Trouver ref

→ dépend de la fct à minimiser, précision souhaitée et de méthode pour calculer le pas opt !

Complexité en $O(N^2 L)$ où N est la dimension de l'espace et L le nb d'it pour atteindre la solution opti.

158
162
215
219
226
229
253

Développement : minimisation de fonctionnelle quadratique.

Théorème : Soit $A \in S_p^{++}(\mathbb{R})$, $b \in \mathbb{R}^p$ et $f : \begin{array}{ccc} \mathbb{R}^p & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle \end{array}$

On définit (x_m) par $x_0 \in \mathbb{R}$ et $\forall m \in \mathbb{N}$, $x_{m+1} = x_m - \rho_m \nabla f(x_m)$ avec $\forall m \in \mathbb{N}$, $\rho_m = \underset{\rho \in \mathbb{R}^*_+}{\operatorname{argmin}} \left(f(x_m - \rho \nabla f(x_m)\right)$. Alors $x_m \to x_*$ où $x_* \in \mathbb{R}^p$ est l'unique minimum global de f.

Dans tout le dvlpt, $\|x\| = \|x\|_2 = (\sum x_i^2)^{\frac{1}{2}}$

On peut sauter l'Etape 1 si pas le temps (les lemmes sont dans le plan) mais savoir les démo.

Démo : Etape 1 : Calcul de ∇f et convexité de f

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

$\forall h \in \mathbb{R}^p \quad f(x+h) = \dots = f(x) + \langle Ax - b, h\rangle + \frac{1}{2}\langle Ah, h\rangle$ d'où $\nabla f(x) = Ax - b$

De plus $H_f = A \in S_p^{++}(\mathbb{R})$ donc f est strictement convexe pk ?

(2) Etape 2 : f est continue et coercive

f est polynomiale en les composantes de $x \in \mathbb{R}^p$, d'où f est continue sur $\mathbb{R}^p$.

CS : $\langle b, x\rangle \leq \|b\| \|x\|$

Lemme 5.10 p 215

De plus, comme $A \in S_m^{++}(\mathbb{R})$, d'après le lemme (annexe) et l'inégalité de Cauchy-Schwarz :

$$\forall x \in \mathbb{R}^p, \quad f(x) \geq \frac{\alpha}{2}\|x\|^2 - \|b\|\|x\| \qquad (\text{où } \alpha = \inf(Sp(A) \setminus \{0\})$$

inut de car $A \in S^{++}$

d'où $f(x) \underset{\|x\| \to +\infty}{\longrightarrow} +\infty$ ie f coercive.

faut-il savoir le démontrer ?

f continue et coercive (sur $\mathbb{R}^p$ fermé !) $\Rightarrow$ existence d'un minimum global sur $\mathbb{R}^p$.

De plus f strictement convexe (sur $\mathbb{R}^p$ convexe !) $\Rightarrow$ unicité du minimum global sur $\mathbb{R}^p$. On le note x_*

Thm 5.12 p 211

En particulier $\nabla f(x_*) = 0$ ie $x_* = A^{-1}b$ ($A \in S^{++}$ donc inv)

si $\exists a$ et b, $a < b$, f m'impose atteint en a et b on a $f'(a) = f'(b) = 0$ mais comp. $f'(a) < f'(b)$ ABS

→ que sur $\mathbb{R}$ pas $\mathbb{R}^p$

(3) Etape 3 : $\forall (x,y) \in (\mathbb{R}^p)^2$, $f(y) - f(x) \geq \langle \nabla f(x), y - x\rangle + \frac{\alpha}{2}\|x - y\|^2$

Soient $x, y \in \mathbb{R}^p$: $f(y) - f(x) = \frac{1}{2}(\langle Ay, y\rangle - \langle Ax, x\rangle) - \langle b, y - x\rangle$

$= \langle Ax - b, y - x\rangle - \langle Ax, y\rangle + \frac{1}{2}(\langle Ay, y\rangle + \langle Ax, x\rangle)$

$= \langle Ax - b, y - x\rangle + \frac{1}{2}\langle A(x-y), (x-y)\rangle$

$\geq \langle \nabla f(x), y - x\rangle + \frac{\alpha}{2}\|x - y\|^2$

On réutilise le lemme

(4) Etape 4 : Construction du pas optimal

Soit $x \in \mathbb{R}^p$ tel que $\nabla f(x) \neq 0$ ($x \neq x_*$). On pose $\varphi_x : \begin{array}{ccc} \mathbb{R}_+ & \longrightarrow & \mathbb{R} \\ \rho & \longmapsto & f(x - \rho\nabla f(x)) \end{array}$. C'est une fonction C^1 (car f C^∞) et comme f est coercive, $\varphi_x(\rho) \underset{\rho \to +\infty}{\longrightarrow} +\infty$. Elle admet donc son minimum en un point $\rho_x \in \mathbb{R}_+$, avec $\rho_x \neq 0$ car $\varphi_x(0) = f(x)$ et $x \neq x_*$.

car f polynomiale

↳ qui annule la dérivée de φ_x.

⚠ À vérifier

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

Donc pour tout $\rho \neq \rho_x$:

(1) avec "y" $= x - \rho \nabla f(x)$ et "x" $= x - \rho_x \nabla f(x)$

$\varphi_x(\rho) - \varphi_x(\rho_x) \geq \underbrace{\langle \nabla f(x - \rho_x \nabla f(x)), (\rho - \rho_x) \nabla f(x) \rangle} + \frac{\alpha}{2} \|(\rho - \rho_x) \nabla f(x)\|^2 > 0$

Donc ρ_x est l'unique minimum global° de φ_x ie $\rho_x = \underset{\rho \in \mathbb{R}_+}{\operatorname{argmin}} (\varphi_x(\rho))$

(5) <u>Etape 5</u> : Construction de la suite $(x_m)_{m \in \mathbb{N}}$ convergent vers x_*

Notre suite (x_m) est ainsi bien définie tant que $x_m \neq x_*$

Si elle atteint x_*, elle est stationnaire en ce point et le résultat est clair

Sinon la suite $(f(x_m))$ est minorée (car f est minorée) et décroissante (appliquer (3) en "x" $= x_{m+1}$ et "y" $= x_m$ et utiliser (4)). D'après le théorème de la limite monotone, la suite $(f(x_m))$ converge

Alors $f(x_m) - f(x_{m+1}) \geq \frac{\alpha}{2} \|x_{m+1} - x_m\|^2$ donc $\|x_{m+1} - x_m\| \xrightarrow[m \to +\infty]{} 0$

Or (x_m) est bornée car f coercive et $(f(x_m))$ convergente. Montrons que (x_m) admet une seule valeur d'adhérence.

Bolzano Weierstrass — On considère une sous-suite $(x_{\varphi(m)})$ convergente vers $x \in \mathbb{R}^p$. Par continuité de ∇f on a $\nabla f(x_{\varphi(m)}) \xrightarrow[m \to \infty]{} \nabla f(x)$.

Et comme $(x_{\varphi(m)+1})$ converge aussi vers x puisque $\|x_{m+1} - x_m\| \xrightarrow[m \to +\infty]{} 0$ on obtient

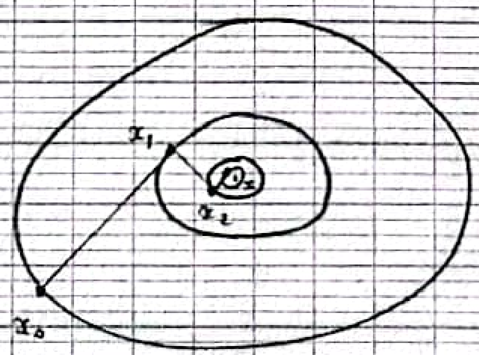
Voir pdf pariot

$$0 \overset{?}{=} \langle \nabla f(x_{\varphi(m)}), \nabla f(x_{\varphi(m)+1}) \rangle \xrightarrow[m \to +\infty]{} \|\nabla f(x)\|^2$$

Ainsi $\nabla f(x) = 0$, ie $x = x_*$.

prop 9 p30 X.G Analyse — Par conséquent $(x_m)_{m \in \mathbb{N}}$ converge vers x_*. (par unicité)

<u>Schéma :</u>



Soit $f : \mathbb{R}^n \to \mathbb{R}$ application continue coercive. Il existe $x \in \mathbb{R}^n$, $f(x) = \inf_{y \in \mathbb{R}^n} f(y)$.

Preuve : Soit $a \in \mathbb{R}^n$, on pose $F = \{x \in \mathbb{R}^n, f(x) \leq f(a)\} = f^{-1}(]-\infty, f(a)])$ fermé car f continue, bornée car f coercive donc F compact. f continue sur F donc atteint son inf.

Si $H_f \in S_n^{++}(\mathbb{R})$ alors f est strictement convexe.

Preuve : Taylor-Lagrange, $\forall x, y \in C$, $\exists z \in]x, y[$

$$f(y) = f(x) + \nabla f(x) \cdot (y-x) + \frac{1}{2}(y-x) \cdot H_f(z)(y-x)$$

Or $\langle (y-x), H_f(z)(y-x) \rangle > 0$ car $H_f \in S_n^{++}(\mathbb{R})$

et donc $f(y) > f(x) + \nabla f(x) \cdot (y-x)$

Soit $y_1, y_2, x \in C$:

$$f(y_1) > f(x) + \nabla f(x) \cdot (y_1 - x) \quad (1)$$

$$f(y_2) > f(x) + \nabla f(x) \cdot (y_2 - x) \quad (2)$$

Soit $\lambda \in]0, 1[$

$(1-\lambda)(1) + \lambda(2) \to (1-\lambda) f(y_1) + \lambda f(y_2) > f(x) + \nabla f(x) \cdot ((1-\lambda) y_1 + \lambda y_2 - x)$

On prend $x = (1-\lambda) y_1 + \lambda y_2$ d'où la stricte convexité de f sur $C = \mathbb{R}^n$.

Remarque : La réciproque est fausse, $f(x) = x^4$ est strictement convexe sur l'axe réel mais $f''(x) = 12x^2$ est nul en $x = 0$.

Théorème : La solution u du système linéaire $Ax = b$ est le vecteur qui réalise le minimum global de la fonctionnelle quadratique f.

Preuve : $\forall x \in \mathbb{R}^n$, $f(x) = \frac{1}{2}\langle A(x-u), x-u \rangle - \frac{1}{2}\langle Au, u \rangle$ et $A \in S_n^{++}(\mathbb{R})$

donc $\forall x \in \mathbb{R}^n \setminus \{0\}$, $f(x) > -\frac{1}{2}\langle Au, u \rangle = \frac{1}{2}\langle Au, u \rangle - \langle b, u \rangle = f(u)$

ce qui entraîne que f atteint son minimum en u et que ce minimum est unique sur $\mathbb{R}^n$.

Réciproquement si f admet un extremum global $x \in \mathbb{R}^n$ alors $\nabla f(x) = 0$ ie $Ax = b$

Théorème : Si $f : \mathbb{R}^n \to \mathbb{R}$ strictement convexe et s'il existe un minimum global, celui-ci est unique.

Preuve : Soit x^* un minimum global de f sur $\mathbb{R}^n$, soit $y \in \mathbb{R}^n \neq x_*$, $\lambda \in]0, 1[$:

$$f(x_*) < f(\underbrace{(1-\lambda) x_* + \lambda y}_{\in \mathbb{R}^n \text{ convexe}}) < (1-\lambda) f(x_*) + \lambda f(y)$$

et $f(x_*) < f(y)$

Lemme : On note $0 < \lambda_1 \leqslant \lambda_2 \leqslant \dots \leqslant \lambda_m$ les valeurs propres de A. Alors :

$$\forall x \in \mathbb{R}^m, \; \langle Ax, x \rangle \geqslant \lambda_1 \|x\|_2^2 .$$

Preuve : La matrice A, symétrique, se diagonalise dans une base orthonormée $\{f_1, \dots, f_m\}$ avec $A f_k = \lambda_k f_k \quad \forall k \in [\![1, m]\!]$. Dans cette base :

$$x = \sum_{k=1}^{m} \alpha_k f_k \qquad \text{et} \quad Ax = \sum_{k=1}^{m} \alpha_k \lambda_k f_k$$

de sorte que $\quad \langle Ax, x \rangle = \sum_{k=1}^{m} \lambda_k \alpha_k^2 \geqslant \lambda_1 \sum_{k=1}^{m} \alpha_k^2 = \lambda_1 \|x\|_2^2 .$

Proposition : Soit (E, d) espace métrique compact et $(x_n) \in E^{\mathbb{N}}$ admettant une et une seule valeur d'adhérence x. Alors (x_n) converge vers x.

Preuve : Supposons le contraire. Alors :

$$\exists \varepsilon > 0, \forall N, \exists n \geqslant N, \quad d(x_n, x) \geqslant \varepsilon$$

On peut donc construire une sous suite $(x_{\varphi(n)})$ de (x_n) telle que $\forall n, \; d(x_{\varphi(n)}, x) \geqslant \varepsilon$.
Comme E est compact, on peut extraire de la sous suite $(x_{\varphi(n)})$ une nouvelle sous suite $(x_{\varphi\circ\psi(n)})$ convergente. Si y est sa limite, on a $d(x, y) \geqslant \varepsilon$ donc $x \neq y$ et y est une valeur d'adhérence de (x_n). C'est absurde, d'où le résultat.