Minimisation de fonctionnelle quadratique

Dorian Cacitti-Holland

2020-2021

Références.

1. Oraux X-ENS Analyse 4 (pas tout à fait)
2. Analyse matricielle de Jean-Etienne Rombaldi (pas tout à fait)

Leçons.

1. 158 Matrices symétriques réelles, matrices hermitiennes
2. 162 Systèmes d'équations linéaires, opérations élémentaires, aspects algorithmiques et conséquences théoriques
3. 215 Applications différentiables définies sur un ouvert de $\mathbb{R}^n$, exemples et applications
4. 219 Extrema, existence, caractéristation, recherche, exemple et applications
5. 226 Suites vectorielles et réelles définies par une relation $u_{n+1} = f(u_n)$, exemples, applications à la résolution approchée d'équations
6. 229 Fonctions monotones, fonctions convexes, exemples et applications
7. 233 Analyse numérique matricielle, résolution de systèmes linéaires, recherche d'éléments propres, exemples
8. 253 Utilisation de la notion de convexité en analyse

Théorême. Soit $A \in S_p^{++}(\mathbb{R}), b \in \mathbb{R}^p$ et

$$f : \begin{array}{rcl} \mathbb{R}^p & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle \end{array}$$

Alors on définit la suite $(x_n)_{n\in\mathbb{N}}$ par

$$\begin{cases} x_0 \in \mathbb{R} \\ \forall n \in \mathbb{N}, x_{n+1} = x_n - \rho_n \nabla f(x_n) \end{cases}$$

avec $\forall n \in \mathbb{N}, \rho_n = \underset{\rho\in\mathbb{R}_+^*}{argmin}(f(x_n - \rho\nabla f(x_n)))$.

Ainsi

$$x_n \underset{n\to+\infty}{\longrightarrow} x_*$$

avec $x_* \in \mathbb{R}^p$ unique minimum global de f.

Démonstration.
Etape 1 : Calcul de ∇f et convexité de f

Soit $x \in \mathbb{R}^p$, alors, comme $A \in S_n^{++}(\mathbb{R})$,

$$\begin{array}{rcl} \forall h \in \mathbb{R}^p, f(x+h) & = & f(x) + \frac{1}{2}\langle Ah, x\rangle + \frac{1}{2}\langle Ax, h\rangle - \langle b, h\rangle + \frac{1}{2}\langle Ah, h\rangle \\ & = & f(x) + \frac{1}{2}\langle h, Ax\rangle + \frac{1}{2}\langle Ax, h\rangle - \langle b, h\rangle + \frac{1}{2}\langle Ah, h\rangle \\ & = & f(x) + \langle Ax - b, h\rangle + \frac{1}{2}\langle Ah, h\rangle \end{array}$$

D'où

$$\nabla f(x) = Ax - b$$

De plus $H_f = A \in S_p^{++}(\mathbb{R})$, donc f est strictement convexe.
Etape 2 : f continue et coercive

En effet f est polynomiale en les composantes de $x \in \mathbb{R}^p$, d'où f est continue sur $\mathbb{R}^p$.
De plus, comme $A \in S_n^{++}(\mathbb{R})$, d'après le premier lemme et l'inégalité de Cauchy-Schwarz,

$$\forall x \in \mathbb{R}^p, f(x) \geq \frac{\alpha}{2}\|x\|^2 - \|b\|\,\|x\|$$

D'où

$$f(x) \underset{\|x\| \longrightarrow +\infty}{\longrightarrow} +\infty$$

Ce qui montre que f est coercive.
Par conséquent f admet son minimum global en un unique point $x_* \in \mathbb{R}^p$.
En particulier $\nabla f(x_*) = 0$, ie $x_* = A^{-1}b$.
Etape 3 : $\forall (x,y) \in (\mathbb{R}^p)^2, f(y) - f(x) \geq \langle \nabla f(x), y - x\rangle + \frac{\alpha}{2}\|x-y\|^2$

Soit $(x,y) \in \mathbb{R}^p$, alors, comme $A \in S_p^{++}(\mathbb{R})$,

$$\begin{array}{rcl} f(y) - f(x) & = & \frac{1}{2}\left(\langle Ay, y\rangle - \langle Ax, x\rangle\right) - \langle b, y-x\rangle \\ & = & \frac{1}{2}\left(\langle Ay, y\rangle + \langle Ax, x\rangle\right) - \langle Ax, x\rangle - \langle b, y-x\rangle \\ & = & \frac{1}{2}\left(\langle Ay, y\rangle + \langle Ax, x\rangle\right) + \langle Ax, y-x\rangle - \langle Ax, y\rangle - \langle b, y-x\rangle \\ & = & \frac{1}{2}\left(\langle Ay, y\rangle + \langle Ax, x\rangle\right) + \langle Ax-b, y-x\rangle - \langle Ax, y\rangle \\ & = & \frac{1}{2}\left(\langle Ay, y\rangle + \langle Ax, x\rangle - 2\langle Ax, y\rangle\right) + \langle Ax-b, y-x\rangle \\ & = & \frac{1}{2}\langle A(x-y), x-y\rangle + \langle Ax-b, y-x\rangle \\ & = & \frac{1}{2}\langle A(x-y), x-y\rangle + \langle \nabla f(x), y-x\rangle \\ & \geq & \frac{\alpha}{2}\|x-y\|^2 + \langle \nabla f(x), y-x\rangle \end{array}$$

Etape 4 : Construction du pas optimal ρ

Soit $x \in \mathbb{R}^p \backslash \{x_*\}$ tel que $\nabla f(x) \neq 0$, on considère

$$\varphi_x : \begin{array}{rcl} \mathbb{R}_+ & \longrightarrow & \mathbb{R} \\ \rho & \longmapsto & f(x - \rho\nabla f(x)) \end{array}$$

Ainsi φ_x est de classe C^1, car f est de classe C^∞, et, comme f est coercive et $\nabla f(x) \neq 0$,

$$\varphi_x(\rho) \underset{\rho \to +\infty}{\longrightarrow} +\infty$$

Donc φ_x admet son minimum en un point $\rho_x \in \mathbb{R}_+$, avec $\rho_x \neq 0$ car $\varphi_x(0) = f(x)$ et $x \neq x_*$.
En particulier

$$0 = \varphi_x'(\rho_x) = df(x-\rho_x \nabla f(x))(-\nabla f(x)) = \langle \nabla f(x-\rho_x \nabla f(x)), -\nabla f(x) \rangle = -\langle \nabla f(x-\rho_x f(x)), \nabla f(x) \rangle$$

D'où, grâce à l'étape 3 appliquée avec "y" $= x - \rho \nabla f(x)$ et "x" $= x - \rho_x \nabla f(x)$,

$$\begin{aligned} \forall \rho \in \mathbb{R}_+ \backslash \{\rho_x\}, \quad & \varphi_x(\rho) - \varphi_x(\rho_x) \\ & = f(x - \rho \nabla f(x)) - f(x - \rho_x \nabla f(x)) \\ & \geq \langle \nabla f(x - \rho_x \nabla f(x)), x - \rho \nabla f(x) - (x - \rho_x \nabla f(x)) \rangle \\ & \quad + \tfrac{\alpha}{2} \|x - \rho_x \nabla f(x) - (x - \rho \nabla f(x))\|^2 \\ & = \langle \nabla f(x - \rho_x \nabla f(x)), (\rho_x - \rho) \nabla f(x) \rangle + \tfrac{\alpha}{2} \|(\rho - \rho_x) \nabla f(x)\|^2 \\ & = 0 + \tfrac{\alpha}{2} (\rho - \rho_x)^2 \|\nabla f(x)\|^2 \\ & > 0 \end{aligned}$$

D'où ρ_x est l'unique minimum global de φ_x, ie

$$\rho_x = \underset{\rho \in \mathbb{R}_+}{argmin}(\varphi_x(\rho))$$

Etape 5 : Construction de la suite $(x_n)_{n\in\mathbb{N}}$ convergent vers x_*
La suite $(x_n)_{n\in\mathbb{N}}$ est bien définie tant que $x_n \neq x_*$, or s'il existe $n \in \mathbb{N}$ tel que $x_n = x_*$ alors la suite stationne en x_* ce qui permet de conclure sur la convergence.
Sinon la suite $(f(x_n))_{n\in\mathbb{N}}$ est minorée (car f est minorée) et décroissante car, d'après l'étape 3 (appliquée en "x" $= x_{n+1}$ et "y" $= x_n$) et un calcul effectué à l'étape 4,

$$\forall n \in \mathbb{N}, f(x_n) - f(x_{n+1}) \geq \langle \nabla f(x_{n+1}), \rho_n \nabla f(x_n) \rangle + \frac{\alpha}{2} \|\rho_n \nabla f(x_n)\|^2 = 0 + \frac{\alpha}{2} \|\rho_n \nabla f(x_n)\|^2 \geq 0$$

Donc, d'après le théorème de la limite monotone, la suite $(f(x_n))_{n\in\mathbb{N}}$ converge.
D'où, en faisant tendre n vers $+\infty$ dans l'inégalité précédente, on obtient

$$\|x_{n+1} - x_n\| = \|\rho_n \nabla f(x_n)\| \underset{n\to+\infty}{\longrightarrow} 0$$

Or $(x_n)_{n\in\mathbb{N}}$ est bornée car f coercive et $(f(x_n))_{n\in\mathbb{N}}$ convergente, et $(x_n)_{n\in\mathbb{N}}$ admet une unique valeur d'adhérence.
En effet on considère une sous-suite $(x_{\psi(n)})_{n\in\mathbb{N}}$ convergeante vers $x \in \mathbb{R}^p$.
Par continuité de ∇f,

$$\nabla f(x_{\psi(n)}) \underset{n\to+\infty}{\longrightarrow} \nabla f(x)$$

De plus

$$x_{\psi(n)+1} = x_{\psi(n)+1} - x_{\psi(n)} + x_{\psi(n)} \underset{n\to+\infty}{\longrightarrow} x$$

D'où

$$0 = \langle \nabla f(x_{\psi(n)}), \nabla f(x_{\psi(n)+1}) \rangle \underset{n\to+\infty}{\longrightarrow} \|\nabla f(x)\|^2$$

Ainsi $\nabla f(x) = 0$, ie $x = x_*$.
Par conséquent $(x_n)_{n\in\mathbb{N}}$ converge vers x_*. $\square$

Lemme. (S'il reste du temps) Soit $A \in S_p^{++}(\mathbb{R})$ et $\alpha = inf(Sp(A)) \in \mathbb{R}_+^*$, alors

$$\forall x \in \mathbb{R}^p, \langle Ax, x\rangle \geq \alpha \left\|x\right\|^2$$

Démonstration. Comme $A \in S_p^{++}(\mathbb{R})$, d'après le théorème spectral il existe $D \in M_p(\mathbb{R})$ diagonale composée des valeurs propres de A comptées avec multiplicité et $P \in O_p(\mathbb{R})$ tel que

$$A = PD\ {}^tP$$

Soit $x \in \mathbb{R}^p$ et $y = {}^tPx$, alors, comme P est orthogonale,

$$\left\|x\right\|^2 = \langle x, x\rangle = \langle Py, Py\rangle = \langle y, y\rangle = \left\|y\right\|^2$$

Et

$$\langle Ax, x\rangle = \langle APy, Py\rangle = \langle\ {}^tPAPy, y\rangle = \langle Dy, y\rangle = \sum_{i=1}^{p} \lambda_i y_i^2 \geq \alpha \left\|y\right\|^2$$

D'où

$$\langle Ax, x\rangle \geq \alpha \left\|x\right\|^2$$

□