

C8, Méthode de Newton et variantes

Rappel, point fixe de contraction, $n = 1$, $f \in C^1(\mathbb{R}, \mathbb{R})$

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

Itération : $k \geq 0$, $x^{(k+1)} = f(x^{(k)})$

On suppose que $\lim_{k \rightarrow +\infty} x^{(k)} = \bar{x}$ (et donc $\bar{x} = f(\bar{x})$)
 $x^{(k)} \neq \bar{x}$ pour tout k

Si $f'(\bar{x}) \neq 0$ la convergence est exactement linéaire

Démonstration :

Théorème des Accroissements Finis, il existe c_k entre $\bar{x}$ et $x^{(k)}$ t.q.

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

$$\lim_{k \rightarrow +\infty} \frac{|x^{(k+1)} - \bar{x}|}{|x^{(k)} - \bar{x}|} = |f'(\bar{x})|$$

C8, point fixe de contraction, convergence quadratique ?

Rappel, point fixe de contraction, $n = 1$, $f \in C^1(\mathbb{R}, \mathbb{R})$

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

Itération : $k \geq 0$, $x^{(k+1)} = f(x^{(k)})$

On suppose que $\lim_{k \rightarrow +\infty} x^{(k)} = \bar{x}$ (et donc $\bar{x} = f(\bar{x})$)
 $x^{(k)} \neq \bar{x}$ pour tout k

Si $f'(\bar{x}) = 0$ et $f \in C^2(\mathbb{R}, \mathbb{R})$

la convergence est au moins quadratique

Démonstration :

Développement de f au deuxième ordre en $\bar{x}$,
il existe c_k entre $\bar{x}$ et $x^{(k)}$ t.q.

$$f(x^{(k)}) = f(\bar{x}) + f'(\bar{x})(x^{(k)} - \bar{x}) + \frac{f''(c_k)}{2}(x^{(k)} - \bar{x})^2,$$
$$|x^{(k+1)} - \bar{x}| = |f(x^{(k)}) - f(\bar{x})| = \frac{|f''(c_k)|}{2}(x^{(k)} - \bar{x})^2$$

$$\lim_{k \rightarrow +\infty} \frac{|x^{(k+1)} - \bar{x}|}{|x^{(k)} - \bar{x}|^2} = \frac{1}{2} f''(\bar{x})$$

Méthode de Newton, 1ere idée pour $n = 1$

$g \in C^3(\mathbb{R}, \mathbb{R})$, on cherche $\bar{x} \in \mathbb{R}$ t.q. $g(\bar{x}) = 0$

$f(x) = x - h(x)g(x)$, avec $h(x) \neq 0$ pour tout x , donc

$$g(x) = 0 \Leftrightarrow x = f(x)$$

Algorithme du point fixe :

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

Itération : $k \geq 0$, $x^{(k+1)} = f(x^{(k)})$

On suppose que la suite $(x^{(k)})_{k \in \mathbb{N}}$ converge, on pose

$$\bar{x} = \lim_{k \rightarrow +\infty} x^{(k)} \text{ (donc } \bar{x} = f(\bar{x}) \text{ et } g(\bar{x}) = 0)$$

Si $f'(\bar{x}) \neq 0$ et $f \in C^2(\mathbb{R}, \mathbb{R})$, alors la convergence est quadratique

$$f'(x) = 1 - h'(x)g(x) - h(x)g'(x), \quad f'(\bar{x}) = 1 - h(\bar{x})g'(\bar{x})$$

Un choix possible est donc $h(x) = \frac{1}{g'(x)}$

(à condition que g' ne s'annule pas)

Algorithme de Newton pour $n = 1$:

Itération : $k \geq 0$, $x^{(k+1)} = x^{(k)} - \frac{g(x^{(k)})}{g'(x^{(k)})}$

Méthode de Newton, autre interprétation pour $n = 1$

$g \in C^2(\mathbb{R}, \mathbb{R})$, on cherche $\bar{x} \in \mathbb{R}$ t.q. $g(\bar{x}) = 0$

Algorithme de Newton, $n = 1$:

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

Itération : $k \geq 0$, $x^{(k+1)} = x^{(k)} - \frac{g(x^{(k)})}{g'(x^{(k)})}$

Autre interprétation par développement de Taylor en $x^{(k)}$:

il existe c_k entre $\bar{x}$ et $x^{(k)}$ t.q.

$$0 = g(\bar{x}) = g(x^{(k)}) + g'(x^{(k)})(\bar{x} - x^{(k)}) + \frac{g''(c_k)}{2}(\bar{x} - x^{(k)})^2$$

En négligeant le terme quadratique, ceci suggère

$$0 = g(x^{(k)}) + g'(x^{(k)})(x^{(k+1)} - x^{(k)}), \text{ c'est-à-dire}$$

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

Un exemple avec $n = 1$, point fixe

$$g(x) = x^2 - 2, \bar{x} = \sqrt{2}.$$

Algorithme du point fixe, $f(x) = x - g(x)$:

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

Itération : $k \geq 0, x^{(k+1)} = f(x^{(k)})$

Question : A-t-on convergence de cet algorithme ?

Réponse : non

car $f'(x) = 1 - 2x, |f'(\bar{x})| = |1 - 2\sqrt{2}| > 1$

Un exemple avec $n = 1$, Newton

Algorithme de Newton :

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

Itération : $k \geq 0$,

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

$$2x^{(k)}(x^{(k+1)} - x^{(k)}) = 2 - (x^{(k)})^2$$

$$x^{(k+1)} = \frac{(x^{(k)})^2 + 2}{2x^{(k)}}$$

Question : A-t-on convergence de cet algorithme ?

Réponse : oui si $x^{(0)} > 0$ (et la convergence est quadratique)

Démonstration : Exercice

Méthode de Newton, $n \geq 1$

$g \in C^1(\mathbb{R}^n, \mathbb{R}^n)$, on cherche $\bar{x} \in \mathbb{R}^n$ t.q. $g(\bar{x}) = 0$

Algorithme de Newton :

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

Itération : $k \geq 0$, $J_g(x^{(k)})(x^{(k+1)} - x^{(k)}) = -g(x^{(k)})$
 $J_g(x^{(k)}) \in \mathcal{M}_n(\mathbb{R})$

Généralisation si E est un e.v.n. dimension finie :

$g \in C^1(E, E)$, on cherche $\bar{x} \in E$ t.q. $g(\bar{x}) = 0$

Algorithme de Newton :

Initialisation : $x^{(0)} \in E$

Itération : $k \geq 0$, $Dg(x^{(k)})(x^{(k+1)} - x^{(k)}) = -g(x^{(k)})$
 $Dg(x^{(k)}) \in \mathcal{L}(E, E)$

En considérant une base de E , on peut se ramener au cas $E = \mathbb{R}^n$

Rappel : Si $E = \mathbb{R}^n$, $Dg(x)(h) = J_g(x)h$ pour tout $x, h \in \mathbb{R}^n$

Méthode de Newton, $n \geq 1$, questions

$g \in C^1(\mathbb{R}^n, \mathbb{R}^n)$, on cherche $\bar{x} \in \mathbb{R}^n$ t.q. $g(\bar{x}) = 0$

Algorithme de Newton :

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

Itération : $k \geq 0$, $J_g(x^{(k)})(x^{(k+1)} - x^{(k)}) = -g(x^{(k)})$

Questions :

1. Existence de la suite $(x^{(k)})_{k \in \mathbb{N}}$?
La matrice $J_g(x^{(k)})$ est-elle inversible ?
2. A-t-on $x^{(k)} \rightarrow \bar{x}$ quand $k \rightarrow +\infty$?
3. A-t-on une convergence au moins quadratique ?

Méthode de Newton, $n \geq 1$, réponse aux questions

$g \in C^3(\mathbb{R}^n, \mathbb{R}^n)$, $\bar{x} \in \mathbb{R}^n$ t.q. $g(\bar{x}) = 0$

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.

Résumé : le seul problème pour Newton est qu'il faut (souvent) que $x^{(0)}$ soit "assez près" de la solution exacte pour avoir ces trois résultats

Petit calcul liminaire

$g \in C^1(\mathbb{R}^n, \mathbb{R}^n)$, $M \in C^1(\mathbb{R}^n, \mathcal{M}_n(\mathbb{R}))$, $\phi(x) = M(x)g(x)$ (donc $\phi \in C^1(\mathbb{R}^n, \mathbb{R}^n)$).

On munit $\mathbb{R}^n$ d'une norme et $\mathcal{M}_n(\mathbb{R})$ de la norme induite

Soit $x \in \mathbb{R}^n$. On cherche à calculer $J_\phi(x)$

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

avec $\lim_{h \rightarrow 0} \varepsilon_i(h) = 0$, $i = 1, 2$. De ces deux formules on déduit

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

On remarque que $R(h) = \varepsilon(h)\|h\|$ avec $\lim_{h \rightarrow 0} \varepsilon(h) = 0$, et donc

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

Newton, existence de la suite $(x^{(k)})_{k \in \mathbb{N}}$

$g \in C^1(\mathbb{R}^n, \mathbb{R}^n)$, l'algorithme de Newton s'écrit

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

Itération : $k \geq 0$, $J_g(x^{(k)})(x^{(k+1)} - x^{(k)}) = -g(x^{(k)})$

Hypothèse : $J_g(\bar{x})$ est inversible

Rappel : l'ensemble des matrices inversibles est un ouvert

Il existe $\alpha_0 > 0$ t.q. $J_g(x)$ est inversible pour tout $x \in B(\bar{x}, \alpha_0)$.

Si $x^{(k)} \in B(\bar{x}, \alpha_0)$, alors $x^{(k+1)}$ est bien défini,

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

Mais a-t-on $x^{(k+1)} \in B(\bar{x}, \alpha_0)$? En général : non...

Pour $x \in B(\bar{x}, \alpha_0)$, on pose $M(x) = (J_g(x))^{-1}$ et $f(x) = x - M(x)g(x)$ (donc $x^{(k+1)} = f(x^{(k)})$), la question devient

Existe-t-il $0 < \alpha \leq \alpha_0$ t.q. $f(x) \in B(\bar{x}, \alpha)$ si $x \in B(\bar{x}, \alpha)$?

Newton, existence de la suite $(x^{(k)})_{k \in \mathbb{N}}$, fin

$g \in C^2(\mathbb{R}^n, \mathbb{R}^n)$, $J_g(\bar{x})$ est inversible

$\alpha_0 > 0$, $J_g(x)$ est inversible pour tout $x \in B(\bar{x}, \alpha_0)$.

Pour $x \in B(\bar{x}, \alpha_0)$, $M(x) = (J_g(x))^{-1}$ et $f(x) = x - M(x)g(x)$

Existe-t-il $0 < \alpha \leq \alpha_0$ t.q. $f(x) \in B(\bar{x}, \alpha)$ si $x \in B(\bar{x}, \alpha)$?

$M \in C^1(\mathbb{R}^n, \mathcal{M}_n(\mathbb{R}))$ ($g \in C^2(\mathbb{R}^n, \mathbb{R}^n)$), le calcul linéaire donne

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

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,

il existe $\alpha \in]0, \alpha_0[$ tel que $\|J_f(z)\| \leq 1/2$ pour tout $z \in B(\bar{x}, \alpha)$

Le TAF donne, si $x \in B(\bar{x}, \alpha)$,

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

avec $I(x, \bar{x}) = \{\bar{x} + t(x - \bar{x}), t \in]0, 1[\}$. Donc $f(x) \in B(\bar{x}, \alpha)$

Newton, convergence de la suite $(x^{(k)})_{k \in \mathbb{N}}$

$g \in C^2(\mathbb{R}^n, \mathbb{R}^n)$, deux écritures de l'algorithme de Newton

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

Itération : $k \geq 0$, $J_g(x^{(k)})(x^{(k+1)} - x^{(k)}) = -g(x^{(k)})$, c'est-à-dire $x^{(k+1)} = f(x^{(k)})$ avec $f(x) = x - M(x)g(x)$, $M(x) = (J_g(x))^{-1}$

On choisit $0 < \alpha < \alpha_0$ t.q. $\|J_f(z)\| \leq 1/2$ pour tout $z \in B(\bar{x}, \alpha)$

On a alors $(x, y \in B(\bar{x}, \alpha))$

1. f envoie $B(\bar{x}, \alpha)$ dans $B(\bar{x}, \alpha)$ ($\|f(x) - \bar{x}\| \leq \frac{1}{2}\|x - \bar{x}\|$)
2. f strictement contractante car $\|f(x) - f(y)\| \leq \frac{1}{2}\|x - y\|$ (par le TAF et $\|J_f(z)\| \leq 1/2$ pour tout $z \in B(\bar{x}, \alpha)$)

Conclusion :

$\lim_{k \rightarrow +\infty} x^{(k)} = \bar{x}$, $\bar{x}$ unique point fixe de f dans $B(\bar{x}, \alpha)$

Newton, convergence quadratique

$g \in C^3(\mathbb{R}^n, \mathbb{R}^n)$. Algorithme de Newton écrit avec f :

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

Itération : $k \geq 0, x^{(k+1)} = f(x^{(k)})$

avec $f(x) = x - M(x)g(x)$, $M(x) = (J_g(x))^{-1}$

On choisit $0 < \alpha < \alpha_0$ t.q. $\|J_f(z)\| \leq 1/2$ pour tout $z \in B(\bar{x}, \alpha)$

$J_f \in C^1(\mathbb{R}^n, \mathcal{M}_n(\mathbb{R}))$ (car g est de classe C^3 , J_g et M de classe C^2 , DM et J_f classe C^1). On note

$$\beta = \max_{c \in B(\bar{x}, \alpha)} \|DJ_f(c)\| \quad (\|\cdot\| \text{ est la norme induite sur } \mathcal{L}(\mathbb{R}^n, \mathcal{M}_n(\mathbb{R}))$$

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

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

$$\|f(x) - \bar{x}\| = \|f(x) - f(\bar{x})\| \leq \left(\sup_{z \in I(x, \bar{x})} \|J_f(z)\| \right) \|x - \bar{x}\| \leq \beta \|x - \bar{x}\|^2$$

Avec $x = x^{(k)}$, $\|x^{(k+1)} - \bar{x}\| \leq \beta \|x^{(k)} - \bar{x}\|^2$. La convergence de l'algorithme de Newton est donc au moins quadratique

Inconvénients éventuels de la méthode de Newton

$g \in C^1(\mathbb{R}^n, \mathbb{R}^n)$, $J_g(\bar{x})$ inversible, on cherche $\bar{x} \in \mathbb{R}^n$ t.q. $g(\bar{x}) = 0$

Algorithme de Newton :

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

Itération : $k \geq 0$, $J_g(x^{(k)})(x^{(k+1)} - x^{(k)}) = -g(x^{(k)})$

Inconvénients éventuels

1. La convergence est quadratique (au moins si $g \in C^3(\mathbb{R}^n, \mathbb{R}^n)$) mais pour avoir la convergence il est parfois nécessaire de partir "assez près" de $\bar{x}$ (inconnu)
2. Coûts de la méthode : Calcul de $J_g(x^{(k)})$ et résolution du système linéaire $J_g(x^{(k)})(x^{(k+1)} - x^{(k)}) = -g(x^{(k)})$

Deux idées pour réduire les coûts :

1. Calculer partiellement $J_g(x^{(k)})$
(réduire le coût de la résolution du système linéaire)
2. Ne pas calculer $J_g(x^{(k)})$

Newton sans calculer toute la jacobienne de g

Exemple lié au projet : $g(u) = Au + R(u) - b$, $u \in \mathbb{R}^n$,
 $A \in \mathcal{M}_n(\mathbb{R})$, $R \in C^1(\mathbb{R}^n; \mathbb{R}^n)$, $b \in \mathbb{R}^n$

La matrice A est issue de la discrétisation de $-u''$, elle n'a que 3 diagonales non nulles. La fonction R est non linéaire. Sa i -ème composante, notée R_i , dépend de toutes les composantes de u mais dépend de u_j surtout pour j “proche” de i

A chaque itération de Newton $J_g(x^{(k)})$ sera une matrice “pleine” (c'est-à-dire à termes probablement tous non nuls)

Suggestion : On ne conserve dans $J_g(x^{(k)})$ que A et les termes $\partial_j R_i(u^{(k)})$ pour j “proche” de i

Par exemple, on remplace $\partial_j R_i(u^{(k)})$ par 0 si $j \notin \{i-1, i, i+1\}$

Avantage : Résolution peu coûteuse du système linéaire

Inconvénient ; Probablement une convergence non quadratique

Newton sans calculer la jacobienne de g , $n = 1$

Méthode de la sécante, $g \in C^1(\mathbb{R}, \mathbb{R})$, $g(\bar{x}) = 0$, $g'(\bar{x}) \neq 0$

On remplace à chaque itération $g'(x^{(k)})$ par $\frac{g(x^{(k)}) - g(x^{(k-1)})}{x^{(k)} - x^{(k-1)}}$
(qui est une approximation de $g'(x^{(k)})$)

Initialisation : $x^{(0)} \in \mathbb{R}$, $x^{(1)} - x^{(0)} = -g(x^{(0)})$

Itération : $k \geq 1$, $\frac{g(x^{(k)}) - g(x^{(k-1)})}{x^{(k)} - x^{(k-1)}}(x^{(k+1)} - x^{(k)}) = -g(x^{(k)})$

Intérêt : Ne pas calculer $g'(x)$ (lorsque ce calcul est difficile)

Inconvénient : Convergence non quadratique, mais la convergence est d'ordre au moins $\frac{1+\sqrt{5}}{2}$ si il y a convergence (exercice)

Newton sans calculer la jacobienne de g , $n > 1$

Méthode de “quasi-Newton” , $g \in C^1(\mathbb{R}^n, \mathbb{R}^n)$, $g(\bar{x}) = 0$

On remplace à chaque itération $J_g(x^{(k)})$ par une matrice $B^{(k)}$ de $\mathcal{M}_n(\mathbb{R})$, en essayant que $B^{(k)}$ se rapproche de $J_g(\bar{x})$

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

Itération : $k \geq 0$, $B^{(k)}(x^{(k+1)} - x^{(k)}) = -g(x^{(k)})$

Conditions sur $B^{(k)}$:

1. Pour tout $k \geq 1$, $B^{(k)}(x^{(k)} - x^{(k-1)}) = g(x^{(k)}) - g(x^{(k-1)})$
2. Pour tout $k \geq 1$, $B^{(k)}$ “proche” de $B^{(k-1)}$

Exemple pour l’item 2 (méthode de Broyden, 1965) :

$B_0 = I$, $B^{(k)}s = B^{(k-1)}s$ pour s t.q. $s \cdot (x^{(k)} - x^{(k-1)}) = 0$

(ceci détermine complètement $B^{(k)}$ si $x^{(k)} \neq x^{(k-1)}$)

La méthode ainsi décrite est peu efficace mais a inspiré d’excellentes méthode si g est le gradient d’une fonction de $\mathbb{R}^n$ de $\mathbb{R}$ dont on cherche le minimum, partie 3 du cours