

Méthode du gradient à pas optimal: Bonis p. 159 219-226-229-233

Soit (P): $Ax = b$ où $A \in \mathbb{S}^{++}(N)$ et $b \in \mathbb{R}^n$; $\bar{x}$ l'unique solution de (P)

On munit $\mathbb{R}^n$ de $\langle \cdot, \cdot \rangle_A$ et $\langle \cdot, \cdot \rangle$ où $\forall x, y \in \mathbb{R}^n$,

$$\langle x, y \rangle_A = \langle Ax, Ay \rangle. \text{ B canonique. } (\langle x, y \rangle = \langle y, x \rangle)$$

$$\text{Soit } \phi: \mathbb{R}^n \rightarrow \mathbb{R} \\ x \mapsto \frac{1}{2} \|Ax - b\|_A^2 \quad ; \text{ alors } \nabla \phi(x) = Ax - b \quad \textcircled{1} \\ = A(x - \bar{x})$$

et ϕ atteint son minimum en $\bar{x}$, et seulement en ce point.

On va montrer que pour $a \in \mathbb{R}^n \setminus \{\bar{x}\}$ et $(\alpha_k)_{k \in \mathbb{N}} \in (\mathbb{R}^+)^{\mathbb{N}}$ définie

$$\text{par: } \begin{cases} x_0 = a \\ x_{k+1} = x_k - \alpha_k \nabla \phi(x_k) \end{cases} \quad \text{où } \alpha_k = \begin{cases} \frac{\|\nabla \phi(x_k)\|^2}{\|\nabla \phi(x_k)\|_A^2}, & \text{si } x_k \neq \bar{x} \\ 0, & \text{sinon} \end{cases}$$

alors on a: $x_k \rightarrow \bar{x}$ et $\forall k \in \mathbb{N}$,

$$\|x_{k+1} - \bar{x}\| \leq \sqrt{\frac{\lambda_1}{\lambda_2}} \left(\frac{\lambda_1 - \lambda_1}{\lambda_2 + \lambda_1} \right)^{k+1} \|x_0 - \bar{x}\| \quad \text{où } \begin{cases} \lambda_2 = \max \{\lambda(A)\} > 0 \\ \lambda_1 = \min \{\lambda(A)\} > 0 \end{cases}$$

Démonstration: Commençons par expliquer la construction de $(\alpha_k)_k$:

la solution de (P) revient à minimiser ϕ ; or, pour $y \in \mathbb{R}^n \setminus \{\bar{x}\}$, $\nabla \phi(y) \neq 0$

$\nabla \phi(y)$ indique la direction selon laquelle ϕ croît le plus vite.

Donc pour x_k construit, on cherche x_{k+1} sur la droite passant par

x_k et dirigée par $-\nabla \phi(x_k)$ (supposé non nul). Si $\nabla \phi(x_k) = 0$ alors

$x_k = \bar{x}$ et on pose $x_{k+1} = x_k$.

Plus précisément: $f: \mathbb{R} \rightarrow \mathbb{R} \\ t \mapsto \phi(x_k - t \nabla \phi(x_k))$ atteint son minimum en α_k

$$\text{En effet: } \phi(x_k + h) = \phi(x_k) + \langle Ax_k - b, h \rangle + \frac{1}{2} \|Ah\|_A^2 \quad \textcircled{2}$$

$$\text{soit } f(t) = \phi(x_k) + \langle Ax_k - b, -t \nabla \phi(x_k) \rangle + \frac{1}{2} \| -t \nabla \phi(x_k) \|_A^2$$

$$= \phi(x_k) - t \|\nabla \phi(x_k)\|^2 + \frac{t^2}{2} \|\nabla \phi(x_k)\|_A^2 \quad \text{car une polynôme}$$

de 2nd degré qui atteint son minimum en α_k .

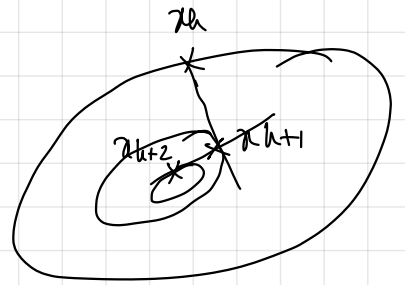
$$\text{d'où le cas: } f'(\alpha_k) = 0$$

$$\text{or } f'(\alpha h) = \frac{2}{h} \left(\phi(x_h - \alpha h \nabla \phi(x_h)) \right)$$

$$= \nabla \phi(x_h - \alpha h \nabla \phi(x_h)) \cdot \nabla \phi(x_h)$$

$$= \left\langle \nabla \phi(\underbrace{x_h - \alpha h \nabla \phi(x_h)}_{x_{h+1}}), \nabla \phi(x_h) \right\rangle$$

$$= \left\langle \nabla \phi(x_{h+1}), \nabla \phi(x_h) \right\rangle = 0 \quad (*)$$



Donc (x_h, x_{h+1}) est tangente en x_{h+1} à la ligne de niveau de ϕ passant par x_{h+1} . (3)

Supposons que $x_p \neq \bar{x}$ $\forall p \in \mathbb{N}$ et notons $g_p = \nabla \phi(x_p)$

$$\begin{aligned} \text{Puis : } \|x_{p+1} - \bar{x}\|_A^2 &= \left\langle A(x_{p+1} - \bar{x}), x_{p+1} - x_p + x_p - \bar{x} \right\rangle \\ &= \left\langle A(x_{p+1} - \bar{x}), x_{p+1} - x_p \right\rangle + \left\langle A(x_{p+1} - \bar{x}), x_p - \bar{x} \right\rangle \\ &= \underbrace{\left\langle g_{p+1}, -\alpha_p g_p \right\rangle}_{=0 \text{ par } (*)} + \left\langle A(x_{p+1} - \bar{x}), x_p - \bar{x} \right\rangle \end{aligned}$$

$$\begin{aligned} \text{Et } \left\langle A(x_{p+1} - \bar{x}), x_p - \bar{x} \right\rangle &= \left\langle A(x_{p+1} - x_p + x_p - \bar{x}), x_p - \bar{x} \right\rangle \\ &= \left\langle A(x_{p+1} - x_p), x_p - \bar{x} \right\rangle + \left\langle A(x_p - \bar{x}), x_p - \bar{x} \right\rangle \\ &= \left\langle x_{p+1} - x_p, A(x_p - \bar{x}) \right\rangle + \|x_p - \bar{x}\|_A^2 \\ &= \left\langle -\alpha_p g_p, g_p \right\rangle + \|x_p - \bar{x}\|_A^2 \\ &= -\frac{\|g_p\|^2}{\|g_p\|_A^2} \|g_p\|^2 + \|x_p - \bar{x}\|_A^2 \end{aligned}$$

$$\begin{aligned} \text{Par ailleurs : } \|x_p - \bar{x}\|_A^2 &= \left\langle A(x_p - \bar{x}), x_p - \bar{x} \right\rangle = \left\langle A(x_p - \bar{x}), A^{-1} A(x_p - \bar{x}) \right\rangle \\ &= \left\langle g_p, A^{-1} g_p \right\rangle = \|g_p\|_{A^{-1}}^2 \end{aligned}$$

$$\text{Donc } \|x_{p+1} - \bar{x}\|_A^2 = \|x_p - \bar{x}\|_A^2 \left[1 - \frac{\|g_p\|^4}{\|g_p\|_A^2} \times \frac{1}{\|g_p\|_{A^{-1}}^2} \right]$$

$$\leq \|x_p - \bar{x}\|_A^2 \left[1 - 4 \frac{\lambda_1 \lambda_2}{(\lambda_1 + \lambda_2)^2} \right] = \|x_p - \bar{x}\|_A^2 \left(\frac{\lambda_2 - \lambda_1}{\lambda_2 + \lambda_1} \right)^2$$

(4) d'après l'inégalité de Kantorovich.

En réinjectant, on obtient $\|x_{p+1} - \bar{x}\|_A \leq \left(\frac{\lambda_2 - \lambda_1}{\lambda_2 + \lambda_1} \right)^{p+1} \|x_0 - \bar{x}\|_A$ (5)

Or $\sqrt{\lambda_1} u \leq u \cdot \|u\|_A \leq \sqrt{\lambda_2} u$ (6) donc :

$$\|x_{p+1} - \bar{x}\| \leq \sqrt{\frac{\lambda_2}{\lambda_1}} \left(\frac{\lambda_2 - \lambda_1}{\lambda_2 + \lambda_1} \right)^{p+1} \|x_0 - \bar{x}\| \quad \square$$

(1) (2)

$$\begin{aligned} \phi(x+u) &= \frac{1}{2} \|x+u\|_A^2 - \langle b, x+u \rangle \\ &= \frac{1}{2} \langle x+u, A(x+u) \rangle - \langle b, x \rangle - \langle b, u \rangle \\ &= \frac{1}{2} \langle x, A(x+u) \rangle + \frac{1}{2} \langle u, A(x+u) \rangle - \langle b, x \rangle - \langle b, u \rangle \\ &= \frac{1}{2} \langle x, Ax \rangle + \frac{1}{2} \langle u, Au \rangle + \frac{1}{2} \langle x, Au \rangle - \langle b, x \rangle - \langle b, u \rangle \\ &= \phi(x) + \langle x, Au \rangle + \frac{1}{2} \|u\|_A^2 - \langle b, u \rangle \\ &= \phi(x) + \langle Ax, u \rangle - \langle b, u \rangle + \frac{1}{2} \|u\|_A^2 \\ &= \phi(x) + \langle Ax - b, u \rangle + \frac{1}{2} \|u\|_A^2 \\ &= \phi(x) + \langle Ax - b, u \rangle + \frac{c}{2} \|u\|^2 \quad (\text{équivalence des normes au d'ordre } \infty) \\ &= \phi(x) + \langle Ax - b, u \rangle + o(\|u\|) \end{aligned}$$

donc $D\phi(x)(u) = \langle Ax - b, u \rangle$

$D\phi(\bar{x}) = A\bar{x} - b = 0$ donc $b = A\bar{x}$

donc $D\phi(x) = Ax - A\bar{x} = A(x - \bar{x})$.

(3) Les lignes de niveau de ϕ sont des sphères pour la norme $\| \cdot \|_A$ ou des ellipsoïdes pour $\| \cdot \|$, de centre $\bar{x}$

(4) Soit $x \in \mathbb{R}^n \setminus \{0\}$, par le théorème spectral, A admet une base (ex...en) de vecteurs propres ; on note λ_i la VP associé à e_i ; qui a aussi vect P de A^{-1} associé à la VP $\frac{1}{\lambda_i}$

Ainsi ; si $x = x_1 e_1 + \dots + x_n e_n$:

$$Ax = x_1 \lambda_1 e_1 + \dots + x_n \lambda_n e_n$$

$$\|x\|_A \|x\|_{A^{-1}} = \sqrt{\langle x, Ax \rangle} \sqrt{\langle x, A^{-1}x \rangle}$$

$$= \sqrt{\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} \frac{1}{d_{\min}} x_1 \\ \vdots \\ \frac{1}{d_{\min}} x_n \end{pmatrix} \right\rangle} \sqrt{\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} \frac{1}{d_{\max}} x_1 \\ \vdots \\ \frac{1}{d_{\max}} x_n \end{pmatrix} \right\rangle}$$

$$= \sqrt{\sum d_i x_i^2} \sqrt{\sum \frac{1}{d_i} x_i^2}$$

$$= \sqrt{d_{\max} \sum \frac{d_i}{d_{\max}} x_i^2} \sqrt{\frac{1}{d_{\min}} \sum \frac{d_{\min}}{d_i} x_i^2}$$

$$= \sqrt{\frac{d_{\max}}{d_{\min}}} \sqrt{\sum \frac{d_i}{d_{\max}} x_i^2} \sqrt{\sum \frac{d_{\min}}{d_i} x_i^2}$$

$$\leq \frac{1}{2} \sqrt{\frac{d_{\max}}{d_{\min}}} \sum_{i=1}^n \left(\frac{d_i}{d_{\max}} + \frac{d_{\min}}{d_i} \right) x_i^2$$

abs $\leq \frac{1}{2} (d_{\max}^2 + d_{\min}^2)$

Or $g: [d_{\min}, d_{\max}] \rightarrow \mathbb{R}$

$$t \mapsto \frac{t}{d_{\max}} + \frac{d_{\min}}{t}$$

et convexe en fait que
somme de deux x appliqué
convexes

comme $g(d_{\min}) = g(d_{\max}) = 1 + \frac{d_{\min}}{d_{\max}} \geq g(t) \forall t$

$$\text{donc } \|x\|_A \|x\|_{A^{-1}} \leq \frac{1}{2} \sqrt{\frac{d_{\max}}{d_{\min}}} \left(1 + \frac{d_{\min}}{d_{\max}} \right) \|x\|^2$$

$$\leq \frac{1}{2} \sqrt{\frac{d_{\max}}{d_{\min}}} \left(\frac{d_{\max} + d_{\min}}{d_{\max} d_{\min}} \right) \|x\|^2$$

$$\text{donc } \frac{\|x\|^4}{\|x\|_A^2 \|x\|_{A^{-1}}^2} \geq 4 \frac{d_{\min}}{d_{\max}} \left(\frac{d_{\max} d_{\min}}{d_{\max} + d_{\min}} \right)^2 \dots ?$$

$$\textcircled{5} \|x_{p+1} - \bar{x}\|_A^2 \leq \left(\frac{d_2 - d_1}{d_2 + d_1} \right)^2 \|x_p - \bar{x}\|_A^2$$

$$\leq \left(\frac{d_2 - d_1}{d_2 + d_1} \right)^2 \left(\frac{d_2 - d_1}{d_2 + d_1} \right)^2 \|x_{p-1} - \bar{x}\|_A^2$$

$$\vdots$$

$$\leq \left[\left(\frac{d_2 - d_1}{d_2 + d_1} \right)^2 \right]^{1+p} \|x_0 - \bar{x}\|_A^2$$

$$\textcircled{6} \|x\|_A = \sqrt{\langle x, Ax \rangle} = \sqrt{\sum d_i x_i^2} = \sqrt{d_{\max}} \sqrt{\sum \frac{d_i}{d_{\max}} x_i^2}$$

$$\leq \sqrt{\sum_{i=1}^n} \sqrt{\sum_{i=1}^n x_i^2} = \sqrt{\sum_{i=1}^n \|x_i\|^2} \quad \sum_{i=1}^n$$

$$\|x\|_A = \sqrt{\sum_{i=1}^n} \sqrt{\sum_{\substack{i=1 \\ \lambda_i \geq a}}^n x_i^2} \geq \sqrt{\sum_{i=1}^n} \|x_i\|.$$