Optimisation et optimisation numérique

Chapitre 3 : Méthodes de Newton et quasi-Newton

Lucie Le Briquer

4 février 2018

Table des matières

1 Introduction **2**

2 Méthodes de quasi-Newton **3**
2.1 Méthode de mise à jour de la métrique 3

3 Gradient conjugué **5**
3.1 Cas quadratique, A définie positive 5

1 Introduction

$$f(x+h) = f(x) + f'(x)h + \frac{1}{2}f''(x)(h,h) + o(|h|^2)$$

$f'(x+h) = 0 = f'(x) + f''(x)h + o(|h|)$
Idée. $h = -f''(x)^{-1}f'(x)$. Problème : inversibilité de $f''(x)$? Point critique ou minimum local ?

Définition 1 ($Q-$convergence)

Soit E un e.v.n., $(x_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$. On note $q_n = \frac{|x_{n+1}-x_*|}{|x_n-x_*|}$ avec la convention $\frac{0}{0} = 0$.

- On dit que (x_n) converge $Q-$linéairement vers x_* si $\overline{\lim_{n\to+\infty}} q_n < 1$
- On dit que (x_n) converge $Q-$superlinéairement vers x_* si $\overline{\lim_{n\to+\infty}} q_n = 0$
- On dit que (x_n) converge $Q-$quadratiquement vers x_* si $q_n = O(|x_n - x_*|)$

Théorème 1

On suppose que f est $\mathcal{C}^2$ elliptique ($f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$). Alors, la méthode de Newton es bien définie et si elle converge, alors elle converge $Q-$superlinéairement. Si de plus f est $\mathcal{C}^3$, alors la convergence, si elle a lieu est $Q-$quadratique.

Preuve.
Si f est $\mathcal{C}^2$ elliptique, il existe $\alpha > 0$ tel que $f''(x)(h,h) \geqslant \alpha|x|^2$. En particulier, $f''(x)$ est inversible $\forall x \in \mathbb{R}^n$ et $|f''(x)^{-1}| < \frac{1}{\alpha}$. Quitte à translater on peut supposer que $x_* = 0$ et on pose $F(x) = x - f''(x)^{-1}f'(x)$. Or $0 = f'(0) = f'(x) - f''(x)x + r(x)$ où $r_(x) = o(|x|)$ si f est $\mathcal{C}^1$ et $= O(|x|^2)$ si f est $\mathcal{C}^3$.

$$f''(x)^{--1}f'(x) - x = -f''(x)^{-1}r(x)$$

□

2 Méthodes de quasi-Newton

Une approximation de l'inverse du hessien :

$$H_k \simeq \underbrace{\nabla^2 f(x_k)^{-1}}_{\text{matrice hessienne}}$$

$$d_k = -H_k \nabla f(x_k)$$

Conditions :

1. H_k est définie positive.

$$\int_0^1 f''(x_k - t(x_{k+1} - x_k))\underbrace{(x_{k+1} - x_k)}_{s_k}dt = f'(x_{k+1}) - f'(x_k)$$

$$\bar{G}_k = \int_0^1 \nabla^2 f(x_k + t(x_{k+1} - x_k))$$

$$\bar{G}_k s_k = \nabla f'(x_{k+1}) - \nabla f'(x_k)$$

2. $H_{k+1}y_k = s_k$ (CQN) Conditions de Quasi-Newton

2.1 Méthode de mise à jour de la métrique

$$H_{k+1} = \underbrace{H_k}_{\text{val. courante}} + \underbrace{B_k}_{\text{correction}}$$

$H_+ = H + B$ avec B de rang faible.
(DFP) Davidan-Fletcher-Powell

$$B = \frac{ss^T}{\langle y, s\rangle}$$

(BFGS) Broyder-Fletcher-Golfarb-Shannno

$$B = -\frac{sy^T H + Hys^T}{\langle y, s\rangle} + \left(1 + \frac{\langle y, Hy\rangle}{\langle y, s\rangle}\right)\frac{ss^T}{\langle y, s\rangle}$$

On vérifie (TD) que :

$$H_+ = \underbrace{\left(I - \frac{sy^T}{\langle y, s\rangle}\right)}_{\pi^T} H \underbrace{\left(I - \frac{ys^Y}{\langle y, s\rangle}\right)}_{\pi} + \frac{ss^T}{\langle y, s\rangle}$$

où $\pi = p_{\mathbb{R}s^T//\mathbb{R}y}$.
On vérifie (TD) que $H_+y = s$, la condition 2 est donc vérifiée.

Théorème 2

Soient $y \neq 0$ et $H > 0$. Alors $H_+ = H+B$ (DFP ou BFGS) est définie positive ssi $\langle y, s\rangle > 0$.

Remarque. Mise à jour de Wolfe.

$$\langle y_k, s_k\rangle = \langle \nabla f(x_{k+1}) - \nabla f(x_k)t_k d_k\rangle = t_k(q'(t_k) - q'(0)) \geqslant t_k(M_2 - 1)q'(0) > 0$$

Convergence ? Ouvert en toute généralité.

Théorème 3

Si $f(x) = \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle + c$ avec A définie positive $\in M_n(\mathbb{R})$. On applique un algorithme quasi-Newton (DFP ou BFGS) ainsi qu'une recherche linéaire exacte. Alors, pour tout $0 \leqslant j < k$ tel que $\nabla f(x_k) \neq 0$, on a :

1. $\langle \nabla f(x_k), s_j\rangle = 0$ (orthogonalité)
2. $H_k y_j = s_j$ (CQN vérifiées)
3. $\langle s_k, s_j\rangle_A = 0$ (les directions d_k sont A conjuguées)

De plus, si $\tau = \inf\{k \geqslant 0 \mid \nabla f(x_k) = 0\}$, alors $\tau \leqslant n$ et si $\tau = n$, $H_n = A^{-1}$.

Preuve.

On peut supposer $x_* = A^{-1}b = 0$ (translation $x \mapsto f(x + x_*)$ sinon) si bien que $\nabla f(x) = Ax$ et $y_j = As_j$. Montrons alors le résultat par récurrence sur $k < \tau$.

- $k = 0$: rien à montrer
- Si $(1), (2), (3)$ vraie au rang k et $\nabla f(x_{k+1}) \neq 0$ montrons qu'elles sont toujours vraies au rang $k + 1$.

  1. Comme on fait une recherche linéaire *exacte*, on a :

  $$\langle \nabla f(x_{k+1}), d_k \rangle = 0 = \langle \nabla f(x_{k+1}), s_k \rangle$$

  car $s_k = t_k d_k$. De plus,

  $$\begin{aligned}\langle \nabla f(x_{k+1}, s_j) \rangle &= \langle \nabla f(x_{j+1}), s_j \rangle + \sum_{h=j+1}^{k} \langle \underbrace{\nabla f(x_{h+1}) - \nabla f(x_h)}_{As_h}, s_j \rangle \\ &= \langle \nabla f(x_{j+1}), s_j \rangle + \sum_{h=j+1}^{k} \langle s_h, s_j \rangle_A \\ &= 0 \qquad \text{par récurrence}\end{aligned}$$

  2. $H_{k+1}y_k = s_k$ est vrai par construction. Étudions le cas DFP.

  $$\begin{aligned}H_{k+1}y_j &= H_k y_j + \frac{s_k s_k^T y_j}{\langle y_k, s_k \rangle} - \frac{H y_k y_k^T H y_j}{\langle s_k, y_k \rangle} \\ &= s_j + \frac{s_k s_k^T y_j}{\langle y_k, s_k \rangle} - \frac{H_k y_k y_k^T H_k y_j}{\langle s_k, y_k \rangle} \\ &\underset{(*)}{=} s_j - \frac{H_k y_k y_k^T H_k y_j}{\langle s_k, y_k \rangle} \underset{(**)}{=} s_j\end{aligned}$$

  $(*)$ car $\langle s_k, y_j \rangle = \langle s_k, As_j \rangle = \langle s_k, s_j \rangle_A = 0$.
  $(**)$ car $y_k^T H_k y_j = \langle H_k y_j, y_k \rangle = \langle s_j, y_k \rangle = -\langle s_j, As_k \rangle = \langle s_j, s_k \rangle_A = 0$

  Donc (2) est vérifiée. Idem pour (BFGS).

  3. $s_{k+1} = t_{k+1}d_{k+1} = -t_{k+1}H_{k+1}\nabla f(x_{k+1})$. Ainsi,

  $$\begin{aligned}\langle s_{k+1}, s_j \rangle_A &= -t_{k+1}\langle H_{k+1}\nabla f(x_{k+1}), \underbrace{As_j}_{y_j} \rangle \\ &= -t_{k+1}\langle \nabla f(x_{k+1}), H_{k+1}y_j \rangle \\ &= -t_{k+1}\langle \nabla f(x_{k+1}), s_j \rangle = 0 \qquad \text{par } (1)\end{aligned}$$

Enfin si $\tau > n - 1$, $(s_j)_{0 \leqslant j < n}$ est un famille $A-$orthogonale de vecteurs non nuls i.e. une base. Comme on a $H_n A s_j = H_n y_j = s_j$ on a $H_n = A^{-1}$. □

Propriété 1

On va supposer que $C = (f \leqslant f(x_0))$ est convexe, f $\mathcal{C}^2$ et que $mI \leqslant \nabla^2 f(x) \leqslant MI$ et que ∇f et $L-$lipschitz. Alors la convergence est quadratique avec Wolfe.

3 Gradient conjugué

Intérêt. Pas de construction d'une approximation de $\nabla^2 f^{-1}$.

$$d_k = -\underbrace{g_k}_{\nabla f(x_k)} + c_{k-1} d_{k-1}$$

où c_{k+1} doit être calculable itérativement.

3.1 Cas quadratique, A définie positive

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

On note $D_k = \mathrm{Vect}(g_0, \dots, d_k)$. Regardons la variété affine $V_k = x_k + D_k$ et prenons comme point suivant $x_k = \mathrm{argmin}_{V_k} f$. On note $\tau = \inf\{k \geqslant 0 \mid \nabla f(x_k) = 0\}$.

Théorème 4

$\forall 0 \leqslant k < \tau$, on a :

1. $\dim D_k = k + 1$
2. $x_{k+1} = x_k + t_k$ où :
$$t_k = -\frac{g_k}{\langle g_k, d_k\rangle_A} = \frac{|g_k|^2}{|d_k|_A^2} > 0$$
et $d_k = -g_k + p_{D_{k-1}}(g_k)$ avec p la projection $A-$orthogonale sur D_{k-1}.
3. Si $k \geqslant 1$,
$$d_k = -g_k + c_{k-1} d_{k-1}$$
avec $c_{k-1} = \frac{\langle g_k, d_{k-1}\rangle_A}{|d_{k-1}|_A^2} = \frac{|g_k|^2}{|g_{k-1}|^2}$.
4. De plus, $\forall 0 \leqslant i < j \leqslant k$,
$$\langle d_i, d_j\rangle_A = \langle g_i, g_j\rangle = 0$$

Preuve.

Supposons que c'est vrai pour tout l $0 \leqslant l < k < \tau$ et montrons que c'est vrai en k.

Par construction, $x_{k+1} = p_{V_k}(x_*)$ (projection orthogonale sur la métrique A). Comme $p_{V_{k-1}} \circ p_{V_k} = p_{V_{k-1}}$, on a $x_k = p_{V_{k-1}}(x_*) = p_{V_{k-1}}(x_{k+1})$ et :

$$s_k = x_{k+1} - x_k = x_{k+1} - p_{V_{k-1}}(x_{k+1}) = p_{D_{k-1}^{\perp_A}}(x_{k+1}) \in D_{k-1}^{\perp_A}$$

- $s_k \neq 0$, en effet :
$$\langle \nabla f(x_{k+1}), w\rangle = 0 \quad \forall w \in D_k \quad \text{(CN1)}$$
et si $s_k = 0$ alors $x_{k+1} = x_k$ et $\nabla f(x_{k+1}) = \nabla f(x_k)$ et $\nabla f(x_k) \in D_k^{\perp_A} \cap D_k = \{0\}$, absurde.
- Montrons que :
$$\mathbb{R}s_k \overset{\perp_A}{\oplus} D_{k-1} = \mathbb{R}g_k \overset{\perp}{\oplus} D_{k-1}$$

En effet, comme $s_k \in D_{k-1}^{\perp_A} \backslash \{0\}$, on a ue $D_{k-1} \not\subseteq \mathbb{R}_{s_k} \overset{\perp_A}{\oplus} D_{k-1} \subset D_k$ donc $\dim D_k \leqslant k+1$. Or par récurrence $\dim D_{k-1} = k$, ainsi $\dim(D_k) = k+1$ et $\mathbb{R}s_k \overset{\perp_A}{\oplus} D_{k-1} = D_k$. Enfin, $g_k \in D_{k-1}^{\perp}$ (CN1), d'où $\mathbb{R}g_k \overset{\perp}{\oplus} D_{k-1} = D_k$.

- Soit $d_k = -p_{D_{k-1}^{\perp_A}}(g_k) = -g_k + p_{D_{k-1}}(g_k)$. On a $D_k = \mathrm{R}d_k \overset{\perp_A}{\oplus} D_{k-1}$ et $d_0, \ldots, d_k$ est obtenue par orthogonalisation de $g_0, \ldots, g_k$ pour la métrique A. En particulier, $d_k \in D_{k-1}^{\perp_A}$, et les directions d_j sont $A-$orthogonales. Et comme $g_k \in D_{k-1}^{\perp}$, on a (4).

- Montrons que $\exists t_k > 0$ tel que $s_k = x_{k-1} - x_k = t_k d_k$. En effet $D_k \cap D_{k-1}^{\perp_A}$ est une droite contenant s_k et d_k ($\neq 0$). $\exists t_k \in \mathbb{R}$ tel que $s_k = t_k d_k$. Or :

$$\begin{aligned} 0 = \langle g_{k+1}, g_k \rangle &= \langle g_{k+1} - g_k \rangle + \langle g_k, g_k \rangle \\ &= \langle A s_k, g_k \rangle - |g_k|^2 \\ &= t_k \langle A d_k, g_k \rangle + |g_k|^2 \end{aligned}$$

avec $t_k = -\frac{|g_k|^2}{\langle d_k, g_k \rangle_A}$.

- Reste le calcul explicite de d_k.

$$d_k = -g_k + p_{D_{k-1}}(g_k)$$

Or pour $l < k-1$,

$$\langle g_k, t_l d_l \rangle_A = \langle g_k, t_l A d_l \rangle = \langle g_k, A s_l \rangle = \langle g_k, g_{l+1} - g_l \rangle = 0 \quad \text{(réc)}$$

d'où $p_{D_{k-1}}(g_k) = c_{k-1} d_{k-1}$. Or $\langle d_k, d_{k-1} \rangle_A = 0$. Ainsi $\langle -g_k + c_{k-1} d_{k-1}, d_{k-1} \rangle_A = 0$. Finalement,

$$c_k = \frac{\langle g_k, d_{k-1} \rangle_A}{\langle d_{k-1}, d_{k-1} \rangle_A}$$

L'autre forme s'obtient en remarquant qu'on peut écrire $d_{k-1} = \frac{s_{k-1}}{t_{k-1}}$:

$$c_k = \frac{\langle g_k, s_{k-1} \rangle_A}{\langle d_{k-1}, s_{k-1} \rangle_A} = \frac{\langle g_k, A s_{k-1} \rangle}{\langle d_{k-1}, A s_{k-1} \rangle} = \frac{\langle g_k, g_k - g_{k-1} \rangle}{\langle d_{k-1}, g_k - g_{k-1} \rangle} = \frac{\langle g_k, g_k \rangle}{\langle d_{k-1}, -g_{k-1} \rangle}$$

Or, $d_{k-1} = -g_{k-1} + P_{D_{k-2}}(g_{k-1})$ donc $\langle d_{k-1}, -g_{k-1} \rangle = |g_{k-1}|^2$

□