} = 1 + \dfrac{\lambda_n}{\lambda_1}$.

On en déduit que

$$\begin{aligned}
\sqrt{\langle Ax, x\rangle \langle A^{-1}x, x\rangle} &\leq \frac{1}{2}\sqrt{\frac{\lambda_1}{\lambda_n}} \sum_{i=1}^{n} \left(1 + \frac{\lambda_n}{\lambda_1}\right) x_i^2 \\
&= \frac{1}{2}\left(\sqrt{\frac{\lambda_1}{\lambda_n}} + \sqrt{\frac{\lambda_n}{\lambda_1}}\right) \|x_i\|^2,
\end{aligned}$$

d'où, finalement :

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

□

Théorème 2. *Soit $A \in \mathcal{S}_n^{++}(\mathbb{R})$ et $b \in \mathbb{R}^n$. On cherche à minimiser $f : x \mapsto \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle$ quand x parcourt $\mathbb{R}^n$.*

Il existe une unique solution $\bar{x}$ à ce problème, et elle est caractérisée par $\nabla f(\bar{x}) = 0$.

De plus, la suite définie par $\begin{cases} x_0 \in \mathbb{R}^n \\ \forall k \in \mathbb{N} \quad x_{k+1} = x_k + t_k\, d_k \end{cases}$ converge vers $\bar{x}$, où $d_k = -\nabla f(x_k)$ et t_k est l'unique réel minimisant la fonction $t \mapsto f(x_k + t\, d_k)$.

Démonstration.

- **Etape 1 : montrons que le minimum de f est atteint en une unique valeur $\bar{x}$, caractérisée par $\nabla f(\bar{x}) = 0$, et déterminons la valeur de ce minimum.**

  La fonction f est polynomiale, donc de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^n$. Par ailleurs, pour tout $x, h \in \mathbb{R}^n$, on a

  $$\begin{aligned}
  f(x+h) &= \frac{1}{2}\langle A(x+h), x+h\rangle + \langle b, x+h\rangle \\
  &= \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + \frac{1}{2}\langle Ah, h\rangle + \langle b, h\rangle + \frac{1}{2}(\langle Ax, h\rangle + \langle Ah, x\rangle)
  \end{aligned}$$

  Comme $A \in \mathcal{S}_n^{++}(\mathbb{R})$, on a $\langle Ax, h\rangle = {}^t h A x = {}^t x\, {}^t A h = {}^t x A h = \langle Ah, x\rangle$. On en déduit que

  $$\begin{aligned}
  f(x+h) &= f(x) + \langle Ax, h\rangle + \langle b, h\rangle + \frac{1}{2}\langle Ah, h\rangle \\
  &= f(x) + \langle Ax + b, h\rangle + \frac{1}{2}\langle Ah, h\rangle \qquad (1)
  \end{aligned}$$

  Par ailleurs, on a $\frac{1}{2}|\langle Ah, h\rangle| \leq \frac{1}{2}\|Ah\| \cdot \|h\|$ d'après l'inégalité de Cauchy-Schwartz donc :

  $$\frac{1}{2}\langle Ah, h\rangle = \underset{\|h\| \to 0}{o}(\|h\|).$$

  Ainsi, on en déduit que $\nabla f(x) \cdot h = \langle Ax + b, h\rangle$ pour tout $h \in \mathbb{R}^n$, i.e $\nabla f(x) = Ax + b$, et la hessienne de f vaut $H(f) = A$ par définition. Comme $A \in \mathcal{S}_n^{++}(\mathbb{R})$, f est donc strictement convexe, donc admet au plus un minimum sur l'espace convexe $\mathbb{R}^n$.

Par ailleurs, si on pose $\overline{x} = -A^{-1}b$, on a d'après (1) :

$$\forall h \in \mathbb{R}^n \quad f(\overline{x}+h) - f(\overline{x}) = \frac{1}{2}\langle Ah, h\rangle = {}^t h A h \geq 0 \quad \text{car} \quad A \in \mathcal{S}_n^{++}(\mathbb{R}).$$

Donc f admet un unique minimum sur $\mathbb{R}^n$, qui vérifie $\nabla f(\overline{x}) = 0$, et ce minimum vaut

$$\begin{aligned} f(\overline{x}) &= \frac{1}{2}\langle -b, -A^{-1}b\rangle + \langle b, -A^{-1}b\rangle \\ &= -\frac{1}{2}\langle A^{-1}b, b\rangle. \end{aligned}$$

- **Etape 2 : calculons t_k.**

  Remarquons déjà que s'il existe $k \in \mathbb{N}$ tel que $d_k = 0$, alors $x_k = \overline{x}$ d'après l'étape 1, donc l'algorithme converge en temps fini. On suppose donc désormais que $d_k \neq 0$ pour tout $k \in \mathbb{N}$.

  On a alors, pour $k \in \mathbb{N}$ et $t \in \mathbb{R}$

$$\begin{aligned} 2f(x_k + t d_k) &= \langle A(x_k + t d_k), x_k + t d_k\rangle + 2\langle b, x_k + t d_k\rangle \\ &= \langle A x_k, x_k\rangle + t\langle A x_k, d_k\rangle + t\langle A d_k, x_k\rangle + t^2\langle A d_k, d_k\rangle + 2\langle b, x_k\rangle + 2t\langle b, d_k\rangle \\ &= \langle A d_k, d_k\rangle \cdot t^2 + 2\langle A x_k + b, d_k\rangle \cdot t + \langle A x_k, x_k\rangle + 2\langle b, x_k\rangle. \\ &= \langle A d_k, d_k\rangle t^2 - 2\|d_k\|^2 t + 2f(x_k) \end{aligned}$$

  Comme $\|d_k\|^2 > 0$, la fonction $t \mapsto 2f(x_k + t d_k)$ est polynomiale de degré 2, et elle admet un minimum en $t_k = \dfrac{\|d_k\|^2}{\langle A d_k, d_k\rangle}$.

- **Etape 3 : cherchons à majorer la quantité $\|x_k - \overline{x}\|$.**

  La matrice A est symétrique définie positive, donc elle est diagonalisable dans une base orthonormée et ses valeurs propres $\lambda_1 \leq \cdots \leq \lambda_n$ sont toutes strictement positives. Pour un vecteur $u \in \mathbb{R}^n$, si $(u_1, \ldots, u_n)$ sont ses coordonnées dans une telle base, on a alors

$$\begin{aligned} \langle Au, u\rangle &= \sum_{i=1}^{n} u_i (Au)_i \\ &= \sum_{i=1}^{n} u_i \lambda_i u_i \\ &\geq \lambda_1 \sum_{i=1}^{n} u_i \cdot u_i \\ &= \lambda_1 \|u\|^2. \end{aligned}$$

  En appliquant ce résultat à $u = \|x_k - \overline{x}\|$, il vient

$$\begin{aligned} \lambda_1 \|x_k - \overline{x}\|^2 &\leq \langle A(x_k - \overline{x}), x_k - \overline{x}\rangle \\ &= \langle A x_k, x_k\rangle - \langle A\overline{x}, x_k\rangle - \langle A x_k, \overline{x}\rangle + \langle A\overline{x}, \overline{x}\rangle \\ &= \langle A x_k, x_k\rangle - 2\langle A(-A^{-1}b), x_k\rangle + \langle A(-A^{-1}b), -A^{-1}b\rangle \\ &= \langle A x_k, x_k\rangle + 2\langle b, x_k\rangle + \langle b, A^{-1}b\rangle. \end{aligned}$$

  On a $\langle A x_k, x_k\rangle + 2\langle b, x_k\rangle = 2f(x_k)$, et $\langle b, A^{-1}b\rangle = \langle A^{-1}b, b\rangle = -2f(\overline{x})$, d'où

$$\|x_k - \overline{x}\|^2 \leq \frac{2}{\lambda_1}(f(x_k) - f(\overline{x})).$$

- **Etape 4 : trouvons une relation de récurrence sur la suite $\boldsymbol{(f(x_k) - f(\overline{x}))_{k\in\mathbb{N}}}$.**

  On écrit

$$
\begin{aligned}
2\,(f(x_{k+1}) - f(\overline{x})) &= 2\,f(x_k + t_k\,d_k) - 2\,f(\overline{x}) \\
&= \left\langle A\,x_k + \frac{\|d_k\|^2}{\langle A d_k, d_k\rangle}\,A\,d_k,\; x_k + \frac{\|d_k\|^2}{\langle A d_k, d_k\rangle}\,d_k \right\rangle + 2\left\langle b,\; x_k + \frac{\|d_k\|^2}{\langle A d_k, d_k\rangle}\,d_k \right\rangle - 2\,f(\overline{x}) \\
&= 2\,f(x_k) - 2\,f(\overline{x}) + \frac{\|d_k\|^2}{\langle A d_k, d_k\rangle}[2\,\langle A x_k, d_k\rangle + 2\,\langle b, d_k\rangle] + \frac{\|d_k\|^4}{\langle A d_k, d_k\rangle}.
\end{aligned}
$$

  On a $\langle A x_k, d_k\rangle + \langle b, d_k\rangle = \langle A x_k + b, d_k\rangle = -\langle d_k, d_k\rangle$, d'où

$$
\begin{aligned}
2\,(f(x_{k+1}) - f(\overline{x})) &= 2\,f(x_k) - 2\,f(\overline{x}) - \frac{\|d_k\|^4}{\langle A d_k, d_k\rangle} \\
&= 2\,(f(x_k) - f(\overline{x}))\left(1 - \frac{\|d_k\|^4}{2\,(f(x_k) - f(\overline{x}))\,\langle A d_k, d_k\rangle}\right).
\end{aligned}
$$

  Par ailleurs,

$$
\begin{aligned}
\langle A^{-1}\,d_k, d_k\rangle &= \langle A^{-1}(A x_k + b), (A x_k + b)\rangle \\
&= \langle\, x_k + A^{-1} b, A x_k + b\rangle \\
&= \langle A x_k, x_k\rangle + \langle b, x_k\rangle + \langle b, x_k\rangle + \langle A^{-1}\,b, b\rangle \\
&= 2\,f(x_k) - 2\,f(\overline{x}).
\end{aligned}
$$

  Et d'après l'inégalité de Kantorovitch, on a

$$
\langle A^{-1}\,d_k, d_k\rangle\,\langle A\,d_k, d_k\rangle \;\leq\; \frac{1}{4}\left(\sqrt{c(A)} + \frac{1}{\sqrt{c(A)}}\right)\|d_k\|^4.
$$

  On déduit de ces deux calculs que

$$
\begin{aligned}
2\,(f(x_{k+1}) - f(\overline{x})) &\leq 2\,(f(x_k) - f(\overline{x}))\left(1 - \frac{1}{\frac{1}{4}\left(\sqrt{c(A)} + \frac{1}{\sqrt{c(A)}}\right)^2}\right) \\
&= 2\,(f(x_k) - f(\overline{x}))\left(1 - \frac{4 c(A)}{(c(A) + 1)^2}\right) \\
&= 2\,(f(x_k) - f(\overline{x}))\left(\frac{c(A) - 1}{c(A) + 1}\right)^2.
\end{aligned}
$$

  On en déduit que :

$$
\forall k \in \mathbb{N} \quad 2\,(f(x_{k+1}) - f(\overline{x})) \;\leq\; 2\,(f(x_0) - f(\overline{x}))\left(\frac{c(A) - 1}{c(A) + 1}\right)^{2k}.
$$

- **Etape 4 : conclusion**

  D'après les résultats des étapes 3 et 4, on a pour $k \in \mathbb{N}$:

$$
\begin{aligned}
\|x_k - \overline{x}\|^2 &\leq 2\,(f(x_k) - f(\overline{x})) \\
&\leq 2\,(f(x_0) - f(\overline{x}))\left(\frac{c(A) - 1}{c(A) + 1}\right)^{2k}.
\end{aligned}
$$

Par ailleurs, on a $\left|\left(\frac{c(A)-1}{c(A)+1}\right)^2\right| < 1$, donc $\lim\limits_{k\to+\infty} \|x_k - \overline{x}\|^2 = 0$. Ainsi, $(x_k)_{k\in\mathbb{N}}$ converge bien vers $\overline{x}$, et ceci conclut la preuve.

$\square$