

Université de Marseille
Licence de Mathématiques, 3ème année, analyse numérique et optimisation
Détails sur la preuve de Newton

Quelques rappels

Soient E et F deux espaces vectoriels normés et f une application de classe C^1 de E dans F .

Soit $x \in E$. La dérivée (ou différentielle) de f au point x est une application linéaire continue de E dans F (on rappelle aussi que si E est de dimension finie, une application linéaire de E dans F est nécessairement continue) et pour tout $h \in E$,

$$f(x+h) = f(x) + Df(x)(h) + \varepsilon(h)\|h\|_E,$$

avec $\lim_{h \rightarrow 0} \varepsilon(h) = 0$. (Noter que $\varepsilon(h) \in F$.)

Si $E = \mathbb{R}^n$ et $F = \mathbb{R}^p$ (avec $n \geq 1, p \geq 1$). L'application linéaire $Df(x)$ est représentée par une matrice notée $J_f(x)$ (appelée matrice jacobienne de f au point x), élément de $\mathcal{M}_{p,n}(\mathbb{R})$.

Pour tout $h \in \mathbb{R}^n$, $Df(x)(h) = J_f(x)h$.

Dans le cas d'espaces généraux E et F de dimension finie, il est possible de choisir des bases de E et F et de représenter alors $Df(x)$ par une matrice jacobienne. Cette matrice jacobienne dépend bien sûr du choix des bases. Dans le cas $E = \mathbb{R}^n, F = \mathbb{R}^p$ la matrice jacobienne $J_f(x)$ correspond à choisir les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^p$.

Petit calcul liminaire

On prend ici une application g de classe C^1 de $\mathbb{R}^n$ dans $\mathbb{R}^n$ et une application M de classe C^1 de $\mathbb{R}^n$ dans $\mathcal{M}_n(\mathbb{R})$.

On pose $\phi(x) = M(x)g(x)$ (l'application ϕ est donc une application de classe C^1 de $\mathbb{R}^n$ dans $\mathbb{R}^n$).

On munit $\mathbb{R}^n$ d'une norme notée $\|\cdot\|$ et $\mathcal{M}_n(\mathbb{R})$ de la norme induite, notée aussi $\|\cdot\|$.

Soit $x \in \mathbb{R}^n$. On cherche à calculer la matrice jacobienne de ϕ au point x . Pour cela, on utilise les formules suivantes, pour tout $h \in \mathbb{R}^n$,

$$g(x+h) = g(x) + J_g(x)h + \varepsilon_1(h)\|h\|,$$

$$M(x+h) = M(x) + DM(x)(h) + \varepsilon_2(h)\|h\|,$$

avec $\lim_{h \rightarrow 0} \varepsilon_i(h) = 0, i = 1, 2$. Noter que $DM(x)$ est une application linéaire (et donc continue) de $\mathbb{R}^n$ dans $\mathcal{M}_n(\mathbb{R})$.

De ces deux formules on déduit

$$\begin{aligned} \phi(x+h) &= M(x+h)g(x+h) = (M(x) + DM(x)(h) + \varepsilon_2(h)\|h\|)(g(x) + J_g(x)h + \varepsilon_1(h)\|h\|) = \\ &\quad \phi(x) + DM(x)(h)g(x) + M(x)J_g(x)h + R(h), \end{aligned}$$

avec

$$R(h) = DM(x)(h)J_g(x)h + (M(x) + DM(x)(h))\varepsilon_1(h)\|h\| + (g(x) + J_g(x)h)\varepsilon_2(h)\|h\| + \varepsilon_2(h)\|h\|\varepsilon_1(h)\|h\|.$$

On remarque que $R(h) = \varepsilon(h)\|h\|$ avec $\lim_{h \rightarrow 0} \varepsilon(h) = 0$. Ceci est dû en particulier au fait que

$$\|DM(x)(h)J_g(x)h\| \leq \|DM(x)(h)\|\|J_g(x)h\| \leq \|DM(x)\|\|h\|\|J_g(x)\|\|h\|,$$

où on a aussi noté $\|\cdot\|$ la norme sur $\mathcal{L}(\mathbb{R}^n, \mathcal{M}_n(\mathbb{R}))$ induite par les normes sur $\mathbb{R}^n$ et $\mathcal{M}_n(\mathbb{R})$. Ceci montre que (pour tout $h \in \mathbb{R}^n$)

$$J_\phi(x)h = DM(x)(h)g(x) + M(x)J_g(x)h.$$

Démonstration de la convergence de l'algorithme de Newton

Pour une fonction g de classe C^1 de $\mathbb{R}^n$ dans $\mathbb{R}^n$, l'algorithme de Newton s'écrit

Initialisation $x^{(0)} \in \mathbb{R}^n$.

Itérations $J_g(x^{(k)})(x^{(k+1)} - x^{(k)}) = -g(x^{(k)})$, pour tout $k \geq 0$.

Théorème 1. Soient g de classe C^3 de $\mathbb{R}^n$ dans $\mathbb{R}^n$ et $\bar{x} \in \mathbb{R}^n$ tel que $g(\bar{x}) = 0$. On suppose que $J_g(\bar{x})$ est inversible. Alors, il existe $\alpha > 0$ tel que, si $x^{(0)} \in B(\bar{x}, \alpha) = \{y \in \mathbb{R}^n, \|y - \bar{x}\| \leq \alpha\}$,

1. l'algorithme de Newton définit bien une suite $(x^{(k)})_{k \in \mathbb{N}}$ (plus précisément, pour chaque k , la matrice $J_g(x^{(k)})$ est inversible),
2. la suite $(x^{(k)})_{k \in \mathbb{N}}$ converge vers $\bar{x}$,
3. la suite $(x^{(k)})_{k \in \mathbb{N}}$ converge vers $\bar{x}$ de manière au moins quadratique.

Démonstration. Comme $J_g(\bar{x})$ est inversible et que l'ensemble des matrices inversibles est un ouvert, il existe $\alpha_0 > 0$ tel que $J_g(x)$ est inversible pour tout $x \in B(\bar{x}, \alpha_0)$.

Pour $x \in B(\bar{x}, \alpha_0)$, on pose $f(x) = x - (J_g(x))^{-1}(g(x))$. L'algorithme de Newton pour g est alors identique à l'algorithme du point fixe de contraction pour f , c'est-à-dire $x^{(k+1)} = f(x^{(k)})$.

Pour montrer les points 1-2 du théorème, il suffit donc de démontrer qu'il existe $\alpha > 0$ tel que f est une application strictement contractante de $B(\bar{x}, \alpha)$ dans $B(\bar{x}, \alpha)$.

Pour démontrer l'existence d'un tel α , on utilise le petit calcul liminaire. Il donne, en notant $M(x) = (J_g(x))^{-1}$ (qui est de classe C^2 car g est de classe C^3), pour tout $x \in B(\bar{x}, \alpha_0)$,

$$J_f(x)h = h - DM(x)(h)g(x) - M(x)J_g(x)h, \text{ pour tout } h \in \mathbb{R}^n.$$

Comme $g(\bar{x}) = 0$ et $M(\bar{x}) = (J_g(\bar{x}))^{-1}$, on a donc $J_f(\bar{x})h = h - h = 0$ pour tout h , c'est-à-dire $J_f(\bar{x}) = 0$.

Comme J_f est continue (et même de classe C^1), il existe $\alpha \in]0, \alpha_0[$ tel que $\|J_f(z)\| \leq 1/2$ pour tout $z \in B(\bar{x}, \alpha)$.

Ceci donne bien les points 1-2. En effet, si $x \in B(\bar{x}, \alpha)$, par le théorème des accroissements finis,

$$\|f(x) - \bar{x}\| = \|f(x) - f(\bar{x})\| \leq \sup_{z \in I(x, \bar{x})} \|J_f(z)\| \|x - \bar{x}\| \leq \frac{\alpha}{2} < \alpha, \quad (1)$$

car $I(x, \bar{x}) = \{tx + (1-t)\bar{x}, t \in]0, 1[\} \subset B(\bar{x}, \alpha)$. Ceci montre que f est une application de $B(\bar{x}, \alpha)$ dans lui-même. Puis si $x, y \in B(\bar{x}, \alpha)$

$$\|f(x) - f(y)\| \leq \sup_{z \in I(x, y)} \|J_f(z)\| \|x - y\| \leq \frac{1}{2} \|x - y\|,$$

car $I(x, y) = \{tx + (1-t)y, t \in]0, 1[\} \subset B(\bar{x}, \alpha)$. Ceci montre que f est strictement contractante.

Il reste à montrer le point 3 du théorème. Pour cela on précise l'inégalité (1) en exploitant le caractère C^1 de J_f . On note

$$\beta = \max_{c \in B(\bar{x}, \alpha)} \|DJ_f(c)\|.$$

Si $z \in B(\bar{x}, \alpha)$, comme $J_f(\bar{x}) = 0$,

$$\|J_f(z)\| = \|J_f(z) - J_f(\bar{x})\| \leq \sup_{c \in I(z, \bar{x})} \|DJ_f(c)\| \|z - \bar{x}\| \leq \beta \|z - \bar{x}\|.$$

L'inégalité (1) donne alors pour tout $x \in B(\bar{x}, \alpha)$ (comme $\|z - \bar{x}\| \leq \|x - \bar{x}\|$ si $z \in I(x, \bar{x})$),

$$\|f(x) - \bar{x}\| \leq \beta \|x - \bar{x}\|^2.$$

Cette inégalité appliquée à $x = x^{(k)}$ donne $\|x^{(k+1)} - \bar{x}\| \leq \beta \|x^{(k)} - \bar{x}\|^2$ et donc la convergence au moins quadratique de l'algorithme de Newton.

□