Méthode de Newton

Référence : [Rou09] p.152-155 ([Gou08]).

Soit $f : [c,d] \longrightarrow \mathbb{R}$ une fonction de classe C^2. On suppose que $c < d$, $f(c) < 0$, $f(d) > 0$ et $f'(x) > 0$ pour tout $x \in [c,d]$.
On considère la suite récurrente $x_{n+1} = F(x_n)$ pour $n \geq 0$ où $F(x) = x - \dfrac{f(x)}{f'(x)}$. Alors :

1. f a un unique zéro a et pour tout $x \in [c,d]$, $\exists z \in [a,x]$ tel que :
$$F(x) - a = \frac{1}{2}\frac{f''(z)}{f'(x)}(x-a)^2$$
2. $\exists C > 0$ tel que $\forall x \in [c,d]$,
$$|F(x) - a| \leq C|x-a|^2$$
et il existe $\alpha > 0$ tel que $I = [a-\alpha, a+\alpha]$ est stable par F et $\forall x_0 \in I$, la suite $(x_n)_n$ a une convergence d'ordre 2 vers a sur I.
3. Supposons de plus que $f''(x) > 0$ pour tout $x \in [c,d]$. Alors $I = [a,d]$ est stable par F et $\forall x_0 \in I$, $(x_n)_n$ est strictement décroissante (ou constante) avec $0 \leq x_{n+1} - a \leq C(x_n - a)^2$ et :
$$x_{n+1} - a \sim \frac{1}{2}\frac{f''(a)}{f'(a)}(x_n - a)^2$$
quand $n \longrightarrow +\infty$.

Explications : Pour résoudre $f(x) = 0$, on cherche à transformer l'équation en un problème équivalent de point fixe de la forme $F(x) = x$. Cela peut se faire de bien des manières, par exemple en prenant $F(x) = x + \lambda(x)f(x)$ où λ est une fonction ne s'annulant pas.
Or la convergence des itérés $x_{n+1} = F(x_n)$ vers la solution a cherchée serait très rapide si ce point est superattractif, ie $F'(a) = 0$. Or $f(a) = 0$, d'où $F'(a) = 1 + \lambda(a)f'(a)$ et ceci incite à choisir $\lambda(x) = -\dfrac{1}{f'(x)}$ et à considérer la suite récurrente : $x_{n+1} = x_n - \dfrac{f(x_n)}{f'(x_n)}$.
C'est la méthode de Newton, à l'interprétation géométrique suivante : l'égalité précédente
$$f(x_n) + f'(x_n)(x_{n+1} - x_n) = 0$$
exprime que $x_{n+1} = F(x_n)$ est l'abscisse de l'intersection avec l'axe (Ox) de la droite
$$y = f(x_n) + f'(x_n)(x - x_n)$$
qui est la tangente au graphe de f au point d'abscisse x_n.
Si $e_n = x_n - a$ est l'erreur commise à la n-ième étape de la récurrence, on voit que e_{n+1} sera ainsi un $o(e_n)$ (écart d'une courbe et de sa tangente), en fait un $O(e_n^2)$ si f est deux fois dérivable.
La méthode de Newton donne effectivement une convergence rapide vers a à condition de partir d'un x_0 suffisamment proche de la solution (cf l'assertion (2) du théorème). On peut assouplir cette contrainte si f est une fonction convexe (cf l'assertion (3) du théorème). Tous ces résultat s'obtiennent à partir de l'assertion (1) du théorème.

Démonstration

Étape 1 : démontrons l'assertion (1). Comme f est continue, $f(c) < 0$ et $f(d) > 0$, alors d'après le théorème des valeurs intermédiaires $\exists a \in]c, d[$ tel que $f(a) = 0$.
Ainsi :

$$\begin{aligned} F(a) &= a - \frac{f(a)}{f'(a)} = a \\ F'(x) &= 1 - \frac{f'(x)^2 - f(x)f''(x)}{f'(x)^2} \\ F'(a) &= 1 - 1 = 0 \end{aligned}$$

On s'attend donc à avoir $F(x) - a$ de l'ordre de $(x-a)^2$ et c'est le cas, en effet : comme $f(a) = 0$, on peut écrire pour $x \in [c, d]$:

$$F(x) - a = x - a - \frac{f(x)}{f'(x)} = \frac{f(a) - f(x) - (a-x)f'(x)}{f'(x)}$$

D'après la formule de Taylor Lagrange appliquée à f, on a $\exists z \in [a, x]$ tel que :

$$f(a) = f(x) + (a-x)f'(x) + \frac{1}{2}(a-x)^2 f''(z)$$

D'où :

$$F(x) - a = \frac{1}{2}(a-x)^2 \frac{f''(z)}{f'(x)}$$

Étape 2 : démontrons l'assertion (2). On pose :

$$C = \frac{\max_{[a,d]} |f''|}{2 \min_{[c,d]} f'}$$

alors $\forall x \in [c, d]$:

$$|F(x) - a| \leq C|x-a|^2$$

Soit $\alpha > 0$ assez petit pour que $C\alpha < 1$ et que l'intervalle $I = [a - \alpha, a + \alpha]$ sont contenu dans $[c, d]$. Alors $x \in I$ entraîne $|F(x) - a| \leq C\alpha^2 < \alpha$, d'où $F(I) \subset I$.
Si $x_0 \in I$, on a donc $\forall n$, $x_n \in I$ et :

$$|x_{n+1} - a| = |F(x_n) - a| \leq C|x_n - a|^2$$

D'où :

$$C|x_n - a| \leq (C|x_0 - a|)^{2^n} \leq (C\alpha)^{2^n}$$

et la convergence d'ordre 2 de x_n vers a puisque $C\alpha < 1$.

Étape 3 : démontrons l'assertion (3). La dérivée f' est croissante (car $f'' > 0$), donc f est convexe sur $]c, d[$. Pour $a \leq x \leq d$, on a $f'(x) > 0$ et $f(x) \leq 0$ d'où :

$$F(x) = x - \frac{f(x)}{f'(x)} \leq x$$

avec inégalité stricte si $x > a$.
D'après l'assertion (1) d'autre part, on a :

$$F(x) - a = \frac{1}{2}\frac{f''(z)}{f'(x)}(x-a)^2 \geq 0$$

strictement si $x > a$ (ce qui est clair géométriquement le graphe de f étant au-dessus de ses tangentes).
Ces deux inégalités montrent que $I = [a, d]$ est stable par F et que pour $a < x_0 \leq d$, les itérés x_n vérifient aussi $a < x_n \leq d$ et forment une suite strictement décroissante. Et si $x_0 = a$, la suite est constante.
La suite $(x_n)_n$ admet donc une limite l qui vérifie $F(l) = l$ (décroissante et minorée par a), donc

$f(l) = 0$ et l ne peut qu'être a.
La convergence de x_n vers a est quadratique, car on a comme en (2) :

$$0 \leq x_{n+1} - a \leq C(x_n - a)^2$$

Enfin cette inégalité est essentiellement optimale, ie si $a < x_0 \leq d$, on a $x_n > a$ pour tout n et :

$$\frac{x_{n+1} - a}{(x_n - a)^2} = \frac{1}{2} \frac{f''(z_n)}{f'(x_n)}$$

d'après l'assertion (1) avec $a < z_n < x_n \longrightarrow a$. La fraction tend donc vers $\dfrac{f''(a)}{2f'(a)}$ lorsque $n \longrightarrow +\infty$.

Lemmes utilisés

Lemme 0.1 (Théorème des valeurs intermédiaires) *Soit I un intervalle de $\mathbb{R}$ et soit $f : I \longrightarrow \mathbb{R}$ une application continue, alors $f(I)$ est un intervalle.*
Autrement dit, si $f(a) \leq f(b)$ avec $a < b$, alors pour tout γ tel que $f(a) \leq \gamma \leq f(b)$, il existe $c \in [a, b]$ tel que $f(c) = \gamma$.

Démonstration Comme les parties connexes de $\mathbb{R}$ sont les intervalles, alors I est connexe et l'image d'un connexe par une application continue est connexe, donc $f(I)$ est un intervalle.

Lemme 0.2 (L'image d'un connexe par une application continue) *Soient (E, d) et (E', d') deux espaces métriques. Soit $f : E \longrightarrow E'$ continue. Si E est connexe, alors $f(E)$ est connexe.*

Démonstration Soit B une partie ouverte et fermée de $f(E)$. Il existe alors un ouvert O et un fermé F de E' tels que $B = O \cap f(E)$ et $B = F \cap f(E)$. On a alors $f^{-1}(B) = f^{-1}(O) = f^{-1}(F)$, donc $f^{-1}(B)$ est ouvert et fermé dans E et E étant connexe, $f^{-1}(B) = \emptyset$ ou E, ie $B = \emptyset$ ou $B = f(E)$.

Lemme 0.3 (Les parties connexes de $\mathbb{R}$) *Les parties connexes de $\mathbb{R}$ sont les intervalles de $\mathbb{R}$.*

Lemme 0.4 (Formule de Taylor Lagrange) *Soit $f : [a, b] \longrightarrow \mathbb{R}$ une application de classe C^n sur $[a, b]$ telle que $f^{(n+1)}$ existe sur $]a, b[$. Alors $\exists c \in]a, b[$ tel que :*

$$f(b) = f(a) + (b - a)f'(a) + \ldots + \frac{(b - a)^n}{n!} f^{(n)}(a) + \frac{(b - a)^{n+1}}{(n + 1)!} f^{(n+1)}(c)$$

Démonstration Considérons l'application :

$$\begin{aligned} \phi : [a, b] &\longrightarrow \mathbb{R} \\ x &\longmapsto f(b) - f(x) - (b - x)f'(x) - \ldots - \frac{(b - x)^n}{n!} f^{(n)}(x) - A\frac{(b - x)^{n+1}}{(n + 1)!} \end{aligned}$$

la constante A étant choisie telle que $\phi(a) = \phi(b) = 0$.
Cette application est continue sur $[a, b]$, dérivable sur $]a, b[$ et $\forall x \in]a, b[$:

$$\phi'(x) = -\frac{(b - x)^n}{n!} f^{(n+1)}(x) + A\frac{(b - x)^n}{n!}$$

donc d'après le théorème de Rolle, il existe $c \in]a, b[$ tel que $\phi'(c) = 0$, ce qui s'écrit $A = f^{(n+1)}(c)$, d'où le résultat car $\phi(a) = 0$.

Lemme 0.5 (Convexité et dérivée croissante) *Soit $f : I \longrightarrow \mathbb{R}$ une application dérivable sur I, alors sont équivalents :*
- *f est convexe ;*
- *f' est croissante ;*
- *la courbe représentative de f est au-dessus de ses tangentes.*

Références

[Gou08] Xavier Gourdon. *Analyse.* Ellipses, 2008.
[Rou09] François Rouvière. *Petit guide du calcul différentiel.* Cassini, 2009.