COURS UM6P
ANNEE 2020-2021

Laurent Dumas

CHAPITRE 1

INTERPOLATION POLYNOMIALE

CHAPITRE 2

RESOLUTION DE SYSTEMES NON LINEAIRES

L'objectif de ce chapitre est de présenter différentes méthodes de résolution approchée d'une équation $f(x) = 0$ où f est une fonction définie sur un sous-ensemble A de $\mathbf{R}^n$ et à valeurs dans $\mathbf{R}^n$. Dans le premier paragraphe, le problème est retraduit en termes de recherche de points fixes pour une fonction auxiliaire g à préciser afin de définir la famille des méthodes itératives. La méthode de Newton peut alors être construite puis étudiée en détail dans le paragraphe suivant. Le paragraphe 3 présente ensuite trois autres méthodes de recherche de zéros d'une fonction réelle. Enfin, le paragraphe 4 propose un procédé général d'accélération de convergence d'une suite pouvant être appliqué aux suites itératives précédentes.

2.1 Méthodes itératives

2.1.1 Théorème du point fixe

Commençons par rappeler le théorème classique suivant dû à Picard :

Théorème 1. *soit A un fermé de $\mathbf{R}^n$ et g une application de A dans $\mathbf{R}^n$ telle que :*
(i) $g(A) \subset A$.
(ii) g est contractante pour une certaine norme de $\mathbf{R}^n$, à savoir :

$$\exists K \in [0,1[, \quad \forall (x,y) \in A^2, \quad ||g(x) - g(y)|| \le K||x - y||.$$

Alors, g admet un unique point fixe $\bar{x} \in A$. De plus, la suite $(x_n)_{n\in\mathbf{N}}$ construite par récurrence :

$$x_0 \in A \quad \text{et} \quad x_{n+1} = g(x_n) \quad (n \in \mathbf{N}) \tag{1}$$

converge vers $\bar{x}$ et

$$\forall k \in \mathbf{N}, \quad \forall l \in \{0, \ldots, k\}, \quad ||x_k - \bar{x}|| \leq \frac{K^{k-l}}{1-K}||x_{l+1} - x_l|| \tag{2}$$

Démonstration. L'unicité du point fixe est immédiate avec l'hypothèse (ii). Pour démontrer son existence, on considère la suite $(x_n)_{n\in\mathbf{N}}$ définie correctement par la relation (1). Grâce à l'hypothèse (ii), on remarque que :

$$\forall n \in \mathbf{N}^*, \ ||x_{n+1} - x_n|| = ||g(x_n) - g(x_{n-1})|| \leq K||x_n - x_{n-1}||\ldots \leq K^n||x_1 - x_0||$$

puis en utilisant l'inégalité triangulaire

$$\forall n \in \mathbf{N}, \ \forall k \in \mathbf{N}, \ k \leq n \Rightarrow ||x_n - x_k|| \leq \sum_{i=0}^{n-k-1} K^{k+i}||x_1 - x_0|| \leq \frac{K^k}{1-K}||x_1 - x_0|| \tag{3}$$

La suite $(x_n)_{n\in\mathbf{N}}$ est donc de Cauchy dans le sous-espace complet A (fermé dans un espace complet) et converge ainsi vers $\bar{x} \in A$. Par passage à la limite dans (1), la fonction g étant continue (car K-Lipschitzienne), on a bien $g(\bar{x}) = \bar{x}$.

En passant ensuite à la limite dans l'inégalité (3) lorsque $n \to +\infty$, on obtient la relation (2) pour $l = 0$. Le cas général $l > 0$ est obtenu par un simple décalage d'indice. □

Le corollaire suivant est souvent utilisé dans la pratique :

Corollaire 1. *En conservant les notations du Théorème 1, on suppose que la fonction g vérifie (i) et la nouvelle proposition (ii)' :*

(ii)' il existe un entier $n_0 > 0$ tel que g^{n_0} (g itérée n_0 fois) soit contractante. Alors, g admet un unique point fixe.

Démonstration. L'unicité d'un point fixe de g provient aisément de l'hypothèse (ii)'. Pour démontrer son existence, on applique le Théorème 1 à la fonction g^{n_0} : soit $\bar{x} \in A$ tel que

$$g^{n_0}(\bar{x}) = \bar{x}.$$

Comme $g(\bar{x})$ est aussi un point fixe de g^{n_0}, on a par unicité de ce dernier :

$$g(\bar{x}) = \bar{x} \quad \square$$

Remarque : L'hypothèse (ii)' est strictement plus faible que l'hypothèse (ii) : Il suffit pour cela de considérer la fonction $g : \begin{pmatrix} \mathbf{R}^2 \to \mathbf{R}^2 \\ (x_1, x_2) \mapsto (2x_2, 0) \end{pmatrix}$. On observe que $g^2 = 0$ est contractante tandis que g ne l'est pas. Par contre, (0,0) est bien l'unique point fixe de g et de g^2.

2.1.2 Notion d'ordre

Le théorème précédent permet de construire par récurrence une suite d'approximations $(x_n)_{n\in\mathbf{N}}$ convergeant vers le point fixe $\bar{x}$ d'une fonction g, stable et contractante sur un ensemble fermé A de $\mathbf{R}^n$. Afin de préciser la vitesse de convergence de cette suite, on note $e_n = x_n - \bar{x}$, l'erreur commise au rang n, et on définit la notion d'ordre de convergence :

Définition 1. *La convergence de la suite $(x_n)_{n\in\mathbf{N}}$ vers $\bar{x}$ est (au moins) d'ordre $p \in \mathbf{N}^*$ si*

$$\exists C \in \mathbf{R}_+^*,\ \exists n_0 \in \mathbf{N}^*,\ \forall n \in \mathbf{N},\ n \geq n_0 \Rightarrow ||e_{n+1}|| \leq C||e_n||^p$$

(avec $C < 1$ si $p = 1$).

On observe que la convergence d'ordre $p > 1$ d'une suite permet de multiplier approximativement par p à chaque étape le nombre de décimales exactes de $\bar{x}$ calculées avec la suite en question (lorsque $p = 1$ on rajoute simplement une décimale exacte après chaque itération).

Dans le cadre général du Théorème 1, la convergence est d'ordre 1 (on dit aussi géométrique). Cet ordre peut être amélioré dans certains cas lorsque g est deux fois différentiable :

Théorème 2. *: les notations et hypothèses (i) et (ii) du Théorème 1 sont conservées. On suppose en outre que $g \in C^2(A, \mathbf{R}^n)$ et $\bar{x}$ appartient à l'intérieur de A. On peut préciser la vitesse de convergence de la suite $(x_n)_{n\in\mathbf{N}}$ dans deux cas :*

(i) si $Dg(\bar{x})$ (différentielle de g en $\bar{x}$) est un isomorphisme de $\mathbf{R}^n$, la convergence est exactement d'ordre 1, à savoir :

$$\exists (C_1, C_2) \in]0,1[^2,\ \exists n_0 \geq 1,\ \forall n \in \mathbf{N},\ n \geq n_0 \Rightarrow C_1||e_n|| \leq ||e_{n+1}|| \leq C_2||e_n||$$

(ii) si $Dg(\bar{x})$ est l'endomorphisme nul, la convergence est au moins d'ordre 2 (on dit aussi quadratique).

Démonstration. On écrit la formule de Taylor-Young à l'ordre 2 entre x_n et $\bar{x}$ pour la fonction g :

$$e_{n+1} = g(x_n) - g(\bar{x}) = Dg(\bar{x}) \cdot e_n + \frac{1}{2} D^2 g(\bar{x})(e_n, e_n) + ||e_n||^2 \varepsilon(e_n)$$

où ε est une fonction continue définie au voisinage de $0 \in \mathbf{R}^n$ et nulle en ce point.

Sachant déjà que e_n tend vers 0, on peut écrire :

$$\exists C > 0, \exists n_0 \in \mathbf{N}^*, \forall n \in \mathbf{N}, n \geq n_0 \Rightarrow ||e_{n+1} - Dg(\bar{x}) \cdot e_n|| \leq C||e_n||^2.$$

Le cas (ii) où $Dg(\bar{x}) \equiv 0$ est ainsi démontré. Dans le cas (i) où $Dg(\bar{x})$ est un isomorphisme, soit correctement $n_1 \in \mathbf{N}$ tel que

$$\forall n \in \mathbf{N}, n \geq n_1 \Rightarrow C||e_n|| \leq \frac{1}{2} \inf_{v \in S(0,1)} ||Dg(\bar{x}) \cdot v||$$

où $S(0,1)$ représente la sphère unité de $\mathbf{R}^n$. Alors, pour $n \geq \max(n_0, n_1)$,

$$\begin{aligned} \frac{||e_{n+1}||}{||e_n||} = \frac{||Dg(\bar{x}) \cdot e_n + (e_{n+1} - Dg(\bar{x}) \cdot e_n)||}{||e_n||} &\geq \inf_{v \in S(0,1)} ||Dg(\bar{x}) \cdot v|| - C||e_n|| \\ &\geq \frac{1}{2} \inf_{v \in S(0,1)} ||Dg(\bar{x}) \cdot v|| > 0. \end{aligned}$$

Il suffit alors de prendre $C_1 = \frac{1}{2} \inf_{v \in S(0,1)} ||Dg(\bar{x}) \cdot v||$ (et $C_2 = K$) pour achever la démonstration de ce cas □

2.1.3 Cas particulier des fonctions réelles

L'ordre de convergence peut être précisé lorsque g est une fonction réelle suffisamment régulière :

Théorème 3. *: soit $g \in C^1(\mathbf{R}, \mathbf{R})$ et $\bar{x}$ un point fixe de g tel que $|g'(\bar{x})| < 1$ (point fixe attractif). Alors, il existe un voisinage V de $\bar{x}$ tel que la suite $(x_n)_{n \in \mathbf{N}}$ des itérées construite à partir de g suivant (1) avec $x_0 \in V$, converge vers $\bar{x}$.*
(i) Si $g'(\bar{x}) \neq 0$, la convergence est exactement géométrique.
(ii) S'il existe un entier $p \geq 2$ et un voisinage V de $\bar{x}$ tel que $g \in C^p(V, \mathbf{R})$ et

$$g'(\bar{x}) = g''(\bar{x}) = \ldots = g^{(p-1)}(\bar{x}) = 0, \quad g^{(p)}(\bar{x}) \neq 0,$$

alors la convergence est exactement d'ordre p.

Démonstration. On remarque tout d'abord que l'hypothèse $|g'(\bar{x})| < 1$ permet de déduire l'existence d'un voisinage fermé V de $\bar{x}$ sur lequel $g_{/V}$ soit stable et contractante (en utilisant la continuité de g' et l'inégalité des accroissements finis). Ceci permet d'assurer grâce au Théorème 1 la convergence de la suite $(x_n)_{n\in\mathbf{N}}$ vers $\bar{x}$ lorsque $x_0 \in V$. Si de plus $g'(\bar{x}) \neq 0$ (cas (i)), on peut en outre supposer que

$$\exists (C_1, C_2) \in]0,1[^2,\ \forall x \in V,\ C_1 \leq |g'(x)| \leq C_2$$

et à nouveau grâce à l'inégalité des accroissements finis,

$$\forall n \in \mathbf{N}, \quad C_1|x_n - \bar{x}| \leq |x_{n+1} - \bar{x}| \leq C_2|x_n - \bar{x}|$$

et la convergence est exactement géométrique. Dans le cas (ii), on écrit l'égalité de Taylor-Lagrange à l'ordre p entre x_n et $\bar{x}$:

$$x_{n+1} - \bar{x} = \frac{g^{(p)}(y_n)}{p!}(x_n - \bar{x})^p \qquad (y_n \in]x_n, \bar{x}[)$$

et la conclusion suit aisément en utilisant la continuité de $g^{(p)}$ □

2.2 Méthode de Newton

2.2.1 Définition générale

Théorème 4. *: on considère une fonction $f \in C^3(A, \mathbf{R}^n)$ (avec A ouvert de $\mathbf{R}^n$) et $\bar{x} \in A$ tel que $f(\bar{x}) = 0$. Si $Df(\bar{x})$ est un isomorphisme, alors il existe un voisinage $V \subset A$ de $\bar{x}$ telle que la suite $(x_n)_{n\in\mathbf{N}}$ construite à partir de la formule (1) pour la fonction $g : \begin{pmatrix} V \to \mathbf{R}^n \\ x \mapsto x - Df(x)^{-1} \cdot f(x) \end{pmatrix}$ et avec $x_0 \in V$ soit bien définie et converge vers $\bar{x}$ de manière quadratique.*

Démonstration. Comme $f \in C^3(A, \mathbf{R}^n)$ et $Df(\bar{x})$ est un isomorphisme, on sait qu'il existe un voisinage fermé V de $\bar{x}$ tel que la fonction g de l'énoncé soit correctement définie sur V. Par construction, on a $g(\bar{x}) = \bar{x}$. De plus, $Dg(\bar{x}) = 0$: en effet,

$$Dg(\bar{x}) = I - D(Df(\bar{x})^{-1})(\bar{x}) \cdot f(\bar{x}) - Df(\bar{x})^{-1} \cdot Df(\bar{x}) = I - 0 - I = 0.$$

On peut donc supposer en outre que V est stable par g. Il suffit alors d'invoquer le Théorème 3 pour en déduire que la suite $(x_n)_{n\in\mathbf{N}}$ est bien définie et converge quadratiquement vers $\bar{x}$ □

2.2.2 Cas particulier des fonctions réelles

Si on se place dans le cas où $n=1$, on peut préciser le Théorème 5 :

Théorème 5. *: soit $f \in C^2(A,\mathbf{R})$ (A ouvert de $\mathbf{R}$) et $\bar{x} \in A$ tel que $f(\bar{x})=0$.*
(i) On suppose que $f'(\bar{x}) \neq 0$. Alors, il existe un voisinage V de $\bar{x}$ telle que la suite $(x_n)_{n\in\mathbf{N}}$ construite par la formule

$$x_0 \in V \quad \text{et} \quad x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad (n \in \mathbf{N}) \tag{4}$$

existe et converge vers $\bar{x}$ de manière quadratique.
(ii) Si $f'(\bar{x})=0$ et $f''(\bar{x}) \neq 0$, la suite $(x_n)_{n\in\mathbf{N}}$ est encore bien définie par la formule (4) dans un voisinage V de $\bar{x}$ (en supposant de plus $x_{n+1}=\bar{x}$ si $x_n=\bar{x}$) et converge géométriquement vers $\bar{x}$.

Démonstration. La partie (i) du théorème concernant l'existence et la convergence de la suite $(x_n)_{n\in\mathbf{N}}$ n'est qu'une récriture du Théorème 4. On peut également remarquer directement que

$$\forall n \in \mathbf{N},\ \exists y_n \in]x_n,\bar{x}[,\ e_{n+1} = x_{n+1}-\bar{x} = e_n^2 \frac{f''(y_n)}{2f'(x_n)} \tag{5}$$

en écrivant la relation de Taylor-Lagrange à l'ordre 2 entre $\bar{x}$ et x_n :

$$0 = f(\bar{x}) = f(x_n) + (\bar{x}-x_n)f'(x_n) + \frac{1}{2}(\bar{x}-x_n)^2 f''(y_n) \quad (y_n \in]x_n,\bar{x}[)$$

et en utilisant la relation de récurrence (4). Dans le cas (ii) où $f'(\bar{x})=0$ et $f''(\bar{x}) \neq 0$, on peut considérer un voisinage fermé V de $\bar{x}$ tel que

$$\forall x \in V,\ x \neq \bar{x} \Rightarrow f'(x) \neq 0$$

et constuire la fonction g sur V telle que

$$g(x) = \begin{cases} \bar{x} & \text{si} \quad x = \bar{x} \\ x - \dfrac{f(x)}{f'(x)} & \text{si} \quad x \neq \bar{x} \end{cases}$$

Ainsi définie, on montre que g est C^1 et $g'(\bar{x}) = \dfrac{1}{2}$: en effet, pour $x \neq \bar{x}$, on a

$$g'(x) = 1 - \frac{f'^2(x) - f(x)f''(x)}{f'^2(x)} = \frac{f(x)f''(x)}{f'^2(x)}$$

et en écrivant les formules de Taylor-Young autour de $\bar{x}$ pour f, f' et f'',

$$\begin{cases} f(x) = \dfrac{(x-\bar{x})^2}{2} f''(\bar{x}) + o((x-\bar{x})^2) \\ f'(x) = (x-\bar{x}) f''(\bar{x}) + o(x-\bar{x}) \\ f''(x) = f''(\bar{x}) + o(1) \end{cases}$$

on obtient bien

$$\lim_{x\to\bar{x}} g'(x) = \frac{1}{2} = \lim_{x\to\bar{x}} \frac{g(x)-g(\bar{x})}{x-\bar{x}}.$$

Le Théorème 3 permet alors de conclure dans ce cas □

2.2.3 Interprétation géométrique

La méthode de Newton peut s'interpréter géométriquement en dimension un : la valeur de x_{n+1} est égale à l'abscisse de l'intersection de l'axe Ox avec la tangente à la courbe représentative de f au point x_n. Autrement dit, f est localement remplacée par sa tangente afin d'obtenir une meilleure approximation de $\bar{x}$.

Il est également aisé de comprendre graphiquement pourquoi la méthode n'est pas forcément convergente si le point initial x_0 n'est pas suffisamment proche de $\bar{x}$. On devra donc procéder au cas par cas pour trouver un voisinage V de $\bar{x}$ convenable. À noter cependant qu'il existe de nombreux résultats donnant des conditions suffisantes sur f pour assurer la convergence de la méthode sur un certain voisinage (voir TD).

2.2.4 Extensions de la méthode

Même lorsque $\bar{x}$ est une racine de f de multiplicité m supérieure ou égale à 2, il est possible dans le cas réel de construire une suite d'approximations convergeant quadratiquement vers $\bar{x}$ en modifiant légèrement la méthode de Newton (voir TD). En particulier, la méthode de Newton améliorée basée sur la relation de récurrence

$$\forall n \in \mathbf{N}, \quad x_{n+1} = h(x_n)$$

avec

$$h(x) = x - \frac{(\frac{f}{f'})}{(\frac{f}{f'})'}(x) = x - \frac{f(x)f'(x)}{f'(x)^2 - f(x)f''(x)}$$

permet de retrouver (avec un supplément de calcul) une convergence quadratique si $m = 1$ ou 2.

Une autre extension possible en dimension quelconque correspond à la possibilité de pouvoir remplacer l'isomorphisme $Df(x)^{-1}$ par une approximation de celui-ci (en contrepartie, la convergence est seulement géométrique). Le théorème correspondant peut s'énoncer ainsi :

Théorème 6. *: soit $f \in C^1(A, \mathbf{R}^n)$ (avec A ouvert de $\mathbf{R}^n$) et $\bar{x} \in A$ tel que $f(\bar{x}) = 0$ et $Df(\bar{x})$ est un isomorphisme. On considère une famille $(A_n)_{n\in\mathbf{N}}$ d'isomorphismes de $\mathbf{R}^n$ telle que*

$$\exists \lambda \in [0, \frac{1}{2}[, \quad \sup_{n\in\mathbf{N}} |||A_n - Df(\bar{x})||| \leq \frac{\lambda}{|||Df(\bar{x})|||}.$$

Il existe un voisinage fermé V de $\bar{x}$ tel que la suite $(x_n)_{n\in\mathbf{N}}$ définie par la formule :

$$x_o \in V \quad \text{et} \quad x_{n+1} = x_n - A_n^{-1} \cdot f(x_n) \quad (n \in \mathbf{N})$$

est bien définie et converge géométriquement vers $\bar{x}$. De plus, la convergence est géométrique :

$$\exists \beta \in [0, 1[, \quad \forall n \in \mathbf{N}, \quad ||x_n - \bar{x}|| \leq \beta^n ||x_o - \bar{x}||.$$

On parle dans ce cas de la méthode de Newton généralisée.

Remarque : On peut dresser un lien entre la méthode de Newton généralisée et les méthodes de gradient (à pas simple, variable ou optimal) destinées à la recherche du minimum d'une fonctionnelle $J : \mathbf{R}^n \to \mathbf{R}$. En effet, ces dernières reviennent à rechercher des zéros de l'application ∇J avec la méthode de Newton généralisée.

2.2.5 Exemples

(i) Calcul approché d'une racine réelle

Soit $a \in \mathbf{R}_+^*$. On peut calculer une valeur approchée de $\sqrt{a}$ en appliquant la méthode de Newton à la fonction $f(x) = x^2 - a$. La suite des approximations est alors construite à partir de la relation de récurrence (appelée aussi algorithme phénicien)

$$x_{n+1} = \frac{1}{2}(x_n + \frac{a}{x_n}) \quad (n \in \mathbf{N})$$

On montre qu'il suffit de prendre $x_0 > 0$ pour que la suite converge vers $\sqrt{a}$ (en décroissant à partir du rang 1).

(ii) Calcul approché des racines d'un polynôme réel scindé

Soit
$$P(x)=\prod_{i=1}^{r}(x-\xi_i)^{m_i}$$
un polynôme réel scindé. On commence par diviser P par le PGCD de P et de P' afin de se ramener à un polynôme à racines simples (on peut donc à présent supposer que $m_i=1$). On applique alors la méthode de Newton avec $f(x)=P(x)$ partant de $x_0 \geq \max_{1\leq i\leq r}(\xi_i)$. On peut montrer que la suite des itérées converge quadratiquement vers la plus grande racine de P, par exemple ξ_r. Deux choix sont ensuite possibles pour déterminer les autres racines : le premier consiste à effectuer d'abord la division euclidienne de P par $X-\xi$ (où ξ est la valeur approchée de ξ_r obtenue) puis à recommencer l'opération précédente. Cette méthode (dite de déflation) est cependant dangereuse numériquement (car instable). Une autre méthode (dite de suppression de zéros et dûe à Maehly) consiste à construire la suite itérative suivante :
$$x_{n+1}=x_n-\frac{P(x_n)}{P'(x_n)-\frac{P(x_n)}{x_n-\xi}}$$
(issue du calcul approché de la dérivée logarithmique du polynôme $\frac{P}{X-\xi_r}$). On montre que cette suite, correctement initialisée, peut converger vers toutes les autres racines du polynôme P.

2.2.6 Implémentation numérique

L'implémentation de la méthode de Newton peut s'avérer coûteuse en dimension $n>1$ car elle requiert à chaque étape l'inversion du Jacobien de f en x_n. On utilise plutôt la méthode de Newton généralisée pour des suites particulières A_n : on peut par exemple prendre (en s'assurant que les conditions du théorème sont bien vérifiées) $A_n=Df^{-1}(x_0)$ (et effectuer une factorisation LU de cette matrice) ou choisir de ne réactualiser $Df^{-1}(x_n)$ que toutes les N itérations. Le gain réalisé en temps de calcul compense en général largement la perte de rapidité de la convergence.

2.3 Autres méthodes

Trois autres méthodes d'approximation des zéros d'une fonction réelle f sont présentées dans ce paragraphe.

2.3.1 Méthode de la sécante

Dans cette variante de la méthode de Newton, la dérivée $f'(x_n)$ est remplacé par le taux d'accroissement $\dfrac{f(x_n)-f(x_{n-1})}{x_n-x_{n-1}}$:

Théorème 7. *: soit $f \in C^2(\mathbf{R},\mathbf{R})$ et $\bar{x}$ un zéro de f tel que $f'(\bar{x}) \neq 0$. Alors, il existe un voisinage V de $\bar{x}$ tel que la suite $(x_n)_{n\in\mathbf{N}}$ construite suivant la formule*

$$(x_0,x_1)\in V^2,\ x_1\neq x_0 \quad \text{et} \quad \forall n\in\mathbf{N}^*, \quad x_{n+1}=x_n-\frac{f(x_n)(x_n-x_{n-1})}{f(x_n)-f(x_{n-1})}$$

est correctement définie (ou prend la valeur $\bar{x}$ et dans ce cas on suppose qu'elle stationne) et converge vers $\bar{x}$. De plus, si $f''(\bar{x})\neq 0$

$$\exists C>0,\ \exists q\in]0,1[,\ \exists n_0\in\mathbf{N},\ \forall n\in\mathbf{N},\ n\geq n_0 \Rightarrow |e_n|\leq Cq^{\alpha^n}$$

avec $\alpha = \dfrac{1+\sqrt{5}}{2}$.

Démonstration. Soit tout d'abord I un voisinage de $\bar{x}$ tel que

$$\exists m\in\mathbf{R}_+^*, \quad \forall x\in I, \quad |f'(x)|\geq m.$$

On note $M=\max_{x\in I}|f''(x)|$ et on considère $J=]\bar{x}-\beta,\bar{x}+\beta[\subset I$ avec β tel que $0<\dfrac{2M\beta}{m}<1$.

On montre que la suite $(x_n)_{n\in\mathbf{N}}$ de l'énoncé avec $(x_0,x_1)\in J^2$ et $x_1\neq x_0$ est bien définie. Pour cela, on démontre par récurrence la propriété suivante :

$$(x_n\in J) \text{ et } (x_n\neq x_{n-1} \text{ ou } x_n=x_{n-1}=\bar{x})$$

Supposons la propriété vraie jusqu'au rang n (elle est vraie pour $n=1$) : il est alors possible de définir x_{n+1} (dans un cas $f(x_n)\neq f(x_{n-1})$ car f est strictement monotone sur J et $x_{n+1}=\bar{x}$ dans l'autre cas). En utilisant deux fois le théorème des accroissements finis :

$$\begin{cases} f(x_n)=(x_n-\bar{x})f'(y_{1,n}), & (y_{1,n}\in]x_n,\bar{x}[) \\ f(x_n)-f(x_{n-1})=(x_n-x_{n-1})f'(y_{2,n}) & (y_{2,n}\in]x_{n-1},x_n[) \end{cases}$$

il vient

$$x_{n+1}=x_n-\frac{(x_n-\bar{x})f'(y_{1,n})(x_n-x_{n-1})}{(x_n-x_{n-1})f'(y_{2,n})}=x_n-(x_n-\bar{x})\frac{f'(y_{1,n})}{f'(y_{2,n})}$$

ou

$$\begin{aligned} x_{n+1}-\bar{x} &= (x_n-\bar{x})\Big[\frac{f'(y_{2,n})-f'(y_{1,n})}{f'(y_{2,n})}\Big] \\ &= (x_n-\bar{x})\frac{f''(y_{3,n})}{f'(y_{2,n})}(y_{2,n}-y_{1,n}) \qquad (y_{3,n}\in]y_{1,n},y_{2,n}[) \end{aligned}$$

Grâce à la propriété de récurrence, on a alors

$$|e_{n+1}| = |x_{n+1}-\bar{x}| \leq |x_n-\bar{x}|\frac{M}{m}2\beta \tag{6}$$

ce qui permet d'affirmer que $x_{n+1} \in J$. Pour démontrer la deuxième partie de la propriété, on observe que si $x_n = \bar{x}$, alors $x_{n+1} = x_n = \bar{x}$ tandis que si $x_n \neq \bar{x}$, on a $x_{n+1} \neq x_n$ (car dans ce cas $x_n \neq x_{n-1}$, $f(x_n) \neq f(x_{n-1})$ et $f(x_n) \neq 0$). L'inégalité (6) permet en outre d'assurer que la suite $(x_n)_{n\in\mathbf{N}}$ converge vers $\bar{x}$ de manière géométrique. Pour améliorer cette estimation, on écrit (en adoptant les notations relatives aux différences divisées, voir chapitre sur l'interpolation polynomiale) :

$$\begin{aligned} e_{n+1} = e_n - \frac{f(x_n)}{f[x_n,x_{n-1}]} &= e_n - \frac{1}{f[x_n,x_{n-1}]}(x_n-\bar{x})f[x_n,\bar{x}] \\ &= e_n\Big(1-\frac{f[x_n,\bar{x}]}{f[x_n,x_{n-1}]}\Big) \\ &= e_n e_{n-1}\frac{f[x_n,x_{n-1},\bar{x}]}{f[x_n,x_{n-1}]}. \end{aligned}$$

En utilisant alors les propriétés classiques suivantes :

$$\lim_{n\to+\infty} f[x_n,x_{n-1}] = f'(\bar{x})$$

et

$$\lim_{n\to+\infty} f[x_n,x_{n-1},\bar{x}] = \frac{f''(\bar{x})}{2},$$

on peut affirmer qu'il existe $M' > \dfrac{M}{2m}$ tel que pour n assez grand,

$$|e_{n+1}| \leq M'|e_n e_{n-1}|$$

ou de manière équivalente $\varepsilon_{n+1} \leq \varepsilon_n + \varepsilon_{n-1}$ en posant $\varepsilon_n = \ln(M'|e_n|)$.

On obtient ensuite par récurrence

$$\varepsilon_n \leq A\Big(\frac{1+\sqrt{5}}{2}\Big)^n + B\Big(\frac{1-\sqrt{5}}{2}\Big)^n$$

avec

$$\begin{cases} A = \dfrac{\varepsilon_1 - \frac{1-\sqrt{5}}{2}\varepsilon_0}{\sqrt{5}} \\ B = \dfrac{\frac{1+\sqrt{5}}{2}\varepsilon_0 - \varepsilon_1}{\sqrt{5}} \end{cases}$$

Il est toujours possible de choisir un voisinage V de $\bar{x}$ inclus dans J tel que ε_0 et ε_1 soient toujours négatifs (prendre $V =]\bar{x} - \gamma, \bar{x} + \gamma[$ avec $\gamma < \min(\dfrac{1}{M'}, \dfrac{m}{2M})$). Dans ce cas, $A < 0$ et pour n assez grand,

$$\varepsilon_n \leq \frac{A}{2}\Big(\frac{1+\sqrt{5}}{2}\Big)^n = \frac{A}{2}\alpha^n$$

soit

$$e_n \leq \frac{1}{M'} e^{\frac{A}{2}\alpha^n}$$

ce qui permet de clôre ensuite aisément la démonstration du théorème □

Remarque : Cette méthode possède l'avantage par rapport à la méthode de Newton de ne pas nécessiter le calcul de f' mais s'avère moins performante qu'elle (car $\alpha < 2$). Comme la méthode de Newton, elle est en outre aisément programmable. À noter qu'il peut survenir cependant un problème de compensation d'erreur dans l'estimation du taux d'accroissement

$$\tau_n = \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}.$$

En effet, si l'on dispose d'une précision machine de ε, il est inutile et vain de calculer τ_n dès que $|x_n - x_{n-1}| \leq \sqrt{\varepsilon}$. Pour atteindre une précision ε sur $\bar{x}$, on doit alors continuer le calcul avec la dernière valeur (constante) de ce taux.

2.3.2 Méthode de dichotomie (ou de bissection)

Soit $f \in C^0(\mathbf{R})$. On suppose connu un intervalle $[a, b]$ de $\mathbf{R}$ sur lequel f ne s'annule qu'une fois (en $\bar{x}$) en changeant de plus de signe (on peut par exemple supposer f continue et strictement monotone sur $[a, b]$ avec $f(a)f(b) < 0$). La méthode de dichotomie se construit ainsi : les trois premiers termes de la suite d'approximations sont pris égaux à

$$\begin{cases} x_0 = a \\ x_1 = b \\ x_2 = \dfrac{x_0 + x_1}{2} \end{cases}$$

Ensuite, parmi les deux réels $f(x_0)f(x_2)$ et $f(x_1)f(x_2)$, il en existe forcément un (par exemple le premier) qui soit strictement négatif (sinon $x_2 = \bar{x}$ et la construction s'arrête). Dans ce cas, on construit alors

$$x_3 = \frac{x_0 + x_2}{2}$$

et le processus se répète indéfiniment (ou s'arrête si pour un certain $n \in \mathbf{N}$, $x_n = \bar{x}$) permettant de construire les termes de la suite $(x_n)_{n\in\mathbf{N}}$. On montre aisément que $\lim\limits_{n\to+\infty} x_n = \bar{x}$ et que la convergence est géométrique :

$$\forall n \in \mathbf{N},\ |e_n| \le \frac{1}{2^{n-1}}|x_1 - x_0|.$$

2.3.3 Méthode de la fausse position (ou regula falsi)

Cette méthode reprend les hypothèses de la méthode de dichotomie (par exemple f continue et strictement monotone sur $[a, b]$ avec $f(a)f(b) < 0$) et construit la suite d'approximations suivante :

$$\begin{cases} x_0 = a \\ x_1 = b \\ x_2 = \dfrac{x_0 f(x_1) - x_1 f(x_0)}{f(x_1) - f(x_0)} \end{cases}$$

(x_2 correspond au point d'intersection de la corde de f entre x_0 et x_1 avec l'axe des abscisses). La construction de la suite $(x_n)_{n\in\mathbf{N}}$ suit alors en tout point celle associée à la méthode de dichotomie (seule change la façon de découper l'intervalle de départ). La convergence est encore en général géométrique mais peut dans certains cas devenir quadratique.

Remarque : L'avantage essentiel des deux dernières méthodes présentées par rapport aux deux premières (dont la convergence est plus rapide) tient au fait que leur convergence est plus facilement assurée.

2.4 Accélération de convergence

2.4.1 Procédé Δ^2 d'Aitken

Le théorème suivant donne un procédé d'accélération de convergence pour certaines suites itératives. Auparavant, on introduit l'opérateur classique aux différences finies progressives Δ :

$$\Delta : \begin{pmatrix} \mathbf{R}^{\mathbf{N}} \to \mathbf{R}^{\mathbf{N}} \\ (x_n)_{n\in\mathbf{N}} \mapsto (y_n = x_{n+1} - x_n)_{n\in\mathbf{N}} \end{pmatrix}.$$

Théorème 8. *: soit $\bar{x} \in \mathbf{R}$ et $(x_n)_{n\in\mathbf{N}}$ une suite réelle ne prenant jamais la valeur $\bar{x}$. On suppose qu'il existe un coefficient $k \in [0,1[$ et une suite auxiliaire $(z_n)_{n\in\mathbf{N}}$ convergeant vers 0 vérifiant*

$$\forall n \in \mathbf{N}, \quad x_{n+1} - \bar{x} = (k + z_n)(x_n - \bar{x}) \tag{7}$$

Alors, la suite (x'_n) telle que

$$x'_n = x_n - \frac{(\Delta x_n)^2}{\Delta^2 x_n} = x_n - \frac{x_{n+1}^2 - 2x_{n+1}x_n + x_n^2}{x_{n+2} - 2x_{n+1} + x_n}$$

est bien définie pour n assez grand et vérifie

$$\lim_{n\to+\infty} \frac{x'_n - \bar{x}}{x_n - \bar{x}} = 0.$$

Démonstration. On note $e_n = x_n - \bar{x}$. La relation (7) se récrit $e_{n+1} = (k + z_n)e_n$. Par conséquent

$$\Delta x_n = \Delta e_n = (k - 1 + z_n)e_n$$

et

$$\begin{aligned} \Delta^2 x_n = \Delta^2 e_n &= (k - 1 + z_{n+1})e_{n+1} - (k - 1 + z_n)e_n \\ &= \big[(k - 1 + z_{n+1})(k + z_n) - (k - 1 + z_n)\big]e_n \\ &= \big[(k-1)^2 + k(z_n + z_{n+1}) + z_n z_{n+1} - 2z_n\big]e_n. \end{aligned}$$

En particulier, on observe qu'avec les hypothèses faites $\Delta^2 x_n \neq 0$ pour n assez grand, ce qui permet de définir la suite (x'_n) à partir d'un certain rang. De plus, on a alors

$$\begin{aligned} x'_n - \bar{x} &= x_n - \bar{x} - \frac{(k - 1 + z_n)^2 e_n^2}{\big[(k-1)^2 + k(z_n + z_{n+1}) + z_n z_{n+1} - 2z_n\big]e_n} \\ &= \Big[\frac{k(z_n + z_{n+1}) + z_n z_{n+1} - 2z_n - 2z_n(k-1) + z_n^2}{(k-1)^2 + k(z_n + z_{n+1}) + z_n z_{n+1} - 2z_n}\Big]e_n \end{aligned}$$

ce qui assure ensuite facilement le résultat annoncé

□

2.4.2 Méthode de Steffensen

Il s'agit du procédé d'Aitken appliqué aux suite itératives du paragraphe 1 du type $x_{n+1} = g(x_n)$. Formellement, on a

$$x'_n = x_n - \frac{(g(x_n) - x_n)^2}{(g \circ g)(x_n) - 2g(x_n) + x_n} = \frac{x_n (g \circ g)(x_n) - g(x_n)^2}{(g \circ g)(x_n) - 2g(x_n) + x_n}.$$

Le théorème suivant précise l'existence et le comportement de la suite (x'_n) :

Théorème 9. *: soit $p \in \mathbf{N}^*$, $g \in C^{p+1}(\mathbf{R}, \mathbf{R})$ et $\bar{x} \in \mathbf{R}$ tel que*

$$\begin{cases} g(\bar{x}) = \bar{x} \\ g'(\bar{x}) = \ldots = g^{(p-1)}(\bar{x}) = 0 \\ g^{(p)}(\bar{x}) \neq 0 \end{cases}$$

(on suppose en outre que $g'(\bar{x}) \neq 1$ si $p = 1$).

Alors, il existe un voisinage V de $\bar{x}$ sur lequel la fonction

$$h : \begin{pmatrix} V \to \mathbf{R} \\ x \mapsto \begin{cases} \dfrac{x(g \circ g)(x) - g(x)^2}{(g \circ g)(x) - 2g(x) + x} & \text{si} \quad x \neq \bar{x} \\ \bar{x} \quad \text{si} \quad x = \bar{x} \end{cases} \end{pmatrix}$$

est correctement définie et où la suite $(x'_n)_{n \in \mathbf{N}}$ construite par la formule

$$x'_0 \in V \quad \text{et} \quad x'_{n+1} = h(x'_n) \quad (n \in \mathbf{N})$$

est également bien définie et converge vers $\bar{x}$ avec un ordre $2p - 1$ si $p > 1$ et 2 si $p = 1$.

Démonstration. : On se ramène tout d'abord facilement au cas $\bar{x} = 0$ en posant $g_1(y) = g(y + \bar{x}) - \bar{x}$. Ensuite, on écrit la formule de Taylor-Young à l'ordre $p + 1$ pour g en 0 :

$$g(x) = \frac{g^{(p)}(0)}{p!} x^p + O(x^{p+1}) \equiv A x^p + O(x^{p+1}).$$

En itérant cette relation, il vient

$$\begin{aligned} (g \circ g)(x) &= A(Ax^p + O(x^{p+1}))^p + O(x^{p(p+1)}) \\ &= \begin{cases} O(x^{p^2}) & \text{si} \quad p > 1 \\ A^2 x + O(x^2) & \text{si} \quad p = 1 \end{cases} \end{aligned}$$

On reporte ces estimations dans l'expression de h (en observant au préalable que $A \neq 1$ si $p = 1$) :

$$h(x) = \begin{cases} \dfrac{O(x^{p^2+1}) - A^2x^{2p} + O(x^{2p+1})}{O(x^{p^2}) - 2Ax^p + O(x^{p+1}) + x} \sim A^2x^{2p-1} \quad \text{si} \quad p > 1 \\ \dfrac{A^2x^2 + O(x^3) - A^2x^2 + O(x^3)}{A^2x + O(x^2) - 2Ax + O(x^2) + x} = \dfrac{O(x^3)}{(A-1)^2x + O(x^2)} = O(x^2) \quad \text{sinon} \end{cases}$$

La démonstration s'achève alors aisément en s'inspirant du Théorème 4 $\square$

Remarque : Cette méthode est intéressante car elle permet de construire une suite d'approximations avec convergence quadratique d'un point fixe $\bar{x}$ d'une fonction g même si $|g'(\bar{x})| > 1$ (point fixe répulsif). Par contre, elle n'a pas d'utilité pratique lorsque $p > 1$ car dans ce cas, la fonction $h(x) = (g \circ g)(x)$ donne une convergence plus rapide (d'ordre $p^2 > 2p - 1$).