Cours L2 Résolution numérique des systèmes d'équations linéaires et non linéaires

Roland Masson

Année 2019-2020

Chapitre II: Résolution numérique des systèmes non linéaires

Définitions

Zéro d'une fonction: soit $\mathbf{f} \in C^0(U, \mathbb{R}^n)$ où U est un ouvert de $\mathbb{R}^n$. On dit que $\bar{\mathbf{x}}$ est un zéro de $\mathbf{f}$ ssi $\mathbf{f}(\bar{\mathbf{x}}) = 0$.

Point fixe d'une fonction: soit $\mathbf{g} \in C^0(U, \mathbb{R}^n)$ où U est un ouvert de $\mathbb{R}^n$. On dit que $\bar{\mathbf{x}}$ est un point fixe de $\mathbf{g}$ ssi $\mathbf{g}(\bar{\mathbf{x}}) = \bar{\mathbf{x}}$.

Définitions

Algorithme du point fixe: soit $\mathbf{g} \in C^0(U, \mathbb{R}^n)$ avec U ouvert de $\mathbb{R}^n$, l'algorithme du point fixe est la suite récurrente définie par $\mathbf{x}^{(0)} \in U$ et

$$\mathbf{x}^{(k+1)} = \mathbf{g}(\mathbf{x}^{(k)}) \text{ pour tout } k \in \mathbb{N}.$$

Propriété: si $\lim_{k\to+\infty} \mathbf{x}^{(k)} = \bar{\mathbf{x}} \in U$, alors $\mathbf{g}(\bar{\mathbf{x}}) = \bar{\mathbf{x}}$.

Preuve: par continuité de $\mathbf{g}$ sur U,

$$\bar{\mathbf{x}} = \lim_{k\to+\infty} \mathbf{x}^{(k+1)} = \lim_{k\to+\infty} \mathbf{g}(\mathbf{x}^{(k)}) = \mathbf{g}(\lim_{k\to+\infty} \mathbf{x}^{(k)}) = \mathbf{g}(\bar{\mathbf{x}}).$$

Définitions

Lien entre zéro et point fixe d'une fonction: soient $\mathbf{f} \in C^0(U, \mathbb{R}^n)$ avec U ouvert de $\mathbb{R}^n$, et $\bar{\mathbf{x}} \in U$ un zéro de $\mathbf{f}$. On considère une fonction $M \in C^0(U, \mathcal{M}_n(\mathbb{R}))$, alors $\bar{\mathbf{x}}$ est un point fixe de

$$\mathbf{g}(\mathbf{x}) = \mathbf{x} - M(\mathbf{x})\mathbf{f}(\mathbf{x})$$

Exemple: **l'algorithme de Newton** est l'algorithme de point fixe obtenu pour

$$\mathbf{g}(\mathbf{x}) = \mathbf{x} - D\mathbf{f}(\mathbf{x})^{-1}\mathbf{f}(\mathbf{x}),$$

il s'écrit:

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - D\mathbf{f}(\mathbf{x}^{(k)})^{-1}\mathbf{f}(\mathbf{x}^{(k)}) \text{ pour tout } k \in \mathbb{N}.$$

Définitions

On considère une suite $\mathbf{x}^{(k)} \in \mathbb{R}^n$, $k \in \mathbb{N}$ telle que $\lim_{k\to+\infty} \mathbf{x}^{(k)} = \bar{\mathbf{x}}$.

Convergence linéaire d'une suite: On dit que la suite converge linéairement (ou à l'ordre 1) ssi il existe $\beta < 1$ tel que

$$\|\mathbf{x}^{(k+1)} - \bar{\mathbf{x}}\| \leq \beta \|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\| \text{ pour tout } k \in \mathbb{N}.$$

Convergence quadratique d'une suite: On dit que la suite converge quadratiquement (ou à l'ordre 2) ssi il existe $\gamma \geq 0$ tel que

$$\|\mathbf{x}^{(k+1)} - \bar{\mathbf{x}}\| \leq \gamma \|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\|^2 \text{ pour tout } k \in \mathbb{N}.$$

Objectifs

Etablir des conditions suffisantes de convergence des algorithmes de point fixe (et de Newton) et évaluer leur ordre de convergence.

On étudiera la **convergence "locale"** des algorithmes de point fixe au sens où:

- le point fixe $\bar{\mathbf{x}}$ est supposé exister,
- le point de départ $\mathbf{x}^{(0)}$ sera supposé dans un voisinage de $\bar{\mathbf{x}}$.

Chapitre II: Résolution numérique des systèmes non linéaires

Section 1: Algorithmes de point fixe et de Newton en dimension 1

Algorithme de point fixe

Soit

$$g \in C^0(U, \mathbb{R}), \;\; U \text{ ouvert de } \mathbb{R}$$

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

Etant donné $x^{(0)} \in U$, la méthode du point fixe est définie par la suite récurrente $x^{(k)}$, $k \in \mathbb{N}$ telle que

$$x^{(k+1)} = g(x^{(k)}).$$

Convergence linéaire locale de l'algorithme de point fixe

Soit $g \in C^1(U, \mathbb{R})$ avec U ouvert de $\mathbb{R}$ et soit $\bar{x} \in U$ tel que $g(\bar{x}) = \bar{x}$. On suppose que $|g'(\bar{x})| < 1$. Alors il existe $\alpha > 0$ et $\beta < 1$ tels que si

$$x^{(0)} \in I_\alpha =]\bar{x} - \alpha, \bar{x} + \alpha[,$$

on a

- $I_\alpha \subset U$,
- $x^{(k)} \in I_\alpha$ pour tout $k \in \mathbb{N}$,
- $|x^{(k+1)} - \bar{x}| \leq \beta |x^{(k)} - \bar{x}|$ pour tout $k \in \mathbb{N}$,
- $|x^{(k)} - \bar{x}| \leq \beta^k \, |x^{(0)} - \bar{x}|$ pour tout $k \in \mathbb{N}$,
- $\lim_{k \to +\infty} x^{(k)} = \bar{x}$.
- $\displaystyle \lim_{k \to +\infty} \frac{x^{(k+1)} - \bar{x}}{x^{(k)} - \bar{x}} = g'(\bar{x})$.

Convergence linéaire locale de l'algorithme de point fixe: preuve

- D'après la formule des accroissements finis pour $g \in C^1(U, \mathbb{R})$, en tenant compte de $g(\bar{x}) = \bar{x}$, il existe $c^{(k)} \in]\bar{x}, x^{(k)}[$ tel que

$$x^{(k+1)} - \bar{x} = g(x^{(k)}) - g(\bar{x}) = g'(c^{(k)})(x^{(k)} - \bar{x}).$$

- Comme $|g'(\bar{x})| < 1$ et que g' est continue sur U alors il existe $\alpha > 0$ et $\beta < 1$ tels que

$$|g'(x)| \leq \beta \text{ pour tout } x \in I_\alpha.$$

- Si $x^{(k)} \in I_\alpha$ on a $c^{(k)} \in I_\alpha$ et donc

$$|x^{(k+1)} - \bar{x}| \leq \beta |x^{(k)} - \bar{x}|.$$

Convergence linéaire locale de l'algorithme de point fixe: preuve

- Montrons par récurrence que si $x^{(0)} \in I_\alpha$, il résulte que $x^{(k)} \in I_\alpha$ pour tout $k \in \mathbb{N}$. Puis ceci impliquera que

$$|x^{(k+1)} - \bar{x}| \leq \beta |x^{(k)} - \bar{x}| \text{ pour tout } k \in \mathbb{N}^*.$$

Preuve: on suppose que $x^{(k)} \in I_\alpha$ et on va montrer que ceci implique $x^{(k+1)} \in I_\alpha$.
Par hypothèse $|x^{(k)} - \bar{x}| < \alpha$. Comme

$$|x^{(k+1)} - \bar{x}| \leq \beta |x^{(k)} - \bar{x}| < \beta\alpha < \alpha.$$

Il en résulte bien que $x^{(k+1)} \in I_\alpha$.

Convergence linéaire locale de l'algorithme de point fixe: Preuve suite

- Montrons par récurrence que

$$|x^{(k)} - \bar{x}| \leq \beta^k |x^{(0)} - \bar{x}|.$$

Pour $k = 0$, l'inégalité se réduit à $|x^{(0)} - \bar{x}| \leq |x^{(0)} - \bar{x}|$ toujours vérifiée.
Supposons la vérifiée pour $k \in \mathbb{N}$, d'où

$$|x^{(k+1)} - \bar{x}| \leq \beta |x^{(k)} - \bar{x}| \leq \beta \beta^k |x^{(0)} - \bar{x}| = \beta^{k+1} |x^{(0)} - \bar{x}|$$

L'inégalité est donc aussi vraie pour $k + 1$.

Convergence linéaire locale de l'algorithme de point fixe: Preuve suite

- Comme $\beta < 1$, on a $\lim_{k\to+\infty} \beta^k = 0$ et donc $\lim_{k\to+\infty} x^{(k)} = \bar{x}$.
- On a montré que $x^{(k+1)} - \bar{x} = g'(c^{(k)})(x^{(k)} - \bar{x})$ avec $c^{(k)} \in]\bar{x}, x^{(k)}[\subset I_\alpha$.

  On a $\lim_{k\to+\infty} c^{(k)} = \bar{x}$ puis par continuité de g' au point $\bar{x} \in U$ on en déduit que

$$\lim_{k\to+\infty} \frac{x^{(k+1)} - \bar{x}}{x^{(k)} - \bar{x}} = \lim_{k\to+\infty} g'(c^{(k)}) = g'(\bar{x}).$$

Convergence linéaire locale de l'algorithme de point fixe: Preuve suite

Remarque: l'hypothèse de contraction locale au voisinage de $\bar{x}$ suivante: $g \in C^0(U, \mathbb{R})$ et il existe $\alpha > 0$ et $\beta < 1$ tels que

$$|g(y) - g(x)| \leq \beta |y - x| \text{ pour tout } (x, y) \in I_\alpha \times I_\alpha,$$

suffit à obtenir la convergence linéaire du point fixe vers $\bar{x}$.

Convergence quadratique locale de l'algorithme de point fixe

Soit $g \in C^2(U, \mathbb{R})$ avec U ouvert de $\mathbb{R}$ et soit $\bar{x} \in U$ tel que $g(\bar{x}) = \bar{x}$. On suppose que $g'(\bar{x}) = 0$. Alors il existe $\alpha > 0$, $\beta < 1$ et $\gamma \geq 0$ avec $\alpha\gamma \leq \beta$ tels que si

$$x^{(0)} \in I_\alpha =]\bar{x} - \alpha, \bar{x} + \alpha[,$$

on a

- $I_\alpha \subset U$,
- $x^{(k)} \in I_\alpha$ pour tout $k \in \mathbb{N}$,
- $|x^{(k+1)} - \bar{x}| \leq \gamma |x^{(k)} - \bar{x}|^2$ pour tout $k \in \mathbb{N}$,
- $|x^{(k)} - \bar{x}| \leq \beta^{(2^k - 1)} \, |x^{(0)} - \bar{x}|$ pour tout $k \in \mathbb{N}$,
- $\lim_{k \to +\infty} x^{(k)} = \bar{x}$.
- $\displaystyle \lim_{k \to +\infty} \frac{x^{(k+1)} - \bar{x}}{(x^{(k)} - \bar{x})^2} = \frac{1}{2} g''(\bar{x})$.

Convergence quadratique locale de l'algorithme de point fixe: Preuve

- D'après la formule de Taylor à l'ordre 2 en tenant compte que $g(\bar{x}) = \bar{x}$ et $g'(\bar{x}) = 0$, il existe $c^{(k)} \in]\bar{x}, x^{(k)}[$ tel que

$$x^{(k+1)} - \bar{x} = g(x^{(k)}) - g(\bar{x}) - g'(\bar{x})(x^{(k)} - \bar{x}) = \frac{1}{2} g''(c^{(k)})(x^{(k)} - \bar{x})^2$$

- Soit $\tilde{\alpha} > 0$ tel que $I_{\tilde{\alpha}} \subset U$. On pose

$$\gamma = \sup_{x \in I_{\tilde{\alpha}}} \frac{1}{2} |g''(x)| \quad \text{et} \quad \alpha = \min\left(\tilde{\alpha}, \frac{\beta}{\gamma}\right).$$

  On a bien $\alpha\gamma \leq \beta$.
- Si $x^{(k)} \in I_\alpha$ on a $c^{(k)} \in I_\alpha$ et donc

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

Convergence quadratique locale de l'algorithme de point fixe: Preuve

- On va montrer par récurrence que si $x^{(0)} \in I_\alpha$, il résulte que $x^{(k)} \in I_\alpha$ pour tout $k \in \mathbb{N}$. Puis cela impliquera que

$$|x^{(k+1)} - \bar{x}| \leq \gamma |x^{(k)} - \bar{x}|^2 \text{ pour tout } k \in \mathbb{N}^*,$$

Preuve: on suppose que $x^{(k)} \in I_\alpha$ et on va montrer que ceci implique $x^{(k+1)} \in I_\alpha$.
On a

$$|x^{(k+1)} - \bar{x}| \leq \gamma |x^{(k)} - \bar{x}|^2 < \gamma \alpha^2 = (\alpha\gamma)\alpha \leq \beta\alpha < \alpha$$

et donc $|x^{(k+1)} - \bar{x}| < \alpha$ ce qui implique que $x^{(k+1)} \in I_\alpha$.

Convergence quadratique locale de l'algorithme de point fixe: Preuve suite

- Montrons par récurrence que

$$|x^{(k)} - \bar{x}| \leq \beta^{(2^k-1)}|x^{(0)} - \bar{x}|.$$

Pour $k = 0$, l'inégalité se réduit à $|x^{(0)} - \bar{x}| \leq |x^{(0)} - \bar{x}|$ toujours vérifiée. Supposons la vérifiée pour $k \in \mathbb{N}$, d'où

$$\begin{aligned} |x^{(k+1)} - \bar{x}| \leq \gamma|x^{(k)} - \bar{x}|^2 \leq \gamma\beta^{(2^{k+1}-2)}|x^{(0)} - \bar{x}|^2 \\ < \gamma\alpha\beta^{(2^{k+1}-2)}|x^{(0)} - \bar{x}| \end{aligned}$$

comme $\alpha\gamma \leq \beta$, l'inégalité est donc aussi vraie pour $k+1$.

Convergence quadratique locale de l'algorithme de point fixe: Preuve suite

- Comme $\beta < 1$, on a $\lim_{k\to+\infty} \beta^{(2^k-1)} = 0$ puis donc $\lim_{k\to+\infty} x^{(k)} = \bar{x}$.

- On a montré que $x^{(k+1)} - \bar{x} = \frac{1}{2} g''(c^{(k)})(x^{(k)} - \bar{x})^2$ avec $c^{(k)} \in]\bar{x}, x^{(k)}[\subset I_\alpha$.

  On a $\lim_{k\to+\infty} c^{(k)} = \bar{x}$ puis par continuité de g'' au point $\bar{x} \in U$ on en déduit que

$$\lim_{k\to+\infty} \frac{x^{(k+1)} - \bar{x}}{(x^{(k)} - \bar{x})^2} = \lim_{k\to+\infty} \frac{1}{2} g''(c^{(k)}) = \frac{1}{2} g''(\bar{x}).$$

Algorithme de Newton en dimension 1

L'algorithme de Newton pour trouver un zéro $\bar{x}$ d'une fonction f est l'algorithme de point fixe obtenu pour $g(x) = x - \dfrac{1}{f'(x)} f(x)$:

$$x^{(k+1)} = x^{(k)} - \frac{1}{f'(x^{(k)})} f(x^{(k)}) \text{ pour tout } k \in \mathbb{N}.$$

On remarque que $g'(\bar{x}) = 1 - \dfrac{1}{f'(\bar{x})} f'(\bar{x}) + \dfrac{f''(\bar{x})}{(f'(\bar{x}))^2} f(\bar{x}) = 0$, en supposant $f'(\bar{x}) \neq 0$.
Le théorème de convergence quadratique locale s'applique donc si $f'(\bar{x}) \neq 0$ et $g \in C^2(U, \mathbb{R})$ (on verra dans la section 3 que $f \in C^2(U, \mathbb{R})$ suffit).

Algorithme de Newton en dimension 1

L'algorithme de Newton s'obtient aussi par approximations successives de f par sa tangente aux points $x^{(k)}$:

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

La solution de $l(x) = 0$ donne bien

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

en supposant que $f'(x^{(k)}) \neq 0$.

Chapitre II: Résolution numérique des systèmes non linéaires

Section 2: Rappels et compléments de calcul différentiel

Rappels sur les fonctions vectorielles: différentielles

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

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

  En utilisant la notation: $o(\mathbf{h}) : \mathbb{R}^n \to \mathbb{R}^m$ pour les applications telles que $\lim_{\mathbf{h} \neq 0 \to 0} \dfrac{\|o(\mathbf{h})\|}{\|\mathbf{h}\|}$ on écrit de façon équivalente que

$$\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) - D\mathbf{f}(\mathbf{x})(\mathbf{h}) = o(\mathbf{h}).$$

Rappels sur les fonctions vectorielles: différentielles

- Remarque: si $\mathbf{f}$ est différentiable au point $\mathbf{x} \in U$ alors $\mathbf{f}$ est continue en $\mathbf{x}$.

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

Rappels sur les fonctions vectorielles: différentielles

Exemples:

- pour $n = 1$ on retrouve la dérivée au sens classique $D\mathbf{f}(x) = \mathbf{f}'(x) \in \mathbb{R}^m$

- pour $m = 1$, $Df(\mathbf{x})$ est une forme linéaire de $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$

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

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

$$\frac{\partial \mathbf{f}}{\partial x_j}(\mathbf{x}) = lim_{h_j \to 0} \frac{\mathbf{f}(\mathbf{x} + h_j \boldsymbol{\epsilon}^{(j)}) - \mathbf{f}(\mathbf{x})}{h_j}.$$

avec $\boldsymbol{\epsilon}^{(j)}$ le j^{ieme} vecteur de la base canonique $\epsilon_i^{(j)} = I_{i,j}$ pour $i = 1, \cdots, n$.

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

- Si $\mathbf{f}$ est différentiable au point $\mathbf{x}$ alors elle admet des dérivées partielles au point $\mathbf{x}$ et

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

- En notant, dans la base canonique de $\mathbb{R}^m$

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{pmatrix}$$

on a donc

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

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

- Matrice représentant l'application linéaire $D\mathbf{f}(\mathbf{x})$:

$$J(\mathbf{x}) \in \mathcal{M}_{m,n}(\mathbb{R})$$

telle que $J_{i,j}(\mathbf{x}) = \dfrac{\partial f_i}{\partial x_j}(\mathbf{x}), i = 1, \cdots, m; j = 1, \cdots, n$

est appelée la matrice Jacobienne de $\mathbf{f}$ au point $\mathbf{x}$, avec

$$D\mathbf{f}(\mathbf{x})(\mathbf{h}) = J(\mathbf{x})\mathbf{h}.$$

$$J(\mathbf{x}) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(\mathbf{x}) & \cdots & \dfrac{\partial f_1}{\partial x_n}(\mathbf{x}) \\ \vdots & \vdots & \vdots \\ \dfrac{\partial f_m}{\partial x_1}(\mathbf{x}) & \cdots & \dfrac{\partial f_m}{\partial x_n}(\mathbf{x}) \end{pmatrix}$$

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

- La réciproque n'est pas vraie: si $\mathbf{f}$ admet des dérivées partielles en $\mathbf{x}$, elle n'est pas nécéssairement différentiable au point $\mathbf{x}$.
  - Exemple: $f(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}) = \frac{x_1 x_2}{\sqrt{x_1^2+x_2^2}}$ a des dérivées partielles nulles en 0 mais n'est pas différentiable en 0.

- Si $\mathbf{f}$ admet des dérivées partielles continues au point $\mathbf{x}$ pour tout $i = 1 \cdots, n$ alors $\mathbf{f}$ est différentiable au point $\mathbf{x}$

- $\mathbf{f}$ est continuement différentiable sur U (ie $D\mathbf{f}$ existe et est continue sur U) ssi $\mathbf{f}$ admet des dérivées partielles continues sur U. On dit que $\mathbf{f}$ est $C^1\left(U, \mathbb{R}^m\right)$.

Rappels sur les fonctions vectorielles: différentiation des fonctions composées

Soient

- $\mathbf{g} : \mathbb{R}^p \to \mathbb{R}^n$ différentiable en $\mathbf{x} \in \mathbb{R}^p$
- $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ différentiable en $g(\mathbf{x}) \in \mathbb{R}^n$.

Alors la fonction $\mathbf{f} \circ \mathbf{g} : \mathbb{R}^p \to \mathbb{R}^m$ est différentiable en $\mathbf{x}$ et on a

$$D(\mathbf{f} \circ \mathbf{g})(\mathbf{x}) = D\mathbf{f}(\mathbf{g}(\mathbf{x}))D\mathbf{g}(\mathbf{x}).$$

avec $D\mathbf{g}(\mathbf{x}) \in \mathcal{L}\left(\mathbb{R}^p, \mathbb{R}^n\right)$, $Df(\mathbf{g}(\mathbf{x})) \in \mathcal{L}\left(\mathbb{R}^n, \mathbb{R}^m\right)$,
$D(\mathbf{f} \circ \mathbf{g})(\mathbf{x}) \in \mathcal{L}\left(\mathbb{R}^p, \mathbb{R}^m\right)$,

Rappels sur les fonctions vectorielles: formule des accroissements finis

Rappel dans le cas $n = m = 1$ (**théorème des accroissements finis**): $f : [a, b] \subset \mathbb{R} \to \mathbb{R}$:

si f est continue sur $[a, b]$ et différentiable sur (a, b) alors il existe $c \in (a, b)$ tel que

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

Rappels sur les fonctions vectorielles: formule des accroissements finis

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

    $$f(\mathbf{b}) - f(\mathbf{a}) = Df(\mathbf{c})(\mathbf{b} - \mathbf{a})$$

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

Rappels sur les fonctions vectorielles: formule des accroissements finis

- Cas général: $n \geq 1$, $m \geq 1$:

  $\mathbf{f} : U \subset \mathbb{R}^n \to \mathbb{R}^m$ différentiable sur U et $[\mathbf{a}, \mathbf{b}] \subset U$, alors

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

Rappels sur les fonctions vectorielles: formule des accroissements finis: Preuve

Soit $\phi(t) = \mathbf{f}((1-t)\mathbf{a} + t\mathbf{b})$.

On a $\phi'(t) = D\mathbf{f}((1-t)\mathbf{a} + t\mathbf{b})(\mathbf{b} - \mathbf{a})$ et

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

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

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

Rappels sur les fonctions vectorielles: différentielle d'ordre 2

Soit $\mathbf{f} \in C^1(U, \mathbb{R}^m)$ avec U ouvert de $\mathbb{R}^n$. On a $D\mathbf{f} \in C^0\Big(U, \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)\Big)$.

On dit que $\mathbf{f}$ admet une différentielle seconde, notée $D^2\mathbf{f}$, en $\mathbf{x} \in U$ ssi $D\mathbf{f}$ est différentiable en $\mathbf{x} \in U$, donc ssi pour tous $\mathbf{h} \in \mathbb{R}^n$ et $\mathbf{k} \in \mathbb{R}^n$ on a

$$D\mathbf{f}(\mathbf{x}+\mathbf{h})(\mathbf{k}) - D\mathbf{f}(\mathbf{x})(\mathbf{k}) - D^2\mathbf{f}(\mathbf{x})(\mathbf{h})\mathbf{k} = o(\mathbf{h})\mathbf{k}.$$

La différentielle seconde $D^2\mathbf{f}(\mathbf{x})$ 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 $\mathcal{L}\Big(\mathbb{R}^n, \mathbb{R}^n; \mathbb{R}^m\Big)$.

Rappels sur les fonctions vectorielles: différentielle d'ordre 2

Soit $\mathbf{f} \in C^1(U, \mathbb{R}^m)$ avec U ouvert de $\mathbb{R}^n$. On a $D\mathbf{f} \in C^0\Big(U, \mathcal{L}(\mathbb{R}^n; \mathbb{R}^m)\Big)$.

Si $D\mathbf{f}$ est continuement différentiable sur U on dit que $\mathbf{f} \in C^2(U, \mathbb{R}^m)$ et on note

$$D^2\mathbf{f} : U \to \mathcal{L}\Big(\mathbb{R}^n, \mathbb{R}^n; \mathbb{R}^m\Big),$$

la différentielle de $D\mathbf{f}$ appelée différentielle seconde de $\mathbf{f}$.

Rappels sur les fonctions vectorielles: différentielle d'ordre 2

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

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

- On a alors le **théorème de Schwarz**: $D^2\mathbf{f}(\mathbf{x})$ est **symétrique** au sens où $D^2\mathbf{f}(\mathbf{x})(\mathbf{h})\mathbf{k} = D^2\mathbf{f}(\mathbf{x})(\mathbf{k})\mathbf{h}$ ou encore

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

- Dans le cas $m = 1$, on appelle $H(\mathbf{x}) \in \mathcal{M}_{n,n}(\mathbb{R})$ la matrice dite Hessienne représentant la forme bilinéaire $D^2 f(\mathbf{x})$ dans la base canonique.

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

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

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

avec

$$\|D^2\mathbf{f}(\mathbf{x})\| = \sup_{\mathbf{u}, \mathbf{v} \neq 0 \in \mathbb{R}^n} \frac{\|D^2\mathbf{f}(\mathbf{x})(\mathbf{u}, \mathbf{v})\|}{\|\mathbf{u}\| \|\mathbf{v}\|}.$$

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

Preuve: soit

$$\varphi(t) = \mathbf{f}(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) - \mathbf{f}(\mathbf{x}) - t\ D\mathbf{f}(\mathbf{x})(\mathbf{y} - \mathbf{x}),$$

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

$$\varphi(1) = \varphi(1) - \varphi(0) = \int_0^1 \Big(D\mathbf{f}(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) - D\mathbf{f}(\mathbf{x})\Big)(\mathbf{y} - \mathbf{x})dt$$

$$\|\varphi(1)\| = \|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}) - D\mathbf{f}(\mathbf{x})(\mathbf{y} - \mathbf{x})\| \le \int_0^1 \|D\mathbf{f}(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) - D\mathbf{f}(\mathbf{x})\| \|\mathbf{y} - \mathbf{x}\| dt$$

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

$$\|D\mathbf{f}(\mathbf{x} + t(\mathbf{y} - \mathbf{x})) - D\mathbf{f}(\mathbf{x})\| \le \sup_{\mathbf{z} \in (\mathbf{x}, \mathbf{x} + t(\mathbf{y} - \mathbf{x}))} \|D^2\mathbf{f}(\mathbf{z})\|\ \|\mathbf{y} - \mathbf{x}\|\ t$$

Chapitre II: Résolution numérique des systèmes non linéaires

Section 3: Algorithmes de point fixe et de Newton en dimension n

Algorithme de point fixe

Soit
$$\mathbf{g} \in C^0(U, \mathbb{R}^n), \; U \text{ ouvert de } \mathbb{R}^n$$
telle qu'il existe $\bar{\mathbf{x}} \in U$ avec $\mathbf{g}(\bar{\mathbf{x}}) = \bar{x}$.

Etant donné $\mathbf{x}^{(0)} \in U$, la méthode du point fixe est définie par la suite récurrente $\mathbf{x}^{(k)}$, $k \in \mathbb{N}$ telle que
$$\mathbf{x}^{(k+1)} = \mathbf{g}(\mathbf{x}^{(k)}).$$

- L'analyse de la méthode du point fixe doit donner des conditions suffisantes sur $\mathbf{g}$ et sur $\mathbf{x}^{(0)}$ pour que la suite $\mathbf{x}^{(k)}$ converge vers $\bar{\mathbf{x}}$ et estimer l'ordre de convergence.

Convergence linéaire locale de l'algorithme de point fixe

Soit $\mathbf{g} \in C^1(U, \mathbb{R}^n)$ avec U ouvert de $\mathbb{R}^n$ et soit $\bar{\mathbf{x}} \in U$ tel que $\mathbf{g}(\bar{\mathbf{x}}) = \bar{\mathbf{x}}$. On suppose que $\|D\mathbf{g}(\bar{\mathbf{x}})\| < 1$. Alors il existe $\alpha > 0$ et $\beta < 1$ tels que si

$$\mathbf{x}^{(0)} \in B(\bar{\mathbf{x}}, \alpha) = \{\mathbf{x} \in \mathbb{R}^n \text{ tels que } \|\mathbf{x} - \bar{\mathbf{x}}\| < \alpha\},$$

on a

- $B(\bar{\mathbf{x}}, \alpha) \subset U$,
- $\mathbf{x}^{(k)} \in B(\bar{\mathbf{x}}, \alpha)$ pour tout $k \in \mathbb{N}$,
- $\|\mathbf{x}^{(k+1)} - \bar{\mathbf{x}}\| \leq \beta \|\mathbf{x}^k - \bar{\mathbf{x}}\|$ pour tout $k \in \mathbb{N}$,
- $\|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\| \leq \beta^k \, \|\mathbf{x}^{(0)} - \bar{\mathbf{x}}\|$ pour tout $k \in \mathbb{N}$,
- $\lim_{k \to +\infty} \mathbf{x}^{(k)} = \bar{\mathbf{x}}$.

Convergence linéaire locale de l'algorithme de point fixe: preuve

- D'après la formule des accroissements finis pour $\mathbf{g} \in C^1(U, \mathbb{R}^n)$ on a en tenant compte de $\mathbf{g}(\bar{\mathbf{x}}) = \bar{\mathbf{x}}$, on a:

$$\begin{aligned} \|\mathbf{x}^{(k+1)} - \bar{\mathbf{x}}\| &= \|\mathbf{g}(\mathbf{x}^{(k)}) - \mathbf{g}(\bar{\mathbf{x}})\| \\ &\leq \sup_{\mathbf{x} \in]\bar{\mathbf{x}}, \mathbf{x}^{(k)}[} \|D\mathbf{g}(\mathbf{x})\| \|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\|. \end{aligned}$$

- Comme $\|D\mathbf{g}(\bar{\mathbf{x}})\| < 1$ et que $D\mathbf{g}$ est continue en $\bar{\mathbf{x}}$, alors il existe $\alpha > 0$ et $\beta < 1$ tels que

$$\|D\mathbf{g}(\mathbf{x})\| \leq \beta \text{ pour tout } \mathbf{x} \in B(\bar{\mathbf{x}}, \alpha).$$

- Si $\mathbf{x}^{(k)} \in B(\bar{\mathbf{x}}, \alpha)$ on a donc

$$\|\mathbf{x}^{(k+1)} - \bar{\mathbf{x}}\| \leq \beta \|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\|.$$

Convergence linéaire locale de l'algorithme de point fixe: preuve

- On va montrer par récurrence que si $\mathbf{x}^{(0)} \in B(\bar{\mathbf{x}}, \alpha)$, il résulte que $\mathbf{x}^{(k)} \in B(\bar{\mathbf{x}}, \alpha)$ pour tout $k \in \mathbb{N}$. Puis ceci impliquera que

$$\|\mathbf{x}^{(k+1)} - \bar{\mathbf{x}}\| \leq \beta \|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\| \text{ pour tout } k \in \mathbb{N}^*.$$

  Preuve: on suppose que $\mathbf{x}^{(k)} \in B(\bar{\mathbf{x}}, \alpha)$ et on va montrer que ceci implique $\mathbf{x}^{(k+1)} \in B(\bar{\mathbf{x}}, \alpha)$.
  Comme
$$\|\mathbf{x}^{(k+1)} - \bar{\mathbf{x}}\| \leq \beta \|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\| < \beta\alpha < \alpha,$$
  il en résulte bien que $\mathbf{x}^{(k+1)} \in B(\bar{\mathbf{x}}, \alpha)$.

Convergence linéaire locale de l'algorithme de point fixe: Preuve suite

- Montrons par récurrence que

$$\|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\| \leq \beta^k \|\mathbf{x}^{(0)} - \bar{\mathbf{x}}\|.$$

Pour $k = 0$, l'inégalité se réduit à $\|\mathbf{x}^{(0)} - \bar{\mathbf{x}}\| \leq \|\mathbf{x}^{(0)} - \bar{\mathbf{x}}\|$ toujours vérifiée.
Supposons la vérifiée pour $k \in \mathbb{N}$, d'où

$$\|\mathbf{x}^{(k+1)} - \bar{\mathbf{x}}\| \leq \beta \|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\| \leq \beta \beta^k \|\mathbf{x}^{(0)} - \bar{\mathbf{x}}\| = \beta^{k+1} \|\mathbf{x}^{(0)} - \bar{\mathbf{x}}\|$$

L'inégalité est donc aussi vraie pour $k + 1$.

La convergence s'en déduit car $\beta < 1$.

Convergence quadratique locale de l'algorithme de point fixe

Soit $\mathbf{g} \in C^2(U, \mathbb{R}^n)$ avec U ouvert de $\mathbb{R}^n$ et soit $\bar{\mathbf{x}} \in U$ tel que $\mathbf{g}(\bar{\mathbf{x}}) = \bar{\mathbf{x}}$. On suppose que $D\mathbf{g}(\bar{\mathbf{x}}) = 0$. Alors il existe $\alpha > 0$, $\beta < 1$ et $\gamma \geq 0$ avec $\alpha\gamma \leq \beta$ tels que si

$$\mathbf{x}^{(0)} \in B(\bar{\mathbf{x}}, \alpha) = \{\mathbf{x} \in \mathbb{R}^n \text{ tels que } \|\mathbf{x} - \bar{\mathbf{x}}\| < \alpha\},$$

on a

- $B(\bar{\mathbf{x}}, \alpha) \subset U$,
- $\mathbf{x}^{(k)} \in B(\bar{\mathbf{x}}, \alpha)$ pour tout $k \in \mathbb{N}$,
- $\|\mathbf{x}^{(k+1)} - \bar{\mathbf{x}}\| \leq \gamma \|\mathbf{x}^k - \bar{\mathbf{x}}\|^2$ pour tout $k \in \mathbb{N}$,
- $\|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\| \leq \beta^{(2^k - 1)} \|\mathbf{x}^{(0)} - \bar{\mathbf{x}}\|$ pour tout $k \in \mathbb{N}$,
- $\lim_{k \to +\infty} \mathbf{x}^{(k)} = \bar{\mathbf{x}}$.

Convergence quadratique locale de l'algorithme de point fixe: Preuve

- D'après la formule de Taylor à l'ordre 2 en tenant compte que $\mathbf{g}(\bar{\mathbf{x}}) = \bar{\mathbf{x}}$ et $D\mathbf{g}(\bar{\mathbf{x}}) = 0$, on a

$$\begin{aligned}\|\mathbf{x}^{(k+1)} - \bar{\mathbf{x}}\| &= \|\mathbf{g}(\mathbf{x}^{(k)}) - \mathbf{g}(\bar{\mathbf{x}}) - D\mathbf{g}(\bar{\mathbf{x}})(\mathbf{x}^{(k)} - \bar{\mathbf{x}})\| \\ &\leq \sup_{\mathbf{x} \in]\bar{\mathbf{x}}, \mathbf{x}^{(k)}[} \frac{1}{2}\|D^2\mathbf{g}(\mathbf{x})\| \|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\|^2.\end{aligned}$$

- Soit $\tilde{\alpha} > 0$ tel que $B(\bar{\mathbf{x}}, \tilde{\alpha}) \subset U$. On pose

$$\gamma = \sup_{\mathbf{x} \in B(\bar{\mathbf{x}}, \tilde{\alpha})} \frac{1}{2}\|D^2\mathbf{g}(\mathbf{x})\| \quad \text{et} \quad \alpha = \min\left(\tilde{\alpha}, \frac{\beta}{\gamma}\right).$$

  On a bien $\alpha\gamma \leq \beta$.
- Si $\mathbf{x}^{(k)} \in B(\bar{\mathbf{x}}, \alpha)$ on a donc

$$\|\mathbf{x}^{(k+1)} - \bar{\mathbf{x}}\| \leq \gamma \|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\|^2.$$

Convergence quadratique locale de l'algorithme de point fixe: Preuve

- On va montrer par récurrence que si $\mathbf{x}^{(0)} \in B(\bar{\mathbf{x}}, \alpha)$, il résulte que $\mathbf{x}^{(k)} \in B(\bar{\mathbf{x}}, \alpha)$ pour tout $k \in \mathbb{N}$. Puis ceci impliquera que

$$\|\mathbf{x}^{(k+1)} - \bar{\mathbf{x}}\| \leq \gamma \|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\|^2 \text{ pour tout } k \in \mathbb{N}^*.$$

Preuve: on suppose que $\mathbf{x}^{(k)} \in B(\bar{\mathbf{x}}, \alpha)$ et on va montrer que ceci implique $\mathbf{x}^{(k+1)} \in B(\bar{\mathbf{x}}, \alpha)$.
Comme

$$\|\mathbf{x}^{(k+1)} - \bar{\mathbf{x}}\| \leq \gamma \|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\|^2 < \gamma \alpha^2 = (\alpha\gamma)\alpha \leq \beta\alpha < \alpha,$$

on a $\|\mathbf{x}^{(k+1)} - \bar{\mathbf{x}}\| < \alpha$ ce qui implique que $\mathbf{x}^{(k+1)} \in B(\bar{\mathbf{x}}, \alpha)$.

Convergence quadratique locale de l'algorithme de point fixe: Preuve suite

- Montrons par récurrence que

$$\|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\| \leq \beta^{(2^k-1)} \|\mathbf{x}^{(0)} - \bar{\mathbf{x}}\|.$$

Pour $k = 0$, l'inégalité se réduit à $\|\mathbf{x}^{(0)} - \bar{\mathbf{x}}\| \leq \|\mathbf{x}^{(0)} - \bar{\mathbf{x}}\|$ toujours vérifiée.
Supposons la vérifiée pour $k \in \mathbb{N}$, d'où

$$\begin{aligned}\|\mathbf{x}^{(k+1)} - \bar{\mathbf{x}}\| \leq \gamma \|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\|^2 \leq \gamma \beta^{(2^{k+1}-2)} \|\mathbf{x}^{(0)} - \bar{\mathbf{x}}\|^2 \\ < \gamma \alpha \beta^{(2^{k+1}-2)} \|\mathbf{x}^{(0)} - \bar{\mathbf{x}}\|\end{aligned}$$

comme $\alpha\gamma \leq \beta$, l'inégalité est donc aussi vraie pour $k + 1$.

La convergence s'en déduit car $\beta < 1$.

Algorithme de Newton

Soit

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

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

Etant donné $\mathbf{x}^{(0)} \in U$, la méthode de Newton est définie par la suite récurrente $\mathbf{x}^{(k)}$, $k \in \mathbb{N}$ telle que

$$D\mathbf{f}(\mathbf{x}^{(k)})(\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}) = -\mathbf{f}(\mathbf{x}^{(k)}), \quad k \in \mathbb{N}.$$

- A chaque itération il faudra donc calculer la différentielle $D\mathbf{f}(\mathbf{x}^{(k)})$, vérifier qu'elle est bien inversible et résoudre un système linéaire.
- L'analyse de la méthode de Newton doit donner des conditions suffisantes sur $\mathbf{f}$ et sur $\mathbf{x}^{(0)}$ pour que $D\mathbf{f}(\mathbf{x}^{(k)})$ soit inversible pour tout $k \in \mathbb{N}$ et pour que la suite $\mathbf{x}^{(k)}$ converge vers $\bar{\mathbf{x}}$.

Algorithme de Newton

Remarque: l'algorithme de Newton équivaut à l'algorithme de point fixe pour la fonction

$$\mathbf{g}(\mathbf{x}) = \mathbf{x} - D\mathbf{f}(\mathbf{x})^{-1}\mathbf{f}(\mathbf{x}).$$

En supposant que $D\mathbf{f}(\bar{\mathbf{x}})$ est **inversible**, on vérifie que $\mathbf{g}(\bar{\mathbf{x}}) = \bar{\mathbf{x}}$ et que $D\mathbf{g}(\bar{\mathbf{x}}) = 0$ qui sont les conditions de convergence quadratique locale de l'algorithme de point fixe associé à $\mathbf{g}$.

Cependant, l'hypothèse $\mathbf{g} \in C^2(U, \mathbb{R}^n)$ est trop forte. On va donc faire une preuve directe de la convergence quadratique locale de l'algorithme de Newton.

Convergence quadratique locale de l'algorithme de Newton

Soit $\mathbf{f} \in C^2(U, \mathbb{R}^n)$ avec U ouvert de $\mathbb{R}^n$ et soit $\bar{\mathbf{x}} \in U$ tel que $\mathbf{f}(\bar{\mathbf{x}}) = 0$. On suppose que $D\mathbf{f}(\bar{\mathbf{x}})$ est **inversible**. Alors il existe $\alpha > 0$, $\beta < 1$ et $\gamma \geq 0$ avec $\alpha\gamma \leq \beta$ tels que si

$$\mathbf{x}^{(0)} \in B(\bar{\mathbf{x}}, \alpha) = \{\mathbf{x} \in \mathbb{R}^n \text{ tels que } \|\mathbf{x} - \bar{\mathbf{x}}\| < \alpha\},$$

on a

- $B(\bar{\mathbf{x}}, \alpha) \subset U$,
- $D\mathbf{f}(\mathbf{x}^{(k)})$ est inversible pour tout $k \in \mathbb{N}$,
- $\mathbf{x}^{(k)} \in B(\bar{\mathbf{x}}, \alpha)$ pour tout $k \in \mathbb{N}$,
- $\|\mathbf{x}^{(k+1)} - \bar{\mathbf{x}}\| \leq \gamma \|\mathbf{x}^k - \bar{\mathbf{x}}\|^2$ pour tout $k \in \mathbb{N}$,
- $\|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\| \leq \beta^{(2^k - 1)} \|\mathbf{x}^{(0)} - \bar{\mathbf{x}}\|$ pour tout $k \in \mathbb{N}$,
- $\lim_{k \to +\infty} \mathbf{x}^{(k)} = \bar{\mathbf{x}}$.

Convergence quadratique locale de l'algorithme de Newton: Preuve

On commence par montrer le lemme suivant:

Soit $\mathbf{f} \in C^2(U, \mathbb{R}^n)$ avec U ouvert de $\mathbb{R}^n$ et $\bar{\mathbf{x}} \in U$ tel que $D\mathbf{f}(\bar{\mathbf{x}})$ est inversible. Alors il existe $\tilde{\alpha} > 0$, $C_1 > 0$, $C_2 > 0$ tels que $B(\bar{\mathbf{x}}, \tilde{\alpha}) \subset U$ et

- $D\mathbf{f}(\mathbf{x})$ inversible et $\|(Df(\mathbf{x}))^{-1}\| \leq C_1$ pour tous $\mathbf{x} \in B(\bar{\mathbf{x}}, \tilde{\alpha})$
- $\|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}) - D\mathbf{f}(\mathbf{x})(\mathbf{y} - \mathbf{x})\| \leq C_2 \|\mathbf{y} - \mathbf{x}\|^2$ pour tous $(\mathbf{x}, \mathbf{y}) \in B(\bar{\mathbf{x}}, \tilde{\alpha})$

Preuve: le point 1 est une application de $D\mathbf{f}(\mathbf{x})^{-1} = \frac{1}{\det(D\mathbf{f}(\mathbf{x}))} {}^t\mathrm{Co}(D\mathbf{f}(\mathbf{x}))$ et de sa continuité en $\bar{\mathbf{x}}$ pour $\det(D\mathbf{f}(\bar{\mathbf{x}})) \neq 0$.
Le point 2 résulte directement de la formule de Taylor d'ordre 2 pour $\mathbf{f} \in C^2(U, \mathbb{R}^n)$.

Convergence quadratique locale de l'algorithme de Newton: Preuve suite

- On pose $\gamma = C_1 C_2$ et $\alpha = \min\left(\tilde{\alpha}, \frac{\beta}{\gamma}\right)$. On a bien $\alpha\gamma \leq \beta$.
- On suppose que $\mathbf{x}^{(k)} \in B(\bar{\mathbf{x}}, \alpha)$. D'après le lemme $D\mathbf{f}(\mathbf{x}^{(k)})$ est inversible donc $\mathbf{x}^{(k+1)}$ est bien défini.

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

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

  d'où

  $$\begin{aligned}\|\mathbf{x}^{(k+1)} - \bar{\mathbf{x}}\| &\leq \|(D\mathbf{f}(\mathbf{x}^{(k)}))^{-1}\| C_2 \|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\|^2 \\ &\leq C_1 C_2 \|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\|^2 = \gamma \|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\|^2.\end{aligned}$$

Convergence quadratique locale de l'algorithme de Newton: Preuve suite

- On va montrer par récurrence que si $\mathbf{x}^{(0)} \in B(\bar{\mathbf{x}}, \alpha)$, il résulte que $\mathbf{x}^{(k)} \in B(\bar{\mathbf{x}}, \alpha)$ pour tout $k \in \mathbb{N}$. Puis ceci impliquera que

$$\|\mathbf{x}^{(k+1)} - \bar{\mathbf{x}}\| \leq \gamma \|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\|^2 \text{ pour tout } k \in \mathbb{N}.$$

- On suppose que $\mathbf{x}^{(k)} \in B(\bar{\mathbf{x}}, \alpha)$ et on va montrer que ceci implique $\mathbf{x}^{(k+1)} \in B(\bar{\mathbf{x}}, \alpha)$.
  Comme

$$\|\mathbf{x}^{(k+1)} - \bar{\mathbf{x}}\| \leq \gamma \|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\|^2 < \gamma \alpha^2 = (\alpha\gamma)\alpha \leq \beta\alpha < \alpha$$

et donc $\|\mathbf{x}^{(k+1)} - \bar{\mathbf{x}}\| < \alpha$ ce qui implique que $\mathbf{x}^{(k+1)} \in B(\bar{\mathbf{x}}, \alpha)$.

Convergence quadratique locale de l'algorithme de Newton: Preuve suite

- Montrons par récurrence que

$$\|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\| \leq \beta^{(2^k-1)}\|\mathbf{x}^{(0)} - \bar{\mathbf{x}}\|.$$

Pour $k = 0$, l'inégalité se réduit à $\|\mathbf{x}^{(0)} - \bar{\mathbf{x}}\| \leq \|\mathbf{x}^{(0)} - \bar{\mathbf{x}}\|$ toujours vérifiée. Supposons la vérifiée pour $k \in \mathbb{N}$, d'où

$$\begin{aligned}\|\mathbf{x}^{(k+1)} - \bar{\mathbf{x}}\| \leq \gamma\|\mathbf{x}^{(k)} - \bar{\mathbf{x}}\|^2 \leq \gamma\beta^{(2^{k+1}-2)}\|\mathbf{x}^{(0)} - \bar{\mathbf{x}}\|^2 \\ < \gamma\alpha\beta^{(2^{k+1}-2)}\|\mathbf{x}^{(0)} - \bar{\mathbf{x}}\|\end{aligned}$$

comme $\alpha\gamma \leq \beta$, l'inégalité est donc aussi vraie pour $k+1$.

La convergence s'en déduit car $\beta < 1$.

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 $\mathbf{r}^k$

$$D\mathbf{f}(\mathbf{x}^{(k)})(\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}) = -\mathbf{f}(\mathbf{x}^{(k)}) + \mathbf{r}^k,$$

tel que

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

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

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

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

Variantes de l'algorithme de Newton: Quasi Newton

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

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

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

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

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

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