Méthode de Newton

Julie Parreaux

2018-2019

Référence du développement : Rouvière [2, p.140].

Leçons où on présente le développement : 223 (Suite numérique) ; 224 (Développement asymptotique) ; 226 (Suite récurrente) ; 228 (Continuité et dérivabilité).
Leçons où il peut être évoqué : 229 (Fonctions monotones, convexes).

1 Introduction

La méthode de Newton est une méthode numérique itérative qui grâce à une suite récurrente résout l'équation $f(x) = 0$ lorsque la fonction f possède de bonnes propriétés. Elle a également l'avantage d'avoir une interprétation géométrique intuitive et intéressante. L'idée principal de cet algorithme (outre d'utilisé une méthode itérative), est de se ramener à une problème équivalent de point fixe $F(x) = x$ pour lesquels on a de jolis théorèmes qui peuvent s'appliquer (théorème de Picard par exemple). Elle permet de bien montrer l'existence d'une telle solution (elle ne se limite pas à montrer que f à un zéro) en exhibant cette solution. Un énoncé plus général que celui que l'on prouve permet de montrer l'existence de cette solution.

Théorème 1 (Méthode de Newton générale [1, p.59])**.** *Soit $x_0 \in I$, où I est un intervalle de $\mathbb{R}$, supposons qu'il existe deux nombres $c \geq 0$ et $\lambda > 0$ tel que*

1. $|f(x_0)| \leq \frac{c}{2\lambda}$;
2. $\forall x, z \in [x_0 - c, x_0 + c] \subseteq I$, *on a*
   (a) $|f'(x)| \geq \frac{1}{\lambda}$;
   (b) $|f'(x) - f(x)| \geq \frac{1}{2\lambda}$.

Dans ces conditions, il existe une unique solution α de l'équation $f(x) = 0$, $\alpha \in [x_0 - c, x_0 + c]$. En outre, si $(z_n)_{n \in \mathbb{N}}$ est une suite quelconque de points de $[x_0 - c, x_0 + c]$, on peut définir une suite $(x_n)_{n \in \mathbb{N}}$ de points de cet intervalle telle que

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

et la limite de cette suite est α.

Cette méthode a de nombreuses applications que ce soit en mathématiques ou en informatique. En effet, à chaque étape elle est d'une complexité raisonnable (si la complexité de f et de sa dérivée l'est aussi) car on ne fait "qu'une division". De plus, à l'aide du Jacobien, nous pouvons étendre facilement la méthode à Newton à $\mathbb{R}^d$ même si on en parle pas ici. Grâce à sa convergence quadratique et sa complexité raisonnable, la méthode de Newton reste une des méthodes les plus utilisée.

Le choix de F est un choix important pour garantir cette vitesse de convergence. Cette transformation peut se faire de bien des manières. Généralement, on pose $F(x) = x + \lambda(x)f(x)$ où λ est une fonction qui ne s'annule pas. De plus, la vitesse de convergence de la suite $F(x_n) = x_{n+1}$ que l'on va considéré pour résoudre le système est bien plus importante si a (la solution de l'équation) est un point fixe superattractif. Dans le cadre de notre problème, on est amenée à choisir, $\lambda(x) = -\frac{1}{f'(x)}$ (d'où la suite récurrente).

2 Méthode de Newton

On donne la version du théorème que l'on souhaite montrer. Dans notre version, l'existence d'une solution est prouvée mais sous des hypothèse bien plus forte que dans le cas de la version générale (un TVI est appliqué).

Théorème 2 (Méthode de Newton développée)**.** *Soit $f : [c,d] \to \mathbb{R}$ une fonction de classe $\mathcal{C}^2$. On suppose $c < d$ et $f(c) < 0 < f(d)$ et $\forall x \in [c,d], f'(x) > 0$. On considère la suite récurrente suivante :*

$$x_{n+1} = F(x_n), n \geq 0 \quad \text{avec} \quad F(x) = x - \tfrac{f(x)}{f'(x)}$$

Alors,

1. *f a un unique zéro en a.*
2. *$\forall x \in [c,d], \exists \alpha > 0$ tel que la suite $(x_n)_{n\in\mathbb{N}}$ a une convergence quadratique vers a dans $I = [a-\alpha, a+\alpha]$.*
3. *De plus, si $f'' > 0$, on a l'équivalent suivant :*

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

(L'ajout de l'hypothèse de convexité à la méthode de Newton, nous permet de nous affranchir de l'intervalle I tout en gardant la même vitesse de convergence.)

Schéma du développement

1. f admet un unique zéro en a.
2. $\forall x \in [c,d], \exists z \in [a,x]$ tel que $F(x) - a = \frac{f''(z)}{f'(x)}(x-a)^2$.
3. Il existe $C > 0$ tel que $\forall x \in [a,d], |F(x) - a| \leq C\,|x-a|^2$.
4. $\forall x \in [c,d], \exists \alpha > 0$ tel que la suite $(x_n)_{n\in\mathbb{N}}$ a une convergence quadratique vers a dans I.
5. On suppose que $f'' > 0$. On a l'équivalent suivant : $x_{n+1} - a \underset{n\to+\infty}{\sim} \frac{1}{2}\frac{f''(a)}{f'(a)}(x_n-a)^2$.

Démonstration. Soit $f : [c,d] \to \mathbb{R}$ une fonction de classe $\mathcal{C}^2$. On suppose $c < d$ et $f(c) < 0 < f(d)$ et $\forall x \in [c,d], f'(x) > 0$. On considère la suite récurrente suivante :

$$x_{n+1} = F(x_n), n \geq 0 \quad \text{avec} \quad F(x) = x - \tfrac{f(x)}{f'(x)}$$

On considère son interprétation géométrique qui est somme toute très intuitive dont voici un exemple (Figure 1).

Étape 1 : Montrons que f admet un unique zéro en a. Comme f est continue sur $[c,d]$ (par hypothèse) et croit strictement ($f'(x) > 0$) de $f(c) < 0$ à $f(d) > 0$ (par hypothèse), elle s'annule en un unique point $a \in]c,d[$ (par le TVI (assure l'existence) et la stricte croissance (assure l'unicité).

Étape 2 : Montrons que $\forall x \in [c,d], \exists z \in [a,x]$ tel que $F(x) - a = \frac{f''(z)}{f'(x)}(x-a)^2$. Par calcul direct, on a

$$F(a) \underbrace{=}_{f'(a)>0 \text{ et définition}} a - \frac{f(a)}{f'(a)} \underbrace{=}_{f(a)=0, \text{ par hypothèse}} a$$

et

$$F'(a) \underbrace{=}_{f'(a)>0 \text{ et déivée}} 1 - \frac{(f'(a))^2 - f''(a)f(a)}{(f'(a))^2} \underbrace{=}_{f(a)=0, \text{ par hypothèse}} 1 - \frac{(f'(a))^2}{(f'(a))^2} = 0$$

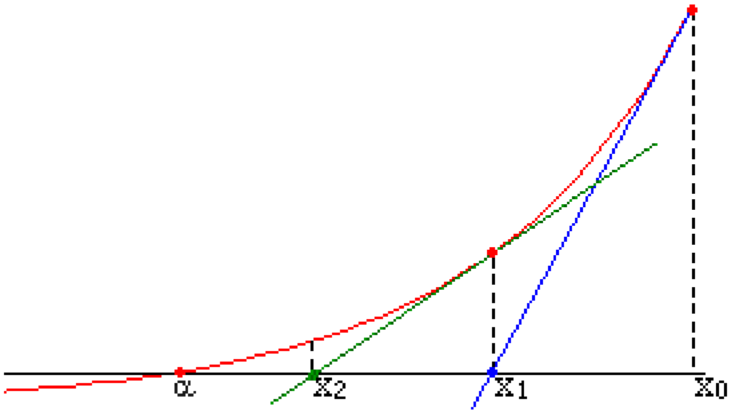

FIGURE 1 – Illustration géométrique de la méthode de Newton sur une fonction convexe.

d'où

$$F(x)-a \underbrace{=}_{\text{définition}} x-a-\frac{f(x)}{f'(x)} \underbrace{=}_{f(a)=0} x-a-\frac{f(x)-f(a)}{f'(x)} \underbrace{=}_{\text{même dénominateur}} \frac{f(a)-f(x)-(a-x)f'(x)}{f'(x)}$$

On applique alors la formule de Taylor–Lagrange à l'ordre deux d'origine x et d'extrémité a, alors il existe $z \in]a,x[$ tel que

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

Étape 3 : Montrons qu'il existe $C>0$ tel que $|F(x)-a| \leq C\,|x-a|^2$, $\forall x \in [a,d]$ et qu'il existe $\alpha>0$ tel que $I=[a-\alpha,a+\alpha]$ soit stable par F. On pose $C=\frac{\max_{x\in[c,d]}|f''(x)|}{2\min_{x\in[c,d]}f'(x)}$. Par ce qui précède (de l'étape 2), on en déduit que

$$\begin{array}{rcll}
|F(x)-a| & = & \left|\frac{1}{2}\frac{f''(z)}{f'(x)}(a-x)^2\right| & \text{(par l'étape 2)}\\
 & = & \left|\frac{1}{2}\frac{f''(z)}{f'(x)}\right|\,|a-x|^2 & \text{(par homogénéité de la valeur absolue)}\\
 & \leq & \frac{1}{2}\frac{\max_{x\in[c,d]}|f''(x)|}{\min_{x\in[c,d]}f'(x)}(a-x)^2 & (|f''(z)| \leq \max_{x\in[c,d]}|f''(x)| \text{ et } \max_{x\in[c,d]}|f'(x)|>0)\\
 & & & (|f'(x)| \geq \min_{x\in[c,d]}f'(x)>0)\\
 & \leq & C\,|a-x|^2 & \text{(définition de C)}
\end{array}$$

Soit $\alpha>0$ tel que $C\alpha<1$ et assez petit pour que $I=[a-\alpha,a+\alpha]\subset[c,d]$. Soit $x\in I$, alors on a :

$$\begin{array}{rcll}
|F(x)-a| & \leq & C\,|x-a|^2 & \text{(par l'inégalité précédent)}\\
 & \leq & C\alpha^2 & (x\in I \Rightarrow |x-a|\leq\alpha)\\
 & < & \alpha & (C\alpha<1, \text{ par hypothèse})
\end{array}$$

Donc, par définition de I, $F(I)\subset I$.

Étape 4 : Montrons que $\forall x\in[c,d]$, $\exists\alpha>0$ tel que la suite $(x_n)_{n\in\mathbb{N}}$ a une convergence quadratique vers a dans I. Par la précédente étape, on a $F(I)\subset I$. On a donc si $x_0\in I$, alors $\forall n$, $x_n\in I$. On va alors en déduire la convergence. Par récurrence sur $n\in\mathbb{N}$, on montre $C\,|x_n-a| \leq (C\,|x_0-a|)^{2^n}$ (en effet, l'étape de récurrence s'écrit comme suit

$$\begin{array}{rcll}
C\,|x_{n+1}| & = & C\,|F(x_n)-a| & \text{égalité de l'étape 3}\\
 & \leq & C\left(C\,|x_n-a|^2\right) & \\
 & = & (C\,|x_n-a|)^2 & \text{inégalité de l'étape 3}\\
 & \leq & (C\,|x_0-a|)^{2^{n+1}} & \text{hypothèse de récurrence}
\end{array}$$

L'initialisation se fait trivialement d'où la récurrence). On en déduit alors,

$$C\,|x_n - a| \underbrace{\leq}_{\text{récurrence}} \left(C\, \underbrace{|x_0 - a|}_{\leq \alpha} \right)^{2^n} \leq \left(C \underbrace{\alpha}_{\leq 1} \right)^{2^n}$$

Comme $C\alpha < 1$, la convergence de $(x_n)_{n\in\mathbb{N}}$ est quadratique, si $x_0 \in I$.

Étape 5 : On suppose de plus que $f'' > 0$. Montrons qu'on a l'équivalent suivant : $x_{n+1} - a \sim \frac{1}{2}\frac{f''(a)}{f'(a)}(x_n - a)^2$ si $n \to +\infty$. Comme $f''(x) > 0$, f' est croissante et f est convexe sur $]c, d[$.

Soit $a \leq x \leq d$. On a $f(x) \leq 0$ (par hypothèse sur a et par croissance stricte de f) et $f'(x) > 0$ (par hypothèse sur f'). On en déduit que

$$F(x) \underbrace{=}_{\text{Définition de } F} x - \frac{f(x)}{f'(x)} \underbrace{\leq}_{\frac{f(x)}{f'(x)} \leq 0} x$$

On a de plus,

$$F(x) - a \underbrace{=}_{\text{égalité de l'étape 2}} \frac{1}{2} \frac{\overbrace{f''(z)}^{>0}}{\underbrace{f'(z)}_{>0}} \underbrace{(x-a)^2}_{>0} \geq 0$$

Comme I est stable par F (par l'étape 4), si $a < x_0 < d$ alors $a < x_n < d$. Donc la suite $(x_n)_{n\in\mathbb{N}}$ est strictement décroissante. Dans le cas $x_0 = a$, la suite $(x_n)_{n\in\mathbb{N}}$ est constante. Donc, dans tous les cas, la suite $(x_n)_{n\in\mathbb{N}}$ admet une limite l telle que $F(l) = l$, c'est-à-dire $f(l) = 0$, donc $l = a$.

La convergence vers a est quadratique et on a $0leqx_{n+1} - a \leq C(x_n - a)^2$ d'après l'étape 3. Enfin, cette inégalité est essentiellement optimale : si $a < x_0 \leq d$, on a $x_n > 0$ pour tout n, on peut donc diviser l'expression de $x_{n+1} - a$ (de l'étape 3) par $(x_n - a)^2$ qui est toujours non nul,

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

d'après 2, avec $a < z_n < x_n$. On a alors $\lim_{n\to\infty} \frac{x_{n+1}-a}{(x_n-a)^2} = \frac{1}{2}\frac{f''(a)}{f(a)}$. D'où le résultat.

Remarque : On est dans le cas, f convexe. Dans ce cas on peut adoucir les hypothèses de la méthode de Newton sans diminuer sa vitesse de convergence.

□

3 Compléments : notion de points fixes

Le théorème de point fixe joue un rôle central en calcul différentiel, comme clef des théorèmes des fonctions inverses, des fonctions implicites et du théorème de Cauchy-Lipschitz [2, p.137]. Ce théorème est puissant (avec une preuve relativement simple) puisqu'il donne existence, unicité et une bonne méthode d'approximation de la solution.

Définition 1. Soient X un espace métrique complet (non vide), d la distance de X, et F une application de X dans lui-même. On dit que F est contractante si et seulement si, il existe une constante positive $k < 1$ telle que pour tout $x, y \in X$, $d(F(x), F(y)) \leq kd(x, y)$.

Théorème 3. *Soient X un espace métrique complet (non vide), d la distance de X, et F une application de X dans lui-même. On suppose F contractante.*

Alors, il existe un unique point fixe $a \in X$ tel que $F(a) = a$ (point fixe de F). De plus, ce point peut s'obtenir comme limite de la suite $(x_n)_{n\in\mathbb{N}}$ des itérés, définie par récurrence à partir d'un point quelconque $x_0 \in X$ selon $x_{n+1} = F(x_n)$. On a plus précisément pour $n \geq 1$, $d(x_n, a) \leq \frac{k^n}{1-k} d(x_0, x_1)$.

Démonstration [2, p.159]. Pour $n \geq 1$ on a $d(x_n, x_{n+1}) = d(F(x_{n-1}), F(x_n)) \leq kd(x_{n-1}, x_n)$ (car F est contractante). Par récurrence, on en déduit que $d(x_n, x_{n+1}) \leq k^n d(x_0, x_1)$. Par inégalité triangulaire, on a alors, pour $n \geq 0$ et $n \geq 0$, $d(x_n, x_{n+p}) \leq (k^n + k^{n+1} + \cdots + k^{n+p})\, d(x_0, x_1) \leq \frac{k^n}{1-k} d(x_0, x_1)$ (sommation d'une série géométrique de raison $k < 1$).

On a donc $d(x_n, x_{n+p}) \leq \epsilon$ pour n assez grand et $p \geq 0$ (par l'inégalité précédente et $k < 1$). Les x_n forment une suite de Cauchy de X qui converge vers un point a. En faisant tendre p vers $+\infty$ dans l'inégalité ci-dessous, on obtient $d(x_n, a) \leq \frac{k^n}{1-k} d(x_0, x_1)$ (caractère Cauchy de la suite).

Enfin, $x_{n+1} = F(x_n)$ tend vers a (par ce qui précède) et vers $F(a)$ (F est continue car contractante). D'où $F(a) = a$ et a est un point fixe de F.

Pour l'unicité, s'il existait un deuxième point fixe b, alors $d(a, b) = d(F(a), F(b)) \leq kd(a, b)$ d'où $d(a, b) = 0$ et par les propriétés de la distance $a = b$. $\square$

Contre-exemple : Soient $X =]0, 1[$ et $F(x) = \frac{x}{2}$ une application contractante de X dans lui-même. Cependant F est sans points fixes car X n'est pas complet.

Contre-exemple : Soient $X = [0, 1]$ un espace complet et $F(x) = \sqrt{x^2 + 1}$ une application contractante. Cependant F est sans points fixes car F ne s'applique pas de X dans lui-même puisque $F(X) = \left[1, \sqrt{2}\right]$.

Contre-exemple : Soient $X = \mathbb{R}$ un espace complet et $F(x) = \sqrt{x^2 + 1}$ de X dans lui-même. Cependant F est sans points fixes car F n'est pas contractante dans X même si $|F(x) - F(y)| < |x - y|$ pour $x \neq y$.

Contre-exemple : Soient $X = \left[0, \frac{\pi}{2}\right]$ un espace complet et $F(x) = \sin x$ de X dans lui-même mais non-contractante. Cependant F possède un unique point fixe dont la suite des itérés converge très lentement.

Contre-exemple : Soient X un espace complet quelconque et $F(x) = x$ de X dans lui-même mais non-contractante. Cependant tout point de X est point fixe de F.

Corollaire 1 ([2, p.159]). *Soient X un espace métrique complet et F une application de X dans lui-même. On suppose qu'une certaine itéré F^p est contractante, où $p \geq 1$.*

Alors, F a un unique point fixe qui est limite de la suite $(F^n(x_0))_{n\in\mathbb{N}}$ avec $x_0 \in X$ quelconque. De plus, ka vitesse de convergence de cette suite est géométrique selon les puissances de $k_p^{\frac{1/p}{den}}$ où k_p est la constante de Lipschitz de F^p.

Démonstration. Le théorème 3 donne a, unique point fixe de F^p et limite de la suite des $(F^{np}(x_0))_{n\geq 0}$ pour tout $x_0 \in X$. Montrons alors que a est l'unique point fixe de F. Comme $F(a) = F(F^p(a)) = F^{p+1}(a) = F^p(F(a))$, le point $F(a)$ est aussi point fixe de F^p. D'où par l'unicité de ce point fixe, $F(a) = a$, soit a est un point fixe de F. Inversement, tout point fixe de F est point fixe de F^p. D'où par l'unicité sur F^p, a est l'unique point fixe de F.

D'après le théorème 3, on a $d\left(F^{np}(x_0), a\right) \leq \frac{k_p^n}{1-k_p} d(x_0, F^p(x_0))$. En remplaçant le point initial x_0 par $F^q(x_0)$ on en déduit $d\left(F^{np+q}(x_0), a\right) \leq \frac{k_p^n}{1-k_p} d\left(F^q(x_0), F^{p+q}(x_0)\right)$ pour $q \in \{0, \ldots, p-1\}$ et $n \geq 0$. Par division euclidienne, tout entier m s'écrit $m = np + q$ avec $0 \leq q \leq p - 1$. Alors, $n > \left(\frac{m}{p}\right) - 1$ et on obtient facilement $d(F^m(x_0), a) \leq Ck_p^{\frac{m}{p}}$, pour $m \in \mathbb{N}$ où C est une constante indépendante de m. La suite des itérés $F^m(x_0)$ à partir d'un x_0 quelconque, converge donc vers le point fixe. La vitesse de convergence est donc géométrique selon les puissances de $k_p^{\frac{1/p}{den}}$. $\square$

Remarque : Ce raffinement du résultat possède quelques applications que nous allons détailler ici (dans une preuve de Cauchy–Lipschitz par exemple).

Application (Point fixe et équations intégrales [2, p.175]) : Soient $I = [a, b]$ un intervalle compact, $K : I \times I \to \mathbb{R}$ une fonction continue et E l'espace des fonctions réelles continues sur I, muni de la norme de la convergence uniforme. On se donne $\varphi \in E$. Pour tout $t \in I$,

1. l'équation de Fredholm $x(t) = \varphi(t) + \int_a^b K(s, t)x(s)ds$ admet une unique solution $x \in E$ si $(b - a) \max_{s,t\in I} |K(s, t)| < 1$.
2. l'équation $x(t) = \varphi(t) + \int_0^1 \lambda x(s)ds$ admet une solution unique si $\lambda \neq 1$ et admet une solution si $\lambda = 1$ et $\int_0^1 \varphi(t)dt = 0$.
3. l'équation de Volterra $x(t) = \varphi(t) + \int_a^t K(s, t)x(s)ds$ admet toujours une unique solution $x \in E$.

Démonstration.

1. Soit $F(x)(t) = \varphi(t) + \int_a^b K(s, t)x(s)ds$ pour $t \in I$. A chaque $x \in E$, l'opérateur F associe la fonction $F(x) \in E$ et le problème s'écrit donc $F(x) = x$. On a
$$|F(x)(t) - F(y)(t)| = \left|\int_a^b K(s, t)(x(s) - y(s))ds\right| \leq (a - b)\, \|K\|_\infty \|x - y\|_\infty$$
D'où $\|F(x) - F(y)\|_\infty \leq (a - b)\, \|K\|_\infty \|x - y\|_\infty$. On en déduit que F est contractante sur E si $(b - a)\|K\|_\infty = (b - a) \max_{s,t\in I} |K(s, t)| < 1$. Par complétude de E, lorsque F est contractante, elle admet un unique point fixe.
2. L'équation proposée entraîne que $x(t) = \varphi(t) + c$ où c est une constante. En intégrant de 0 à 1 l'intégrale donnée, il vient que $\int_0^1 x(t)dt = \int_0^1 \varphi(t)dt + \lambda \int_0^1 x(s)ds$, d'où en reportant $x = \varphi + c$, $(1 - \lambda)c = \lambda \int_0^1 \varphi(t)dt$.
   — Si $\lambda \neq 1$, il y a une unique solution quelle que soit $\varphi \in E$: $x(t) = \varphi(t) + \frac{\lambda}{1-\lambda} \int_0^1 \varphi(s)ds$.

— Si $\lambda = 1$, il y a une solution si et seulement si $\int_0^1 \varphi(t)dt = 0$: $x(t) = \varphi(t) + c$ où c est une constante arbitraire.

3. Soit $F(x)(t) = \varphi(t) + \int_a^t K(s,t)x(s)ds$, l'équation intégrale devient donc $F(x) = x$. On a pour $x, y \in E$ et $a \leq t \leq b$,
$$|F(x)(t) - F(y)(t)| = \left|\int_a^b K(s,t)(x(s) - y(s))ds\right| \leq \|K\|_\infty \int_a^b |x(s) - y(s)|\, ds \leq (t-a)\|K\|_\infty \|x - y\|_\infty$$
On en déduit par récurrence sur $p \geq 1$, $|F^p(x)(t) - F^p(y)(t)| \leq \frac{((t-a)\|K\|_\infty)^p}{p!}\|x-y\|_\infty$ et finalement $\|F^p(x)(t) - F^p(y)(t)\|_\infty \leq \frac{((b-a)\|K\|_\infty)^p}{p!}\|x-y\|_\infty$.
Si p est choisi suffisamment grand, F^p est contractante sur E donc admet un unique point fixe x, qui est aussi l'unique point fixe de F.

□

Références

[1] J. Dieudonné. *Calcul infinitésimal*. Hermann éditeurs des sciences et des arts, 1992.

[2] F. Rouvière. *Petit guide de calcul différentiel à l'usage de la licence et de l'agrégation*. Cassini, 3^{me} edition, 2009.