Cours d'analyse numérique de licence de mathématiques

Roland Masson

2013

Solveurs non linéaires: plan

- Rappels de calculs différentiels pour des fonctions vectorielles de variable vectorielle.
- Algorithme de Newton pour résoudre $f(x) = 0$ avec $f \in C^1(U, \mathbb{R}^n)$ avec U ouvert de $\mathbb{R}^n$.
- Convergence quadratique de l'algorithme de Newton pour $f \in C^2(U, \mathbb{R}^n)$.

Rappels sur les fonctions vectorielles: différentielles

- Application linéaire tangente: soit U un ouvert de $\mathbb{R}^n$ et $f : U \to \mathbb{R}^m$, on dit que f est différentiable au point $x \in U$ ssi il existe une application linéaire notée $f'(x) \in \mathcal{L}\Big(\mathbb{R}^n, \mathbb{R}^m\Big)$ telle que

$$lim_{h \neq 0 \to 0} \frac{\|f(x+h) - f(x) - f'(x)(h)\|}{\|h\|} = 0$$

- Si f est différentiable au point $x \in U$ alors f est continue en x.
- Différentielle: si f est différentiable pour tout $x \in U$, on note $x \to f'(x)$ l'application de U dans $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ appelée différentielle de f

Rappels sur les fonctions vectorielles: différentielles

Exemples:

- pour $n = 1$ on retrouve la dérivée au sens classique $f'(x) \in \mathbb{R}^m$
- pour $m = 1$, $f'(x)$ est une forme linéaire de $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$
  - Gradient $\nabla f(x) \in \mathbb{R}^n$: grace à la structure euclidienne de $\mathbb{R}^n$, il existe un unique vecteur $\nabla f(x)$ appelé gradient de f au point x tel que $(\nabla f(x), v) = f'(x)(v)$ pour tout $v \in \mathbb{R}^n$ et dont la définition dépend du produit scalaire.

Méthodes de Newton pour résoudre $f(x) = 0$

Soit

$$f : U \subset \mathbb{R}^n \to \mathbb{R}^n,$$

telle qu'il existe $\bar{x} \in U$ avec $f(\bar{x}) = 0$.

Etant donné $x^1 \in U$, la méthode de Newton calcule une suite x^k, $k = 2, \cdots$, par linéarisations successives de f ie on approche l'équation $f(x^{k+1}) = 0$ par sa linéarisation au voisinage de x^k:

$$f'(x^k)(x^{k+1} - x^k) = -f(x^k).$$

- A chaque itération il faudra donc calculer la dérivée $f'(x^k)$ et résoudre un système linéaire.
- L'analyse de la méthode de Newton doit donner des conditions suffisantes sur f et sur x^1 pour que $f'(x^k)$ soit inversible pour tout $k = 1, \cdots$, et pour que la suite x^k converge vers $\bar{x}$.

Rappels sur les fonctions vectorielles $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$

Dérivées partielles: $\frac{\partial f}{\partial x_j}(x)$ est la dérivée (si elle existe) de f selon la direction e_j au point x ie

$$\frac{\partial f}{\partial x_j}(x) = lim_{h_j \to 0} \frac{f(x + h_j e_j) - f(x)}{h_j}.$$

- Si f est différentiable au point x alors elle admet des dérivées partielles au point x et

  $$f'(x)(h) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) h_j \text{ pour tout } h \in \mathbb{R}^n.$$

  Matrice représentant l'application linéaire $f'(x)$:

  $$J(x) \in \mathcal{M}_{m,n} \text{ telle que } J_{i,j}(x) = \frac{\partial f_i}{\partial x_j}(x), i = 1, \cdots, m; j = 1, \cdots, n$$

  est appelée la matrice Jacobienne de f au point x.

Rappels sur les fonctions vectorielles $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$

- La réciproque n'est pas vraie: si f admet des dérivées partielles en x, elle n'est pas nécéssairement différentiable au point x.
  - Exemple: $f(x_1, x_2) = \frac{x_1 x_2}{\sqrt{x_1^2 + x_2^2}}$ a des dérivées partielles nulles en 0 mais n'est pas différentiable en 0.
- Si f admet des dérivées partielles continues au point x pour tout $i = 1 \cdots, n$ alors f est différentiable au point x
- f est continuement différentiable sur U (ie f' existe et est continue sur U) ssi f admet des dérivées partielles continues sur U. On dit que f est $C^1\Big(U, \mathbb{R}^n\Big)$.

Rappels sur les fonctions vectorielles: formule des accroissements finis

- Rappel dans le cas $n = m = 1$ (théorème des accroissements finis):
  $f : [a, b] \subset \mathbb{R} \to \mathbb{R}^m$: si f est continue sur $[a, b]$ et différentiable sur (a, b) alors il existe $c \in (a, b)$ tel que

  $$f(b) - f(a) = f'(c)(b - a).$$

- Extension au cas $m = 1$, $n \geq 1$
  $[a, b] = \{(1 - t)a + tb, t \in [0, 1]\} \subset U \subset \mathbb{R}^n$,
  $(a, b) = \{(1 - t)a + tb, t \in (0, 1)\}$.
  Si f est différentiable sur U, alors il existe $c \in (a, b)$ tel que

  $$f(b) - f(a) = f'(c)(b - a) = (\nabla f(c), b - a) \text{ aussi noté } \nabla f(c) \cdot (b - a)$$

  Preuve: on applique le théorème des accroissements finis à $\varphi(t) = f((1 - t)a + tb)$

Rappels sur les fonctions vectorielles: formule des accroissements finis

- Cas général: $n \geq 1$, $m \geq 1$: $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ différentiable sur U et $[a, b] \subset U$, alors

$$\|f(b) - f(a)\| \leq \sup_{x \in (a,b)} \|f'(x)\| \|b - a\|.$$

- Preuve:
  - Soit $\phi(t) = f((1-t)a + tb)$, on a $\phi'(t) = f'((1-t)a + tb)(b-a)$ et

$$\|\phi'(t)\| \leq \sup_{x \in (a,b)} \|f'(x)\| \|b - a\|,$$

on conclut par $f(b) - f(a) = \displaystyle\int_0^1 \phi'(t)dt$, d'où

$$\|f(b) - f(a)\| \leq \int_0^1 \|\phi'(t)\| dt \leq \sup_{x \in (a,b)} \|f'(x)\| \|b - a\|.$$

Rappels sur les fonctions vectorielles: différentielle d'ordre 2

Soit $f \in C^1(U, \mathbb{R}^m)$ avec U ouvert de $\mathbb{R}^n$. On a $f' \in C^0\Big(U, \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)\Big)$. Si f' est continuement différentiable sur U on dit que $f \in C^2(U, \mathbb{R}^m)$ et on note f'' sa différentielle appelée différentielle seconde de f. La différentielle seconde $f''(x)$ est un élément de $\mathcal{L}\Big(\mathbb{R}^n; \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)\Big)$ isomorphe à l'ensemble des applications bilinéaires $\mathcal{L}\Big(\mathbb{R}^n, \mathbb{R}^n; \mathbb{R}^m\Big)$.

Dérivées partielles d'ordre 2: si f est différentiable d'ordre 2 en x alors elle admet des dérivées partielles d'ordre 2 au point x notées $\dfrac{\partial^2 f(x)}{\partial x_i \partial x_j}$ avec

$$f''(x)(h,k) = \sum_{i,j=1}^{n} \frac{\partial^2 f(x)}{\partial x_i \partial x_j} h_i k_j \text{ pour tous } h, k \in \mathbb{R}^n$$

On a alors le théorème de Schwarz: $f''(x)$ est symétrique au sens où

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{\partial^2 f(x)}{\partial x_j \partial x_i} \text{ pour tous } i, j = 1, \cdots, n.$$

Dans le cas $m = 1$, on appelle $H(x) \in \mathcal{M}_{n,n}$ la matrice dite Hessienne représentant la forme bilinéaire $f''(x)$ dans la base canonique.

Rappels sur les fonctions vectorielles: formule de Taylor à l'ordre 2 dans le cas $f \in C^2\Big(U, \mathbb{R}^m\Big)$

Soit $f \in C^2(U, \mathbb{R}^m)$ avec U ouvert de $\mathbb{R}^n$ et $a, b \in U$ tels que $[a, b] \subset U$. Alors on a

$$\|f(b) - f(a) - f'(a)(b-a)\| \leq \frac{1}{2} \sup_{x \in (a,b)} \|f''(x)\| \|b-a\|^2,$$

avec

$$\|f''(x)\| = \sup_{u,v \neq 0 \in \mathbb{R}^n} \frac{\|f''(x)(u,v)\|}{\|u\|\|v\|}.$$

Rappels sur les fonctions vectorielles: formule de Taylor à l'ordre 2 dans le cas $f \in C^2\Big(U, \mathbb{R}^m\Big)$

Preuve: soit

$$\varphi(t) = f(x + t(y-x)) - f(x) - t\ f'(x)(y-x),$$

on a $\varphi'(t) = \Big(f'(x+t(y-x)) - f'(x)\Big)(y-x)$ et donc

$$\varphi(1) = \varphi(1) - \varphi(0) = \int_0^1 \Big(f'(x+t(y-x)) - f'(x)\Big)(y-x)dt$$

$$\|\varphi(1)\| = \|f(y) - f(x) - f'(x)(y-x)\| \leq \int_0^1 \|f'(x+t(y-x)) - f'(x)\| \|y-x\| dt$$

on conclut par la formule des accroissements finis sur $f' \in C^1(U, \mathbb{R}^m)$:

$$\|f'(x+t(y-x)) - f'(x)\| \leq \sup_{z \in (x+t(y-x), y)} \|f''(z)\|\ \|y-x\|\ t$$

Algorithme de Newton

Soit

$$f \in C^1(U, \mathbb{R}^n), \; U \text{ ouvert de } \mathbb{R}^n$$

telle qu'il existe $\bar{x} \in U$ avec $f(\bar{x}) = 0$.

Etant donné $x^1 \in U$, la méthode de Newton calcule une suite x^k, $k = 2, \cdots$, par linéarisations successives de f ie on approche l'équation $f(x^{k+1}) = 0$ par sa linéarisation au voisinage de x^k:

$$f'(x^k)(x^{k+1} - x^k) = -f(x^k).$$

- A chaque itération il faudra donc calculer la dérivée $f'(x^k)$ et résoudre un système linéaire.
- L'analyse de la méthode de Newton doit donner des conditions suffisantes sur f et sur x^1 pour que $f'(x^k)$ soit inversible pour tout $k = 1, \cdots$, et pour que la suite x^k converge vers $\bar{x}$.

Convergence quadratique de l'algorithme de Newton

Soit $f \in C^2(U, \mathbb{R}^n)$ avec U ouvert de $\mathbb{R}^n$ et $\bar{x} \in U$ tel que $f(\bar{x}) = 0$. On suppose que $f'(\bar{x})$ est inversible. Alors il existe $\alpha > 0$ et $\beta > 0$ tels que

- $B(\bar{x}, \alpha) = \{x \mid \|x - \bar{x}\| < \alpha\} \subset U$
- Si $x^1 \in B(\bar{x}, \alpha)$ alors la suite $x^k, k \in \mathbb{N}$ est bien définie et $x^k \in B(\bar{x}, \alpha)$ pour tout $k \in \mathbb{N}$
- Si $x^1 \in B(\bar{x}, \alpha)$ alors la suite $x^k, k \in \mathbb{N}$ converge vers $\bar{x}$ et

$$\|x^{k+1} - \bar{x}\| \leq \beta \|x^k - \bar{x}\|^2 \text{ (convergence quadratique)}$$

Convergence quadratique de l'algorithme de Newton: Preuve

On commence par montrer le lemme suivant:
Soit $f \in C^2(U, \mathbb{R}^n)$ avec U ouvert de $\mathbb{R}^n$ et $\bar{x} \in U$ tel que $f'(\bar{x})$ est inversible.
Alors il existe $\gamma > 0,\ C_1 > 0,\ C_2 > 0$ tels que $B(\bar{x}, \gamma) \subset U$ et

- $f'(x)$ inversible et $\|(f'(x))^{-1}\| \leq C_1$ pour tous $x \in B(\bar{x}, \gamma)$
- $\|f(y) - f(x) - f'(x)(y - x)\| \leq C_2 \|y - x\|^2$ pour tous $(x, y) \in B(\bar{x}, \gamma)$

Preuve: le point 1 est une application de $\|(I - A)^{-1}\| \leq \frac{1}{(1 - \|A\|)}$ pour $\|A\| < 1$ et le point 2 résulte directement de la formule de Taylor d'ordre 2.

Convergence quadratique de l'algorithme de Newton: Preuve suite

Soit $\alpha = \min\left(\gamma, \frac{1}{C_1 C_2}\right)$. On suppose que $x^k \in B(\bar{x}, \alpha)$ (et donc $f'(x^k)$ inversible), on va montrer que ceci implique $x^{k+1} \in B(\bar{x}, \alpha)$.

Comme $f'(x^k)(x^{k+1} - x^k) + f(x^k) = 0$, on a

$$f'(x^k)(x^{k+1} - \bar{x}) = f(\bar{x}) - f(x^k) - f'(x^k)(\bar{x} - x^k)$$

d'où

$$\|x^{k+1} - \bar{x}\| \leq \|(f'(x^k))^{-1}\| C_2 \|x^k - \bar{x}\|^2 \leq C_1 C_2 \|x^k - \bar{x}\|^2 \leq \alpha.$$

On a donc $x^{k+1} \in B(\bar{x}, \alpha)$ et on donc a montré par récurrence que c'est vrai pour tout $k \in \mathbb{N}$ si $x^1 \in B(\bar{x}, \alpha)$. On a ensuite

$$\|x^{k+1} - \bar{x}\| \leq (C_1 C_2)^{2^k - 1} \|x^1 - \bar{x}\|^{2^k}$$

Par ailleurs comme $x^1 \in B(\bar{x}, \alpha)$, on a $\|x^1 - \bar{x}\| < \frac{1}{C_1 C_2}$, d'où la convergence de la suite vers $\bar{x}$. Elle est quadratique avec $\beta = C_1 C_2$.

Variantes de l'algorithme de Newton: Inexact Newton

Si le système linéaire est résolu avec une méthode itérative on veut ajuster le critère d'arrêt du solveur linéaire pour préserver la convergence quadratique de l'algorithme de Newton à moindre coût.
Ceci revient à résoudre le système linéaire de façon approchée avec un résidu r^k

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

tel que

$$\|r^k\| \leq \eta_k \|f(x^k)\|.$$

Différentes stratégies existent pour ajuster η_k de façon à préserver la convergence quadratique sans trop résoudre le système linéaire: par exemple

$$\eta_k = \min\left(\eta_{max}, \frac{|\|f(x^k)\| - \|f(x^{k-1}) + f'(x^{k-1})(x^k - x^{k-1})\||}{\|f(x^{k-1})\|}\right),$$

avec $\eta_{max} = 0.1$. Ce choix prend en compte la fiabilité de l'approximation tangentielle de f.

Variantes de l'algorithme de Newton: Quasi Newton

On n'a pas toujours en pratique accès au calcul exact de la Jacobienne de f. Il existe des méthodes itératives pour l'approcher comme par exemple l'algorithme de Broyden suivant:

- Initialisation: x^0, $x^1 \in U$, $B^0 \in \mathcal{M}_n$
- Itérations
  - On pose $\delta^k = x^k - x^{k-1}$ et $y^k = f(x^k) - f(x^{k-1})$
  - Mise à jour de rang 1 de la Jacobienne approchée:

$$B^k = B^{k-1} + \left(\frac{y^k - B^{k-1}\delta^k}{(\delta^k)^t \delta^k}\right)(\delta^k)^t$$

  - On résoud le système linéaire $B^k \left(x^{k+1} - x^k\right) = -f(x^k)$

La correction de rang 1 de la Jacobienne approchée est construite pour vérifier la condition dite de la sécante:

$$B^k(x^k - x^{k-1}) = f(x^k) - f(x^{k-1}).$$