Chapitre 2

Systèmes non linéaires

Dans le premier chapitre, on a étudié quelques méthodes de résolution de systèmes linéaires en dimension finie. L'objectif est maintenant de développer des méthodes de résolution de systèmes non linéaires, toujours en dimension finie. On se donne $g \in C(\mathbb{R}^N, \mathbb{R}^N)$ et on cherche x dans $\mathbb{R}^N$ solution de :

$$\left\{ \begin{array}{l} x \in \mathbb{R}^N \\ g(x) = 0. \end{array} \right. \tag{2.0.1}$$

Au Chapitre I on a étudié des méthodes de résolution du système (2.0.1) dans le cas particulier $g(x) = Ax - b$, $A \in \mathcal{M}_N(\mathbb{R})$, $b \in \mathbb{R}^N$. On va maintenant étendre le champ d'étude au cas où g n'est pas forcément affine. On étudiera deux familles de méthodes pour la résolution approchée du système (2.0.1) :
– les méthodes de point fixe : point fixe de contraction et point fixe de monotonie
– les méthodes de type Newton [1].

2.1 Les méthodes de point fixe

2.1.1 Point fixe de contraction

Soit $g \in C(\mathbb{R}^N, \mathbb{R}^N)$, on définit la fonction $f \in C(\mathbb{R}^N, \mathbb{R}^N)$ par $f(x) = x + g(x)$. On peut alors remarquer que $g(x) = 0$ si et seulement si $f(x) = x$. Résoudre le système non linéaire (2.0.1) revient donc à trouver un point fixe de f. Encore faut-il qu'un tel point fixe existe…

Théorème 2.1 (Point fixe) *Soit E un espace métrique complet, d la distance sur E, et $f : E \to E$ une fonction strictement contractante, c'est–à–dire telle qu'il existe $k \in]0, 1[$ tel que $d(f(x), f(y)) \leq kd(x, y)$ pour tout $x, y \in E$.*
Alors il existe un unique point fixe $\bar{x} \in E$ qui vérifie $f(\bar{x}) = \bar{x}$. De plus si $x^{(0)} \in E$, et $x^{(n+1)} = f(x^{(n)})$ $\forall n \geq 0$, alors $x^{(n)} \to \bar{x}$ quand $n + \infty$.

Démonstration :

Etape 1 : Existence de $\bar{x}$ et convergence de la suite

Soit $x^{(0)} \in E$ et $(x^{(n)})_{n \in \mathbb{N}}$ la suite définie par $x^{(n+1)} = f(x^{(n)})$ pour $n \geq 0$. On va montrer que :

1. $(x^{(n)})_n$ est de Cauchy (donc convergente car E est complet),

1. Isaac Newton (1643 - 1727, né d'une famille de fermiers, est un philosophe, mathématicien, physicien, alchimiste et astronome anglais. Figure emblématique des sciences, il est surtout reconnu pour sa théorie de la gravitation universelle et la création, en concurrence avec Leibniz, du calcul infinitésimal.

2. $\lim_{n\to+\infty} x^{(n)} = \bar{x}$ est point fixe de f.

Par hypothèse, on sait que pour tout $n \geq 1$,

$$d(x^{(n+1)}, x^{(n)}) = d(f(x^{(n)}), f(x^{(n-1)})) \leq kd(x^{(n)}, x^{(n-1)}).$$

Par récurrence sur n, on obtient que

$$d(x^{(n+1)}, x^{(n)}) \leq k^n d(x^{(1)}, x^{(0)}), \ \forall n \geq 0.$$

Soit $n \geq 0$ et $p \geq 1$, on a donc :

$$\begin{aligned} d(x^{(n+p)}, x^{(n)}) &\leq d(x^{(n+p)}, x^{(n+p-1)}) + \cdots + d(x^{(n+1)}, x^{(n)}) \\ &\leq \sum_{q=1}^{p} d(x^{(n+q)}, x^{(n+q-1)}) \\ &\leq \sum_{q=1}^{p} k^{n+q-1} d(x^{(1)}, x^{(0)}) \\ &\leq d(x^{(1)}, x^{(0)}) k^n (1 + k + \ldots + k^{p-1}) \\ &\leq d(x^{(1)}, x^{(0)}) \frac{k^n}{1-k} \longrightarrow 0 \text{ quand } n \to +\infty \text{ car } k < 1. \end{aligned}$$

La suite $(x^{(n)})_{n\in\mathbb{N}}$ est donc de Cauchy, *i.e.* :

$$\forall \varepsilon > 0, \ \ \exists n_\varepsilon \in \mathbb{N}\,; \ \ \forall n \geq n_\varepsilon, \ \ \forall \ \ p \geq 1 \ \ d(x^{(n+p)}, x^{(n)}) \leq \varepsilon.$$

Comme E est complet, on a donc $x^{(n)} \longrightarrow \bar{x}$ dans E quand $n \to +\infty$.
Comme la fonction f est strictement contractante, elle est continue, donc on a aussi $f(x^{(n)}) \longrightarrow f(\bar{x})$ dans E quand $n + \infty$. En passant à la limite dans l'égalité $x^{(n+1)} = f(x^{(n)})$, on en déduit que $\bar{x} = f(\bar{x})$.

Etape 2 : Unicité

Soit $\bar{x}$ et $\bar{y}$ des points fixes de f, qui satisfont donc $\bar{x} = f(\bar{x})$ et $\bar{y} = f(\bar{y})$. Alors $d(f(\bar{x}), f(\bar{y})) = d(\bar{x}, \bar{y}) \leq kd(\bar{x}, \bar{y})$; comme $k < 1$, ceci est impossible sauf si $\bar{x} = \bar{y}$.

■

Remarque 2.2

1. *Sous les hypothèses du théorème 2.1,* $d(x^{(n+1)}, \bar{x}) = d(f(x^{(n)}), f(\bar{x})) \leq kd(x^{(n)}, \bar{x})$; *donc si* $x^{(n)} \neq \bar{x}$ *alors* $\frac{d(x^{(n+1)}, \bar{x})}{d(x^{(n)}, \bar{x})} \leq k \ \ (< 1)$. *La convergence est donc au moins linéaire (même si de fait, cette méthode converge en général assez lentement).*
2. *On peut généraliser le théorème du point fixe en remplaçant l'hypothèse "f strictement contractante" par " il existe $n > 0$ tel que $f^{(n)} = \underbrace{f \circ f \circ \ldots \circ f}_{n \text{ fois}}$ est strictement contractante " (reprendre la démonstration du théorème pour le vérifier).*

La question qui vient alors naturellement est : que faire si f n'est pas strictement contractante ? Soit $g \in C(\mathbb{R}^N, \mathbb{R}^N)$ telle que $f(x) = x + g(x)$. On aimerait déterminer les conditions sur g pour que f soit strictement contractante. Plus généralement si $\omega \neq 0$, on définit $f_\omega(x) = x + \omega g(x)$, et on remarque que x est solution du système (2.0.1) si et seulement si x est point fixe de $f_\omega(x)$.
On aimerait dans ce cas avoir des conditions pour que f_ω soit strictement contractante.

Théorème 2.3 (Point fixe de contraction avec relaxation)
On désigne par $|.|$ la norme euclidienne sur $\mathbb{R}^N$. Soit $g \in C(\mathbb{R}^N, \mathbb{R}^N)$ telle que

$$\exists \alpha > 0 \text{ tel que } (g(x) - g(y)) \cdot (x - y) \leq -\alpha |x - y|^2, \forall x, y \in \mathbb{R}^N, \tag{2.1.2}$$

$$\exists M > 0 \text{ tel que } |g(x) - g(y)| \leq M|x - y|, \forall x, y \in \mathbb{R}^N. \tag{2.1.3}$$

Alors la fonction f_ω est strictement contractante si $0 < \omega < \dfrac{2\alpha}{M^2}$.
Il existe donc un et un seul $\bar{x} \in \mathbb{R}^N$ tel que $g(\bar{x}) = 0$ et $x^{(n)} \longrightarrow \bar{x}$ quand $n + \infty$ avec $x^{(n+1)} = f_\omega(x^{(n)}) = x^{(n)} + \omega g(x^{(n+1)})$.

Remarque 2.4 *Le théorème 2.3 permet de montrer que sous les hypothèses* (2.1.2) *et* (2.1.3)*, et pour $\omega \in]0, \frac{2\alpha}{M^2}[$, on peut obtenir la solution de* (2.0.1) *en construisant la suite :*

$$\begin{cases} x^{(n+1)} = x^{(n)} + \omega g(x^{(n)}) \;\; n \geq 0, \\ x^{(0)} \in \mathbb{R}^N. \end{cases} \tag{2.1.4}$$

Or on peut aussi écrire cette suite de la manière suivante :

$$\begin{cases} \tilde{x}^{(n+1)} &= f(x^{(n)}), \;\; \forall n \geq 0 \\ x^{(n+1)} &= \omega \tilde{x}^{(n+1)} + (1 - \omega) x^{(n)}, \quad x^{(0)} \in \mathbb{R}^N. \end{cases} \tag{2.1.5}$$

En effet si $x^{(n+1)}$ est donné par la suite (2.1.5)*, alors $x^{(n+1)} = \omega \tilde{x}^{(n+1)} - (1-\omega) x^{(n)} = \omega f(x^{(n)}) + (1-\omega) x^{(n)} = \omega g(x^{(n)}) + x^{(n)}$. Le procédé de construction de la suite* (2.1.5) *est l'algorithme de relaxation sur f.*

Démonstration du théorème 2.3
Soit $0 < \omega < \dfrac{2\alpha}{M^2}$. On veut montrer que f est strictement contractante, c.à.d. qu'il existe $k < 1$ tel que $|f_\omega(x) - f_\omega(y)| \leq k|x - y| \;\forall (x, y) \in (\mathbb{R}^N)^2$. Soit $(x, y) \in (\mathbb{R}^N)^2$, alors, par définition de la norme euclidienne,

$$\begin{aligned} |f_\omega(x) - f_\omega(y)|^2 &= \big(x - y + \omega(g(x) - g(y))\big) \cdot \big(x - y + \omega(g(x) - g(y))\big) \\ &= |x - y|^2 + 2(x - y) \cdot (\omega(g(x) - g(y))) + \omega^2 |g(x) - g(y)|^2. \end{aligned}$$

Donc grâce aux hypothèses (2.1.2) et (2.1.3), on a : $|f_\omega(x) - f_\omega(y)|^2 \leq (1 - 2\omega\alpha + \omega^2 M^2) \; |x - y|^2$, et donc la fonction f_ω est strictement contractante si $1 - 2\omega\alpha + \omega^2 M^2 < 1$ ce qui est vérifié si $0 < \omega < \dfrac{2\alpha}{M^2}$.

■

Remarque 2.5 (Quelques rappels de calcul différentiel)
Soit $h \in C^2(\mathbb{R}^N, \mathbb{R})$. La fonction h est donc en particulier différentiable, c.à.d. que pour tout $x \in \mathbb{R}^N$, il existe $Dh(x) \in \mathcal{L}(\mathbb{R}^N, \mathbb{R})$ telle que $h(x + y) = h(x) + Dh(x)(y) + |y|\varepsilon(y)$ où $\varepsilon(y) \underset{y \to 0}{\to} 0$. On a dans ce cas, par définition du gradient, $Dh(x)(y) = \nabla h(x) \cdot y$ où $\nabla h(x) = (\partial_1 h(x), \cdots, \partial_N h(x))^t \in \mathbb{R}^N$ est le gradient de h au point x (on désigne par $\partial_i h$ la dérivée partielle de f par rapport à sa i-ème variable).
Comme on suppose $h \in C^2(\mathbb{R}^N, \mathbb{R})$, on a donc $g = \nabla h \in C^1(\mathbb{R}^N, \mathbb{R}^N)$, et g est continûment différentiable, c'est–à–dire

$$Dg(x) \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^N), \text{ et } g(x + y) = g(x) + Dg(x)(y) + |y|\varepsilon(y),$$

où $\varepsilon(y) \underset{y \to 0}{\to} 0$.
Comme $Dg(x) \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^N)$, on peut représenter $Dg(x)$ par une matrice de $\mathcal{M}_N(\mathbb{R})$, on confond alors l'application linéaire et la matrice qui la représente dans la base canonique, et on écrit par abus de notation $Dg(x) \in \mathcal{M}_N(\mathbb{R})$. On peut alors écrire, grâce à cet abus de notation, $Dg(x)(y) = Dg(x)y$ avec $(Dg(x)y)_i = \sum_{i,j=1,N} \partial^2_{i,j} h_j(x)$ où $\partial^2_{i,j} h = \partial_i(\partial_j h)(x)$.
Comme h est de classe C^2, la matrice $Dg(x)$ est symétrique. Pour $x \in \mathbb{R}^N$, on note $(\lambda_i(x))_{1 \leq i \leq N}$ les valeurs propres de $Dg(x)$, qui sont donc réelles.

La proposition suivante donne une condition suffisante pour qu'une fonction vérifie les hypothèses (2.1.2) et (2.1.3).

Proposition 2.6 *Soit $h \in C^2(\mathbb{R}^N, \mathbb{R})$, et $(\lambda_i)_{i=1,N}$ les valeurs propres de la matrice hessienne de h. On suppose qu'il existe des réels strictement positifs β et γ tels que $-\beta \le \lambda_i(x) \le -\gamma$, $\forall i \in \{1 \ldots N\}$, $\forall x \in \mathbb{R}^N$. (Notons que cette hypothèse est plausible puisque les valeurs propres de la matrice hessienne sont réelles). Alors la fonction $g = \nabla h$ (gradient de h) vérifie les hypothèses* (2.1.2) *et* (2.1.3) *du théorème 2.3 avec $\alpha = \gamma$ et $M = \beta$.*

Démonstration de la proposition 2.6
Montrons d'abord que l'hypothèse (2.1.2) est vérifiée. Soit $(x, y) \in (\mathbb{R}^N)^2$, on veut montrer que $(g(x) - g(y)) \cdot (x - y) \le -\gamma|x - y|^2$. On introduit pour cela la fonction $\varphi \in C^1(\mathbb{R}, \mathbb{R}^N)$ définie par :

$$\varphi(t) = g(x + t(y - x)).$$

On a donc $\varphi(1) - \varphi(0) = g(y) - g(x) = \int_0^1 \varphi'(t)dt$. Or $\varphi'(t) = Dg(x + t(y - x))(y - x)$. Donc $g(y) - g(x) = \int_0^1 Dg(x + t(y - x))(y - x)dt$. On en déduit que :

$$(g(y) - g(x)) \cdot (y - x) = \int_0^1 (Dg(x + t(y - x))(y - x) \cdot (y - x))\, dt.$$

Comme $\lambda_i(x) \in [-\beta, -\gamma]$ $\forall i \in \{1, \ldots, N\}$, on a donc $-\beta|y|^2 \le Dg(z)y \cdot y \le -\gamma|y|^2$. On a donc : $(g(y) - g(x)) \cdot (y - x) \le \int_0^1 -\gamma|y - x|^2 dt = -\gamma|y - x|^2$ ce qui termine la démonstration de (2.1.2).
Montrons maintenant que l'hypothèse (2.1.3) est vérifiée. On veut montrer que $|g(y) - g(x)| \le \beta|y - x|$. On peut écrire :

$$g(y) - g(x) = \int_0^1 Dg(x + t(y - x))(y - x)dt,$$

et donc

$$\begin{aligned} |g(y) - g(x)| &\le \int_0^1 |Dg(x + t(y - x))(y - x)|dt \\ &\le \int_0^1 |Dg(x + t(y - x))||y - x|dt, \end{aligned}$$

où $|.|$ est la norme sur $\mathcal{M}_N(\mathbb{R})$ induite par la norme euclidienne sur $\mathbb{R}^N$.
Or, comme $\lambda_i(x) \in [-\beta, -\gamma]$ pour tout $i = 1, \ldots, N$, la matrice $-Dg(x + t(y - x))$ est symétrique définie positive et donc, d'après l'exercice 5 page 37,

$$|Dg(x + t(y - x)| = \rho(Dg(x + t(y - x)) \le \beta.$$

On a donc ainsi montré que : $|g(y) - g(x)| \le \beta|y - x|$, ce qui termine la démonstration. ■

Remarque 2.7 *Dans de nombreux cas, le problème de résolution d'un problème non linéaire apparaît sous la forme $Ax = R(x)$ où A est une matrice carrée d'ordre N inversible, et $R \in C(\mathbb{R}^N, \mathbb{R}^N)$. On peut le réécrire sous la forme $x = A^{-1}R(x)$. On peut donc appliquer l'algorithme de point fixe sur la fonction $f = A^{-1}R$, ce qui donne comme itération : $x^{(n+1)} = A^{-1}R(x^{(n)})$. Si on pratique un point fixe avec relaxation, avec paramètre de relaxation $\omega > 0$, alors l'itération s'écrit : $\tilde{x}^{(n+1)} = A^{-1}R(x^{(n)})$, $x^{(n+1)} = \omega\tilde{x}^{(n+1)} + (1 - \omega)x^{(n)}$.*

2.1.2 Point fixe de monotonie

Théorème 2.8 (Point fixe de monotonie)
Soient $A \in \mathcal{M}_N(\mathbb{R})$ et $R \in C(\mathbb{R}^N, \mathbb{R}^N)$. On suppose que :

1. $\forall x \in \mathbb{R}^N$, $Ax \geq 0 \Rightarrow x \geq 0$, *c'est-à-dire* $\big((Ax)_i \geq 0,\ \forall i = 1,\ldots,N\big) \Rightarrow \big(x_i \geq 0,\ \forall i = 1,\ldots,N\big)$.
2. R *est monotone, c.à.d. que si* $x \geq y$ *(composante par composante) alors* $R(x) \geq R(y)$ *(composante par composante).*
3. 0 *est une sous-solution du problème, c'est-à-dire que* $R(0) \geq 0$ *et il existe* $\tilde{x} \in \mathbb{R}^N$ *;* $\tilde{x} \geq 0$ *tel que* $\tilde{x}$ *est une sur-solution du problème, c'est-à-dire que* $A\tilde{x} \geq R(\tilde{x})$.

On pose $x^{(0)} = 0$ *et* $Ax^{(n+1)} = R(x^{(n)})$. *On a alors :*

1. $0 \leq x^{(n)} \leq \tilde{x},\ \ \forall n \in \mathbb{N}$,
2. $x^{(n+1)} \geq x^{(n)},\ \ \forall n \in \mathbb{N}$,
3. $x^{(n)} \longrightarrow \bar{x}$ *quand* $n \to +\infty$ *et* $A\bar{x} = R(\bar{x})$.

Démonstration du théorème 2.8
Comme A est inversible la suite $(x^{(n)})_{n\in\mathbb{N}}$ vérifiant

$$\begin{cases} x^{(0)} = 0, \\ Ax^{(n+1)} = R(x^{(n)}),\ \ n \geq 0 \end{cases}$$

est bien définie. On va montrer par récurrence sur n que $0 \leq x^{(n)} \leq \tilde{x}$ pour tout $n \geq 0$ et que $x^{(n)} \leq x^{(n+1)}$ pour tout $n \geq 0$.

1. Pour $n = 0$, on a $x^{(0)} = 0$ et donc $0 \leq x^{(0)} \leq \tilde{x}$ et $Ax^{(1)} = R(0) \geq 0$. On en déduit que $x^{(1)} \geq 0$ grâce aux hypothèses 1 et 3 et donc $x^{(1)} \geq x^{(0)} = 0$.
2. On suppose maintenant (hypothèse de récurrence) que $0 \leq x^{(p)} \leq \tilde{x}$ et $x^{(p)} \leq x^{(p+1)}$ pour tout $p \in \{0,\ldots,n-1\}$.
   On veut montrer que $0 \leq x^{(n)} \leq \tilde{x}$ et que $x^{(n)} \leq x^{(n+1)}$. Par hypothèse de récurrence pour $p = n-1$, on sait que $x^{(n)} \geq x^{(n-1)}$ et que $x^{(n-1)} \geq 0$. On a donc $x^{(n)} \geq 0$. Par hypothèse de récurrence, on a également que $x^{(n-1)} \leq \tilde{x}$ et grâce à l'hypothèse 2, on a donc $R(x^{(n-1)}) \leq R(\tilde{x})$. Par définition de la suite $(x^{(n)})_{n\in\mathbb{N}}$, on a $Ax^{(n)} = R(x^{(n-1)})$ et grâce à l'hypothèse 3, on sait que $A\tilde{x} \geq R(\tilde{x})$. On a donc : $A(\tilde{x} - x^{(n)}) \geq R(\tilde{x}) - R(x^{(n-1)}) \geq 0$. On en déduit alors (grâce à l'hypothèse 1) que $x^{(n)} \leq \tilde{x}$.
   De plus, comme $Ax^{(n)} = R(x^{(n-1)})$ et $Ax^{(n+1)} = R(x^{(n)})$, on a $A(x^{(n+1)} - x^{(n)}) = R(x^{(n)}) - R(x^{(n-1)}) \geq 0$ par l'hypothèse 2, et donc grâce à l'hypothèse 1, $x^{(n+1)} \geq x^{(n)}$.

On a donc ainsi montré (par récurrence) que

$$0 \leq x^{(n)} \leq \tilde{x},\quad \forall n \geq 0$$

$$x^{(n)} \leq x^{(n+1)},\quad \forall n \geq 0.$$

Ces inégalités s'entendent composante par composante, c.à.d. que si $x^{(n)} = (x_1^{(n)} \ldots x_N^{(n)})^t \in \mathbb{R}^N$ et $\tilde{x} = (\tilde{x}_1 \ldots \tilde{x}_N)^t \in \mathbb{R}^N$, alors $0 \leq x_i^{(n)} \leq \tilde{x}_i$ et $x_i^{(n)} \leq x_i^{(n+1)}$, $\forall i \in \{1,\ldots,N\}$, et $\forall n \geq 0$.

Soit $i \in \{1,\ldots,N\}$; la suite $(x_i^{(n)})_{n\in\mathbb{N}} \subset \mathbb{R}$ est croissante et majorée par $\tilde{x}_i$ donc il existe $\bar{x}_i \in \mathbb{R}$ tel que $\bar{x}_i = \lim\limits_{n\to+\infty} x_i^{(n)}$. Si on pose $\bar{x} = (\bar{x}_1 \ldots \bar{x}_N)^t \in \mathbb{R}^N$, on a donc $x^{(n)} \longrightarrow \bar{x}$ quand $n \to +\infty$.

Enfin, comme $Ax^{(n+1)} = R(x^{(n)})$ et comme R est continue, on obtient par passage à la limite lorsque $n \to +\infty$ que $A\bar{x} = R(\bar{x})$ et que $0 \leq \bar{x} \leq \tilde{x}$.

■

L'hypothèse 1 du théorème 2.8 est souvent appelée “principe du maximum". Elle est vérifiée par exemple par les matrices A qu'on a obtenues par discrétisation par différences finies des opérateurs $-u''$ sur l'intervalle $]0,1[$ (voir page 23) et Δu sur $]0,1[\times]0,1[$ (voir page 27). Le principe du maximum est aussi caractérisé de la manière suivante (plus difficile à utiliser en pratique) :

Proposition 2.9 (CNS de monotonie) *L'hypothèse 1 du théorème 2.8 est vérifiée si et seulement si A inversible et A^{-1} a des coefficients ≥ 0.*

Démonstration :
Supposons d'abord que l'hypothèse 1 du théorème 2.8 est vérifiée et montrons que A inversible et que A^{-1} a des coefficients ≥ 0. Si x est tel que $Ax = 0$, alors $Ax \geq 0$ et donc, par hypothèse, $x \geq 0$. Mais on a aussi $Ax \leq 0$, soit $A(-x) \geq 0$ et donc par hypothèse, $x \leq 0$. On en déduit $x = 0$, ce qui prouve que A est inversible.

L'hypothèse 1 donne alors que $y \geq 0 \Rightarrow A^{-1}y \geq 0$. En prenant $y = e_1$ on obtient que la première colonne de A^{-1} est positive, puis en prenant $y = e_i$ on obtient que la i-ème colonne de A^{-1} est positive, pour $i = 2, \ldots, N$. Donc A^{-1} a tous ses coefficients positifs.

Supposons maintenant que A est inversible et que A^{-1} a des coefficients positifs. Soit $x \in \mathbb{R}^N$ tel que $Ax = y \geq 0$, alors $x = A^{-1}y \geq 0$. Donc A vérifie l'hypothèse 1.

■

Théorème 2.10 (Généralisation du précédent)
Soit $A \in \mathcal{M}_N(\mathbb{R})$, $R \in C^1(\mathbb{R}^N, \mathbb{R}^N)$, $R = (R_1, \ldots, R_N)^t$ tels que

1. *Pour tout $\beta \geq 0$ et pour tout $x \in \mathbb{R}^N$, $Ax + \beta x \geq 0 \Rightarrow x \geq 0$*
2. $\dfrac{\partial R_i}{\partial x_j} \geq 0$, $\forall i, j$ *t.q.* $i \neq j$ *(R_i est monotone croissante par rapport à la variable x_j si $j \neq i$) et $\exists \gamma > 0$,* $-\gamma \leq \dfrac{\partial R_i}{\partial x_i} \leq 0, \forall x \in \mathbb{R}^N, \ \ \forall i \in \{1, \ldots, N\}$ *(R_i est monotone décroissante par rapport à la variable x_i).*
3. $0 \leq R(0)$ *(0 est sous-solution) et $\exists \tilde{x} \geq 0$ tel que $A(\tilde{x}) \geq R(\tilde{x})$ ($\tilde{x}$ est sur-solution).*

Soient $x^{(0)} = 0$, $\beta \geq \gamma$, et $(x^{(n)})_{n \in \mathbb{N}}$ la suite définie par $Ax^{(n+1)} + \beta x^{(n+1)} = R(x^{(n)}) + \beta x^{(n)}$. Cette suite converge vers $\bar{x} \in \mathbb{R}^N$ et $A\bar{x} = R(\bar{x})$. De plus, $0 \leq x^{(n)} \leq \tilde{x}$ $\forall n \in \mathbb{N}$ et $x^{(n)} \leq x^{(n+1)}$, $\forall n \in \mathbb{N}$.

Démonstration : On se ramène au théorème précédent avec $A + \beta Id$ au lieu de A et $R + \beta$ au lieu de R. ■

Remarque 2.11 (Point fixe de Brouwer) *On s'est intéressé ici uniquement à des théorèmes de point fixe "constructifs", i.e. qui donnent un algorithme pour le déterminer. Il existe aussi un théorème de point fixe dans $\mathbb{R}^N$ avec des hypothèses beaucoup plus générales (mais le théorème est non constructif), c'est le théorème de Brouwer[2] : si f est une fonction continue de la boule unité de $\mathbb{R}^N$ dans la boule unité, alors elle admet un point fixe dans la boule unité.*

2.1.3 Vitesse de convergence

Définition 2.12 (Vitesse de convergence) *Soit $(x^{(n)})_{n \in \mathbb{N}} \in \mathbb{R}^N$ et $\bar{x} \in \mathbb{R}^N$. On suppose que $x^{(n)} \to \bar{x}$ lorsque $n \to +\infty$, avec $x^{(n)} \neq \bar{x}$ pour tout $n \in \mathbb{N}$. On s'intéresse à la "vitesse de convergence" de la suite $(x^{(n)})_{n \in \mathbb{N}}$. On dit que :*

1. *la convergence est* **au moins linéaire** *s'il existe $\beta \in]0, 1[$ et il existe $n_0 \in \mathbb{N}$ tels que si $n \geq n_0$ alors $\|x^{(n+1)} - \bar{x}\| \leq \beta \|x^{(n)} - \bar{x}\|$.*
2. *La convergence* **est linéaire** *si il existe $\beta \in]0, 1[$ tel que*

$$\frac{\|x^{(n+1)} - \bar{x}\|}{\|x^{(n)} - \bar{x}\|} \longrightarrow \beta \text{ quand } n \to +\infty,$$

2. Luitzen Egbertus Jan Brouwer (1881-1966), mathématicien néerlandais.

3. *La convergence est* **super linéaire** *si*

$$\frac{\|x^{(n+1)} - \bar{x}\|}{\|x^{(n)} - \bar{x}\|} \longrightarrow 0 \text{ quand } n \to +\infty,$$

4. *La convergence est* **au moins quadratique** *si il existe* $\beta \in]0,1[$ *et il existe* $n_0 \in \mathbb{N}$ *tels que si* $n \geq n_0$ *alors* $\|x^{(n+1)} - \bar{x}\| \leq \beta \|x^{(n)} - \bar{x}\|^2$,
5. *La convergence est* **quadratique** *si*

$$\exists \beta > 0 \quad \frac{\|x^{(n+1)} - \bar{x}\|}{\|x^{(n)} - \bar{x}\|^2} \longrightarrow \beta \text{ quand } n \to +\infty.$$

Remarque 2.13 *La convergence quadratique est évidemment plus "rapide" que la convergence linéaire.*

Proposition 2.14 *Soit* $f \in C^1(\mathbb{R}, \mathbb{R})$ *; on suppose qu'il existe* $\bar{x} \in \mathbb{R}$ *tel que* $f(\bar{x}) = \bar{x}$*. On construit la suite*

$$\begin{aligned} & x^{(0)} \in \mathbb{R} \\ & x^{(n+1)} = f(x^{(n)}). \end{aligned}$$

1. *Si on suppose que* $f'(\bar{x}) \neq 0$ *et* $|f'(\bar{x})| < 1$*, alors il existe* $\alpha > 0$ *tel que si* $x^{(0)} \in I_\alpha = [\bar{x} - \alpha, \bar{x} + \alpha]$ *alors* $x^{(n)} \to \bar{x}$ *lorsque* $n \to +\infty$*, et si* $x^{(n)} \neq \bar{x}$*, alors* $\dfrac{|x^{(n+1)} - \bar{x}|}{|x^{(n)} - \bar{x}|} \to |f'(\bar{x})| = \beta$*, où* $\beta \in]0,1[$*. La convergence est donc linéaire.*
2. *Si on suppose maintenant que* $f'(\bar{x}) = 0$ *et* $f \in C^2(\mathbb{R}, \mathbb{R})$*,alors il existe* $\alpha > 0$ *tel que si* $x^{(0)} \in I_\alpha = [\bar{x} - \alpha, \bar{x} + \alpha]$*, alors* $x^{(n)} \to \bar{x}$ *quand* $n + \infty$*, et si* $x^{(n)} \neq \bar{x}$*,* $\forall n \in \mathbb{N}$ *alors*

$$\frac{|x^{(n+1)} - \bar{x}|}{|x^{(n)} - \bar{x}|^2} \to \beta = \frac{1}{2}|f''(\bar{x})|.$$

   La convergence est donc au moins quadratique.

Démonstration
1. Supposons que $|f'(\bar{x})| < 1$, et montrons qu'il existe $\alpha > 0$ tel que si $x^{(0)} \in I_\alpha$ alors $x^{(n)} \to \bar{x}$. Comme $f \in C^1(\mathbb{R}, \mathbb{R})$ il existe $\alpha > 0$ tel que $\gamma = \max_{x \in I_\alpha} |f'(x)| < 1$ (par continuité de f').
On va maintenant montrer que $f : I_\alpha \to I_\alpha$ est strictement contractante, on pourra alors appliquer le théorème du point fixe à $f_{|I_\alpha}$, (I_α étant fermé), pour obtenir que $x^{(n)} \to \bar{x}$ où $\bar{x}$ est l'unique point fixe de $f_{|I\alpha}$.
Soit $x \in I_\alpha$; montrons d'abord que $f(x) \in I_\alpha$: comme $f \in C^1(\mathbb{R}, \mathbb{R})$, il existe $\xi \in]x, \bar{x}[$ tel que $|f(x) - \bar{x}| = |f(x) - f(\bar{x})| = |f'(\xi)||x - \bar{x}| \leq \gamma |x - \bar{x}| < \alpha$, ce qui prouve que $f(x) \in I_\alpha$.
On vérifie alors que $f_{|I_\alpha}$ est strictement contractante en remarquant que pour tous $x, y \in I_\alpha$, $x < y$, il existe $\xi \in]x, y[(\subset I_\alpha)$ tel que $|f(x) - f(y)| = |f'(\xi)||x - y| \leq \gamma |x - y|$ avec $\gamma < 1$.

On a ainsi montré que $x^{(n)} \to \bar{x}$ si $x^{(0)} \in I_\alpha$.

Cherchons maintenant la vitesse de convergence de la suite. Supposons que $f'(\bar{x}) \neq 0$ et $x^{(n)} \neq \bar{x}$ pour tout $n \in \mathbb{N}$. Comme $x^{(n+1)} = f(x^{(n)})$ et $\bar{x} = f(\bar{x})$, on a $|x^{(n+1)} - \bar{x}| = |f(x^{(n)}) - f(\bar{x})|$. Comme $f \in C^1(\mathbb{R}, \mathbb{R})$, il existe $\xi_n \in]x^{(n)}, \bar{x}[$ ou $]\bar{x}, x^{(n)}[$, tel que $f(x^{(n)}) - f(\bar{x}) = f'(\xi_n)(x^{(n)} - \bar{x})$. On a donc

$$\frac{|x^{(n+1)} - \bar{x}|}{|x^{(n)} - \bar{x}|} = |f'(\xi_n)| \longrightarrow |f'(\bar{x})| \text{ car } x^{(n)} \to \bar{x} \text{ et } f' \text{ est continue.}$$

On a donc une convergence linéaire.

2. Supposons maintenant que $f'(\bar{x}) = 0$ et $f \in C^2(\mathbb{R}, \mathbb{R})$. On sait déjà par ce qui précède qu'il existe $\alpha > 0$ tel que si $x^{(0)} \in I_\alpha$ alors $x^{(n)} \to\to \bar{x}$ lorsque $n \to +\infty$. On veut estimer la vitesse de convergence. On suppose pour cela que $x^{(n)} \neq \bar{x}$ pour tout $n \in \mathbb{N}$.
Comme $f \in C^2(\mathbb{R}, \mathbb{R})$, il existe $\xi_n \in]x^{(n)}, \bar{x}[$ tel que $f(x^{(n)}) - f(\bar{x}) = f'(\bar{x})(x^{(n)} - \bar{x}) + \dfrac{1}{2}f''(\xi_n)(x^{(n)} - \bar{x})^2$.
On a donc : $x^{(n+1)} - \bar{x} = \frac{1}{2}f''(\xi_n)(x^{(n)} - \bar{x})^2$ ce qui entraîne que

$$\frac{|x^{(n+1)} - \bar{x}|}{|x^{(n)} - \bar{x}|^2} = \frac{1}{2}|f''(\xi_n)| \longrightarrow \frac{1}{2}|f''(\bar{x})| \text{ quand } n \to +\infty.$$

La convergence est donc quadratique.

■

On étudie dans le paragraphe suivant la méthode de Newton pour la résolution d'un système non linéaire. Donnons l'idée de la méthode de Newton dans le cas $N = 1$ à partir des résultats de la proposition précédente. Soit $g \in C^3(\mathbb{R}, \mathbb{R})$ et $\bar{x} \in \mathbb{R}$ tel que $g(\bar{x}) = 0$. On cherche une méthode de construction d'une suite $(x^{(n)})_n \in \mathbb{R}^N$ qui converge vers $\bar{x}$ de manière quadratique. On pose

$$f(x) = x + h(x)g(x) \text{ avec } h \in C^2(\mathbb{R}, \mathbb{R}) \text{ tel que } h(x) \neq 0 \ \forall x \in \mathbb{R}.$$

on a donc

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

Si par miracle $f'(\bar{x}) = 0$, la méthode de point fixe sur f va donner (pour $x^{(0)} \in I_\alpha$ donné par la proposition 2.14) $(x^{(n)})_{n\in\mathbb{N}}$ tel que $x^{(n)} \to \bar{x}$ de manière au moins quadratique. Or on a $f'(x) = 1 + h'(x)g(x) + g'(x)h(x)$ et donc $f'(\bar{x}) = 1 + g'(\bar{x})h(\bar{x})$. Il suffit donc de prendre h tel que $h(\bar{x}) = -\dfrac{1}{g'(\bar{x})}$. Ceci est possible si $g'(\bar{x}) \neq 0$.

En résumé, si $g \in C^3(\mathbb{R}, \mathbb{R})$ est telle que $g'(\overline{x}) \neq 0 \, \forall x \in \mathbb{R}$ et $g(\bar{x}) = 0$, on peut construire, pour x assez proche de $\bar{x}$, la fonction $f \in C^2(\mathbb{R}, \mathbb{R})$ définie par

$$f(x) = x - \frac{g(x)}{g'(x)}.$$

Grâce à la proposition 2.14, il existe $\alpha > 0$ tel que si $x^{(0)} \in I_\alpha$ alors la suite définie par $x^{(n+1)} = f(x^{(n)}) = x^{(n)} - \frac{g(x^{(n)})}{g'(x^{(n)})}$ converge vers $\bar{x}$ de manière au moins quadratique.
Remarquons que la construction de la suite de Newton s'écrit encore (dans le cas $N = 1$) $g'(x^{(n)})(x^{(n+1)} - x^{(n)}) = -g(x^{(n)})$ ou encore $g(x^{(n)}) + g'(x^{(n)})(x^{(n+1)} - x^{(n)}) = 0$.

2.2 Méthode de Newton

2.2.1 Construction et convergence de la méthode

On a vu ci-dessus comment se construit la méthode de Newton à partir du point fixe de monotonie en dimension $N = 1$. On va maintenant étudier cette méthode dans le cas N quelconque. Soient $g \in C^1(\mathbb{R}^N, \mathbb{R}^N)$ et $\bar{x} \in \mathbb{R}^N$ tels que $g(\bar{x}) = 0$.
On cherche une méthode de construction d'une suite $(x^{(n)})_n \in \mathbb{R}^N$ qui converge vers $\bar{x}$ de manière quadratique. L'algorithme de Newton de construction d'une telle suite s'écrit :

$$\left\{ \begin{array}{l} x^{(0)} \in \mathbb{R}^N \\ Dg(x^{(n)})(x^{(n+1)} - x^{(n)}) = -g(x^{(n)}), \ \forall n \geq 0. \end{array} \right. \tag{2.2.6}$$

(On rappelle que $Dg(x^{(n)})\in \mathcal{M}_N(\mathbb{R})$ est la matrice représentant la différentielle de g en $x^{(n)}$.)

Pour chaque $n \in \mathbb{N}$, il faut donc effectuer les opérations suivantes :

1. Calcul de $Dg(x^{(n)})$,
2. Résolution du système linéaire $Dg(x^{(n)})(x^{(n+1)} - x^{(n)}) = -g(x^{(n)})$.

Remarque 2.15 *Si la fonction g dont on cherche un zéro est linéaire, i.e. si g est définie par $g(x) = Ax - b$ avec $A \in \mathcal{M}_N(\mathbb{R})$ et $b \in \mathbb{R}^N$, alors la méthode de Newton revient à résoudre le système linéaire $Ax = b$. En effet $Dg(x^{(n)}) = A$ et donc (2.2.6) s'écrit $Ax^{(n+1)} = b$.*

Pour assurer la convergence et la qualité de la méthode, on va chercher maintenant à répondre aux questions suivantes :

1. la suite $(x^{(n)})_n$ est–elle bien définie ? A–t–on $Dg(x^{(n)})$ inversible ?
2. A–t–on convergence $x^{(n)} \to \bar{x}$ quand $n + \infty$?
3. La convergence est-elle au moins quadratique ?

Théorème 2.16 (Convergence de la méthode de Newton, I) *Soient $g \in C^2(\mathbb{R}^N, \mathbb{R}^N)$ et $\bar{x} \in \mathbb{R}^N$ tels que $g(\bar{x}) = 0$. On munit $\mathbb{R}^N$ d'une norme $\|\cdot\|$. On suppose que $Dg(\bar{x})$ est inversible. Alors il existe $b > 0$, et $\beta > 0$ tels que*

1. *si $x^{(0)} \in B(\bar{x}, b) = \{x \in \mathbb{R}^N, \|x - \bar{x}\| < b\}$ alors la suite $(x^{(n)})_{n\in\mathbb{N}}$ est bien définie par (2.2.6) et $x^{(n)} \in B(\bar{x}, b)$ pour tout $n \in \mathbb{N}$,*
2. *si $x^{(0)} \in B(\bar{x}, b)$ et si la suite $(x^{(n)})_{n\in\mathbb{N}}$ est définie par (2.2.6) alors $x^{(n)} \to \bar{x}$ quand $n \to +\infty$,*
3. *si $x^{(0)} \in B(\bar{x}, b)$ et si la suite $(x^{(n)})_{n\in\mathbb{N}}$ est définie par (2.2.6) alors $\|x^{(n+1)} - \bar{x}\| \le \beta\|x^{(n)} - \bar{x}\|^2\ \forall n \in \mathbb{N}$.*

Pour démontrer ce théorème, on va commencer par démontrer le théorème suivant, qui utilise des hypothèses plus faibles mais pas très faciles à vérifier en pratique :

Théorème 2.17 (Convergence de la méthode de Newton, II)
Soient $g \in C^1(\mathbb{R}^N, \mathbb{R}^N)$ et $\bar{x} \in \mathbb{R}^N$ tels que $g(\bar{x}) = 0$. On munit $\mathbb{R}^N$ d'une norme $\|\cdot\|$ et $\mathcal{M}_N(\mathbb{R})$ de la norme induite. On suppose que $Dg(\bar{x})$ est inversible. On suppose de plus qu'il existe $a,\ a_1,\ a_2 \in \mathbb{R}_+^*$ tels que :

1. *si $x \in B(\bar{x}, a)$ alors $Dg(x)$ est inversible et $\|Dg(x))^{-1}\| \le a_1$;*
2. *si $x, y \in B(\bar{x}, a)$ alors $\|g(y) - g(x) - Dg(x)(y - x)\| \le a_2\|y - x\|^2$.*

Alors, si on pose : $b = \min\left(a, \dfrac{1}{a_1a_2}\right) > 0$, $\beta = a_1a_2$ et si $x^{(0)} \in B(\bar{x}, b)$, on a :

1. *$(x^{(n)})_{n\in\mathbb{N}}$ est bien définie par (2.2.6),*
2. *$x^{(n)} \to \bar{x}$ lorsque $n \to +\infty$,*
3. *$\|x^{(n+1)} - \bar{x}\| \le \beta\|x^{(n)} - \bar{x}\|^2\ \ \forall n \in \mathbb{N}$.*

Démonstration du théorème 2.17
Soit $x^{(0)} \in B(\bar{x}, b) \subset B(\bar{x}, a)$ où $b \le a$. On va montrer par récurrence sur n que $x^{(n)} \in B(\bar{x}, b)\ \forall n \in \mathbb{N}$ (et que $(x^{(n)})_{n\in\mathbb{N}}$ est bien définie). L'hypothèse de récurrence est que $x^{(n)}$ est bien défini, et que $x^{(n)} \in B(\bar{x}, b)$. On veut montrer que $x^{(n+1)}$ est bien défini et $x^{(n+1)} \in B(\bar{x}, b)$.
Comme $b \le a$, la matrice $Dg(x^{(n)})$ est inversible et $x^{(n+1)}$ est donc bien défini ; on a : $x^{(n+1)} - x^{(n)} = Dg(x^{(n)})^{-1}(-g(x^{(n)}))$. Pour montrer que $x^{(n+1)} \in B(\bar{x}, b)$ on va utiliser le fait que $b \le \dfrac{1}{a_1a_2}$.
Par hypothèse, on sait que si $x, y \in B(\bar{x}, a)$, on a

$$\|g(y) - g(x) - Dg(x)(y - x)\| \le a_2\|y - x\|^2.$$

Prenons $y = \bar{x}$ et $x = x^{(n)} \in B(\bar{x}, a)$ dans l'inégalité ci-dessus. On obtient alors :

$$\|g(\bar{x}) - g(x^{(n)}) - Dg(x^{(n)})(\bar{x} - x^{(n)})\| \leq a_2\|\bar{x} - x^{(n)}\|^2.$$

Comme $g(\bar{x}) = 0$ et par définition de $x^{(n+1)}$, on a donc :

$$\|Dg(x^{(n)})(x^{(n+1)} - x^{(n)}) - Dg(x^{(n)})(\bar{x} - x^{(n)})\| \leq a_2\|\bar{x} - x^{(n)}\|^2,$$

et donc

$$\|Dg(x^{(n)})(x^{(n+1)} - \bar{x})\| \leq a_2\|\bar{x} - x^{(n)}\|^2. \tag{2.2.7}$$

Or $x^{(n+1)} - \bar{x} = [Dg(x^{(n)})]^{-1}(Dg(x^{(n)}))(x^{(n+1)} - \bar{x})$, et donc

$$\|x^{(n+1)} - \bar{x}\| \leq \|Dg(x^{(n)})^{-1}\| \ \|Dg(x^{(n)})(x^{(n+1)} - \bar{x})\|.$$

En utilisant (2.2.7), les hypothèses 1 et 2 et le fait que $x^{(n)} \in B(\bar{x}, b)$, on a donc

$$\|x^{(n+1)} - \bar{x}\| \leq a_1a_2\|x^{(n)} - \bar{x}\|^2 < a_1a_2b^2. \tag{2.2.8}$$

Or $a_1a_2b^2 < b$ car $b \leq \dfrac{1}{a_1a_2}$. Donc $x^{(n+1)} \in B(\bar{x}, b)$.
On a ainsi montré par récurrence que la suite $(x^{(n)})_{n\in\mathbb{N}}$ est bien définie et que $x^{(n)} \in B(\bar{x}, b)$ pour tout $n \geq 0$.
Pour montrer la convergence de la suite $(x^{(n)})_{n\in\mathbb{N}}$ vers $\bar{x}$, on repart de l'inégalité (2.2.8) :

$$a_1a_2\|x^{(n+1)} - \bar{x}\| \leq (a_1a_2)^2\|\bar{x} - x^{(n)}\|^2 = (a_1a_2\|x^{(n)} - \bar{x}\|)^2, \ \forall n \in \mathbb{N},$$

et donc par récurrence $a_1a_2\|x^{(n)} - \bar{x}\| \leq (a_1a_2\|x^{(0)} - \bar{x}\|)^{2^n}$ $\forall n \in \mathbb{N}$. Comme $x^{(0)} \in B(\bar{x}, b)$ et $b \leq \frac{1}{a_1a_2}$, on a $(a_1a_2\|x^{(0)} - \bar{x}\|) < 1$ et donc $\|x^{(n)} - \bar{x}\| \to 0$ quand $n \to +\infty$.

La convergence est au moins quadratique car l'inégalité (2.2.8) s'écrit : $\|x^{(n+1)} - \bar{x}\| \leq \beta\|x^{(n)} - x\|^2$ avec $\beta = a_1a_2$. ■

Démonstration du théorème 2.16
Soient $g \in C^2(\mathbb{R}^N, \mathbb{R}^N)$ et $\bar{x} \in \mathbb{N}$ tels que $g(\bar{x}) = 0$. Par hypothèse, $Dg(\bar{x})$ est inversible. Il suffit de démontrer (pour se ramener au théorème 2.17) qu'il existe $a, a_1, a_2 \in \mathbb{R}_+^*$ tels que

1. si $x \in B(\bar{x}, a)$ alors $Dg(x)$ est inversible et $\|(Dg(x))^{-1}\| \leq a_1$,
2. si $x, y \in B(\bar{x}, a)$ alors $\|g(y) - g(x) - Dg(x)(y - x)\| \leq a_2\|y - x\|^2$.

Remarquons d'abord que $Dg(x) = Dg(\bar{x}) - Dg(\bar{x}) + Dg(x) = Dg(\bar{x})(Id + S)$ où $S = Dg(\bar{x})^{-1}(Dg(x) - Dg(\bar{x}))$. Or si $\|S\| < 1$, la matrice $(Id + S)$ est inversible et $\|(Id + S)^{-1}\| \leq \frac{1}{1-\|S\|}$. Nous allons donc essayer de majorer $\|S\|$. Par définition de S, on a :

$$\|S\| \leq \|Dg(\bar{x})^{-1}\| \ \ \|Dg(x) - Dg(\bar{x}).\|$$

Comme $g \in C^2(\mathbb{R}^N, \mathbb{R}^N)$, on a $Dg \in C^1(\mathbb{R}^N, \mathcal{M}_N(\mathbb{R})))$; donc par continuité de Dg, pour tout $\varepsilon \in \mathbb{R}_+^*$, il existe $a \in \mathbb{R}_+^*$ tel que si $\|x - \overline{x}\| \leq a$ alors $\|Dg(x) - Dg(\overline{x})\| \leq \varepsilon$. En prenant $\varepsilon = \frac{1}{2\|Dg(\bar{x})^{-1}\|}$, il existe donc $a > 0$ tel que si $x \in B(\bar{x}, a)$ alors $\|Dg(x) - Dg(\bar{x})\| \leq \frac{1}{2\|Dg(\bar{x})^{-1}\|}$, et donc si $x \in B(\bar{x}, a)$, alors $\|S\| \leq \dfrac{1}{2}$. On en déduit que si $x \in B(\bar{x}, a)$ alors $Id + S$ est inversible et donc que $Dg(x) = Dg(\bar{x})(Id + S)$ est inversible (on rappelle que $Dg(\bar{x})$ est inversible par hypothèse). De plus, si $x \in B(\bar{x}, a)$ on a : $\|(Id + S)^{-1}\| \leq \frac{1}{1-\|S\|} \leq 2$ et comme $(Id + S)^{-1} = (Dg(\bar{x}))^{-1}Dg(x)$, on a $\|Dg(x)^{-1}Dg(\bar{x})\| \leq 2$, et donc $\|Dg(x)^{-1}\| \leq \|(Dg(\overline{x}))^{-1}\|\|(Dg(x))^{-1}Dg(\overline{x})\| \leq 2\|(Dg(\overline{x}))^{-1}\|$.

En résumé, on a donc prouvé l'existence de a et de $a_1 = 2\|Dg(\bar{x})^{-1}\|$ tels que si $x \in B(\bar{x}, a)$ alors $Dg(x)$ est inversible et $\|Dg(x)^{-1}\| \leq a_1$. Il reste maintenant à trouver a_2 tel que si $x, y \in B(\bar{x}, a)$ alors $\|g(y) - g(x) - Dg(x)(y - x)\| \leq a_2\|y - x\|^2$.
Comme $g \in C^2(\mathbb{R}^N, \mathbb{R}^N)$, on a donc $Dg \in C^1(\mathbb{R}^N, \mathcal{M}_N(\mathbb{R})))$ (remarquons que jusqu'à présent on avait utilisé uniquement le caractère C^1 de g). On définit la fonction $\varphi \in C^1(\mathbb{R}, \mathbb{R}^N)$ par $\varphi(t) = g(x + t(y - x)) - g(x) - tDg(x)(y - x)$. On a donc $\varphi(1) = g(y) - g(x) - Dg(x)(y - x)$ (c'est le terme qu'on veut majorer en norme) et $\varphi(0) = 0$. On écrit maintenant que φ est l'intégrale de sa dérivée :

$$\varphi(1) - \varphi(0) = \int_0^1 \varphi'(t)dt = \int_0^1 Dg(x + t(y - x))(y - x) - Dg(x)(y - x)dt.$$

On a donc

$$\begin{aligned} \|\varphi(1) - \varphi(0)| &= \|g(y) - g(x) - Dg(x)(y - x)\| \\ &\leq \int_0^1 \|Dg(x + t(y - x))(y - x) - Dg(x)(y - x)\|dt \\ &\leq \|y - x\| \int_0^1 \|Dg(x + t(y - x)) - Dg(x)\|dt. \end{aligned} \tag{2.2.9}$$

Pour majorer $\|Dg(x + t(y - x)) - Dg(x))\|$, on utilise alors le théorème des accroissements finis [3] (parfois aussi appelé "théorème de la moyenne") appliqué à Dg ; de l'inégalité (2.2.9), on tire donc que pour $x, y \in B(\bar{x}, a)$ et $t \in]0, 1[$:

$$\|Dg(x + t(y - x)) - Dg(x)\| \leq \qquad t\|y - x\| \sup_{c \in B(\bar{x}, a)} \|D(Dg)(c)\|_{\mathcal{L}(\mathbb{R}^N, \mathcal{M}_N(\mathbb{R}))}. \tag{2.2.10}$$

Comme $D(Dg) = D^2g$ est continue par hypothèse, et comme $B(\bar{x}, a)$ est inclus dans un compact, on a

$$a_2 = \sup_{c \in B(\bar{x}, a)} \|D(Dg)(c)\|_{\mathcal{L}(\mathbb{R}^N, \mathcal{M}_N(\mathbb{R}))} < +\infty.$$

De plus, $t < 1$ et on déduit de (2.2.10) que :

$$\|Dg(x + t(y - x) - Dg(x))\| \leq a_2\|y - x\|,$$

et de l'inégalité (2.2.9) on déduit ensuite que

$$\|g(y) - g(x) - Dg(x)(y - x)\| \leq \int_0^1 a_2\|y - x\|dt\, \|y - x\| = a_2\|y - x\|^2.$$

On a donc ainsi démontré que g vérifie les hypothèses du théorème 2.17, ce qui termine la démonstration du théorème 2.16.

Remarque 2.18 *On ne sait pas bien estimer b dans le théorème 2.16, et ceci peut poser problème lors de l'implantation numérique : il faut choisir l'itéré initial $x^{(0)}$ "suffisamment proche" de $\overline{x}$ pour avoir convergence.*

3. **Théorème des accroissements finis :** Soient E et F des espaces vectoriels normés, soient $h \in C^1(E, F)$ et $(x, y) \in E^2$. On définit $]x, y[= \{tx + (1 - t)y, t \in]0, 1[\}$. Alors : $\|h(y) - h(x)\| \leq \|y - x\| \sup_{z \in]x, y[} \|Dh(z)\|_{\mathcal{L}(E,F)}$.
(On rappelle que si $T \in \mathcal{L}(E, F)$ alors $\|T\|_{[\mathcal{L}(E,F)} = \sup_{x \in E, \|x\|_E = 1} \|Tx\|_F|$.)
Attention : Si $\dim F > 1$, on ne peut pas dire, comme c'est le cas en dimension 1, que :$\exists \xi \in]x, y[$ t.q. $h(y) - h(x) = Dh(\xi)(y - x)$.

2.2.2 Variantes de la méthode de Newton

L'avantage majeur de la méthode de Newton par rapport à une méthode de point fixe par exemple est sa vitesse de convergence d'ordre 2. On peut d'ailleurs remarquer que lorsque la méthode ne converge pas, par exemple si l'itéré initial $x^{(0)}$ n'a pas été choisi "suffisamment proche" de $\overline{x}$, alors la méthode diverge très vite…

L'inconvénient majeur de la méthode de Newton est son coût : on doit d'une part calculer la matrice jacobienne $Dg(x^{(n)})$ à chaque itération, et d'autre part la factoriser pour résoudre le système linéaire $Dg(x^{(n)})(x^{(n+1)} - x^{(n)}) = -g(x^{(n)})$. (On rappelle que pour résoudre un système linéaire, il ne faut pas calculer l'inverse de la matrice, mais plutôt la factoriser sous la forme LU par exemple, et on calcule ensuite les solutions des systèmes avec matrice triangulaires faciles á inverser, voir Chapitre 1.) Plusieurs variantes ont été proposées pour tenter de réduire ce coût.

"Faux quasi Newton"

Soient $g \in C^1(\mathbb{R}^N, \mathbb{R}^N)$ et $\bar{x} \in \mathbb{R}$ tels que $g(\bar{x}) = 0$. On cherche à calculer $\bar{x}$. Si on le fait par la méthode de Newton, l'algorithme s'écrit :

$$\left\{ \begin{array}{l} x^{(0)} \in \mathbb{R}^N, \\ Dg(x^{(n)})(x^{(n+1)} - x^{(n)}) = -g(x^{(n)}), \;\; n \geq 0. \end{array} \right.$$

La méthode du "Faux quasi-Newton" (parfois appelée quasi-Newton) consiste à remplacer le calcul de la matrice jacobienne $Dg(x^{(n)})$ à chaque itération par un calcul toutes les "quelques" itérations. On se donne une suite $(n_i)_{i\in\mathbb{N}}$, avec $n_0 = 0$ et $n_{i+1} > n_i \; \forall i \in \mathbb{N}$, et on calcule la suite $(x^{(n)})_{n\in\mathbb{N}}$ de la manière suivante :

$$\left\{ \begin{array}{l} x^{(0)} \in \mathbb{R}^N \\ Dg(x^{(n_i)})(x^{(n+1)} - x^{(n)}) = -g(x^{(n)}) \text{ si } n_i \leq n < n_{i+1}. \end{array} \right. \tag{2.2.11}$$

Avec cette méthode, on a moins de calculs et de factorisations de la matrice jacobienne $Dg(x)$ à effectuer, mais on perd malheureusement la convergence quadratique : cette méthode n'est donc pas très utilisée en pratique.

Newton incomplet

On suppose que g s'écrit sous la forme : $g(x) = Ax + F_1(x) + F_2(x)$, avec $A \in \mathcal{M}_N(\mathbb{R})$ et $F_1, F_2 \in C^1(\mathbb{R}^N, \mathbb{R}^N)$. L'algorithme de Newton (2.2.6) s'écrit alors :

$$\left\{ \begin{array}{l} x^{(0)} \in \mathbb{R}^N \\ \big(A + DF_1(x^{(n)}) + DF_2(x^{(n)})\big)\,(x^{(n+1)} - x^{(n)}) = \\ \qquad\qquad -Ax^{(n)} - F_1(x^{(n)}) - F_2(x^{(n)}). \end{array} \right.$$

La méthode de Newton incomplet consiste à ne pas tenir compte de la jacobienne de F_2.

$$\left\{ \begin{array}{l} x^{(0)} \in \mathbb{R}^N \\ (A + DF_1(x^{(n)}))(x^{(n+1)} - x^{(n)}) = -Ax^{(n)} - F_1(x^{(n)}) - F_2(x^{(n)}). \end{array} \right. \tag{2.2.12}$$

On dit qu'on fait du "Newton sur F_1" et du "point fixe sur F_2". Les avantages de cette procédure sont les suivants :
– La méthode ne nécessite pas le calcul de $DF_2(x)$, donc on peut l'employer si $F_2 \in C(\mathbb{R}^N, \mathbb{R}^N))$ n'est pas dérivable.
– On peut choisir F_1 et F_2 de manière à ce que la structure de la matrice $A + DF_1(x^{(n)})$ soit "meilleure" que celle de la matrice $A + DF_1(x^{(n)}) + DF_2(x^{(n)})$; si par exemple A est la matrice issue de la discrétisation du Laplacien, c'est une matrice creuse. On peut vouloir conserver cette structure et choisir F_1 et F_2 de manière à ce que la matrice $A + DF_1(x^{(n)})$ ait la même structure que A.

– Dans certains problèmes, on connaît a priori les couplages plus ou moins forts dans les non-linéarités : un couplage est dit fort si la variation d'une variable entraîne une variation forte du terme qui en dépend. Donnons un exemple : Soit f de $\mathbb{R}^2$ dans $\mathbb{R}^2$ définie par $f(x,y) = (x + \sin(10^{-5}y), \exp(x) + y)$, et considérons le système non linéaire $f(x,y) = (a,b)^t$ où $(a,b)^t \in \mathbb{R}^2$ est donné. Il est naturel de penser que pour ce système, le terme de couplage de la première équation en la variable y sera faible, alors que le couplage de deuxième équation en la variable x sera fort.
  On a alors intérêt à mettre en oeuvre la méthode de Newton sur la partie "couplage fort" et une méthode de point fixe sur la partie "couplage faible".

L'inconvénient majeur est la perte de la convergence quadratique. La méthode de Newton incomplet est cependant assez souvent employée en pratique en raison des avantages énumérés ci-dessus.

Remarque 2.19 *Si $F_2 = 0$, alors la méthode de Newton incomplet est exactement la méthode de Newton. Si $F_1 = 0$, la méthode de Newton incomplet s'écrit $A(x^{(n+1)} - x^{(n)}) = -Ax^{(n)} - F_2(x^{(n)})$, elle s'écrit alors $Ax^{(n+1)} = -F_2(x^{(n)})$, ou encore $x^{(n+1)} = -A^{-1}F_2(x^{(n)})$ si A inversible. C'est donc dans ce cas la méthode du point fixe sur la fonction $-A^{-1}F_2$.*

Méthode de la sécante

La méthode de la sécante est une variante de la méthode de Newton dans le cas de la dimension 1 d'espace. On suppose ici $N = 1$ et $g \in C^1(\mathbb{R}, \mathbb{R})$. La méthode de Newton pour calculer $\overline{x} \in \mathbb{R}$ tel que $g(\overline{x}) = 0$ s'écrit :

$$\begin{cases} x^{(0)} \in \mathbb{R} \\ g'(x^{(n)})(x^{(n+1)} - x^{(n)}) = -g(x^{(n)}), \quad \forall n \geq 0. \end{cases}$$

On aimerait simplifier le calcul de $g'(x^{(n)})$, c'est–à–dire remplacer $g'(x^{(n)})$ par une quantité "proche" sans calculer g'. Pour cela, on remplace la dérivée par un quotient différentiel. On obtient la méthode de la sécante :

$$\begin{cases} x^{(0)}, x^{(1)} \in \mathbb{R} \\ \dfrac{g(x^{(n)}) - g(x^{(n-1)})}{x^{(n)} - x^{(n-1)}}(x^{(n+1)} - x^{(n)}) = -g(x^{(n)}) \quad n \geq 1. \end{cases} \tag{2.2.13}$$

Remarquons que dans la méthode de la sécante, $x^{(n+1)}$ dépend de $x^{(n)}$ et de $x^{(n-1)}$: on a une méthode à deux pas ; on a d'ailleurs besoin de deux itérés initiaux $x^{(0)}$ et $x^{(1)}$. L'avantage de cette méthode est qu'elle ne nécessite pas le calcul de g'. L'inconvénient est qu'on perd la convergence quadratique. On peut toutefois montrer (voir exercice 64 page 105) que si $g(\bar{x}) = 0$ et $g'(\bar{x}) \neq 0$, il existe $\alpha > 0$ tel que si $x^{(0)}, x^{(1)} \in [\bar{x} - \alpha.\bar{x} + \alpha] = I_\alpha$, la suite $(x^{(n)})_{n\in\mathbb{N}}$ construite par la méthode de la sécante (2.2.13) est bien définie, que $(x^{(n)})_{n\in\mathbb{N}} \subset I_\alpha$ et que $x^{(n)} \to \bar{x}$ quand $n \to +\infty$. De plus, la convergence est super linéaire, *i.e.* si $x^{(n)} \neq \bar{x}$ pour tout $n \in \mathbb{N}$, alors $\dfrac{x^{(n+1)} - \bar{x}}{x^{(n)} - \bar{x}} \to 0$ quand $n \to +\infty$. On peut même montrer (voir exercice 64 page 105) que la méthode de la sécante est convergente d'ordre d, où d est le nombre d'or.

Méthodes de type "Quasi Newton"

On veut généraliser la méthode de la sécante au cas $N > 1$. Soient donc $g \in C^1(\mathbb{R}^N, \mathbb{R}^N)$. Pour éviter de calculer $Dg(x^{(n)})$ dans la méthode de Newton (2.2.6), on va remplacer $Dg(x^{(n)})$ par $B^{(n)} \in \mathcal{M}_N(\mathbb{R})$ "proche de $Dg(x^{(n)})$". En s'inspirant de la méthode de la sécante en dimension 1, on cherche une matrice $B^{(n)}$ qui, $x^{(n)}$ et $x^{(n-1)}$ étant connus, vérifie la condition :

$$B^{(n)}(x^{(n)} - x^{(n-1)}) = g(x^{(n)}) - g(x^{(n-1)}) \tag{2.2.14}$$

Dans le cas où $N = 1$, cette condition détermine entièrement $B^{(n)}$: $B^{(n)} = \dfrac{g(x^{(n)}) - g(x^{(n-1)})}{x^{(n)} - x^{(n-1)}}$. Si $N > 1$, ceci ne permet pas de déterminer complètement $B^{(n)}$. Il y a plusieurs façons possibles de choisir $B^{(n)}$, nous en verrons en particulier dans le cadre des méthodes d'optimisation (voir chapitre 4, dans ce cas la fonction g est un gradient), nous donnons ici la méthode de Broyden[4].. Celle-ci consiste à choisir $B^{(n)}$ de la manière suivante : à $x^{(n)}$ et $x^{(n-1)}$ connus, on pose $\delta^{(n)} = x^{(n)} - x^{(n-1)}$ et $y^{(n)} = g(x^{(n)}) - g(x^{(n-1)})$; on suppose $B^{(n-1)} \in \mathcal{M}_N(\mathbb{R})$ connue, et on cherche $B^{(n)} \in \mathcal{M}_N(\mathbb{R})$ telle que

$$B^{(n)}\delta^{(n-1)} = y^{(n-1)} \tag{2.2.15}$$

(c'est la condition (2.2.14), qui ne suffit pas à déterminer $B^{(n)}$ de manière unique) et qui vérifie également :

$$B^{(n)}\xi = B^{(n-1)}\xi, \quad \forall \xi \in \mathbb{R}^N \text{ tel que } \xi \perp \delta^{(n)}. \tag{2.2.16}$$

Proposition 2.20 (Existence et unicité de la matrice de Broyden)
Soient $y^{(n)} \in \mathbb{R}^N$, $\delta^{(n)} \in \mathbb{R}^N$ et $B^{(n-1)} \in \mathcal{M}_N(\mathbb{R})$. Il existe une unique matrice $B^{(n)} \in \mathcal{M}_N(\mathbb{R})$ vérifiant (2.2.15) *et* (2.2.16) *; la matrice $B^{(n)}$ s'exprime en fonction de $y^{(n)}$, $\delta^{(n)}$ et $B^{(n-1)}$ de la manière suivante :*

$$B^{(n)} = B^{(n-1)} + \frac{y^{(n)} - B^{(n-1)}\delta^{(n)}}{\delta^{(n)} \cdot \delta^{(n)}} (\delta^{(n)})^t. \tag{2.2.17}$$

Démonstration : L'espace des vecteurs orthogonaux à $\delta^{(n)}$ est de dimension $N-1$. Soit $(\gamma_1, \dots, \gamma_{N-1})$ une base de cet espace, alors $(\gamma_1, \dots, \gamma_{N-1}, \delta^{(n)})$ est une base de $\mathbb{R}^N$ et si $B^{(n)}$ vérifie (2.2.15) et (2.2.16), les valeurs prises par l'application linéaire associée à $B^{(n)}$ sur chaque vecteur de base sont connues, ce qui détermine l'application linéaire et donc la matrice $B^{(n)}$ de manière unique. Soit $B^{(n)}$ définie par (2.2.17), on a :

$$B^{(n)}\delta^{(n)} = B^{(n-1)}\delta^{(n)} + \frac{y^{(n)} - B^{(n-1)}\delta^{(n)}}{\delta^{(n)} \cdot \delta^{(n)}} (\delta^{(n)})^t \delta^{(n)} = y^{(n)},$$

et donc $B^{(n)}$ vérifie (2.2.15). Soit $\xi \in \mathbb{R}^N$ tel que $\xi \perp \delta^{(n)}$, alors $\xi \cdot \delta^{(n)} = (\delta^{(n)})^t \xi = 0$ et donc

$$B^{(n)}\xi = B^{(n-1)}\xi + \frac{(y^{(n)} - B^{(n-1)}\delta^{(n)})}{\delta^{(n)} \cdot \delta^{(n)}} (\delta^{(n)})^t \xi = B^{(n-1)}\xi, \ \forall \xi \perp \delta^{(n)}.$$

■

L'algorithme de Broyden s'écrit donc :

$$\left\{\begin{array}{l}
\text{Initialisation : } x^{(0)},\ x^{(1)} \in \mathbb{R}^N \ \ B_0 \in \mathcal{M}_N(\mathbb{R}) \\
\text{A } x^{(n)},\ x^{(n-1)} \text{ et } B^{(n-1)} \text{ connus, on pose} \\
\qquad\qquad \delta^{(n)} = x^{(n)} - x^{(n-1)} \text{ et } y^{(n)} = g(x^{(n)}) - g(x^{(n-1)}); \\
\text{Calcul de } B^{(n)} = B^{(n-1)} + \frac{y^{(n)} - B^{(n-1)}\delta^{(n)}}{\delta^{(n)} \cdot \delta^{(n)}} (\delta^{(n)})^t, \\
\text{Itération } n \text{ : résolution de : } B^{(n)}(x^{(n+1)} - x^{(n)}) = -g(x^{(n)}).
\end{array}\right.$$

Une fois de plus, l'avantage de cette méthode est de ne pas nécessiter le calcul de $Dg(x)$, mais l'inconvénient est la perte du caractère quadratique de la convergence .

4. C. G. Broyden, "A Class of Methods for Solving Nonlinear Simultaneous Equations." *Math. Comput.* 19, 577-593, 1965

2.3 Exercices

Exercice 46 (Calcul différentiel) *Suggestions en page 105, corrigé détaillé en page 108*

Soit $f \in C^2(\mathbb{R}^N, \mathbb{R})$.
1. Montrer que pour tout $x \in \mathbb{R}^N$, il existe un unique vecteur $a(x) \in \mathbb{R}^N$ tel que $Df(x)(h) = a(x) \cdot h$ pour tout $h \in \mathbb{R}^N$.
Montrer que $(a(x))_i = \partial_i f(x)$.

2. On pose $\nabla f(x) = (\partial_1 f(x), \ldots, \partial_1 f(x))^t$. Soit φ l'application définie de $\mathbb{R}^N$ dans $\mathbb{R}^N$ par $\varphi(x) = \nabla f(x)$. Montrer que $\varphi \in C^1(\mathbb{R}^N, \mathbb{R}^N)$ et que $D\varphi(x)(y) = A(x)y$, où $A(x))_{i,j} = \partial^2_{i,j} f(x)$.

Exercice 47 (Calcul différentiel, suite)

1. Soit $f \in C^2(\mathbb{R}^2, \mathbb{R})$ la fonction définie par $f(x_1, x_2) = ax_1 + bx_2 + cx_1x_2$, où $a, b,$ et c sont trois réels fixés. Donner la définition et l'expression de $Df(x)$, $\nabla f(x)$,Df, $D^2 f(x)$, $H_f(x)$.
2. Même question pour la fonction $f \in C^2(\mathbb{R}^3, \mathbb{R})$ définie par $f(x_1, x_2, x_3) = x_1^2 + x_1^2 x_2 + x_2 \sin(x_3)$.

Exercice 48 (Méthode de monotonie) *Suggestions en page 105, corrigé détaillé en page 108.*

On suppose que $f \in C^1(\mathbb{R}, \mathbb{R})$, $f(0) = 0$ et que f est croissante. On s'intéresse, pour $\lambda > 0$, au système non linéaire suivant de N équations à N inconnues (notées $u_1, \ldots, u_N$) :

$$\begin{aligned} &(Au)_i = \alpha_i f(u_i) + \lambda b_i \ \forall i \in \{1, \ldots, N\}, \\ &u = (u_1, \ldots, u_N)^t \in \mathbb{R}^N, \end{aligned} \tag{2.3.18}$$

où $\alpha_i > 0$ pour tout $i \in \{1, \ldots, N\}$, $b_i \geq 0$ pour tout $i \in \{1, \ldots, N\}$ et $A \in \mathcal{M}_N(\mathbb{R})$ est une matrice vérifiant

$$u \in \mathbb{R}^N, \ Au \geq 0 \Rightarrow u \geq 0. \tag{2.3.19}$$

On suppose qu'il existe $\mu > 0$ t.q. (2.3.18) ait une solution, notée $u^{(\mu)}$, pour $\lambda = \mu$. On suppose aussi que $u^{(\mu)} \geq 0$. Soit $0 < \lambda < \mu$. On définit la suite $(v^{(n)})_{n \in \mathbb{N}} \subset \mathbb{R}^N$ par $v^{(0)} = 0$ et, pour $n \geq 0$,

$$(Av^{(n+1)})_i = \alpha_i f(v_i^{(n)}) + \lambda b_i \ \forall i \in \{1, \ldots, N\}. \tag{2.3.20}$$

Montrer que la suite $(v^{(n)})_{n \in \mathbb{N}}$ est bien définie, convergente (dans $\mathbb{R}^N$) et que sa limite, notée $u^{(\lambda)}$, est solution de (2.3.18) (et vérifie $0 \leq u^{(\lambda)} \leq u^{(\mu)}$).

Exercice 49 (Point fixe amélioré) *Suggestions en page 106,*

Soit $g \in C^3(\mathbb{R}, \mathbb{R})$ et $\overline{x} \in \mathbb{R}$ tels que $g(\overline{x}) = 0$ et $g'(\overline{x}) \neq 0$.
On se donne $\varphi \in C^1(\mathbb{R}, \mathbb{R})$ telle que $\varphi(\overline{x}) = \overline{x}$.
On considère l'algorithme suivant :

$$\left\{ \begin{array}{l} x_0 \in \mathbb{R}, \\ \\ x_{n+1} = h(x_n), n \geq 0. \end{array} \right. \tag{2.3.21}$$

avec $h(x) = x - \dfrac{g(x)}{g'(\varphi(x))}$

1) Montrer qu'il existe $\alpha > 0$ tel que si $x_0 \in [\overline{x} - \alpha, \overline{x} + \alpha] = I_\alpha$, alors la suite donnée par l'algorithme (2.3.21) est bien définie ; montrer que $x_n \to \overline{x}$ lorsque $n \to +\infty$.

On prend maintenant $x_0 \in I_\alpha$ où α est donné par la question 1.

2) Montrer que la convergence de la suite $(x_n)_{n\in\mathbb{N}}$ définie par l'algorithme (2.3.21) est au moins quadratique.

3) On suppose que φ' est lipschitzienne et que $\varphi'(\overline{x}) = \dfrac{1}{2}$. Montrer que la convergence de la suite $(x_n)_{n\in\mathbb{N}}$ définie par (2.3.21) est au moins cubique, c'est-à-dire qu'il existe $c \in \mathbb{R}_+$ tel que

$$|x_{n+1} - \overline{x}| \le c|x_n - \overline{x}|^3, \qquad \forall n \ge 1.$$

4) Soit $\beta \in \mathbb{R}_+^*$ tel que $g'(x) \neq 0 \quad \forall x \in I_\beta =]\overline{x} - \beta, \overline{x} + \beta[$; montrer que si on prend φ telle que :

$$\varphi(x) = x - \frac{g(x)}{2g'(x)} \qquad \text{si } x \in I_\beta,$$

alors la suite définie par l'algorithme (2.3.21) converge de manière cubique.

Exercice 50 (Méthode de Newton pour un système 2×2)

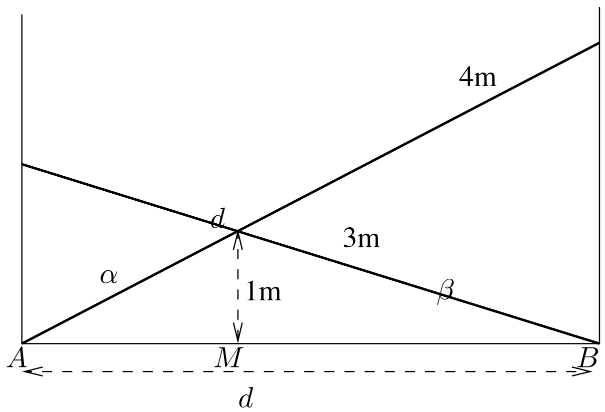
1. Ecrire la méthode de Newton pour la résolution du système suivant :

$$x^2 + 2xy = 0, \qquad (2.3.22)$$
$$xy + 1 = 0. \qquad (2.3.23)$$

2. Calculer les solutions de ce système.
3. Soit $(\overline{x}, \overline{y})$ une solution du problème (2.3.22)-(2.3.23). Montrer qu'il existe $\varepsilon > 0$ tel que si (x_0, y_0) est dans la boule B_ε de centre $(\overline{x}, \overline{y})$ et de rayon ε, alors la suite $(x_n, y_n)_{n\in\mathbb{N}}$ construite par la méthode de Newton converge vers $(\overline{x}, \overline{y})$ lorsque n tends vers $+\infty$.

Exercice 51 (Newton et les échelles…) *Corrigé en page 2.5 page 110*

Soient deux échelles de longueurs respectives 3 et 4 m, posées contre deux murs verticaux selon la figure ci-contre. On sait que les échelles se croisent à 1m du sol, et on cherche à connaître la distance d entre les deux murs.



1. Montrer que le problème revient à déterminer x et y tels que

$$16x^2 = (x^2 + 1)(x + y)^2 \qquad (2.3.24)$$
$$9y^2 = (y^2 + 1)(x + y)^2. \qquad (2.3.25)$$

2. Ecrire l'algorithme de Newton pour la résolution du système (2.3.24)-(2.3.25).

3. Calculer les premiers itérés $x^{(1)} = 1$ et $y^{(1)} = 1$ construits par la méthode de Newton en partant de $x^{(0)} = 1$ et $y^{(0)} = 1$.

Exercice 52 (Nombre d'itérations fini pour Newton) *Corrigé détaillé en page 110*

1. Soit f la fonction de $\mathbb{R}$ dans $\mathbb{R}$ définie par : $f(x) = e^x - 1$. Pour $x^{(0)} \in \mathbb{R}$, on note $(x^{(n)})_{n\in\mathbb{N}}$ la suite des itérés construits par la méthode de Newton pour la recherche d'un point où f s'annule.

1.1 Montrer que pour tout $x^{(0)} \in \mathbb{R}$, la suite $(x^{(n)})_{n\in\mathbb{N}}$ est bien définie.

1.2 Montrer que si $x^{(0)} \neq 0$, alors $x^{(n+1)} \neq x^{(n)}$ pour tout $n \in \mathbb{N}$. En déduire que la méthode de Newton converge en un nombre fini d'opérations si et seulement si $f(x^{(0)}) = 0$.

1.3 Montrer que :
1.3 (a) si $x^{(0)} < 0$ alors $x^{(1)} > 0$.
1.3 (b) si $x^{(0)} > 0$ alors $0 < x^{(1)} < x^{(0)}$.

1.4 Montrer que la suite $(x^{(n)})_{n\in\mathbb{N}}$ converge lorsque n tend vers l'infini et donner sa limite.

2. Soit $F : \mathbb{R}^N \to \mathbb{R}^N$ une fonction continûment différentiable et strictement convexe ($N \geq 1$) et dont la différentielle ne s'annule pas. Soit $x^{(0)} \in \mathbb{R}^N$ le choix initial (ou itéré 0) dans la méthode de Newton.
Montrer que la méthode de Newton converge en un nombre fini d'opérations si et seulement si $F(x^{(0)}) = 0$.

Exercice 53 (Newton et logarithme) *Suggestions en page 106*

Soit f la fonction de $\mathbb{R}_+^*$ dans $\mathbb{R}$ définie par $f(x) = \ln(x)$. Montrer que la méthode de Newton pour la recherche de $\overline{x}$ tel que $f(\overline{x}) = 0$ converge si et seulement si le choix initial $x^{(0)}$ est tel que $x^{(0)} \in]0, e[$.

Exercice 54 (Méthode de Newton pour le calcul de l'inverse) *Corrigé en page 110*

1. Soit $a > 0$. On cherche à calculer $\dfrac{1}{a}$ par l'algorithme de Newton.
   (a) Montrer que l'algorithme de Newton appliqué à une fonction g (dont $\frac{1}{a}$ est un zéro) bien choisie s'écrit :
   $$\begin{cases} x^{(0)} \text{ donné}, \\ x^{(n+1)} = x^{(n)}(2 - ax^{(n)}). \end{cases} \tag{2.3.26}$$
   (b) Montrer que la suite $(x^{(n)})_{n\in\mathbb{N}}$ définie par (2.3.26) vérifie
   $$\lim_{n\to+\infty} x^{(n)} = \begin{cases} \dfrac{1}{a} \text{ si } x^{(0)} \in]0, \dfrac{2}{a}[, \\ -\infty \text{ si } x^{(0)} \in]-\infty, 0[\cup]\dfrac{2}{a}, +\infty[\end{cases}$$
2. On cherche maintenant à calculer l'inverse d'une matrice par la méthode de Newton. Soit donc A une matrice carrée d'ordre N inversible, dont on cherche à calculer l'inverse.
   (a) Montrer que l'ensemble $GL_N(\mathbb{R})$ des matrices carrées inversibles d'ordre N (où $N \geq 1$) est un ouvert de l'ensemble $\mathcal{M}_N(\mathbb{R})$ des matrices carrées d'ordre N.
   (b) Soit T l'application définie de $GL_N(\mathbb{R})$ dans $GL_N(\mathbb{R})$ par $T(B) = B^{-1}$. Montrer que T est dérivable, et que $DT(B)H = -B^{-1}HB^{-1}$.
   (c) Ecrire la méthode de Newton pour calculer A^{-1} en cherchant le zéro de la fonction g de $\mathcal{M}_N(\mathbb{R})$ dans $\mathcal{M}_N(\mathbb{R})$ définie par $g(B) = B^{-1} - A$. Soit $B^{(n)}$ la suite ainsi définie.

(d) Montrer que la suite $B^{(n)}$ définie dans la question précédente vérifie :

$$Id - AB^{(n+1)} = (Id - AB^{(n)})^2.$$

En déduire que la suite $(B^{(n)})_{n\in\mathbb{N}}$ converge vers A^{-1} si et seulement si $\rho(Id - AB^{(0)}) < 1$.

Exercice 55 (Méthode de Newton pour le calcul de la racine)

1. Soit $\lambda \in \mathbb{R}_+$ et f_λ la fonction de $\mathbb{R}$ dans $\mathbb{R}$ définie par $f_\lambda(x) = x^2 - \lambda$.
   1.1 Soit $x^{(0)} \in \mathbb{R}$ fixé. Donner l'algorithme de Newton pour la résolution de l'équation $f_\lambda(x) = 0$.
   1.2 On suppose $x^{(0)} > 0$.
      (i) Montrer que la suite $(x^{(k)})_{k\geq 1}$ est minorée par $\sqrt{\lambda}$.
      (ii) Montrer que la suite $(x^{(k)})_{k\geq 0}$ converge et donner sa limite.

Soit $N \in \mathbb{N}^*$ et soit $A \in \mathcal{M}_N(\mathbb{R})$ une matrice diagonalisable dans $\mathbb{R}$; on note $\lambda_i, i = 1, \ldots, N$ les valeurs propres de A. On suppose que $\lambda_i > 0$ pour tout $i = 1, \ldots, N$.

2. Montrer qu'il existe au moins une matrice $B \in \mathcal{M}_N(\mathbb{R})$ telle que $B^2 = A$. Calculer une telle matrice B dans le cas où $N = 2$ et $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$.
3. On suppose de plus que A est symétrique définie positive. Montrer qu'il existe une unique matrice symétrique définie positive B telle que $B^2 = A$. Montrer par un contre exemple que l'unicité n'est pas vérifiée si on ne demande pas que B soit symétrique définie positive.

Soit F l'application de $\mathcal{M}_N(\mathbb{R})$ dans $\mathcal{M}_N(\mathbb{R})$ définie par $F(X) = X^2 - A$.

4. Montrer que F est différentiable en tout $X \in \mathcal{M}_N(\mathbb{R})$, et déterminer $DF(X)H$ pour tout $H \in \mathcal{M}_N(\mathbb{R})$.
   Dans la suite de l'exercice, on considère la méthode de Newton pour déterminer B.
5. On suppose maintenant $N \geq 1$. On note $(X^{(k)})_{k\in\mathbb{N}}$ la suite (si elle existe) donnée par l'algorithme de Newton à partir d'un choix initial $X^{(0)} = I$, où I est la matrice identité de $\mathcal{M}_N(\mathbb{R})$.
   5.1 Donner le procédé de construction de $X^{(k+1)}$ en fonction de $X^{(k)}$, pour $k \geq 0$.
   5.2 On note $\lambda_1 \leq \cdots \leq \lambda_N$ les valeurs propres de A (dont certaines peuvent être égales) et P la matrice orthogonale telle que $A = P^{-1}\mathrm{diag}(\lambda_1, \cdots, \lambda_N)P$.
      (i) Montrer que pour tout $k \in \mathbb{N}$, $X^{(k)}$ est bien définie et vaut $X^{(k)} = P^{-1}\mathrm{diag}(\mu_1^{(k)}, \cdots, \mu_N^{(k)})P$ où $\mu_i^{(k)}$ est le $k^{ième}$ terme de la suite de Newton associé à la fonction f_{λ_i} avec comme choix initial $\mu_i^{(0)} = 1$.
      (ii) En déduire que la suite $X^{(k)}$ converge vers B quand $k \to +\infty$.

Exercice 56 (Valeurs propres et méthode de Newton)
Suggestions en page 106, corrigé détaillé en page 113

Soit $A \in \mathcal{M}_N(\mathbb{R})$ une matrice symétrique. Soient $\overline{\lambda}$ une valeur propre simple de A et $\overline{x} \in \mathbb{R}^N$ un vecteur propre associé t.q. $\overline{x} \cdot \overline{x} = 1$. Pour calculer $(\overline{\lambda}, \overline{x})$ on applique la méthode de Newton au système non linéaire (de $\mathbb{R}^{N+1}$ dans $\mathbb{R}^{N+1}$) suivant :

$$\begin{aligned} Ax - \lambda x &= 0, \\ x \cdot x &= 1. \end{aligned} \tag{2.3.27}$$

Montrer que la méthode est localement convergente.

Exercice 57 (Modification de la méthode de Newton) *Suggestions en page 106.*

Soient $f \in C^1(\mathbb{R}^N, \mathbb{R}^N)$ et $\overline{x} \in \mathbb{R}^N$ t.q. $f(\overline{x}) = 0$. On considère, pour $\lambda > 0$ donné, la méthode itérative suivante :
– Initialisation : $x^{(0)} \in \mathbb{R}^N$.
– Iterations : pour $n \geq 0$,

$$x^{(n+1)} = x^{(n)} - [Df(x^{(n)})^t Df(x^{(n)}) + \lambda Id]^{-1} Df(x^{(n)})^t f(x^{(n)}).$$

[Noter que, pour $\lambda = 0$, on retrouve la méthode de Newton.]

1. Montrer que la suite $(x^{(n)})_{n\in\mathbb{N}}$ est bien définie.
2. On suppose, dans cette question, que $N = 1$ et que $f'(\overline{x}) \neq 0$. Montrer que la méthode est localement convergente en $\overline{x}$.
3. On suppose que le rang de $Df(\overline{x})$ est égal à N. Montrer que la méthode est localement convergente en $\overline{x}$. [Noter que cette question redonne la question précédente si $N = 1$.]

Exercice 58 (Convergence de la méthode de Newton si $f'(\overline{x}) = 0$)
Suggestions en page 106, corrigé détaillé en page 113

Soient $f \in C^2(\mathbb{R}, \mathbb{R})$ et $\overline{x} \in \mathbb{R}$ t.q. $f(\overline{x}) = 0$.

1. Rappel du cours. Si $f'(\overline{x}) \neq 0$, la méthode de Newton est localement convergente en $\overline{x}$ et la convergence est au moins d'ordre 2.
2. On suppose maintenant que $f'(\overline{x}) = 0$ et $f''(\overline{x}) \neq 0$. Montrer que la méthode de Newton est localement convergente (en excluant le cas $x_0 = \overline{x}\ldots$) et que la convergence est d'ordre 1. Si on suppose f de classe C^3, donner une modification de la méthode de Newton donnant une convergence au moins d'ordre 2.

Exercice 59 (Point fixe et Newton)

Soit $g \in C^3(\mathbb{R}, \mathbb{R})$ et $\overline{x} \in \mathbb{R}$ tels que $g(\overline{x}) = 0$ et $g'(\overline{x}) \neq 0$ et soit $f \in C^1(\mathbb{R}, \mathbb{R})$ telle que $f(\overline{x}) = \overline{x}$.
On considère l'algorithme suivant :

$$\begin{cases} x_0 \in \mathbb{R}, \\ x_{n+1} = h(x_n), n \geq 0. \end{cases} \tag{2.3.28}$$

avec $h(x) = x - \dfrac{g(x)}{g' \circ f(x))}$.

1. Montrer qu'il existe $\alpha > 0$ tel que si $x_0 \in [\overline{x} - \alpha, \overline{x} + \alpha] = I_\alpha$, alors la suite donnée par l'algorithme (2.3.28) est bien définie ; montrer que $x_n \to \overline{x}$ lorsque $n \to +\infty$ (on pourra montrer qu'on peut choisir α de manière à ce que $|h'(x)| < 1$ si $x \in I_\alpha$).
   On prend maintenant $x_0 \in I_\alpha$ où α est donné par la question 1.
2. Montrer que la convergence de la suite $(x_n)_{n\in\mathbb{N}}$ définie par l'algorithme (2.3.28) est au moins quadratique.
3. On suppose de plus que f est deux fois dérivable et que $f'(\overline{x}) = \dfrac{1}{2}$. Montrer que la convergence de la suite $(x_n)_{n\in\mathbb{N}}$ définie par (1) est au moins cubique, c'est-à-dire qu'il existe $c \in \mathbb{R}_+$ tel que
$$|x_{n+1} - \overline{x}| \leq c|x_n - \overline{x}|^3, \qquad \forall n \geq 1.$$
4. Soit $\beta \in \mathbb{R}_+^*$ tel que $g'(x) \neq 0 \quad \forall x \in I_\beta =]\overline{x} - \beta, \overline{x} + \beta[$; montrer que si on prend $f \in C^1(\mathbb{R}, \mathbb{R})$ telle que :
$$f(x) = x - \frac{g(x)}{2g'(x)} \qquad \text{si } x \in I_\beta,$$
alors la suite définie par l'algorithme (1) converge de manière cubique.

Exercice 60 (Variante de la méthode de Newton)
Corrigé détaillé en page 114

Soit $f \in C^1(\mathbb{R},\mathbb{R})$ et $\bar{x} \in \mathbb{R}$ tel que $f(\bar{x}) = 0$. Soient $x_0 \in \mathbb{R}$, $c \in \mathbb{R}_+^*$, $\lambda \in \mathbb{R}_+^*$. On suppose que les hypothèses suivantes sont vérifiées :

(i) $\bar{x} \in I = [x_0 - c, x_0 + c]$,

(ii) $|f(x_0)| \le \frac{c}{2\lambda}$,

(iii) $|f'(x) - f'(y)| \le \frac{1}{2\lambda}$, $\forall (x,y) \in I^2$

(iv) $|f'(x)| \ge \frac{1}{\lambda}\ \forall x \in I$.

On définit la suite $(x^{(n)})_{n\in\mathbb{N}}$ par :

$$\begin{aligned} & x^{(0)} = x_0, \\ & x^{(n+1)} = x^{(n)} - \frac{f(x^{(n)})}{f'(y)}, \end{aligned} \tag{2.3.29}$$

où $y \in I$ est choisi arbitrairement.

1. Montrer par récurrence que la suite définie par (2.3.29) satisfait $x^{(n)} \in I$ pour tout $n \in \mathbb{N}$.
   (On pourra remarquer que si $x^{(n+1)}$ est donné par (2.3.29) *alors $x^{(n+1)} - x_0 = x^{(n)} - x_0 - \frac{f(x^{(n)}) - f(x_0)}{f'(y)} - \frac{f(x_0)}{f'(y)}$.)*
2. Montrer que la suite $(x^{(n)})_{n\in\mathbb{N}}$ définie par (2.3.29) vérifie $|x^{(n)} - \bar{x}| \le \frac{c}{2^n}$ et qu'elle converge vers $\bar{x}$ de manière au moins linéaire.
3. On remplace l'algorithme (2.3.29) par

$$\begin{aligned} & x^{(0)} = x_0, \\ & x^{(n+1)} = x^{(n)} - \frac{f(x^{(n)})}{f'(y^{(n)})}, \end{aligned} \tag{2.3.30}$$

   où la suite $(y^{(n)})_{n\in\mathbb{N}}$ est une suite donnée d'éléments de I. Montrer que la suite $(x^{(n)})_{n\in\mathbb{N}}$ converge vers $\bar{x}$ de manière au moins linéaire, et que cette convergence devient super-linéaire si $f'(y_n) \to f'(\bar{x})$ lorsque $n \to +\infty$.
4. On suppose maintenant que $N \ge 1$ et que $f \in C^1(\mathbb{R}^N, \mathbb{R}^N)$. La méthode définie par (2.3.29) ou (2.3.30) peut-elle se généraliser, avec d'éventuelles modifications des hypothèses, à la dimension N ?

Exercice 61 (Méthode de Newton pour un système semi-linéaire) *Suggestions en page 106.*

On suppose que $f \in C^2(\mathbb{R},\mathbb{R})$ et que f est croissante. On s'intéresse au système non linéaire suivant de N équations à N inconnues (notées $u_1, \ldots, u_N$) :

$$\begin{aligned} & (Au)_i + \alpha_i f(u_i) = b_i\ \forall i \in \{1,\ldots,N\}, \\ & u = (u_1,\ldots,u_N)^t \in \mathbb{R}^N, \end{aligned} \tag{2.3.31}$$

où $A \in \mathcal{M}_N(\mathbb{R})$ est une matrice symétrique définie positive, $\alpha_i > 0$ pour tout $i \in \{1,\ldots,N\}$ et $b_i \in \mathbb{R}$ pour tout $i \in \{1,\ldots,N\}$.

On admet que (2.3.31) admet au moins une solution (ceci peut être démontré mais est difficile).

1. Montrer que (2.3.31) admet une unique solution.
2. Soit u la solution de (2.3.31). Montrer qu'il existe $a > 0$ t.q. la méthode de Newton pour approcher la solution de (2.3.31) converge lorsque le point de départ de la méthode, noté $u^{(0)}$, vérifie $|u - u^{(0)}| < a$.

Exercice 62 (Méthode de Steffensen)
Suggestions en page 107, corrigé détaillé en page 116

Soient $f \in C^2(\mathbb{R}, \mathbb{R})$ et $\overline{x} \in \mathbb{R}$ t.q. $f(\overline{x}) = 0$ et $f'(\overline{x}) \neq 0$. On considère la méthode itérative suivante :
– Initialisation : $x^{(0)} \in \mathbb{R}^N$.
– Itérations : pour $n \geq 0$, si $f(x^{(n)} + f(x^{(n)})) \neq f(x^{(n)})$,

$$x^{(n+1)} = x^{(n)} - \frac{(f(x^{(n)}))^2}{f(x^{(n)} + f(x^{(n)})) - f(x^{(n)})}, \tag{2.3.32}$$

et si $f(x^{(n)} + f(x^{(n)})) = f(x^{(n)})$, $x^{(n+1)} = x^{(n)}$.

1. Montrer qu'il existe $\alpha > 0$ tel que si $x^{(n)} \in B(\overline{x}, \alpha)$, alors $f(x^{(n)} + f(x^{(n)})) \neq f(x^{(n)})$ si $x^{(n)} \neq \overline{x}$. En déduire que si $x_0 \in B(\overline{x}, \alpha)$, alors toute la suite $(x^{(n)})_{n \in \mathbb{N}}$ vérifie (2.3.32) pourvu que $x^{(n)} \neq \overline{x}$ pour tout $n \in \mathbb{N}$.
2. Montrer par des développements de Taylor avec reste intégral qu' il existe une fonction a continue sur un voisinage de $\overline{x}$ telle que si $x_0 \in B(\overline{x}, \alpha)$, alors
$$x^{(n+1)} - \overline{x} = a(x^{(n)})(x^{(n)} - \overline{x}), \text{ pour tout } n \in \mathbb{N} \text{ tel que } x^{(n)} \neq \overline{x}. \tag{2.3.33}$$
3. Montrer que la méthode est localement convergente en $\overline{x}$ et la convergence est au moins d'ordre 2.

Exercice 63 (Méthode de Newton-Tchebycheff)

1. Soit $f \in C^3(\mathbb{R}, \mathbb{R})$ et soit $\bar{x} \in \mathbb{R}$ tel que $\bar{x} = f(\bar{x})$ et $f'(\bar{x}) = f''(\bar{x}) = 0$. Soit $(x_n)_{n \in \mathbb{N}}$ la suite définie par :

$$\begin{cases} x_0 \in \mathbb{R}, \\ x_{n+1} = f(x_n). \end{cases} \tag{PF}$$

(a) Justifier l'appellation "(PF)" de l'algorithme.
(b) Montrer qu'il existe $a > 0$ tel que si $|y - \overline{x}| \leq a$ alors $|f'(y)| \leq \frac{1}{2}$.
(c) Montrer par récurrence sur n que si $x_0 \in B(\bar{x}, a)$ alors $x_n \in B(\bar{x}, \frac{a}{2^n})$.
(d) En déduire que la suite construite par (PF) converge localement, c'est–à–dire qu'il existe un voisinage V de $\bar{x}$ tel que si $x_0 \in V$ alors $x_n \to \bar{x}$ lorsque $n \to +\infty$. .
(e) Montrer que la vitesse de convergence de la suite construite par (PF) est au moins cubique (c'est-à-dire qu' il existe $\beta \in \mathbb{R}_+$ tel que $|x_{n+1} - \bar{x}| \leq \beta |x_n - \bar{x}|^3$ si la donnée initiale x_0 est choisie dans un certain voisinage de $\bar{x}$. On pourra utiliser un développement de Taylor-Lagrange).

2. Soit $g \in C^3(\mathbb{R}, \mathbb{R})$, et soit $\bar{x} \in \mathbb{R}$ tel que $g(\bar{x}) = 0$ et $g'(\bar{x}) \neq 0$. Pour une fonction $h \in C^3(\mathbb{R}, \mathbb{R})$ à déterminer, on définit $f \in C^3(\mathbb{R}, \mathbb{R})$ par $f(x) = x + h(x)g(x)$. Donner une expression de $h(\bar{x})$ et $h'(\bar{x})$ en fonction de $g'(\bar{x})$ et de $g''(\bar{x})$ telle que la méthode (PF) appliquée à la recherche d'un point fixe de f converge localement vers $\bar{x}$ avec une vitesse de convergence au moins cubique.

3. Soit $g \in C^5(\mathbb{R}, \mathbb{R})$, et soit $\bar{x} \in \mathbb{R}$ tel que $g(\bar{x}) = 0$ et $g'(\bar{x}) \neq 0$. On considère la modification suivante (dûe à Tchebychev) de la méthode de Newton :

$$x_{n+1} = x_n - \frac{g(x_n)}{g'(x_n)} - \frac{g''(x_n)[g(x_n)]^2}{2[g'(x_n)]^3}. \tag{2.3.34}$$

Montrer que la méthode (2.3.34) converge localement et que la vitesse de convergence est au moins cubique. [On pourra commencer par le cas où g' ne s'annule pas].

Exercice 64 (Méthode de la sécante) *Corrigé en page 2.5 page 118*

Soient $f \in C^2(\mathbb{R}, \mathbb{R})$ et $\overline{x} \in \mathbb{R}$ t.q. $f(\overline{x}) = 0$ et $f'(\overline{x}) \neq 0$. Pour calculer $\overline{x}$, on considère la méthode itérative suivante (appelée "méthode de la sécante") :

– Initialisation : $x_0, x_1 \in \mathbb{R}$.

– Itérations : pour $n \geq 1$, $x_{n+1} = x_n - \dfrac{f(x_n)(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})}$ si $f(x_n) \neq 0$ et $x_{n+1} = x_n$ si $f(x_n) = 0$.

Si la suite $(x_n)_{n\in\mathbb{N}}$ est bien définie par cette méthode, on pose $e_n = |x_n - \overline{x}|$ pour tout $n \in \mathbb{N}$.

1. Montrer qu'il existe $\varepsilon > 0$ t.q. pour $x_0, x_1 \in]\overline{x} - \varepsilon, \overline{x} + \varepsilon[$, $x_0 \neq x_1$, la méthode de la sécante définit bien une suite $(x_n)_{n\in\mathbb{N}}$ et l'on a $e_{n+1} \leq (1/2)e_n$ pour tout $n \geq 1$. [Raisonner par récurrence : on suppose $x_n, x_{n-1} \in]\overline{x} - \varepsilon, \overline{x} + \varepsilon[$, $x_n \neq x_{n-1}$ et $x_n \neq \overline{x}$. Montrer, grâce à un choix convenable de ε, que $f(x_n) \neq f(x_{n-1})$ et que $f(x_n) \neq 0$. En déduire que x_{n+1} est bien défini et $x_n \neq x_{n+1}$. Puis, toujours grâce à un choix convenable de ε, que $e_{n+1} \leq (1/2)e_n$. Conclure.]

   Dans les questions suivantes, on suppose que $x_0, x_1 \in]\overline{x} - \varepsilon, \overline{x} + \varepsilon[$, $x_0 \neq x_1$ (ε trouvé à la première question) et que la suite $(x_n)_{n\in\mathbb{N}}$ donnée par la méthode de la sécante vérifie $x_n \neq \overline{x}$ pour tout $n \in \mathbb{N}$. On pose $d = (1 + \sqrt{5})/2$ et on démontre que la convergence est en général d'ordre d.

2. Pour $x \neq \overline{x}$, on définit $\mu(x)$ comme la moyenne de f' sur l'intervalle dont les extrémités sont x et $\overline{x}$.
   (a) Montrer que $e_{n+1} = e_n e_{n-1} M_n$, pour tout $n \geq 1$, avec $M_n = |\frac{\mu(x_n) - \mu(x_{n-1})}{f(x_n) - f(x_{n-1})}|$.
   (b) Montrer que la fonction μ est dérivable sur $\mathbb{R} \setminus \{\overline{x}\}$. Calculer $\lim_{x\to\overline{x}} \mu'(x)$.
   (c) Calculer la limite de la suite $(M_n)_{n\geq1}$ lorsque $n \to \infty$. En déduire que la suite $(M_n)_{n\geq1}$ est bornée.

3. Soit $M > 0$, $M \geq M_n$ pour tout $n \geq 1$ (M_n donné à la question précédente). On pose $a_0 = Me_0$, $a_1 = Me_1$ et $a_{n+1} = a_n a_{n-1}$ pour $n \geq 1$.
   (a) Montrer que $Me_n \leq a_n$ pour tout $n \in \mathbb{N}$.
   (b) Montrer qu'il existe $\varepsilon_1 \in]0, \varepsilon]$ t.q. la suite $(a_n)_{n\geq1}$ tend en décroissant vers 0 lorsque $n \to +\infty$, si $x_0, x_1 \in]\overline{x} - \varepsilon_1, \overline{x} + \varepsilon_1[$.
   (c) Dans cette question, on prend $x_0, x_1 \in]\overline{x} - \varepsilon_1, \overline{x} + \varepsilon_1[$. Montrer qu'il existe $\alpha > 0$ et $\beta \in]0, 1[$ t.q $Me_n \leq a_n \leq \alpha(\beta)^{d^n}$ pour tout $n \in \mathbb{N}$ (ceci correspond à une convergence d'ordre au moins d). [On pourra utiliser la relation de récurrence $\ln a_{n+1} = \ln a_n + \ln a_{n-1}$ pour $n \geq 1$].
   (d) (Question plus difficile) Si $f''(\overline{x}) \neq 0$, $e_{n+1} = e_n e_{n-1} M_n$, montrer que $M_n \to \overline{M}$, quand $n \to \infty$, avec $\overline{M} > 0$. En déduire qu'il existe $\gamma > 0$ t.q. $\frac{e_{n+1}}{(e_n)^d} \to \gamma$ quand $n \to \infty$ (ceci signifie que la convergence est exactement d'ordre d). [Considérer, par exemple, $\beta_n = \ln e_{n+1} - d \ln e_n$ et montrer que β_n converge dans $\mathbb{R}$ quand $n \to \infty$.]
   (e) Comparer l'ordre de convergence de la méthode de la sécante à celui de la méthode de Newton.

2.4 Suggestions

Exercice 46 page 98 (Calcul différentiel)

1. Utiliser le fait que $Df(x)$ est une application linéaire et le théorème de Riesz. Appliquer ensuite la différentielle à un vecteur h bien choisi.
2. Mêmes idées...

Exercice 48 page 98 (Méthode de monotonie)

Pour montrer que la suite $(v^{(n)})_{n\in\mathbb{N}}$ est bien définie, remarquer que la matrice A est inversible. Pour montrer qu'elle est convergente, montrer que les hypothèses du théorème du point fixe de monotonie vu en cours sont vérifiées.

Exercice 49 page 98 (Point fixe amélioré)

1) Montrer qu'on peut choisir α de manière à ce que $|h'(x)| < 1$ si $x \in I_\alpha$, et en déduire que $g'(\varphi(x_n) \neq 0$ si x_0 est bien choisi.
2) Remarquer que

$$|x_{n+1} - \overline{x}| = (x_n - \overline{x})(1 - \frac{g(x_n) - g(\overline{x})}{(x_n - \overline{x})g'(\varphi(x_n))}. \tag{2.4.35}$$

En déduire que

$$|x_{n+1} - \overline{x}| \leq \frac{1}{\varepsilon}|x_n - \overline{x}|^2 \sup_{x \in I_\alpha} |\varphi'(x)| \sup_{x \in I_\alpha} |g''(x)|.$$

3) Reprendre le même raisonnement avec des développements d'ordre supérieur.

4) Montrer que φ vérifie les hypothèses de la question 3).

Exercice 53 page 100 (Newton et logarithme)

Etudier les variations de la fonction φ définie par : $\varphi(x) = x - x \ln x$.

Exercice 56 page 101 (Valeurs propres et méthode de Newton)

Ecrire le système sous la forme $F(x, \lambda) = 0$ où F est une fonction de $\mathbb{R}^{N+1}$ dans $\mathbb{R}^{N+1}$ et montrer que $DF(\overline{\lambda}, \overline{x})$ est inversible.

Exercice 57 page 102 (Modification de la méthode de Newton)

1. Remarquer que si $A \in \mathcal{M}_N(\mathbb{R})$ et $\lambda > 0$, alors $A^t A + \lambda Id$ est symétrique définie positive.
2. En introduisant la fonction φ définie par $\varphi(t) = f(tx_n + (1-t)\overline{x})$, montrer que $f(x_n) = (x_n - \overline{x})g(x_n)$, où $g(x) = \int_0^1 f'(tx + (1-t)\overline{x})dt$. Montrer que g est continue.
Montrer que la suite $(x_n)_{n\in\mathbb{N}}$ vérifie $x_{n+1} - \overline{x} = a_n(x_n - \overline{x})$, où

$$a_n = 1 - \frac{f'(x_n)g(x_n)}{f'(x_n)^2 + \lambda},$$

et qu'il existe α tel que si $x_n \in B(\overline{x}, \alpha)$, alors $a_n \in]0, 1[$. Conclure.

3. Reprendre la même méthode que dans le cas $N = 1$ pour montrer que la suite $(x_n)_{n\in\mathbb{N}}$ vérifie $x_{n+1} - \overline{x} = D(x_n)(x_n - \overline{x})$, où $D \in \mathcal{C}(\mathbb{R}^N, \mathcal{M}_N(\mathbb{R}))$. Montrer que $D(\overline{x})$ est symétrique et montrer alors que $\|D(\overline{x})\|_2 < 1$ en calculant son rayon spectral. Conclure par continuité comme dans le cas précédent.

Exercice 58 page 102 (Convergence de la méthode de Newton si $f'(\overline{x}) = 0$)

Supposer par exemple que $f''(\overline{x}) > 0$ et montrer que si x_0 est "assez proche" de $\overline{x}$ la suite $(x_n)_{n\in\mathbb{N}}$ est croissante majorée ou décroissante minorée et donc convergente. Pour montrer que l'ordre de la méthode est 1, montrer que

$$\frac{\|x_{n+1} - \overline{x}\|}{\|x_n - \overline{x}\|} \to \frac{1}{2} \text{ lorsque } n \to +\infty.$$

Exercice 61 page 103 (Méthode de Newton)

1. Pour montrer l'unicité, utiliser la croissance de f et le caractère s.d.p. de A.
2. Utiliser le théorème de convergence du cours.

Exercice 62 page 104 (Méthode de Steffensen))

1. Utiliser la monotonie de f dans un voisinage de $\overline{x}$.
2. Développer le dénominateur dans l'expression de la suite en utilisant le fait que $f(x_n + f(x_n)) - f(x_n) = \int_0^1 \psi'(t)dt$ où $\psi(t) = f(x_n + tf(x_n))$, puis que $f'(x_n + tf(x_n)) = \int_0^t \xi'(s)ds$ où $\xi(t) = f'(x_n + tf(x_n))$. Développer ensuite le numérateur en utilisant le fait que $-f(x_n) = \int_0^1 \varphi'(t)dt$ où $\varphi(t) = f(t\overline{x} + (1-t)x_n)$, et que $f'(t\overline{x} + (1-t)x_n) = \int_0^1 \chi(s)ds + \chi(0)$, où $\chi(t) = f(\overline{x} + (1-t)_n)$.
3. La convergence locale et l'ordre 2 se déduisent des résultats de la question 2.

2.5 Corrigés

Exercice 46 page 98

1. Par définition, $T = Df(x)$ est une application linéaire de $\mathbb{R}^N$ dans $\mathbb{R}^N$, qui s'écrit donc sous la forme : $T(h) = \sum_{i=1}^N a_i h_i = a \cdot h$. Or l'application T dépend de x, donc le vecteur a aussi.
Montrons maintenant que $(a(x))_i = \partial_i f(x)$, pour $1 \leq i \leq N$ Soit $h^{(i)} \in \mathbb{R}^N$ défini par $h_j^{(i)} = h\delta_{i,j}$, où $h > 0$ et $\delta_{i,j}$ désigne le symbole de Kronecker, i.e. $\delta_{i,j} = 1$ si $i = j$ et $\delta_{i,j} = 0$ sinon. En appliquant la définition de la différentielle avec $h^{(i)}$, on obtient :

$$f(x + h^{(i)}) - f(x) = Df(x)(h(i)) + \|h^{(i)}\|\varepsilon(h^{(i)}),$$

c'est–à–dire :

$$f(x_1, \ldots, x_{i-1}, x_i + h, x_{i-1}, \ldots, x_N) - f(x_1, \ldots, x_N) = (a(x))_i h + h\varepsilon(h^{(i)}).$$

En divisant par h et en faisant tendre h vers 0, on obtient alors que $(a(x))_i = \partial_i f(x)$.

2. Comme $f \in C^2(\mathbb{R}^N, \mathbb{R})$, on a $(\partial_i f(x) \in C^1(\mathbb{R}^N, \mathbb{R})$, et donc $\varphi \in C^1(\mathbb{R}^N, \mathbb{R}^N)$. Comme $D\varphi(x)$ est une application linéaire de $\mathbb{R}^N$ dans $\mathbb{R}^N$, il existe une matrice $A(x)$ carrée d'ordre N telle que $D\varphi(x)(y) = A(x)y$ pour tout $y \in \mathbb{R}^N$. Il reste à montrer que $(A(x))_{i,j} = \partial^2_{i,j} f(x)$. Soit $h^{(i)} \in \mathbb{R}^N$ défini à la question préceédente, pour $i,\ j = 1, \ldots, N$, on a

$$(D\varphi(x)(h^{(j)}))_i = (A(x)h^{(j)})_i = \sum_{k=1}^N a_{i,k}(x)h^{(j)})_k = ha_{i,j}(x).$$

Or par définition de la différentielle,

$$\varphi_i(x + h^{(j)}) - \varphi_i(x) = (D\varphi(x)(h^{(j)}))_i + \|h^{(j)}\|\varepsilon_i(h^{(j)}),$$

ce qui entraîne, en divisant par h et en faisant tendre h vers 0 : $\partial_j \varphi_i(x) = a_{i,j}(x)$. Or $\varphi_i(x) = \partial_i f(x)$, et donc $(A(x))_{i,j} = a_{i,j}(x) = \partial^2_{i,j} f(x)$.

Corrigé de l'exercice 48 page 98 (Méthode de monotonie)

Montrons que la suite $v^{(n)}$ est bien définie. Supposons $v^{(n)}$ connu ; alors $v^{(n+1)}$ est bien défini si le système

$$Av^{(n+1)} = d^{(n)},$$

où $d^{(x)}$ est défini par : $d_i^{(n)} = \alpha_i f(v_i^{(n)}) + \lambda b_i$ pour $i = 1, \ldots, N$, admet une solution. Or, grâce au fait que $Av \geq 0 \Rightarrow v \geq 0$, la matrice A est inversible, ce qui prouve l'existence et l'unicité de $v^{(n+1)}$.
Montrons maintenant que les hypothèses du théorème de convergence du point fixe de monotonie sont bien satisfaites.
On pose $R_i^{(\lambda)}(u) = \alpha_i f(u_i) + \lambda b_i$. Le système à résoudre s'écrit donc :

$$Au = R^{(\lambda)}(u)$$

Or 0 est sous–solution car $0 \leq \alpha_i f(0) + \lambda b_i$ (grâce au fait que $f(0) = 0$, $\lambda > 0$ et $b_i \geq 0$).
Cherchons maintenant une sur–solution, c'est–à–dire $\tilde{u} \in \mathbb{R}^N$ tel que

$$\tilde{u} \geq R^{(\lambda)}(\tilde{u}).$$

Par hypothèse, il existe $\mu > 0$ et $u^{(\mu)} \geq 0$ tel que

$$(Au^{(\mu)})_i = \alpha f(u_i^{(\mu)}) + \mu b_i.$$

Comme $\lambda < \mu$ et $b_i \geq 0$, on a

$$(Au^{(\mu)})_i \geq \alpha_i f(u_i^{(\mu)}) + \lambda b_i = R_i^{(\lambda)}(u^{(\mu)}).$$

Donc $u^{(\mu)}$ est sur–solution. Les hypothèses du théorème dont bien vérifiées, et donc $v^{(n)} \to \bar{u}$ lorsque $n \to +\infty$, où $\bar{u}$ est tel que $A\bar{u} = R(\bar{u})$.

Corrigé de l'exercice 50 page 99 (Méthode de Newton pour un système 2×2)

En cours de rédaction

Corrigé de l'exercice 51 page 99 (Newton et les échelles...)

En cours de rédaction

Corrigé de l'exercice 52 page 100 (Nombre d'itérations fini pour Newton)

1.1 Comme f' est définie sur tout $\mathbb{R}$ par $f'(x) = e^x$ et ne s'annule pas, on en déduit que la suite construite par la méthode de Newton, qui s'écrit :

$$x^{(n+1)} = x^{(n)} - \frac{f(x^{(n)})}{f'(x^{(n)})} = x^{(n)} - \frac{e^{x^{(n)}} - 1}{e^{x^{(n)}}}$$

est bien définie.

1.2 Par définition de la suite, on a $x^{(n+1)} - x^{(n)} = -\frac{e^{x^{(n)}}-1}{e^{x^{(n)}}} = 0$ ssi $x^{(n)} = 0$. Donc par récurrence sur n, si $x^{(0)} \neq 0$, on a $x^{(n+1)} \neq x^{(n)}$ pour tout $n \in \mathbb{N}$. De plus, si $f(x^{(0)}) = 0$ (c.à.d. si $x^{(0)} = 0$), la suite est stationnaire. On en déduit que la méthode de Newton converge en un nombre fini d'opérations si et seulement si $f(x^{(0)}) = 0$.

1.3 Par définition, on a : $x^{(1)} = x^{(0)} - \frac{e^{x^{(0)}}-1}{e^{x^{(0)}}}$. Par le théorème de accroissements finis, on a donc : $x^{(1)} = x^{(0)}(1 - e^{\theta - x^{(0)}})$, avec $\theta \in]x^{(0)}, 0[$ si $x^{(0)} < 0$ et $\theta \in]0, x^{(0)}[$ si $x^{(0)} > 0$. Si $x^{(0)} < 0$, on a $e^{\theta - x^{(0)}} > 1$ et donc $x^{(1)} > 0$. En revanche, si $x^{(0)} > 0$, on a $e^{-x^{(0)}} < e^{\theta - x^{(0)}} < 1$ et donc $0 < x^{(1)} < x^{(0)}$.

1.4 On a vu à la question 1.2 que si $x^{(0)} = 0$ la suite est stationnaire et égale à 0. On a vu à la question 1.3 que si $x^{(0)} < 0$ alors $x^{(1)} > 0$. Il suffit donc d'étudier le cas $x^{(0)} > 0$. Or si $x^{(0)} > 0$, on a $0 < x^{(1)} < x^{(0)}$. Par récurrence sur n, on en déduit que si $x^{(0)} > 0$, la suite $(x^{(n)})_{n\in\mathbb{N}}$ est décroissante et minorée par 0, donc elle converge. La limite ℓ de la suite vérifie : $\ell = \ell - \frac{e^{\ell}-1}{e^{\ell}}$, soit encore $\ell = 0$ (unique solution de l'équation $f(x) = 0$).

2. Soient $x^{(n)}$ et $x^{(n+1)}$ deux itérés successifs donnés par la méthode de Newton, tels que $F(x^{(n)}) \neq 0$. On a donc :

$$DF(x^{(n)})(x^{(n+1)} - x^{(n)}) = -F(x^{(n)}), \tag{2.5.36}$$

et en particulier, $x^{(n+1)} \neq x^{(n)}$. Or, la condition de stricte convexité pour une fonction continûment différentiable entraï¿½ne que :

$$DF(x^{(n)})(x^{(n+1)} - x^{(n)}) < F(x^{(n+1)}) - F(x^{(n)}),$$

et donc, avec (2.5.36), $F(x^{(n+1)}) > 0$. On montre ainsi, par récurrence sur n, que si $F(x^{(0)}) \neq 0$, alors $F(x^{(n)}) > 0$ pour tout $n > 0$, ce qui montre que la méthode de Newton converge en un nombre fini d'opérations si et seulement si $F(x^{(0)}) = 0$.

Corrigé de l'exercice 54 page 100 (Méthode de Newton pour le calcul de l'inverse)

1. (a) Soit g la fonction définie de $\mathbb{R}^*$ dans $\mathbb{R}$ par $g(x) = \frac{1}{x} - a$. Cette fonction est continue et dérivable pour tout $x \neq 0$, et on a : $g'(x) = -\frac{1}{x^2}$. L'algorithme de Newton pour la recherche d'un zéro de cette fonction s'écrit donc bien :

$$\begin{cases} x^{(0)} \text{ donné}, \\ x^{(n+1)} = x^{(n)}(2 - ax^{(n)}). \end{cases} \tag{2.5.37}$$

(b) Soit $(x^{(n)})_{n\in\mathbb{N}}$ définie par (2.3.26). D'après le théorème du cours, on sait que la suite $(x^{(n)})_{n\in\mathbb{N}}$ converge de localement (de manière quadratique) dans un voisinage de $\frac{1}{a}$. On veut déterminer ici l'intervalle de convergence précisément. On a $x^{(n+1)} = \varphi(x^{(n)})$ où φ est la fonction définie par de $\mathbb{R}$ dans $\mathbb{R}$ par $\varphi(x) = x(2 - ax)$. Le tableau de variation de la fonction φ est le suivant :

x		0	$\frac{1}{a}$	$\frac{2}{a}$	
$\varphi'(x)$		$+$	0	$-$	
$\varphi(x)$	$-\infty$	$\nearrow$	$\frac{1}{a}$	$\searrow$	$-\infty$

(2.5.38)

Il est facile de remarquer que l'intervalle $]0, \frac{1}{a}[$ est stable par φ et que $\varphi(]\frac{1}{a}, \frac{2}{a}[) \; =]0, \frac{1}{a}[$ Donc si $x^{(0)} \in]\frac{1}{a}, \frac{2}{a}[$ alors $x^{(1)} \in]0, \frac{1}{a}[$, et on se ramène au cas $x^{(0)} \in]0, \frac{1}{a}[$.

On montre alors facilement que si $x^{(0)} \in]0, \frac{1}{a}[$, alors $x^{(n+1)} \geq x^{(n)}$ pour tout n, et donc la suite $(x^{(n)})_{n\in\mathbb{N}}$ est croissante. Comme elle est majorée (par $\frac{1}{a}$), elle est donc convergente. Soit ℓ sa limite, on a $\ell = \ell(2 - a\ell)$, et comme $\ell \geq x^{(0)} > 0$, on a $\ell = \frac{1}{a}$.

Il reste maintenant à montrer que si $x^{(0)} \in] - \infty, 0[\cup]\frac{2}{a}, +\infty[$ alors $\lim_{n\to+\infty} x^{(n)} = -\infty$. On montre d'abord facilement que si $x^{(0)} \in] - \infty, 0[$, la suite $(x_n)_{n\in\mathbb{N}}$ est décroissante. Elle admet donc une limite finie ou infinie. Appelons ℓ cette limite. Celle-ci vérifie : $\ell = \ell(2-a\ell)$. Si ℓ est finie, alors $\ell = 0$ ou $\ell = \frac{1}{a}$ ce qui est impossible car $\ell \leq x^{(0)} < 0$. On en déduit que $\ell = -\infty$.

Enfin, l'étude des variations de la fonction φ montre que si $x^{(0)} \in]\frac{2}{a}, +\infty[$, alors $x^{(1)}] - \infty, 0[$, et on est donc ramené au cas pécédent.

2. (a) L'ensemble $GL_N(\mathbb{R})(\mathbb{R})$ est ouvert car image réciproque de l'ouvert $\mathbb{R}^*$ par l'application continue qui a une matrice associe son déterminant.

(b) L'application T est clairement définie de $GL_N(\mathbb{R})(\mathbb{R})$ dans $GL_N(\mathbb{R})(\mathbb{R})$. Montrons qu'elle est dérivable. Soit $H \in GL_N(\mathbb{R})(\mathbb{R})$ telle que $B + H$ soit inversible. Ceci est vrai si $\|H\|\|B^{-1}\| < 1$, et on a alors, d'après le cours :

$$(B + H)^{-1} = \big(B(Id + B^{-1}H)\big)^{-1} = \sum_{k=0}^{+\infty}(-B^{-1}H)^k B^{-1}.$$

On a donc :

$$\begin{aligned} T(B + H) - T(B) &= \sum_{k=0}^{+\infty}(B^{-1}H)^k B^{-1} - B^{-1} \\ &= (Id + \sum_{k=1}^{+\infty}(-B^{-1}H)^k - Id)B^{-1} \\ &= \sum_{k=1}^{+\infty}(-B^{-1}H)^k B^{-1}. \end{aligned}$$

On en déduit que

$$T(B + H) - T(B) + B^{-1}HB^{-1} = \sum_{k=2}^{+\infty}(-B^{-1}H)^k B^{-1}.$$

L'application qui à H associe $-B^{-1}HB^{-1}$ est clairement linéaire, et de plus,

$$\|T(B + H) - T(B) + B^{-1}HB^{-1}\| \leq \|B^{-1}\| \sum_{k=2}^{+\infty}(\|B^{-1}\|\|H\|)^k.$$

Or $\|B^{-1}\|\|H\| < 1$ par hypothèse. On a donc

$$\begin{aligned} \frac{\|T(B+H)-T(B)-B^{-1}HB^{-1}}{\|H\|}\| &\le \|B^{-1}\|^3\|H\|\sum_{k=0}^{+\infty}(\|B^{-1}\|\|H\|)^k \\ &\to 0 \text{ lorsque } \|H\| \to 0. \end{aligned}$$

On en déduit que l'application T est différentiable et que $DT(B)(H) = -B^{-1}HB^{-1}$.

(c) La méthode de Newton pour la recherche d'un zéro de la fonction g s'écrit :

$$\begin{cases} B^0 \in GL_N(\mathbb{R})(\mathbb{R}), \\ Dg(B^n)(B^{n+1}-B^n) = -g(B^n). \end{cases}$$

Or, d'apres la question précédente, $Dg(B^n)(H) = -(B^n)^{-1}H(B^n)^{-1}$. On a donc

$$Dg(B^n)(B^{n+1}-B^n) = -(B^n)^{-1}(B^{n+1}-B^n)(B^n)^{-1}.$$

La méthode de Newton sécrit donc :

$$\begin{cases} B^0 \in GL_N(\mathbb{R})(\mathbb{R}), \\ -(B^{n+1}-B^n) = (Id - B^nA)B^n. \end{cases} \tag{2.5.39}$$

soit encore

$$\begin{cases} B^0 \in GL_N(\mathbb{R})(\mathbb{R}), \\ B^{n+1} = 2B^n - B^nAB^n. \end{cases} \tag{2.5.40}$$

(d) Par définition, on a :

$$Id - AB^{n+1} = Id - A(2B^n - B^nAB^n) = Id - 2AB^n + AB^nAB^n.$$

Comme les matrices Id et AB^n commutent, on a donc :

$$Id - AB^{n+1} = (Id - AB^n)^2.$$

Une récurrence immédiate montre alors que $Id - AB^n = (Id - AB^0)^{2^n}$. On en déduit que la suite $Id - AB^n$ converge (vers la matrice nulle) lorsque $n \to +\infty$ ssi $\rho(Id - AB^0) < 1$, et ainsi que la suite B^n converge vers A^{-1} si et seulement si $\rho(Id - AB^0) < 1$.

Corrigé de l'exercice 56 page 101 (Valeurs propres et méthode de Newton)

On écrit le système sous la forme $F(x,\lambda)=0$ où F est une fonction de $\mathbb{R}^{N+1}$ dans $\mathbb{R}^{N+1}$ définie par

$$F(y)=F(x,\lambda)=\begin{pmatrix} Ax-\lambda x \\ x\cdot x-1 \end{pmatrix},$$

et on a donc

$$DF(\overline{\lambda},\overline{x})(z,\nu)=\begin{pmatrix} Az-\bar{\lambda}z-\nu\bar{x} \\ 2\bar{x}\cdot z \end{pmatrix},$$

Supposons que $DF(\overline{x},\overline{\lambda})(z,\nu)=0$, on a alors $Az-\bar{\lambda}z-\nu\bar{x}=0$ et $2\bar{x}\cdot z=0$. En multipliant la première équation par $\bar{x}$ et en utilisant le fait que A est symétrique, on obtient :

$$z\cdot A\bar{x}-\bar{\lambda}z\cdot\bar{x}-\nu\bar{x}\cdot\bar{x}=0, \tag{2.5.41}$$

et comme $A\bar{x}=\bar{\lambda}\bar{x}$ et $\bar{x}\cdot\bar{x}=1$, ceci entraîne que $\nu=0$. En revenant à (2.5.41) on obtient alors que $Ax-\bar{\lambda}x=0$, c.à.d. que $x\in Ker(A-\bar{\lambda}Id)=\mathbb{R}\bar{x}$ car $\bar{\lambda}$ est valeur propre simple. Or on a aussi $\bar{x}\cdot z=0$, donc $z\perp\bar{x}$ ce qui n'est possible que si $z=0$. On a ainsi montré que $Df(\bar{x},\bar{\lambda})$ est injective, et comme on est en dimension finie, $Df(\bar{x},\bar{\lambda})$ est bijective. Dnc, d'après le théorème du cours, la méthode de Newton est localement convergente.

Corrigé de l'exercice 58 page 102 (Convergence de la méthode de Newton si $f'(\overline{x})=0$)

Comme $f''(\overline{x})\neq 0$, on peut supposer par exemple $f''(\overline{x})>0$; par continuité de f'', il existe donc $\eta>0$ tel que $f'(x)<0$ si $x\in]\overline{x}-\eta,\overline{x}[$ et $f'(x)>0$ si $x\in]\overline{x},\overline{x}+\eta[$, et donc f est décroissante sur $]\overline{x}-\eta,\overline{x}[$ (et croissante sur $]\overline{x},\overline{x}+\eta[$).
Supposons $x_0\in]\overline{x},\overline{x}+\eta[$, alors $f'(x_0)>0$ et $f''(x_0)>0$.
On a par définition de la suite $(x_n)_{n\in\mathbb{N}}$,

$$\begin{aligned} f'(x_0)(x_1-x_0) &= -f(x_0) \\ &= f(\overline{x})-f(x_0) \\ &= f'(\xi_0)(\overline{x}-x_0),\ \text{où } \xi_0\in]\overline{x},x_0[\end{aligned}$$

Comme f' est strictement croissante sur $]\overline{x},\overline{x}+\eta[$, on a $f'(\xi_0)<f'(x_0)$ et donc $x_1\in]\overline{x},x_0[$.
On montre ainsi par récurrence que la suite $(x_n)_{n\in\mathbb{N}}$ vérifie

$$x_0>x_1>x_2\ldots>x_n>x_{n+1}>\ldots>\overline{x}.$$

La suite $(x_n)_{n\in\mathbb{N}}$ est donc décroissante et minorée, donc elle converge. Soit x sa limite ; comme

$$f'(x_n)(x_{n+1}-x_n)=-f(x_n) \text{ pour tout } n\in\mathbb{N},$$

on a en passant à la limite : $f(x)=0$, donc $x=\overline{x}$.
Le cas $f''(\overline{x})<0$ se traite de la même manière.
Montrons maintenant que la méthode est d'ordre 1. Par définition, la méthode est d'ordre 1 si

$$\frac{\|x_{n+1}-\overline{x}\|}{\|x_n-\overline{x}\|}\to\beta\in\mathbb{R}_+^*.$$

Par définition de la suite $(x_n)_{n\in\mathbb{N}}$, on a :

$$f'(x_n)(x_{n+1}-x_n)=-f(x_n) \tag{2.5.42}$$

Comme $f \in \mathcal{C}^2(\mathbb{R})$ et $f'(\overline{x}) = 0$, il existe $\xi_n \in]\overline{x}, x_n[$ et $\eta_n \in]\overline{x}, x_n[$ tels que $f'(x_n) = f''(\xi_n)(x_n - \overline{x})$ et $-f(x_n) = -\dfrac{1}{2}f''(\eta_n)(\overline{x} - x_n)^2$. On déduit donc de (2.5.42) que

$$\begin{aligned} f''(\xi_n)(x_{n+1} - x_n) &= -\frac{1}{2}f''(\eta_n)(x_n - \overline{x}), \\ \text{soit } f''(\xi_n)(x_{n+1} - \overline{x}) &= (-\frac{1}{2}f''(\eta_n) + f''(\xi_n))(x_n - \overline{x}) \end{aligned}$$

$$\text{Donc } \frac{\|x_{n+1} - \overline{x}\|}{\|x_n - \overline{x}\|} = |1 - \frac{1}{2}\frac{f''(\eta_n)}{f''(\xi_n)}| \to \frac{1}{2} \text{ lorsque } n \to +\infty.$$

La méthode est donc d'ordre 1.
On peut obtenir une méthode d'ordre 2 en appliquant la méthode de Newton à f'.

Corrigé de l'exercice 60 page 103 (Variante de la méthode de Newton)

1. On a évidemment $x^{(0)} = x_0 \in I$. Supposons que $x^{(n)} \in I$ et montrons que $x^{(n+1)} \in I$. Par définition, on peut écrire :

$$x^{(n+1)} = x^{(n)} - \frac{f(x^{(n)}) - f(x_0)}{f'(y)} - \frac{f(x_0)}{f'(y)}.$$

Donc

$$x^{(n+1)} - x_0 = x^{(n)} - x_0 - \frac{f'(\xi_n)(x_n^{(n)} - x_0) - f(x_0)}{f'(y)}, \text{ où } \xi_n \in [x_0, x^{(n)}].$$

On en déduit que

$$x^{(n+1)} - x_0 = (1 - \frac{f'(\xi_n)}{f'(y)})(x_n^{(n)} - x_0) - \frac{f(x_0)}{f'(y)}.$$

Ceci entraîne :

$$\begin{aligned} |x^{(n+1)} - x_0| &= \frac{1}{|f'(y)|}|f'(\xi_n) - f'(y)||x^{(n)} - x_0| + \frac{|f(x_0)|}{f'(y)} \\ &\le \lambda\frac{1}{2\lambda}c + \frac{c}{2\lambda}\lambda = c. \end{aligned}$$

Donc $x^{(n+1)} \in I$.

2. On a :

$$\begin{aligned} x^{(n+1)} - \bar{x} &= x^{(n)} - \bar{x} - \frac{f(x^{(n)}) - f(\bar{x})}{f'(y)} - \frac{f(\bar{x})}{f'(y)}. \\ \text{Donc}|x^{(n+1)} - \bar{x}| &\le |x^{(n)} - \bar{x}||f'(y) - f'(\eta_n)|\frac{1}{|f'(y)|} \text{ où } \eta_n \in [\bar{x}, x^{(n)}]; \end{aligned}$$

Par hypothèse, on a donc

$$\begin{aligned} |x^{(n+1)} - \bar{x}| &\le |x^{(n)} - \bar{x}|\frac{1}{2\lambda}\lambda \\ &\le \frac{c}{2}|x^{(n)} - \bar{x}|. \end{aligned}$$

On en déduit par récurrence que

$$|x^{(n)} - \bar{x}| \le \frac{c}{2^n}|x^{(0)} - \bar{x}|.$$

Ceci entraîne en particulier que

$$\begin{array}{ll} x^{(n)} & \to \bar{x} \\ n & \to +\infty. \end{array}$$

Il reste à montrer que la convergence est au moins linéaire. On a :

$$\frac{|x^{(n+1)} - \bar{x}|}{|x^{(n)} - \bar{x}|} = |f'(y) - f'(x^{(n)})| \frac{1}{|f'(y)|}$$

$$\text{Donc} \quad \begin{array}{l} \dfrac{|x^{(n+1)} - \bar{x}|}{|x^{(n)} - \bar{x}|} \quad \to |1 - \dfrac{f'(\bar{x})}{f'(y)}| = \beta \geq 0 \\ \qquad\qquad n \to +\infty \end{array}$$

La convergence est donc au moins linéaire, elle est linéaire si $f'(\bar{x}) \neq f'(y)$ et super–linéaire si $f'(\bar{x}) = f'(y)$.

3. Le fait de remplacer y par $y^{(n)}$ ne change absolument rien à la preuve de la convergence de $x^{(n)}$ vers $\bar{x}$. Par contre, on a maintenant :

$$\begin{aligned} \frac{|x^{(n+1)} - \bar{x}|}{|x^{(n)} - \bar{x}|} &= |f'(y_n) - f'(\eta_n)| \frac{1}{|f'(y_n)|} \\ &= |1 - \frac{f'(\eta_n)}{f'(y_n)}| \end{aligned}$$

Or $f'(\eta_n) \underset{n\to+\infty}{\to} f'(\bar{x})$ et donc si $f'(y_n) \underset{n\to+\infty}{\to} f'(\bar{x})$ la convergence devient superlinéaire.

4. L'algorithme pour $N \geq 1$ se généralise en :

$$\left\{ \begin{array}{l} x^{(0)} = x_0 \\ x^{(n+1)} = x^{(n)} - (DF(y))^{-1} f(x^{(n)}). \end{array} \right.$$

On a donc

$$x^{(n+1)} - x_0 = x^{(n)} - x_0 - (DF(y))^{-1}(f(x^{(n)}) - f(x_0)) + (DF(y))^{-1} f(x_0).$$

Or $f(x^{(n)}) - f(x^{(0)}) = \varphi(1) - \varphi(0) = \displaystyle\int_0^1 \varphi'(t)dt$, où

$$\varphi(t) = f(tx^{(n)} + (1-t)x^{(0)})$$

et donc

$$\varphi'(t) = Df(tx^{(n)} + (1-t)x^{(0)})(x^{(n)} - x^{(0)}).$$

Donc

$$\begin{array}{l} x^{(n+1)} - x^{(0)} = \\ (x^{(n)} - x^{(0)}) \left(1 - (Df(y))^{-1} \displaystyle\int_0^1 Df(tx^n) + (1-t)x^{(0)})dt\right) \\ +(Df(y))^{-1} f(x_0) = \\ (x^{(n)} - x^{(0)})(Df(y))^{-1} \left(\displaystyle\int_0^1 \left(Df(y) - Df(tx^{(n)} + (1-t)x^{(0)}\right) dt\right) \\ +(Df(y))^{-1} f(x_0). \end{array}$$

On en déduit que :

$$\begin{array}{l} \|x^{(n+1)} - x^{(0)}\| \leq \\ \qquad \|x^{(n)} - x^{(0)}\| \|(Df(y))^{-1}\| \displaystyle\int_0^1 \|Df(y) - Df(tx^{(n)} + (1-t)x^{(0)})\| dt \\ \qquad +\|(Df(y))^{-1}\| \|f(x_0)\|. \end{array}$$

Si on suppose que $x^{(n)} \in I$, alors $tx^{(n)} + (1-t)x^{(0)} \in I$. L'hypothèse (iii) généralisée à la dimension N s'écrit :

$$\|Df(x) - Df(y)\| \leq \frac{1}{2\lambda} \ \forall (x,y) \in I^2,$$

si on suppose de plus que

(ii) $\|f(x_0)\| \leq \dfrac{c}{2\lambda}$ et

(iv) $\|(Df(x))^{-1}\| \leq \lambda \ \forall x \in I$, alors 2.5.43 donne que

$$\begin{aligned} \|x^{(n+1)} - x^{(0)}\| &\leq \|x^{(n)} - x^{(0)}\|\lambda\frac{1}{2\lambda} + \lambda\frac{c}{2\lambda} \\ &\leq c. \end{aligned}$$

ce qui prouve que $x^{(n+1)} \in I$.

On montre alors de la même manière que $x^{(n)}_{n\to\infty} \to \bar{x}$, (car $\|x^{(n+1)} - \bar{x}\| \leq \frac{1}{2}\|x^{(n)} - \bar{x}\|$).

Corrigé de l'exercice 62 page 104 (Méthode de Steffensen)

1. Comme $f'(\overline{x}) \neq 0$, il existe $\bar{\alpha} > 0$ tel que f soit strictement monotone sur $B(\overline{x}, \bar{\alpha})$; donc si $f(x) = 0$ et $x \in B(\overline{x}, \bar{\alpha})$ alors $x = \overline{x}$. De plus, comme $x + f(x) \to \overline{x}$ lorsque $x \to \overline{x}$, il existe α tel que si $x \in B(\overline{x}, \alpha)$, alors $f(x + f(x) \in B(\overline{x}, \bar{\alpha})$. Or si $x \in B(x, \alpha)$, on a :$f(x) \neq 0$ si $x \neq \overline{x}$, donc $x + f(x) \neq x$ et comme $x + f(x) \in B(\overline{x}, \bar{\alpha})$ où f est strictement monotone, on a $f(x) \neq f(x + f(x))$ si $x \neq \overline{x}$. On en déduit que si $x_n \in B(\overline{x}, \alpha)$, alors $f(x_n + f(x_n)) \neq f(x_n)$ (si $x_n \neq \overline{x}$) et donc x_{n+1} est défini par $x_{n+1} = \dfrac{(f(x_n))^2}{f(x_n + f(x_n)) - f(x_n)}$. Ceci est une forme de stabilité du schéma).
2. Montrons maintenant que la suite $(x_n)_{n\in\mathbb{N}}$ vérifie :

$$x_{n+1} - \overline{x} = a(x_n)(x_n - \overline{x})^2 \quad \text{si } x_n \neq \overline{x} \quad \text{et } x_0 \in B(\overline{x}, \alpha),$$

où a est une fonction continue. Par définition de la suite $(x_n)_{n\in\mathbb{N}}$, on a :

$$x_{n+1} - \overline{x} = x_n - \overline{x} - \frac{(f(x_n))^2}{f(x_n + f(x_r)) - f(x_n)}. \tag{2.5.43}$$

Soit $\psi_n : [0,1] \to \mathbb{R}$ la fonction définie par :

$$\psi_n(t) = f(x_n + tf(x_n))$$

On a $\psi_n \in \mathcal{C}^2(]0,1[, \mathbb{R})$, $\psi_n(0) = f(x_n)$ et $\psi_n(1) = f(x_n + f(x_n))$.

On peut donc écrire :

$$f(x_n + f(x_n)) - f(x_n) = \psi_n(1) - \psi_n(0) = \int_0^1 \psi_n'(t)dt$$

Ceci donne :

$$f(x_n + f(x_n)) - f(x_n) = \int_0^1 f'(x_n + tf(x_n))f(x_n)dt$$

On pose maintenant $\xi_n(t) = f'(x_n + tf(x_n))$, et on écrit que $\xi_n(t) = \int_0^t \xi_n'(s)ds + \xi_n(0)$.

On obtient alors :

$$f(x_n + f(x_n)) - f(x_n) = f(x_n)\left[f(x_n)\int_0^1\int_0^t f''(x_n + sf(x_n))ds + f'(x_n)\right]. \quad (2.5.44)$$

Soit $b \in \mathcal{C}(\mathbb{R},\mathbb{R})$ la fonction définie par :

$$b(x) = \int_0^1 \left(\int_0^t f''(x + sf(x))ds\right) dt.$$

Comme $f \in \mathcal{C}(\mathbb{R},\mathbb{R})$, on a $b(x) \to \frac{1}{2}f''(\overline{x})$ lorsque $x \to \overline{x}$

L'égalité (2.5.44) s'écrit alors :

$$f(x_n + f(x_n)) - f(x_n) = (f(x_n))^2 b(x_n) + f(x_n)f'(x_n). \quad (2.5.45)$$

Comme $x_0 \in B(\overline{x},\alpha)$, on a $x_n \in B(\overline{x},\alpha)$ et donc $f(x_n) \neq 0$ si $x_n \neq \overline{x}$.

Donc pour $x_n \neq \overline{x}$, on a grâce à (2.5.43) et (2.5.45) :

$$x_{n+1} - \overline{x} = x_n - \overline{x} - \frac{f(x_n)}{f(x_n)b(x_n) + f'(x_n))} \quad (2.5.46)$$

On a maintenant $-f(x_n) = f(\overline{x}) - f(x_n) = \displaystyle\int_0^1 \varphi'(t)dt$ où $\varphi \in \mathcal{C}^2(\mathbb{R},\mathbb{R})$ est définie par $\varphi(t) = f(t\overline{x} + (1-t)x_n)$.

Donc

$$-f(x_n) = \int_0^1 f'(t\overline{x} + (1-t)x_n)(\overline{x} - x_n)dt.$$

Soit $\chi \in \mathcal{C}^1(\mathbb{R},\mathbb{R})$ la fonction définie par $\chi(t) = f'(t\overline{x} + (1-t)x_n)$,

on a $\chi(0) = f'(x_n)$ et donc :

$$-f(x_n) = \int_0^1 \left[\int_0^t \left(f''(s\overline{x} + (1-s)x_n)(\overline{x} - x_n) + f'(x_n)\right) ds(\overline{x} - x_n)\right] dt$$

Soit $c \in \mathcal{C}(\mathbb{R},\mathbb{R})$ la fonction définie par

$$c(x) = \int_0^1 \left(\int_0^t f''(s\overline{x} + (1-s)x)ds\right) dt,$$

on a $c(x) \to \frac{1}{2}f''(\overline{x})$ lorsque $x \to \overline{x}$ et :

$$-f(x_n) = c(x)(\overline{x} - x_n)^2 + f'(x_n)(\overline{x} - x_n) \quad (2.5.47)$$

De (2.5.47) et(1.6.62), on obtient :

$$\begin{aligned} x_{n+1} - \overline{x} &= (x_n - \overline{x})\left[1 + \frac{c(x_n)(x_n - \overline{x}) - f'(x_n)}{f(x_n)b(x_n) + f'(x_n)}\right] \\ &= \frac{(x_n - \overline{x})}{f(x_n)b(x_n) + f'(x_n)}\left(-c(x_n)(\overline{x} - x_n)^2 b(x_n)\right. \\ &\quad \left. - f'(x_n)(\overline{x} - x_n)b(x_n) + f'(x_n) + c(x_n)(x_n - \overline{x}) - f'(x_n)\right) \end{aligned}$$

On en déduit :

$$(x_{n+1} - \overline{x}) = (x_n - \overline{x})^2 a(x_n) \tag{2.5.48}$$

où

$$a(x) = \frac{c(x)b(x)(x-\overline{x}) + f'(x)b(x)b + c(x)}{f(x) + f'(x)}$$

La fonction a est continue en tout point x tel que

$$D(x) = f(x)b(x) + f'(x) \neq 0.$$

Elle est donc continue en $\overline{x}$ puisque $D(\overline{x}) = f(\overline{x})b(\overline{x}) + f'(\overline{x}) = f'(\overline{x}) \neq 0$.
De plus, comme f, f' et b sont continues, il existe un voisinage de $\overline{x}$ sur lequel D est non nulle et donc a continue.

3. Par continuité de a, pour tout $\varepsilon > 0$, il existe $\eta_\varepsilon > 0$ tel que si $x \in B(\overline{x}, \eta_\varepsilon)$ alors
(7) $|a(x) - a(\overline{x})| \leq \varepsilon$.
Calculons

$$\begin{aligned} a(\overline{x}) &= \frac{f'(\overline{x})b(\overline{x}) + c(\overline{x})}{f'(\overline{x})} \\ &= \frac{1}{2}f''(\overline{x})\frac{1 + f'(\overline{x})}{f'(\overline{x})} = \beta. \end{aligned}$$

Soit $\gamma = \min(\eta_1, \frac{1}{2(\beta+1)})$; si $x \in B(\overline{x}, \gamma)$, alors $|a(x)| \leq \beta + 1$ grâce à (7), et $|x - \overline{x}| \leq \frac{1}{2(\beta+1)}$.
On déduit alors de (6) que si $x_n \in B(\overline{x}, \gamma)$, alors

$$|x_{n+1} - \overline{x}| \leq \frac{1}{2}|x_n - \overline{x}|.$$

Ceci entraîne d'une part que $x_{n+1} \in B(\overline{x}, \gamma)$ et d'autre part, par récurrence, la convergence de la suite $(x_n)_{n\in\mathbb{N}}$ vers $\overline{x}$.
Il reste à montrer que la convergence est d'ordre 2.
Grâce à (6), on a :

$$\frac{|x_{n+1} - \overline{x}|}{|x_n - \overline{x}|^2} = |a(x_n)|.$$

Or on a montré à l'étape 3 que a est continue et que $a(x) \to \beta \in \mathbb{R}$. On a donc une convergence d'ordre au moins 2.

Corrigé de l'exercice 64 page 105 (Méthode de la sécante)

1. Supposons x_{n-1} et x_n connus.
Pour que x_{n+1} soit bien défini, il faut et il suffit que $f(x_n) \neq f(x_{n-1})$. Or par hypothèse, $f'(\bar{x}) \neq 0$. On en déduit qu'il existe un voisinage de $\bar{x}$ sur lequel f' est monotone, donc bijective. Donc il existe ε_1 tel que si $x_n, x_{n-1} \in]\overline{x} - \varepsilon_1, \overline{x} + \varepsilon_1[$, $x_n \neq x_{n-1}$ et $x_n \neq \overline{x}$, alors $f(x_n) \neq f(x_{n-1})$. De même, toujours par injectivité de f sur $]\overline{x} - \varepsilon_1, \overline{x} + \varepsilon_1[$, on a $f(x_n) \neq 0$.
En choisissant x_0 et x_1 dans l'intervalle $]\overline{x} - \varepsilon_1, \overline{x} + \varepsilon_1[$, on a par une récurrence immédiate que la suite $(x_n)_{n\in\mathbb{N}}$ est bien définie.
Par définition, si $f(x_n) \neq 0$, on a :

$$x_{n+1} - \bar{x} = x_n - \bar{x} - \frac{f(x_n) - f(\bar{x})}{x_n - \bar{x}}(x_n - \bar{x})\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}.$$

En notant $I(a,b)$ l'intervalle d'extrémités a et b, il existe donc $\theta_n \in I(\bar{x}, x_n)$ et $\zeta_n \in I(x_{n-1}, x_n)$ tels que $x_{n+1} - \bar{x} = (x_n - \bar{x})(1 - \dfrac{f'(\theta_n)}{f'(\zeta_n)})$, , et donc : $e_{n+1} = |1 - \dfrac{f'(\theta_n)}{f'(\zeta_n)}| e_n$.

Or f' est continue, il existe ε_2 tel que $x_n,\, x_{n-1} \in]\overline{x} - \varepsilon_2, \overline{x} + \varepsilon_2[$, alors $1 - \dfrac{f'(\theta_n)}{f'(\zeta_n)} \leq 1/2$, et donc $e_{n+1} \leq \frac{1}{2} e_n$.

En posant $\varepsilon = \min(\varepsilon_1, \varepsilon_2)$, on a donc par récurrence le fait que si x_0 et x_1 appartiennent à l'intervalle $]\overline{x} - \varepsilon, \overline{x} + \varepsilon[$, la suite $(x_n)_{n\in\mathbb{N}}$ est bien définie et la méthode de la sécante est localement convergente.

2. (a) Par définition,

$$e_{n+1} = e_n - \frac{f(x_n) - f(\overline{x})}{f(x_n) - f(x_{n-1})}(x_n - x_{n-1}).$$

Donc :

$$\begin{aligned}
(f(x_n) - f(x_{n-1})e_{n+1} &= e_n f(x_n) - e_n f(x_{n-1}) - f(x_n)e_n + f(x_n)e_{n-1} && (2.5.49)\\
&= -e_n f(x_{n-1}) + f(x_n)e_{n-1} && (2.5.50)\\
&= e_n e_{n-1}\left(\frac{f(x_n)}{e_n} - \frac{f(x_{n-1})}{e_{n-1}}\right). && (2.5.51)
\end{aligned}$$

Or $\frac{f(x_n)}{e_n} = \frac{f(x_n)-f(\overline{x})}{e_n}$ (resp. $\frac{f(x_{n-1}}{e_{n-1}} = \frac{f(x_{n-1}-f(\overline{x})}{e_{n-1}}$)est la valeur moyenne de f' sur l'intervalle d'extrémités $\overline{x}, x_n$ (resp. $\overline{x}, x_{n-1}$). On en déduit que $(f(x_n) - f(x_{n-1})e_{n+1} = e_n e_{n-1}(\mu_n - \mu_{n-1})$, d'où le résultat.

(b) Si $x > \overline{x}$, la fonction μ vérifie :

$$(x - \overline{x})\mu(x) = \int_{\overline{x}}^{x} f'(t)dt,$$

on en déduit que la fonction μ est continue et dérivable et sa dérivée μ' vérifie :

$$(x - \overline{x})\mu'(x) + \mu(x) = f'(x),\ \forall x > \overline{x}.$$

soit encore

$$\mu'(x) = \frac{f'(x) - \mu(x)}{x - \overline{x}},\ \forall x > \overline{x}. \tag{2.5.52}$$

Or

$$\begin{aligned}
\mu(x) &= \frac{1}{x - \overline{x}}(f(x) - f(\overline{x})) && (2.5.53)\\
&= \frac{1}{x - \overline{x}}(f(x) - (f(x) + (\overline{x} - x)f'(x) + \frac{1}{2}(\overline{x} - x)^2 f''(x)(\overline{x} - x)^3\varepsilon(x). && (2.5.54)
\end{aligned}$$

On en déduit que

$$\mu(x) = f'(x) + \frac{1}{2}(x - \overline{x})f''(x) + (x - \overline{x})^2\varepsilon(x).$$

Et finalement, en reportant dans (2.5.52) :

$$\mu'(x) = \frac{1}{2}f''(x) + (x - \overline{x})\varepsilon(x),\ \forall x > \overline{x}. \tag{2.5.55}$$

On en déduit que μ' admet une limite lorsque x tend vers $\overline{x}$ par valeurs positives. Le même raisonnement pour $x < \overline{x}$ donne le même résultat.

Enfin, comme $f \in C^2(\mathbb{R}, \mathbb{R})$, on peut passer à la limite dans (2.5.55) et on obtient :

$$\lim_{x\to\overline{x}} \mu'(x) = \frac{1}{2}f''(\overline{x}). \tag{2.5.56}$$

(c) Par définition, on a

$$M_n = |\frac{\mu(x_n) - \mu(x_{n-1})}{x_n - x_{n-1}} \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}| = \frac{\mu'(\zeta_n)}{f'(\xi_n)},$$

où ζ_n et ξ_n sont compris entre x_{n-1} et x_n (par le théorème des accroissements finis). Comme la suite $(x_n)_{n\in\mathbb{N}}$ tend vers $\overline{x}$, comme f' est continue et grâce à (2.5.56), on a :

$$\lim_{n\to+\infty} M_n = \frac{1}{2}\frac{f''(\overline{x})}{f'(\overline{x})}.$$

Notons que cette limite est finie car $f'(\overline{x}) \neq 0$ par hypothèse. On en conclut que la suite $(M_n)_{n\in\mathbb{N}}$ est bornée.

3. (a) La relation à démontrer est vérifiée pour $n = 0$ et $n = 1$. Supposons-la vérifiée jusqu'au rang n. On a par définition : $a_{n+1} = a_n a_{n-1} \geq Me_n Me_{n-1}$. Or par la question 2a, $e_n e_{n-1} = M_n e_{n+1} \leq Me_{n+1}$. On en déduit que la relation est encore vérifiée au rang $n + 1$.

(b) Par définition, $a_i = Me_i = M(x_i - \overline{x})$, pour $i = 0, 1$, donc si $x_0, x_1 \in]\overline{x} - \varepsilon_1, \overline{x} + \varepsilon_1[$ avec $\varepsilon_1 < 1/M$, alors $a_0 < 1$ et $a_1 < 1$. On en déduit alors facilement par récurrence que la suite $a_n < 1$, et donc que la suite $(a_n)_{n\in\mathbb{N}}$ est strictement décroissante. Elle converge donc vers une limite $\bar{a}$ qui vérifie $\bar{a} = \bar{a}^2$ et $\bar{a} < 1$. On en déduit que la limite est nulle.

(c) On pose $b_n = \ln a_n$ on a donc

$$b_{n+1} = b_n + b_{n-1}, \; \forall n \geq 1 \tag{2.5.57}$$

L'ensemble de suites $(b_n)_{n\in\mathbb{N}}$ vérifiant (2.5.57) est un espace vectoriel de dimension 2. Pour trouver une base de cet espace vectoriel, on cherche des éléments de cet espace sous la forme $b_n = r^n$, $n \geq 0$. Une telle suite vérifie (2.5.57) si et seulement si $r^2 = r + 1$, c.à.d. $r = \frac{1\pm\sqrt{5}}{2}$. Si la suite $(b_n)_{n\in\mathbb{N}}$ vérifie (2.5.57), il existe donc $C \in \mathbb{R}$ et $D \in \mathbb{R}$ tels que

$$b_n = \ln(a_n) = C(\frac{1+\sqrt{5}}{2})^n + D(\frac{1-\sqrt{5}}{2})^n.$$

On en déduit que $a_n \leq \alpha\beta^{d^n}$, avec $d = \frac{1\pm\sqrt{5}}{2}$, $\alpha = e^{|D|}$ et $\beta = e^C$. Notons qu'on a bien $0 < \beta < 1$ car $C < 0$ puisque $\ln(a_n) < 0$, pour tout $n \in \mathbb{N}$.

(d) Par la question 2(c) et l'hypothèse $f''(\overline{x}) \neq 0$, on déduit que $\overline{M} > 0$. Comme $e_{n+1} = M_n e_n e_{n-1}$, on a $\ln e_{n+1} = \ln M_n + \ln e_n + \ln e_{n-1}$; si on pose $\beta_n = \ln e_{n+1} - d \ln e_n$ (pour $n \geq 0$, on a donc

$$\begin{aligned}
\beta_n &= (1-d)\ln e_n + \ln e_{n-1} + \ln M_n \\
&= (1-d)(\beta_{n-1} + d\ln e_{n-1}) + \ln e_{n-1} + \ln M_n \\
&= (1-d)(\beta_{n-1} + (1-d)d\ln e_{n-1}) + \ln e_{n-1} + \ln M_n.
\end{aligned}$$

Or $(1-d)d = -1$ car d est racine de l'équation : $d^2 - d - 1 = 0$. On obtient donc finalement

$$\beta_n = (1-d)\beta_{n-1} + \ln M_n.$$

On pose maintenant $\beta_n = C_n(1-d)^n$ (obtenu par “variation de la constante" C pour la solution de l'équation homogène $\beta_n = (1-d)\beta_{n-1}$). On obtient alors

$$C_n(1-d)^n = (1-d)C_{n-1}(1-d)^{n-1} + \ln M_n.$$

Ceci entraîne :

$$C_n = C_{n-1} + \frac{\ln M_n}{(1-d)^n}.$$

Donc
$$C_n = C_0 + \sum_{p=1}^{n} \frac{\ln M_p}{(1-d)^p},$$
et comme la suite $(M_n)_{n\in\mathbb{N}}$ est bornée, la série de terme général $\frac{\ln M_p}{(1-d)^p}$ est convergente. Comme $(1-d)^n$ tend vers 0 lorsque n tend vers l'infini, on en déduit que $\beta_n \to 0$. On a donc $\ln e_{n+1} - d \ln e_n \to 0$, i.e. $\frac{e_{n+1}}{e_n^d} \to 1$ lorsque $n \to +\infty$.

(e) L'ordre de convergence de la méthode de la sécante est $d = \frac{1+\sqrt{5}}{2} < 2$, donc plus petit que l'ordre de convergence de la méthode de Newton.