Faculté des Sciences
Aix-Marseille Université

CTES
Centre de Télé-Enseignement Sciences

LICENCE 3 MATHEMATIQUES – INFORMATIQUE. MATHEMATIQUES GENERALES. L3MiMG.

Expédition dans la semaine n°	Etape	Code UE	N° d'envoi de l'UE
5	**2L3MAT**	**SMI5U4T**	**4**

Nom de l'UE : Analyse numérique et optimisation

Le cours contient 3 chapitres (systèmes linéaires, systèmes non linéaires, optimisation). Pour chaque semaine, il est proposé d'étudier une partie du cours, de faire des exercices (corrigés) et, éventuellement, de réaliser un TP en python. Les TP sont conseillés mais non obligatoires. Deux devoirs sont à rendre afin de bénéficier d'une note de contrôle continu.
note finale=max(note-examen, 1/3(2 note-examen + note-contrôle-continu)).

Contenu de l'envoi : Polycopié, section 2.3. Corrigé du premier devoir

- Guide du travail à effectuer

Semaine 1 :
Etudier le paragraphe 2.2.2 (point fixe de monotonie, contenu dans l'envoi 3)
Exercice proposé (avec corrigé, contenu dans l'envoi 3), 84 (méthode de monotonie)

Semaine 2 :
Etudier le paragraphe 2.2.3 (vitesse de convergence, contenu dans l'envoi 3)
Exercice proposé (avec corrigé, contenu dans l'envoi 3) : 85 (point fixe amélioré)

Semaine 3 :
Etudier le paragraphe 2.2.4 (Newton dans R, contenu dans l'envoi 3)
Exercices proposés (avec corrigés) :
86 (Newton pour la fonction log) et 88 (Condition initiale), dans cet envoi

Semaine 4 :
Etudier le paragraphe 2.3 (Newton dans Rn)
Exercices proposés (avec corrigés) : 87, 89, 92 et 99 (sur Newton)
L'exercice 109 (Racine cubique et Newton) fait partie du deuxième devoir (à rendre en mars)

-Coordonnées de l'enseignant responsable de l'envoi

T. Gallouet, CMI, 39 rue Joliot Curie, 13453 marseille cedex 13
email : thierry.gallouet@univ-amu.fr

Vous pouvez aussi consulter la page web:http://www.i2m.univ-amu.fr/~gallouet/tele.d/anum.d
et me poser des questions par email

2.3 Méthode de Newton dans $\mathbb{R}^n$

2.3.1 Construction et convergence de la méthode

On a vu ci-dessus comment se construit la méthode de Newton à partir du point fixe de contraction 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 $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ tels que $g(\bar{\boldsymbol{x}}) = 0$.
On généralise la méthode vue en 1D en remplaçant dans (2.10) la dérivée $g'(x^{(k)})$ par la matrice jacobienne de g au point $x^{(k)}$, qu'on note $Dg(\boldsymbol{x}^{(k)})$. La méthode s'écrit :

$$\begin{cases} \boldsymbol{x}^{(0)} \in \mathbb{R}^n \\ Dg(\boldsymbol{x}^{(k)})(\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}) = -g(\boldsymbol{x}^{(k)}),\ \forall k \geq 0. \end{cases} \tag{2.20}$$

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

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

Remarque 2.18. *Si la fonction g dont on cherche un zéro est linéaire, i.e. si g est définie par $g(\boldsymbol{x}) = A\boldsymbol{x} - b$ 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 $A\boldsymbol{x} = b$. En effet $Dg(\boldsymbol{x}^{(k)}) = A$ et donc (2.20) s'écrit $A\boldsymbol{x}^{(k+1)} = \boldsymbol{b}$.*

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

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

Théorème 2.19 (Convergence de la méthode de Newton, $g \in C^3$)**.** *Soient $g \in C^3(\mathbb{R}^n, \mathbb{R}^n)$ et $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ tels que $g(\bar{\boldsymbol{x}}) = 0$. On munit $\mathbb{R}^n$ d'une norme $\|\cdot\|$. On suppose que $Dg(\bar{\boldsymbol{x}})$ est inversible. Alors la méthode de Newton converge localement, et la convergence est au moins quadratique. Plus précisément, il existe $b > 0$, et $\beta > 0$ tels que*

1. *si $\boldsymbol{x}^{(0)} \in B(\bar{\boldsymbol{x}}, b) = \{\boldsymbol{x} \in \mathbb{R}^n, \|\boldsymbol{x} - \bar{\boldsymbol{x}}\| \leq b\}$ alors la suite $(\boldsymbol{x}^{(k)})_{k\in\mathbb{N}}$ est bien définie par* (2.20) *et $\boldsymbol{x}^{(k)} \in B(\bar{\boldsymbol{x}}, b)$ pour tout $n \in \mathbb{N}$,*
2. *si $\boldsymbol{x}^{(0)} \in B(\bar{\boldsymbol{x}}, b)$ et si la suite $(\boldsymbol{x}^{(k)})_{k\in\mathbb{N}}$ est définie par* (2.20) *alors $\boldsymbol{x}^{(k)} \to \bar{\boldsymbol{x}}$ quand $n \to +\infty$,*
3. *si $\boldsymbol{x}^{(0)} \in B(\bar{\boldsymbol{x}}, b)$ et si la suite $(\boldsymbol{x}^{(k)})_{k\in\mathbb{N}}$ est définie par* (2.20) *alors $\|\boldsymbol{x}^{(k+1)} - \bar{\boldsymbol{x}}\| \leq \beta\|\boldsymbol{x}^{(k)} - \bar{\boldsymbol{x}}\|^2$ $\forall k \in \mathbb{N}$.*

DÉMONSTRATION – Montrons d'abord que la suite converge si $\boldsymbol{x}^{(0)}$ est suffisamment proche de $\bar{\boldsymbol{x}}$. Pour cela on va utiliser le théorème du point fixe : Soit f la fonction définie sur un voisinage de $\bar{\boldsymbol{x}}$ (et à valeurs dans $\mathbb{R}^n$) par $\boldsymbol{x} \mapsto \boldsymbol{x} - (Dg(\boldsymbol{x}))^{-1}g(\boldsymbol{x})$. On a

$$Df(\bar{\boldsymbol{x}}) = \mathrm{Id} - (Dg(\bar{\boldsymbol{x}}))^{-1}(Dg(\bar{\boldsymbol{x}})) = 0.$$

Comme $g \in C^2(\mathbb{R}, \mathbb{R})$, la fonction f est de classe C^1 et donc par continuité de Df, il existe $b > 0$ tel que $\|Df(\boldsymbol{x})\| \leq \frac{1}{2}$ pour tout $\boldsymbol{x} \in B = B(\bar{\boldsymbol{x}}, b)$. Si on montre que $f(B) \subset B$, alors la fonction f est strictement contractante de B dans B, et donc par le théorème du point fixe, la suite définie par (2.20) converge. Soit $\boldsymbol{x} = \boldsymbol{x}^{(k)} \in B$, et soit $y = \boldsymbol{x}^{(k+1)} = f(\boldsymbol{x}^{(k)})$. Grâce au théorème des acroissements finis dans des espaces vectoriels normés [5], on a :

$$\|\boldsymbol{y} - \bar{\boldsymbol{x}}\| = \|f(\boldsymbol{x}) - f(\bar{\boldsymbol{x}})\| \leq \sup_{\boldsymbol{z}\in B} \|Df(\boldsymbol{z})\| \|\boldsymbol{x} - \bar{\boldsymbol{x}}\|, \tag{2.21}$$

5. **Théorème des accroissements finis :** Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés, soient $h \in C^1(E, F)$ et $(\boldsymbol{x}, y) \in E^2$. 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 piège ! ! : 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. $h(y) - h(x) = Dh(\xi)(y - x)$.

et donc

$$\|\boldsymbol{y} - \bar{\boldsymbol{x}}\| \leq \frac{1}{2}\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|.$$

On en déduit que $\boldsymbol{y} \in B$. La suite $(\boldsymbol{x}^{(k)})_{k\in\mathbb{N}}$ définie par (2.20) est donc bien convergente. Pour montrer le caractère quadratique de la convergence, on applique à nouveau l'inégalité des accroissements finis, cette fois-ci à $Df(\boldsymbol{z})$ dans (2.21). En effet, comme $Df \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ (on utilise ici que g est de classe C^3), on a

$$\begin{aligned} \|Df(\boldsymbol{z})\| &= \|Df(\boldsymbol{z}) - Df(\bar{x})\| \\ &\leq \sup_{\xi\in B} \|D^2 f(\xi)\|\|\boldsymbol{z} - \bar{\boldsymbol{x}}\| && (2.22) \\ &\leq \beta\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|. && (2.23) \end{aligned}$$

En reportant cette majoration de $\|Df(\boldsymbol{z})\|$ dans (2.21), on obtient alors (avec $\beta = \sup_{\xi\in B} \|D^2 f(\xi)\|$) :

$$\|y - \bar{\boldsymbol{x}}\| \leq \beta\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|^2$$

ce qui donne la convergence locale au moins quadratique. ■

La condition $g \in C^3(\mathbb{R}^n, \mathbb{R}^n)$ est une condition suffisante mais non nécessaire. Si $g \in C^1(\mathbb{R}^n, \mathbb{R}^n)$, on peut encore démontrer la convergence, mais sous des hypothèses pas très faciles à vérifier en pratique :

Théorème 2.20 (Convergence de la méthode de Newton, $g \in C^1$)**.**
Soient $g \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ et $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ tels que $g(\bar{\boldsymbol{x}}) = 0$. On munit $\mathbb{R}^n$ d'une norme $\|\cdot\|$ et $\mathcal{M}_n(\mathbb{R})$ de la norme induite. On suppose que $Dg(\bar{\boldsymbol{x}})$ est inversible. On suppose de plus qu'il existe $a,\ a_1,\ a_2 \in \mathbb{R}_+^*$ tels que :

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

Alors, si on pose : $b = \min\left(a, \dfrac{1}{a_1 a_2}\right) > 0$, $\beta = a_1 a_2$ et si $\boldsymbol{x}^{(0)} \in B(\bar{x}, b)$, on a :

1. *$(\boldsymbol{x}^{(k)})_{k\in\mathbb{N}}$ est bien définie par* (2.20)*,*
2. *$\boldsymbol{x}^{(k)} \to \bar{\boldsymbol{x}}$ lorsque $n \to +\infty$,*
3. *$\|\boldsymbol{x}^{(k+1)} - \bar{\boldsymbol{x}}\| \leq \beta\|\boldsymbol{x}^{(k)} - \bar{\boldsymbol{x}}\|^2 \ \forall k \in \mathbb{N}$.*

DÉMONSTRATION – Soit $\boldsymbol{x}^{(0)} \in B(\bar{\boldsymbol{x}}, b) \subset B(\bar{\boldsymbol{x}}, a)$ où $b \leq a$. On va montrer par récurrence sur k que $\boldsymbol{x}^{(k)} \in B(\bar{\boldsymbol{x}}, b)$ $\forall k \in \mathbb{N}$ (et que $(\boldsymbol{x}^{(k)})_{k\in\mathbb{N}}$ est bien définie). L'hypothèse de récurrence est que $\boldsymbol{x}^{(k)}$ est bien défini, et que $\boldsymbol{x}^{(k)} \in B(\bar{\boldsymbol{x}}, b)$. On veut montrer que $\boldsymbol{x}^{(k+1)}$ est bien défini et $\boldsymbol{x}^{(k+1)} \in B(\bar{\boldsymbol{x}}, b)$. Comme $b \leq a$, la matrice $Dg(\boldsymbol{x}^{(k)})$ est inversible et $\boldsymbol{x}^{(k+1)}$ est donc bien défini ; on a :

$$\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)} = Dg(\boldsymbol{x}^{(k)})^{-1}(-g(\boldsymbol{x}^{(k)}))$$

Pour montrer que $\boldsymbol{x}^{(k+1)} \in B(\bar{\boldsymbol{x}}, b)$ on va utiliser le fait que $b \leq \dfrac{1}{a_1 a_2}$. Par hypothèse, on sait que si $\boldsymbol{x}, \boldsymbol{y} \in B(\bar{\boldsymbol{x}}, a)$, on a

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

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

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

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

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

et donc

$$\|Dg(\boldsymbol{x}^{(k)})(\boldsymbol{x}^{(k+1)} - \bar{\boldsymbol{x}})\| \leq a_2\|\bar{\boldsymbol{x}} - \boldsymbol{x}^{(k)}\|^2. \tag{2.24}$$

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

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

En utilisant (2.24), les hypothèses 1 et 2 et le fait que $\boldsymbol{x}^{(k)} \in B(\bar{\boldsymbol{x}}, b)$, on a donc

$$\|\boldsymbol{x}^{(k+1)} - \bar{\boldsymbol{x}}\| \leq a_1 a_2\|\boldsymbol{x}^{(k)} - \bar{\boldsymbol{x}}\|^2 < a_1 a_2 b^2. \tag{2.25}$$

Or $a_1a_2b^2 < b$ car $b \leq \dfrac{1}{a_1a_2}$. Donc $\boldsymbol{x}^{(k+1)} \in B(\bar{\boldsymbol{x}}, b)$.

On a ainsi montré par récurrence que la suite $(\boldsymbol{x}^{(k)})_{k\in\mathbb{N}}$ est bien définie et que $\boldsymbol{x}^{(k)} \in B(\bar{\boldsymbol{x}}, b)$ pour tout $k \geq 0$.

Pour montrer la convergence de la suite $(\boldsymbol{x}^{(k)})_{k\in\mathbb{N}}$ vers $\bar{\boldsymbol{x}}$, on repart de l'inégalité (2.25) :

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

et donc par récurrence

$$a_1a_2\|\boldsymbol{x}^{(k)} - \bar{\boldsymbol{x}}\| \leq \left(a_1a_2\|\boldsymbol{x}^{(0)} - \bar{\boldsymbol{x}}\|\right)^{2^k}, \quad \forall k \in \mathbb{N}$$

Comme $\boldsymbol{x}^{(0)} \in B(\bar{\boldsymbol{x}}, b)$ et $b \leq \frac{1}{a_1a_2}$, on a $a_1a_2\|\boldsymbol{x}^{(0)} - \bar{\boldsymbol{x}}\| < 1$ et donc $\|\boldsymbol{x}^{(k)} - \bar{\boldsymbol{x}}\| \to 0$ quand $k \to +\infty$.

La convergence est au moins quadratique car l'inégalité (2.25) s'écrit :

$$\|\boldsymbol{x}^{(k+1)} - \bar{\boldsymbol{x}}\| \leq \beta\|\boldsymbol{x}^{(k)} - \boldsymbol{x}\|^2 \text{ avec } \beta = a_1a_2.$$

■

Le théorème 2.19 peut aussi se démontrer comme corollaire du théorème 2.20. En effet, sous les hypothèses du théorème 2.19 (il est même suffisant de supposer g de classe C^2 au lieu de C^3), on peut démontrer qu'il existe $a, a_1, a_2 \in \mathbb{R}_+^*$ tels que

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

et donc appliquer le théorème 2.20, voir exercice 102 page 178.

Remarque 2.21 (Choix de l'itéré initial)**.** *On ne sait pas bien estimer b dans le théorème 2.19, et ceci peut poser problème lors de l'implantation numérique : il faut choisir l'itéré initial $\boldsymbol{x}^{(0)}$ "suffisamment proche" de $\overline{\boldsymbol{x}}$ pour avoir convergence.*

2.3.2 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 $\boldsymbol{x}^{(0)}$ n'a pas été choisi "suffisamment proche" de $\overline{\boldsymbol{x}}$, 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 $Dg(\boldsymbol{x}^{(k)})$ à chaque itération, et d'autre part la factoriser pour résoudre le système linéaire $Dg(\boldsymbol{x}^{(k)})(\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}) = -g(\boldsymbol{x}^{(k)})$. (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 à 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 $\bar{\boldsymbol{x}} \in \mathbb{R}$ tels que $g(\bar{\boldsymbol{x}}) = 0$. On cherche à calculer $\bar{\boldsymbol{x}}$. Si on le fait par la méthode de Newton, l'algorithme s'écrit :

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

La méthode du "Faux quasi-Newton" (parfois appelée quasi-Newton) consiste à remplacer le calcul de la matrice jacobienne $Dg(\boldsymbol{x}^{(k)})$ 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 $(\boldsymbol{x}^{(k)})_{n\in\mathbb{N}}$ de la manière suivante :

$$\begin{cases} \boldsymbol{x}^{(0)} \in \mathbb{R}^n \\ Dg(\boldsymbol{x}^{(n_i)})(\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}) = -g(\boldsymbol{x}^{(k)}) \text{ si } n_i \leq k < n_{i+1}. \end{cases} \tag{2.26}$$

Avec cette méthode, on a moins de calculs et de factorisations de la matrice jacobienne $Dg(\boldsymbol{x})$ 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 :

$$g(\boldsymbol{x}) = A\boldsymbol{x} + F_1(\boldsymbol{x}) + F_2(\boldsymbol{x}), \text{ avec } A \in \mathcal{M}_n(\mathbb{R}) \text{ avec } F_1, F_2 \in C^1(\mathbb{R}^n, \mathbb{R}^n).$$

L'algorithme de Newton (2.20) s'écrit alors :

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

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

$$\left\{ \begin{array}{l} \boldsymbol{x}^{(0)} \in \mathbb{R}^n \\ (A + DF_1(\boldsymbol{x}^{(k)}))(\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}) = -A\boldsymbol{x}^{(k)} - F_1(\boldsymbol{x}^{(k)}) - F_2(\boldsymbol{x}^{(k)}). \end{array} \right. \tag{2.27}$$

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 $DF_2(\boldsymbol{x})$, 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 $A + DF_1(\boldsymbol{x}^{(k)})$ soit "meilleure" que celle de la matrice $A + DF_1(\boldsymbol{x}^{(k)}) + DF_2(\boldsymbol{x}^{(k)})$; 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 $A + DF_1(\boldsymbol{x}^{(k)})$ 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 $f(x, y) = (a, b)$ où $(a, b) \in \mathbb{R}^2$ 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 2.22. *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*

$$A(\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}) = -A\boldsymbol{x}^{(k)} - F_2(\boldsymbol{x}^{(k)}),$$

En supposant A inversible, on a alors $\boldsymbol{x}^{(k+1)} = -A^{-1}F_2(\boldsymbol{x}^{(k)})$. 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 $\overline{\boldsymbol{x}} \in \mathbb{R}$ tel que $g(\overline{\boldsymbol{x}}) = 0$ s'écrit :

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

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

$$\left\{ \begin{array}{l} x^{(0)}, x^{(1)} \in \mathbb{R} \\ \dfrac{g(x^{(k)}) - g(x^{(k-1)})}{x^{(k)} - x^{(k-1)}}(x^{(k+1)} - x^{(k)}) = -g(x^{(k)}) \quad n \geq 1. \end{array} \right. \tag{2.28}$$

Remarquons que dans la méthode de la sécante, $x^{(k+1)}$ dépend de $x^{(k)}$ et de $x^{(k-1)}$: 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 108 page 181) que si $g(\bar{x}) = 0$ et $g'(\bar{x}) \neq 0$, il existe $\alpha > 0$ tel que si $x^{(0)}, x^{(1)} \in [\bar{x} - \alpha.\bar{x} + \alpha] = I_\alpha$, $x^{(0)} \neq x^{(1)}$, la suite $(x^{(k)})_{n\in\mathbb{N}}$ construite par la méthode de la sécante (2.28) est bien définie, que $(x^{(k)})_{n\in\mathbb{N}} \subset I_\alpha$ et que $x^{(k)} \to \bar{x}$ quand $n \to +\infty$. De plus, la convergence est super linéaire, *i.e.* si $x^{(k)} \neq \bar{x}$ pour tout $n \in \mathbb{N}$, alors $\dfrac{x^{(k+1)} - \bar{x}}{x^{(k)} - \bar{x}} \to 0$ quand $n \to +\infty$. On peut même montrer (voir exercice 108 page 181) 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 $Dg(x^{(k)})$ dans la méthode de Newton (2.20), on va remplacer $Dg(x^{(k)})$ par $B^{(k)} \in \mathcal{M}_n(\mathbb{R})$ "proche de $Dg(x^{(k)})$". En s'inspirant de la méthode de la sécante en dimension 1, on cherche une matrice $B^{(k)}$ qui, $x^{(k)}$ et $x^{(k-1)}$ étant connus (et différents), vérifie la condition :

$$B^{(k)}(x^{(k)} - x^{(k-1)}) = g(x^{(k)}) - g(x^{(k-1)}) \tag{2.29}$$

Dans le cas où $n = 1$, cette condition détermine entièrement $B^{(k)}$;car on peut écrire : $B^{(k)} = \dfrac{g(x^{(k)}) - g(x^{(k-1)}}{x^{(k)} - x^{(k-1)}}$.
Si $n > 1$, la condition (2.29) ne permet pas de déterminer complètement $B^{(k)}$. Il y a plusieurs façons possibles de choisir $B^{(k)}$, 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[6]. Celle-ci consiste à choisir $B^{(k)}$ de la manière suivante : à $x^{(k)}$ et $x^{(k-1)}$ connus, on pose $\delta^{(k)} = x^{(k)} - x^{(k-1)}$ et $y^{(k)} = g(x^{(k)}) - g(x^{(k-1)})$; on suppose $B^{(k-1)} \in \mathcal{M}_n(\mathbb{R})$ connue (et $\delta^{(k)} \neq 0$), et on cherche $B^{(k)} \in \mathcal{M}_n(\mathbb{R})$ telle que

$$B^{(k)}\delta^{(k)} = y^{(k)} \tag{2.30}$$

(c'est la condition (2.29), qui ne suffit pas à déterminer $B^{(k)}$ de manière unique) et qui vérifie également :

$$B^{(k)}\xi = B^{(k-1)}\xi, \quad \forall \xi \in \mathbb{R}^n \text{ tel que } \xi \perp \delta^{(k)}. \tag{2.31}$$

Proposition 2.23 (Existence et unicité de la matrice de Broyden)**.**
Soient $y^{(k)} \in \mathbb{R}^n$, $\delta^{(k)} \in \mathbb{R}^n$, $\delta^{(k)} \neq 0$, et $B^{(k-1)} \in \mathcal{M}_n(\mathbb{R})$. Il existe une unique matrice $B^{(k)} \in \mathcal{M}_n(\mathbb{R})$ vérifiant (2.30) *et* (2.31) *; la matrice $B^{(k)}$ s'exprime en fonction de $y^{(k)}$, $\delta^{(k)}$ et $B^{(k-1)}$ de la manière suivante :*

$$B^{(k)} = B^{(k-1)} + \frac{y^{(k)} - B^{(k-1)}\delta^{(k)}}{\delta^{(k)} \cdot \delta^{(k)}}(\delta^{(k)})^t. \tag{2.32}$$

DÉMONSTRATION – L'espace des vecteurs orthogonaux à $\delta^{(k)}$ est de dimension $n-1$. Soit $(\gamma_1, \ldots, \gamma_{n-1})$ une base de cet espace, alors $(\gamma_1, \ldots, \gamma_{n-1}, \delta^{(k)})$ est une base de $\mathbb{R}^n$ et si $B^{(k)}$ vérifie (2.30) et (2.31), les valeurs prises par l'application linéaire associée à $B^{(k)}$ sur chaque vecteur de base sont connues, ce qui détermine l'application linéaire et donc la matrice $B^{(k)}$ de manière unique. Soit $B^{(k)}$ définie par (2.32), on a :

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

et donc $B^{(k)}$ vérifie (2.30). Soit $\xi \in \mathbb{R}^n$ tel que $\xi \perp \delta^{(k)}$, alors $\xi \cdot \delta^{(k)} = (\delta^{(k)})^t\xi = 0$ et donc

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

■

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

L'algorithme de Broyden s'écrit donc :

$$\begin{cases} \text{Initialisation : } x^{(0)},\ x^{(1)} \in \mathbb{R}^n,\ x^{(0)}, \neq x^{(1)},\ B^{(0)} \in \mathcal{M}_n(\mathbb{R}) \\ \text{Itération k : } x^{(k)},\ x^{(k-1)} \text{ et } B^{(k-1)} \text{ connus, on pose} \\ \qquad \delta^{(k)} = x^{(k)} - x^{(k-1)} \text{ et } y^{(k)} = g(x^{(k)}) - g(x^{(k-1)}); \\ \text{Calcul de } B^{(k)} = B^{(k-1)} + \frac{y^{(k)} - B^{(k-1)}\delta^{(k)}}{\delta^{(k)}\cdot\delta^{(k)}}(\delta^{(k)})^t, \\ \text{résolution de } B^{(k)}(x^{(k+1)} - x^{(k)}) = -g(x^{(k)}). \end{cases}$$

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 .

2.3.3 Exercices (méthode de Newton)

Exercice 86 (Newton et logarithme)**.** *Suggestions en page 182 Corrigé en page 183*
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 87 (Newton pour un système linéaire)**.** *Corrigé en page 183* Soit f l'application définie sur $\mathbb{R}^n$ par $f(x) = Ax - b$ où A est une matrice inversible et $b \in \mathbb{R}^n$. Ecrire l'algorithme de Newton pour la résolution de l'équation $f(x) = 0$ et montrer qu'il converge pour toute condition initiale $x^0 \in \mathbb{R}^n$.

Exercice 88 (Condition initiale et Newton)**.** *Corrigé en page 184* L'algorithme de Newton pour $F(x, y) = (\sin(x) + y, xy)^t$ est-il bien défini pour la condition initiale $(\frac{\pi}{2}, 0)$?

Exercice 89 (Newton dans $\mathbb{R}$ et $\mathbb{R}^2$)**.** *Corrigé en page 184* Soit $a \in \mathbb{R}$ tel que $|a| < 1$ et $(x_0, y_0) \in \mathbb{R}^2$. On définit l'application

$$F : \mathbb{R}^2 \to \mathbb{R}^2$$
$$\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} x - x_0 - ay \\ y - y_0 - a\sin x \end{bmatrix}$$

.

1. Montrer qu'il existe une fonction $f : \mathbb{R} \to \mathbb{R}$, que l'on déterminera, telle que $F(x, y) = (0, 0)$ si et seulement si $x = x_0 + ay$ et $f(y) = 0$.
2. Montrer que pour tout triplet (a, x_0, y_0), il existe un unique couple $(\bar{x}, \bar{y}) \in \mathbb{R}^2$ tel que $F(\bar{x}, \bar{y}) = (0, 0)$.
3. Ecrire l'algorithme de Newton pour f et montrer que l'algorithme de Newton converge au voisinage de $\bar{y}$.
4. Ecrire l'algorithme de Newton pour la fonction F. Montrer que l'algorithme converge au voisinage de $(\bar{x}, \bar{y})$.

Exercice 90 (Méthode de Newton pour un système 2×2)**.** *Corrigé en page 184*

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

$$-5x + 2\sin x + 2\cos y = 0, \tag{2.33}$$
$$2\cos x + 2\sin y - 5y = 0. \tag{2.34}$$

et montrer que la suite définie par cet algorithme est toujours bien définie.

2. Soit $(\overline{x}, \overline{y})$ une solution du problème (2.33)-(2.34). 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 $+\infty$.

3. Montrer qu'il existe au moins une solution $(\overline{x}, \overline{y})$ au problème (2.33)-(2.34).

Exercice 91 (Méthode de Newton pour un autre système 2×2)**.**

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

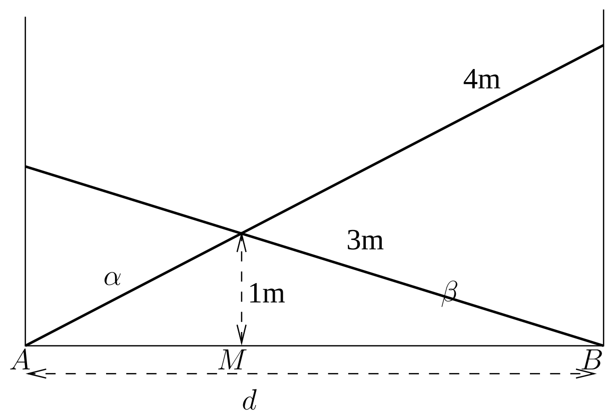
$$x^2 + 2xy = 0, \tag{2.35}$$
$$xy + 1 = 0. \tag{2.36}$$

2. Calculer les solutions de ce système.

3. Soit $(\overline{x}, \overline{y})$ une solution du problème (2.35)-(2.36). 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 92 (Newton et les échelles...). *Corrigé en page 2.3.3 page 185*

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

$$16x^2 = (x^2 + 1)(x + y)^2 \tag{2.37}$$
$$9y^2 = (y^2 + 1)(x + y)^2. \tag{2.38}$$

2. Ecrire l'algorithme de Newton pour la résolution du système (2.37)-(2.38).

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

Exercice 93 (Newton dans $\mathcal{M}_2(\mathbb{R})$). *Corrigé en page 186*

On considère l'application $f : \mathcal{M}_2(\mathbb{R}) \to \mathcal{M}_2(\mathbb{R})$ définie par $f(X) = X^2 - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. L'objectif de cet exercice est de trouver les solutions de $f(X) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

1. Réécrire l'application f comme une application F de $\mathbb{R}^4$ dans $\mathbb{R}^4$.
2. Trouver l'ensemble des solutions de $f(X) = 0$.
3. Ecrire le premier itéré X_1 de l'algorithme de Newton pour l'application f partant de la donnée initiale $X_0 = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$ (On pourra passer par l'application F). Montrer que la suite $(X_k)_k$ définie par cet algorithme est définie par tout k et que l'on peut écrire sous la forme $X_k = \lambda_k \mathrm{Id}$ où $(\lambda_k)_k$ est une suite réelle dont on étudiera la convergence.
4. L'algorithme de Newton converge-t-il au voisinage de $X_* = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$?

Exercice 94 (Recherche d'un point fixe)**.**
On définit la fonction f de $\mathbb{R}$ dans $\mathbb{R}$ par $f(x) = e^{(x^2)} - 4x^2$.

1. Montrer que f s'annule en 4 points de $\mathbb{R}$ et qu'un seul de ces points est entre 0 et 1.
2. On pose $g(x) = (1/2)\sqrt{e^{(x^2)}}$ (pour x dans $\mathbb{R}$).
   Montrer que la méthode du point fixe appliquée à g, initialisée avec un point de l'intervalle $]0, 1[$, est convergente et converge vers le point de $]0, 1[$ annulant f.
   Quel est l'ordre de convergence de cette méthode ?
3. Donner la méthode de Newton pour rechercher les points annulant f.
   Entre cette méthode et la méthode de la question précédente, quelle méthode vous semble *a priori* la plus efficace ?

Exercice 95 (Nombre d'itérations fini pour Newton)**.** *Corrigé détaillé en page 187*

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 $(x^{(k)})_{n\in\mathbb{N}}$ 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 $(x^{(k)})_{n\in\mathbb{N}}$ est bien définie.

1.2 Montrer que si $x^{(0)} \neq 0$, alors $x^{(k+1)} \neq x^{(k)}$ 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 $(x^{(k)})_{n\in\mathbb{N}}$ converge lorsque n tend vers l'infini et donner sa limite.

2. Soit $F : \mathbb{R}^n \to \mathbb{R}$ 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 96 (Méthode de Newton pour un système 2×2)**.** *Corrigé en page 187*

1. On considère l'application $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^2$. Ecrire la méthode de Newton pour calculer la solution de $f(x) = 0$ et montrer qu'elle converge quel que soit le choix initial $x_0 \in \mathbb{R}$.
2. On considère maintenant l'application $F : \mathbb{R}^2 \to \mathbb{R}^2$ définie par
$$F(x, y) = \begin{bmatrix} x^2 - y \\ y^2 \end{bmatrix}$$
   (a) Déterminer l'ensemble des solutions de $F(x, y) = (0, 0)$.
   (b) Ecrire l'algorithme de Newton pour la résolution, et montrer que l'algorithme est bien défini pour tous les couples (x_0, y_0) tels que $x_0 \neq 0$ et $y_0 > 0$.
   (c) Soit $(x_0, y_0) = (1, 1)$. On note $(x_k, y_k), k \in \mathbb{N}$ les itérés de Newton.
      i. Expliciter y_k et en déduire que la suite $(y_k)_{k\in\mathbb{N}}$ converge.
      ii. Montrer que $x_k \geq 2^{-\frac{k}{2}}$ pour tout $k \in \mathbb{N}$ et en déduire que $x_{k+1} \leq x_k$ pour tout $k \in \mathbb{N}$.
      iii. En déduire que la méthode de Newton converge vers une solution de $F(x, y) = (0, 0)$.

Exercice 97 (Méthode de Newton pour le calcul de l'inverse)**.** *Corrigé en page 188*

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 :
$$\begin{cases} x^{(0)} \text{ donné,} \\ x^{(k+1)} = x^{(k)}(2 - ax^{(k)}). \end{cases} \tag{2.39}$$

(b) Montrer que la suite $(x^{(k)})_{n\in\mathbb{N}}$ définie par (2.39) vérifie

$$\lim_{n\to+\infty} x^{(k)} = \begin{cases} \dfrac{1}{a} \text{ si } x^{(0)} \in]0, \dfrac{2}{a}[, \\ -\infty \text{ si } x^{(0)} \in]-\infty, 0[\cup]\dfrac{2}{a}, +\infty[\end{cases}$$

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

$$\text{Id} - AB^{(k+1)} = (\text{Id} - AB^{(k)})^2.$$

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

Exercice 98 (Méthode de Newton pour le calcul de la racine)**.** *Suggestions en page 182, corrigé en page 189*

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 $A = P\text{diag}(\lambda_1, \cdots, \lambda_n)P^{-1}$.
      (i) Montrer que pour tout $k \in \mathbb{N}$, $X^{(k)}$ est bien définie et est donné par

$$X^{(k)} = P\text{diag}(\mu_1^{(k)}, \cdots, \mu_n^{(k)})P^{-1},$$

      où $\mu_i^{(k)}$ est le $k^{ième}$ terme de la suite de Newton pour la résolution de $f_{\lambda_i}(x) = (x - \lambda_i)^2 = 0$ avec comme choix initial $\mu_i^{(0)} = 1$.

(ii) En déduire que la suite $X^{(k)}$ converge vers B quand $k \to +\infty$.

Exercice 99 (Valeurs propres et méthode de Newton)**.**
Suggestions en page 182, corrigé détaillé en page 193

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 :

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

Montrer que la méthode est localement convergente.

Exercice 100 (Modification de la méthode de Newton)**.** *Suggestions en page 182. Corrigé en page 194*

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 $k \geq 0$,

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

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

1. Montrer que la suite $(x^{(k)})_{k \in \mathbb{N}}$ 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 101 (Méthode de Newton pour un système semi-linéaire)**.** *Suggestions en page 183. Corrigé en page 200*
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 $u_1, \ldots, u_n$) :

$$\begin{aligned} &(Au)_i + \alpha_i f(u_i) = b_i \ \forall i \in \{1, \ldots, n\}, \\ &u = (u_1, \ldots, u_n)^t \in \mathbb{R}^n, \end{aligned} \tag{2.40}$$

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 (2.40) admet au moins une solution (ceci peut être démontré mais est difficile).

1. Montrer que (2.40) admet une unique solution.
2. Soit u la solution de (2.40). Montrer qu'il existe $a > 0$ t.q. la méthode de Newton pour approcher la solution de (2.40) converge lorsque le point de départ de la méthode, noté $u^{(0)}$, vérifie $|u - u^{(0)}| < a$.

Exercice 102 (Autre démonstration de la convergence locale de Newton)**.** *Suggestions en page 182. Corrigé en page 192* On se place sous les sous les hypothèses du théorème 2.19 avec g de classe C^2 au lieu de C^3. Montrer qu'il existe $a, a_1, a_2 \in \mathbb{R}_+^*$ tels que

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

et qu'on peut donc appliquer le théorème 2.20 pour obtenir le résultat de convergence locale du théorème 2.19.

Exercice 103 (Convergence de la méthode de Newton si $f'(\overline{x}) = 0$)**.** *Suggestions en page 183, corrigé détaillé en page 195*

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 104 (Point fixe et Newton)**.** *Corrigé en page 196*

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 :

$$\begin{cases} x_0 \in \mathbb{R}, \\ x_{n+1} = h(x_n), n \geq 0. \end{cases} \tag{2.41}$$

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 (2.41) 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 (2.41) 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 105 (Variante de la méthode de Newton)**.**
Corrigé détaillé en page 198

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) $|f(x_0)| \leq \frac{c}{2\lambda}$,
(iii) $|f'(x) - f'(y)| \leq \frac{1}{2\lambda}$, $\forall (x, y) \in I^2$
(iv) $|f'(x)| \geq \frac{1}{\lambda} \forall x \in I$.

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

$$\begin{aligned} x^{(0)} &= x_0, \\ x^{(k+1)} &= x^{(k)} - \frac{f(x^{(k)})}{f'(y)}, \end{aligned} \tag{2.42}$$

où $y \in I$ est choisi arbitrairement.

1. Montrer par récurrence que la suite définie par (2.42) satisfait $x^{(k)} \in I$ pour tout $n \in \mathbb{N}$.
   (On pourra remarquer que si $x^{(k+1)}$ est donné par (2.42) *alors $x^{(k+1)} - x_0 = x^{(k)} - x_0 - \frac{f(x^{(k)})-f(x_0)}{f'(y)} - \frac{f(x_0)}{f'(y)}$.)*
2. Montrer que la suite $(x^{(k)})_{n\in\mathbb{N}}$ définie par (2.42) vérifie $|x^{(k)} - \bar{x}| \leq \frac{c}{2^n}$ et qu'elle converge vers $\bar{x}$ de manière au moins linéaire.
3. On remplace l'algorithme (2.42) par
$$\begin{aligned} &x^{(0)} = x_0, \\ &x^{(k+1)} = x^{(k)} - \frac{f(x^{(k)})}{f'(y^{(k)})}, \end{aligned} \tag{2.43}$$
   où la suite $(y^{(k)})_{n\in\mathbb{N}}$ est une suite donnée d'éléments de I. Montrer que la suite $(x^{(k)})_{n\in\mathbb{N}}$ 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 $n \geq 1$ et que $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$. La méthode définie par (2.42) ou (2.43) peut-elle se généraliser, avec d'éventuelles modifications des hypothèses, à la dimension n ?

Exercice 106 (Méthode de Steffensen)**.**
Suggestions en page 183, corrigé détaillé en page 201

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 $f(x^{(k)} + f(x^{(k)})) \neq f(x^{(k)})$,
$$x^{(k+1)} = x^{(k)} - \frac{(f(x^{(k)}))^2}{f(x^{(k)} + f(x^{(k)})) - f(x^{(k)})}, \tag{2.44}$$
et si $f(x^{(k)} + f(x^{(k)})) = f(x^{(k)})$, $x^{(k+1)} = x^{(k)}$.

1. Montrer qu'il existe $\alpha > 0$ tel que si $x^{(k)} \in B(\overline{x}, \alpha)$, alors $f(x^{(k)} + f(x^{(k)})) \neq f(x^{(k)})$ si $x^{(k)} \neq \overline{x}$. En déduire que si $x_0 \in B(\overline{x}, \alpha)$, alors toute la suite $(x^{(k)})_{n\in\mathbb{N}}$ vérifie (2.44) pourvu que $x^{(k)} \neq \overline{x}$ 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
$$x^{(k+1)} - \overline{x} = a(x^{(k)})(x^{(k)} - \overline{x}), \text{ pour tout } n \in \mathbb{N} \text{ tel que } x^{(k)} \neq \overline{x}. \tag{2.45}$$
3. Montrer que la méthode est localement convergente en $\overline{x}$ et la convergence est au moins d'ordre 2.

Exercice 107 (Méthode de Newton-Tchebycheff)**.** *Corrigé en page 2.3.3 page 203*

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 :

$$x_{n+1} = x_n - \frac{g(x_n)}{g'(x_n)} - \frac{g''(x_n)[g(x_n)]^2}{2[g'(x_n)]^3}. \tag{2.46}$$

Montrer que la méthode (2.46) 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 108 (Méthode de la sécante)**.** *Corrigé en page 2.3.3 page 204*
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 $(M_n)_{n\geq 1}$ lorsque $n \to \infty$. En déduire que la suite $(M_n)_{n\geq 1}$ 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 $(a_n)_{n\geq 1}$ 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.

Exercice 109 (Algorithme de Newton pour calculer une racine cubique)**.**
Soient $n \geq 1$ et $A \in \mathcal{M}_n(\mathbb{R})$ une matrice symétrique définie positive. On note $\{e_1, \ldots, e_n\}$ une base orthonormée de $\mathbb{R}^n$ formée de vecteurs propres de A. On a donc $Ae_i = \lambda_i e_i$, avec $\lambda_i > 0$, $e_i \cdot e_j = 0$ pour $i \neq j$ et $e_i \cdot e_i = 1$. Pour $B \in \mathcal{M}_n(\mathbb{R})$, on pose $F(B) = B^3 - A$.

1. Montrer que la matrice X définie par $Xe_i = \mu_i e_i$ pour $i = 1, \ldots, n$, avec $\mu_i^3 = \lambda_i$, est une racine cubique de A (c'est-à-dire une solution de $F(X) = 0$). Montrer que X est symétrique.

Dans la suite de l'exercice, on s'intéresse à l'algorithme de Newton pour calculer X.

2. Soit $B \in \mathcal{M}_n(\mathbb{R})$. La différentielle de F au point B, notée $DF(B)$, est donc une application linéaire de $\mathcal{M}_n(\mathbb{R})$ dans $\mathcal{M}_n(\mathbb{R})$. Donner, pour tout $H \in \mathcal{M}_n(\mathbb{R})$, l'expression de $DF(B)H$.
3. Montrer que $DF(X)$ est une application inversible.
   [On pourra, par exemple, calculer $DF(X)He_i \cdot e_j$ et montrer que $DF(X)H = 0$ implique $H = 0$.]
4. Donner l'algorithme de Newton pour calculer X et montrer que l'algorithme de Newton donne une suite convergente vers X si l'initialisation de l'algorithme est faite avec une matrice suffisamment proche de X.

Suggestions

Exercice 86 page 174 (Newton et logarithme) Etudier les variations de la fonction φ définie par : $\varphi(x) = x - x\ln x$.

Exercice 98 page 177 (Méthode de Newton pour le calcul de la racine) 1.1 (ii) Montrer que la suite $(x^{(k)})_{k\geq 1}$ est décroissante.
4. Ecrire la définition de la différentiel en faisant attention que le produit matriciel n'est pas commutatif.
5. Ecrire l'algorithme de Newton dans la base des vecteurs propres associés aux valeurs propres de A.

Exercice 102 page 178 (Autre démonstration de la convergence locale de Newton)
Soit $S = Dg(\bar{x})^{-1}(Dg(x) - Dg(\bar{x}))$. Remarquer que

$$Dg(\boldsymbol{x}) = Dg(\bar{\boldsymbol{x}}) - Dg(\bar{\boldsymbol{x}}) + Dg(\boldsymbol{x}) = Dg(\bar{\boldsymbol{x}})(Id + S)$$

et démontrer que $\|S\| < 1$.
En déduire que $Dg(x) = Dg(\bar{x})(Id + S)$ est inversible, et montrer alors 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 $\|Dg(x)^{-1}\| \leq a_1$.
Pour montrer l'existence de a_2, introduire la fonction $\varphi \in C^1(\mathbb{R}, \mathbb{R}^n)$ définie par

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

Exercice 99 page 178 (Valeurs propres et méthode de Newton) Écrire 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 100 page 178 (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 103 page 179 (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 101 page 178 (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 106 page 180 (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.

Corrigés des exercices

Corrigé de l'exercice 86 page 174 (Newton et logarithme) La méthode de Newton pour résoudre $\ln(x) = 0$ s'écrit :

$$x^{(k+1)} - x^{(k)} = -x^{(k)}\ln(x^{(k)}).$$

Pour que la suite $(x^{(k)})_{k\in\mathbb{N}}$ soit bien définie, il faut que $x^{(0)} > 0$. Montrons maintenant que :

1. si $x^{(0)} > e$, alors $x^{(1)} < 0$,
2. si $x^{(0)} \in]1, e[$, alors $x^{(1)} \in]0, 1[$,
3. si $x^{(0)} = 1$, alors $x^{(k)} = 1$ pour tout k,
4. si $x^{(0)} \in]0, 1[$ alors la suite $(x^{(k)})_{k\in\mathbb{N}}$ est strictement croissante et majorée par 1.

Le plus simple pour montrer ces propriétés est d'étudier la fonction φ définie par : $\varphi(x) = x - x\ln x$, dont la dérivée est : $\varphi'(x) = -\ln x$. Le tableau de variation de φ est donc :

	0		1		e		$+\infty$
$\varphi'(x)$		$+$	0	$-$	e	$-$	
			1				
$\varphi(x)$		$\nearrow$		$\searrow$		$\searrow$	
	0	$+$		$+$	0	$-$	

Grâce à ce tableau de variation, on a immédiatement les propriétés 1. à 4. La convergence de la suite $(x^{(k)})_{k\in\mathbb{N}}$ vers 1 peut s'obtenir en passant à la limite dans le schéma.

Corrigé de l'exercice 87 page 174 (Méthode de Newton pour un système linéaire) Pour tout $x \in \mathbb{R}^n$, la matrice $Df(x)$ est inversible ; donc l'algorithme de Newton est bien défini et s'écrit $Df(x)x^{(1)} = b$. Il converge donc en une itération, qui demande la résolution du sysème linéaire $Ax = b$...

Corrigé de l'exercice 88 page 174 (Condition initiale et Newton) On vérifie que

$$DF(x,y) = \begin{bmatrix} \cos(x) & 1 \\ y & x \end{bmatrix}$$

et par conséquent

$$DF\left(\frac{\pi}{2}, 0\right) = \begin{bmatrix} 0 & 1 \\ 0 & \frac{\pi}{2} \end{bmatrix}$$

n'est pas inversible. La matrice $DF(x,y)$ est inversible pour $x = 0$ $y = 1$ par exemple.

Corrigé de l'exercice 89 page 174 (Méthode de Newton pour un système 2×2)

1. On cherche (x,y) tel que $x = x_0 + ay$ et $y = y_0 + a\sin(x) = y_0 + \sin(x_0 + ay)$. Soit encore $f(y) = 0$ avec $f(y) = y - y_0 - a\sin(x_0 + ay)$. La réciproque est immédiate.
2. La fonction f est continue et telle que $\lim_{-\infty} f = -\infty$ et $\lim_{+\infty} f = +\infty$. Par conséquent, comme f est continue, le théorème des valeurs intermédiaires nous assure que f est surjective. De plus, on vérifie que si $|a| < 1$, alors $f'(y) > 0$ et donc f est strictement croissante. Elle est donc injective. Ceci assure l'existence et l'unicité d'un unique $\bar{y}$ tel que $f(\bar{y}) = 0$. Par suite, il existe un unique couple $(\bar{x} = x_0 + a\bar{y}, \bar{y})$ qui vérifie $F(\bar{x}, \bar{y}) = (0, 0)$.
3. On a
$$y^{(k+1)} = y^{(k)} - \frac{f(y^{(k)})}{f'(y^{(k)})} = h(y^{(k)}) \text{ avec } h(y) = y - \frac{y - y_0 - a\sin(x_0 + ay)}{1 - a^2\cos(x_0 + ay)}$$
Pour la convergence, on applique le théorème du cours qui assure si $f'(\bar{y}) \neq 0$ l'existence d'un voisinage de $\bar{y}$ sur lequel l'algorithme de Newton converge.
4. On a
$$\begin{aligned}
\begin{bmatrix} x^{(k+1)} \\ y^{(k+1)} \end{bmatrix} &= \begin{bmatrix} x^{(k)} \\ y^{(k)} \end{bmatrix} - \begin{bmatrix} 1 & -a \\ -a\cos(x^{(k)} & 1 \end{bmatrix}^{-1} \begin{bmatrix} x^{(k)} - x_0 - ay^{(k)} \\ y^{(k)} - y_0 - a\sin(x^{(k)}) \end{bmatrix} \\
&= \begin{bmatrix} x^{(k)} \\ y^{(k)} \end{bmatrix} - \frac{1}{1 - a^2\cos(x^{(k)})} \begin{bmatrix} 1 & a \\ a\cos(x^{(k)} & 1 \end{bmatrix} \begin{bmatrix} x^{(k)} - x_0 - ay^{(k)} \\ y^{(k)} - y_0 - a\sin(x^{(k)}) \end{bmatrix} \\
&= \begin{bmatrix} x^{(k)} \\ y^{(k)} \end{bmatrix} - \frac{1}{1 - a^2\cos(x^{(k)})} \begin{bmatrix} x^{(k)} - x_0 - ay_0 - a^2\sin(x^{(k)}) \\ a\cos(x^{(k)})(x^{(k)} - x_0 - ay^{(k)}) + y^{(k)} - y_0 - a\sin(x^{(k)}) \end{bmatrix}
\end{aligned}$$
Pour la convergence, on applique le théorème du cours qui assure si $DF(\bar{x}, \bar{y})$ est inversible, l'existence d'un voisinage de $(\bar{x}, \bar{y})$ sur lequel l'algorithme de Newton converge.

Corrigé de l'exercice 90 page 174 (Méthode de Newton pour un système 2×2) 1. Soit F l'application de $\mathbb{R}^2$ dans $\mathbb{R}^2$ définie par $F(x,y) = (-5x + 2\sin x + 2\cos y, 2\cos x + 2\sin y - 5y)$. Calculons la matrice jacobienne de F au point (x,y) :

$$DF(x,y) = \begin{bmatrix} -5 + 2\cos x & -2\sin y \\ -2\sin x & 2\cos y - 5 \end{bmatrix}$$

$Det(DF)(x,y) = (-5 + 2\cos x)(-5 + 2\cos y) - 4\sin x \sin y = 25 + 4\cos x\cos y - 10\cos x - 10\cos y - 4\sin x\sin y$.

• Si $\cos x \cos y \geq 0,$ alors $Det(DF)(x,y) \geq 1 > 0,$

• Sinon, alors $\cos x > 0$ et $\cos y < 0$ (ou $\cos y > 0$ et $\cos x < 0$), et $Det(DF)(x,y) \geq 25 - 4 - 10\cos x - 4\sin x \sin y > 0$.

On a donc $Det(DF)(x,y) > 0$ pour tout $(x,y) \in \mathbb{R}^2$, ce qui prouve que la Jacobienne $DF(x,y)$ est toujours inversible. On peut donc toujours écrire l'alorithme de Newton :

$$DF(x^{(k)}, y^{(k)}) \begin{bmatrix} x^{(k+1)} - x^{(k)} \\ y^{(k+1)} - y^{(k)} \end{bmatrix} = -F(x^{(k)}, y^{(k)})$$

2. Il est facile de voir que la fonction F est deux fois continûment différentiable. Il suffit alors d'appliquer le théorème de convergence du cours.

3. Soit f la fonction définie par $f(x) = -5x + 2\sin x$. Elle est strictement décroissante et c'est une bijection de $\mathbb{R}$ dans $\mathbb{R}$. On note g sa réciproque. La fonction g est donc continue. L'équation (2.33) donne $x = g(-2\cos y)$. On cherche donc y t.q. $h(y) = -5y + 2\sin y + 2\cos(g(-2\cos y)) = 0$. La fonction h est continue et $\lim_{\pm\infty} = \pm\infty$. Donc par le théorème des valeurs intérmédiaires, h s'annule.

Corrigé de l'exercice 91 page 174 (Méthode de Newton pour un système 2×2) En cours de rédaction

Corrigé de l'exercice 92 page 175 (Newton et les échelles...) 1. Notons A et B les deux pieds des murs, P le point de croisement des échelles et M sa projection sur le plan horizontal, comme indiqué sur la figure. Soient $x = d(A, M)$ la distance de A à M, $y = d(B, M)$, $\alpha = d(A, P)$ et $\beta = d(B, P)$.
Par le théorème de Pythagore, $x^2 + 1 = \alpha^2$ et $y^2 + 1 = \beta^2$. Par le théorème de Thalès, $\frac{x}{d} = \frac{\alpha}{4}$ et $\frac{y}{d} = \frac{\beta}{3}$. En éliminant α et β, on en déduit que x et y sont solutions du système non linéaire :

$$\begin{aligned} 16x^2 &= (x^2+1)(x+y)^2 \\ 9y^2 &= (y^2+1)(x+y)^2, \end{aligned}$$

qui est bien le système (2.37)-(2.38).

2. Le système précédent s'écrit sous la forme $F(X) = 0$, où F est la fonction de $\mathbb{R}^2$ dans $\mathbb{R}^2$ définie par

$$F(X) = F\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} (x^2+1)(x+y)^2 - 16x^2 \\ (y^2+1)(x+y)^2 - 9y^2 \end{bmatrix}.$$

On a donc

$$DF(X) = \begin{bmatrix} 4x^3 + 2xy^2 + 6x^2y + 2x + 2y - 32x & 2x^2y + 2x^3 + 2y + 2x \\ 2xy^2 + 2y^3 + 2x + 2y & 4y^3 + 2yx^2 + 6y^2x + 2y + 2x - 18y \end{bmatrix}$$

L'algorithme de Newton pour la résolution du système (2.37)-(2.38) s'écrit donc, pour $X_k = \begin{bmatrix} x_k \\ y_k \end{bmatrix}$ connu,

$$DF(X_k)(X_{k+1} - X_k) = -F(X_k).$$

A l'étape k, on doit donc résoudre le système linéaire

$$[DF(X_k)\begin{bmatrix} s \\ t \end{bmatrix} = -\begin{bmatrix} (x_k^2+1)(x_k+y_k)^2 - 16x_k^2 \\ (y_k^2+1)(x_k+y_k)^2 - 9y_k^2 \end{bmatrix} \tag{2.47}$$

avec

$$DF(X_k) = \begin{bmatrix} 4x_k^3 + 2x_ky_k^2 + 6x_k^2y_k + 2x_k + 2y_k - 32x_k & 2x^2y_k + 2x_k^3 + 2y_k + 2x_k \\ \\ 2x_ky_k^2 + 2y_k^3 + 2x_k + 2y_k & 4y_k^3 + 2y_kx_k^2 + 6y_k^2x + 2y_k + 2x_k - 18y_k \end{bmatrix}$$

et on obtient le nouvel itéré $X_{k+1} = \begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix} = \begin{bmatrix} x_k + s \\ y_k + t \end{bmatrix}$.

3. La première itération nécessite donc la résolution du système linéaire (2.47) pour $k = 0$, soit, pour $x_0 = 1$ et $y_0 = 1$,

$$\begin{bmatrix} -16 & 8 \\ 8 & -2 \end{bmatrix}\begin{bmatrix} s \\ t \end{bmatrix} = \begin{bmatrix} 8 \\ 1 \end{bmatrix}$$

dont la solution est $s = \frac{3}{4}$ et $t = \frac{5}{2}$. On en déduit que $x_1 = 1.75$ et $y_1 = 3.5$ construits par la méthode de Newton en partant de $x^{(0)} = 1$ et $y^{(0)} = 1$. La distance d à la première itération est donc $d = 5.25$.

Corrigé de l'exercice 93 page 175 (Newton dans $\mathcal{M}_{2,2}(\mathbb{R})$)

1. Notons $X = \begin{bmatrix} x & y \\ z & t \end{bmatrix}$, on a alors

$$\begin{aligned} f(X) &= \begin{bmatrix} x^2+yz & y(x+t) \\ (x+t)z & yz+t^2 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} x^2+yz-1 & y(x+t) \\ z(x+t) & yz+t^2-1 \end{bmatrix} \end{aligned}$$

On a donc $F(x,y,z,t) = \begin{bmatrix} x^2+yz-1 \\ y(x+t) \\ z(x+t) \\ yz+t^2-1 \end{bmatrix}$.

2. On voit que soit $y = z = 0$ et $x^2 = t^2 = 1$ soit $x = -t$ et $yz = 1 - t^2$. Autrement dit on a une infinité de solutions.

3. On a

$$DF(x,y,z,t) = \begin{bmatrix} 2x & z & y & 0 \\ y & x+t & 0 & y \\ z & 0 & x+t & z \\ 0 & z & y & 2t \end{bmatrix}.$$

Par conséquent,

$$DF(4,0,0,4) = \begin{bmatrix} 8 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 8 \end{bmatrix}.$$

On en déduit

$$\begin{aligned} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \\ t_1 \end{bmatrix} &= \begin{bmatrix} 4 \\ 0 \\ 0 \\ 4 \end{bmatrix} - \begin{bmatrix} 8 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 8 \end{bmatrix}^{-1} \begin{bmatrix} 15 \\ 0 \\ 0 \\ 15 \end{bmatrix} \\ &= \begin{bmatrix} 4 \\ 0 \\ 0 \\ 4 \end{bmatrix} - \begin{bmatrix} \frac{15}{8} \\ 0 \\ 0 \\ \frac{15}{8} \end{bmatrix} = \begin{bmatrix} \frac{17}{8} \\ 0 \\ 0 \\ \frac{17}{8} \end{bmatrix} \end{aligned}$$

On voit que $X_1 = \frac{17}{8}\mathrm{Id}$. Par récurrence on vérifie que

$$\lambda_{k+1} = \lambda_k - \frac{1}{2\lambda_k}(\lambda_k^2 - 1) = \frac{\lambda_k^2+1}{2\lambda_k} = \phi(\lambda_k).$$

On vérifie que $\phi'(\lambda) = \frac{1}{2}\left(1 - \frac{1}{\lambda^2}\right)$ et donc $\phi' \geq 0$ sur $[1,+\infty[$ avec $\phi(1) = 1$ donc $\phi : [1,+\infty[\to [1,+\infty[$ et $|\phi'_{|[1,\infty[}| \leq \frac{1}{2}$. On en déduit la convergence de la suite λ_k vers 1, qui est l'unique point fixe de ϕ dans cet intervalle.

4. On vérifie que

$$DF(-1,0,0,-1) = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}.$$

est inversible. Le théorème de Newton Raphson assure la convergence locale de l'algorithme.

Corrigé de l'exercice 95 page 176 (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 :

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

est bien définie.

1.2 Par définition de la suite, on a $x^{(k+1)} - x^{(k)} = -\frac{e^{x^{(k)}}-1}{e^{x^{(k)}}} = 0$ ssi $x^{(k)} = 0$. Donc par récurrence sur n, si $x^{(0)} \neq 0$, on a $x^{(k+1)} \neq x^{(k)}$ 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 $(x^{(k)})_{n \in \mathbb{N}}$ 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 $x^{(k)}$ et $x^{(k+1)}$ deux itérés successifs donnés par la méthode de Newton, tels que $F(x^{(k)}) \neq 0$. On a donc :

$$DF(x^{(k)})(x^{(k+1)} - x^{(k)}) = -F(x^{(k)}), \tag{2.48}$$

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

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

et donc, avec (2.48), $F(x^{(k+1)}) > 0$. On montre ainsi, par récurrence sur n, que si $F(x^{(0)}) \neq 0$, alors $F(x^{(k)}) > 0$ 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 96 page 176 ([Méthode de Newton pour un système 2×2)

1. Lla méthode de Newton pour calculer la solution de $f(x) = 0$ s'écrit : $x_{k+1} = \frac{1}{2} x_k$, et donc x_k converge vers 0 pour tout x_0.
2. (a) L'ensemble des solutions de $F(x, y) = (0, 0)$.est $\{(0, 0)\}$.

   (b) La matrice jacobienne s'écrit : $DF(x, y) = \begin{bmatrix} 2x & -1 \\ 0 & 2y \end{bmatrix}$, et donc l'algorithme de Newton s'écrit :

   $$\begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ y_k \end{bmatrix} + \begin{bmatrix} \delta x \\ \delta y \end{bmatrix} \text{ avec } \begin{bmatrix} 2x_k & -1 \\ 0 & 2y_k \end{bmatrix} \begin{bmatrix} \delta x \\ \delta y \end{bmatrix} = - \begin{bmatrix} x_k^2 - y_k \\ y_k^2 \end{bmatrix}$$

   Cet algorithme est bien défini dès que la matrice jacobienne est inversible, c.à.d dès que $x_k y_k \neq 0$. C'est vrai pour $k = 0$ par hypothèse. Montrons que c'est vrai pour tout $k \in \mathbb{N}$ par récurrence sur k. Supposons qu'on a $x_p \neq 0$ et $y_p > 0$ pour tout $p \leq k$, et montrons que dans ce cas on a encore $x_{k+1} \neq 0$ et $y_{k+1} > 0$. Par définition de l'algorithme de Newton, on a $y_{k+1} = \frac{y_k}{2} > 0$ et $x_{k+1} = \frac{x_k}{2} + \frac{y_k}{4x_k} = \frac{1}{2x_k}(x_k^2 + \frac{y_k}{2}) \neq 0$, car $x_k^2 + \frac{y_k}{2} > 0$.

   (c) i. Comme $y_0 = 1$ et que $y_{k+1} = \frac{y_k}{2}$, on a $y_k = 2^{-k}$

   ii. On écrit que $x_{k+1} = g_k(x_k)$ avec $g_k(x) = \frac{x}{2} + 2^{-k-2}\frac{1}{x}$. L'étude de la fonction g_k montre que $\min g_k = 2^{-\frac{k+1}{2}}$.
   Comme $x_k \geq 2^{-\frac{k}{2}} > 0$ et que $y_k = 2^{-k}$, on en déduit que $x_{k+1} - x_k = \frac{1}{2x_k}(-x_k^2 + \frac{y_k}{2} \leq 0$.

   iii. La suite $(x_k)_{k \in \mathbb{N}}$ est décroissante minorée par 0, et donc elle converge. En passant à la limite sur l'expression de x_{k+1}, on en déduit que la suite x_k converge vers 0. Par l'expression de y_k, on sait qu'elle converge également vers 0. D'où la conclusion.

Corrigé de l'exercice 97 page 176 (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 :

$$\left\{ \begin{array}{l} x^{(0)} \text{ donné}, \\ x^{(k+1)} = x^{(k)}(2 - ax^{(k)}). \end{array} \right. \tag{2.49}$$

   (b) Soit $(x^{(k)})_{n\in\mathbb{N}}$ définie par (2.39). D'après le théorème du cours, on sait que la suite $(x^{(k)})_{n\in\mathbb{N}}$ converge de localement (de manière quadratique) dans un voisinage de $\frac{1}{a}$. On veut déterminer ici l'intervalle de convergence précisément. On a $x^{(k+1)} = \varphi(x^{(k)})$ 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 :

$$\begin{array}{c|ccccc} x & & 0 & \frac{1}{a} & \frac{2}{a} & \\ \hline \varphi'(x) & & + & 0 & - & \\ \hline \varphi(x) & -\infty & \nearrow & \frac{1}{a} & \searrow & -\infty \end{array} \tag{2.50}$$

   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 $x^{(k+1)} \geq x^{(k)}$ pour tout n, et donc la suite $(x^{(k)})_{n\in\mathbb{N}}$ 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 $\lim_{n\to+\infty} x^{(k)} = -\infty$. 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 $GL_n(\mathbb{R})$ 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 $GL_n(\mathbb{R})$ dans $GL_n(\mathbb{R})$. Montrons qu'elle est dérivable. Soit $H \in GL_n(\mathbb{R})$ 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

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

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 :

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

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 :

$$\begin{cases} B^0 \in \ GL_n(\mathbb{R}), \\ -(B^{n+1} - B^n) = (Id - B^nA)B^n. \end{cases} \tag{2.51}$$

soit encore

$$\begin{cases} B^0 \in \ GL_n(\mathbb{R}), \\ B^{n+1} = 2B^n - B^nAB^n. \end{cases} \tag{2.52}$$

(d) Par définition, on a :

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

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 75 (Méthode de Newton pour le calcul de la racine)

1.1 Par définition, l'algorithme de Newton s'écrit :

$$x^{(k+1)} = x^{(k)} - \frac{f_\lambda(x^{(k)})}{f'_\lambda(x^{(k)})} = x^{(k)} - \frac{(x^{(k)})^2 - \lambda}{2x^{(k)}}$$

1.2 (i) Par définition,

$$x^{(1)} - \sqrt{\lambda} = x^{(0)} - \frac{(x^{(0)})^2 - \lambda}{2x^{(0)}} - \sqrt{\lambda} = \frac{(x^{(0)} - \sqrt{\lambda})^2}{2x^{(0)}} \geq 0.$$

car $x^{(0)} > 0$. On a donc bien $x^{(1)} \geq \sqrt{\lambda}$. Supposons alors que $x^{(\ell)} \geq \sqrt{\lambda}$ pour tout $\ell \leq k$; on a alors :

$$x^{(k+1)} - \sqrt{\lambda} = x^{(k)} - \frac{(x^{(k)})^2 - \lambda}{2x^{(k)}} - \sqrt{\lambda} = \frac{(x^{(k)} - \sqrt{\lambda})^2}{2x^{(k)}} \geq 0.$$

On a ainsi montré par récurrence sur k que $x^{(k} \geq \sqrt{\lambda}$ pour tout $k \in \mathbb{N}$.
1.2 (ii) Par définition,

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

Par la question 1, $(x^{(k)})^2 \geq \lambda$, et $x^{(k)} \geq \sqrt{\lambda} \geq 0$. On en déduit que $x^{(k+1)} \leq x^{(k)}$. La suite $(x^{(k)})_{k\in\mathbb{N}}$ est minorée et décroissante, elle est donc convergente. Sa limite ℓ vérifie :

$$0 = -\frac{(\ell)^2 - \lambda}{2\ell} \text{ et } \ell \geq \sqrt{\lambda}$$

(car $x^{(k)} \geq \sqrt{\lambda}$), ce qui entraîne $x^{(k)} = \sqrt{\lambda}$.

2. Comme A est diagonalisable dans $\mathbb{R}$ de valeurs propres strictement positives, il existe une base orthonormée $(\boldsymbol{u}_i)_{i=1,\ldots,n}$ de vecteurs propres, tels que

$$A\boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i \text{ pour tout } i = 1, \ldots, n.$$

Soit $\mu_i = \sqrt{\lambda_i}$, pour $i = 1, \ldots, n$ (notons que les μ_i sont des réels positifs, puisque, par hypothèse, $\lambda_i \in \mathbb{R}$). Soit B la matrice définie par $B\boldsymbol{u}_i = \mu_i \boldsymbol{u}_i$ pour tout $i = 1, \ldots, n$, et P dont les colonnes sont les vecteurs $\boldsymbol{u}_i$, qui est aussi la matrice de passage de la base canonique à la base $(\boldsymbol{u}_i)_{i=1,\ldots,n}$. On a $A = P^{-1}DP$ et $B = P^{-1}\sqrt{D}P$, où D est la matrice diagonale de coefficients λ_i et $\sqrt{D}$ la matrice diagonale de coefficients μ_i. Et on vérifie que

$$B^2 = P^{-1}\sqrt{D}PP^{-1}\sqrt{D}P = P^{-1}DP = A..$$

Dans le cas particulier de la matrice 2×2 de l'énoncé, qui est triangulaire supérieure, les valeurs propres de A sont 1 et 2, et la matrice A vérifie donc bien les hypothèses de l'énoncé. On peut facilement trouver les vecteurs propres en remarquant que

$$\boldsymbol{c}_1(A) = A\boldsymbol{c}_1(A) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ et donc } \boldsymbol{u}_1 = \boldsymbol{c}_1(A) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

On remarque ensuite que

$$\boldsymbol{c}_1(A) + \boldsymbol{c}_2(A) = 2\begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ et donc } \boldsymbol{u}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

(On peut aussi procéder par une identification des coefficients, si on a envie de se fatiguer...) On obtient donc

$$P = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \text{ et } B = \begin{bmatrix} 1 & \sqrt{2}-1 \\ 0 & \sqrt{2} \end{bmatrix}.$$

3. Comme A et B sont supposées symétriques définies positives, elles sont diagonalisables dans $\mathbb{R}$:
soit $(\lambda_i)_{i=1,\ldots,n}$ (pas forcément distinctes, car elles peuvent être de multiplicité[7] supérieure à 1) et $\boldsymbol{u}_i$ les vecteurs propres associés à A. On a $B^2\boldsymbol{u}_i = A\boldsymbol{u}_i = \lambda_i^2\boldsymbol{u}_i$, et donc $(B^2 - \lambda_i^2\mathrm{Id})\boldsymbol{u}_i = 0$, pour tout $i = 1, \ldots, n$. Comme les matrices B et Id commutent, on a donc

$$(B + \lambda_i\mathrm{Id})(B - \lambda_i\mathrm{Id})\boldsymbol{u}_i = 0 \text{ pour tout } i = 1, \ldots, n$$

Or les valeurs $-\lambda_i$ sont strictement négatives et ne peuvent don pas être valeur propre de B, qui est symétrique définie positive. On en conclut que $B - \lambda_i\mathrm{Id})\boldsymbol{u}_i = 0$, et donc

$$B\boldsymbol{u}_i = \lambda_i\boldsymbol{u}_i \text{ pour tout } i = 1, \ldots, n,$$

ce qui détermine entièrement B (car les $\boldsymbol{u}_i$ forment une base). On a ainsi montré que B est déterminée de manière unique.

7. Ici la multiplicité algébrique est égale à la multiplicité géométrique vu que la matrice est diagonalisable.

Les contre exemples à l'unicité sont nombreux si on ne demande pas que B soit symétrique définie positive. Il suffit de considérer par exemple $A = Id$ et de vérifier que $B = Id$ et $B = -Id$ vérifient toutes deux $B^2 = A$.

4. Soit $H \in M_n(\mathbb{R})$; par définition de F,

$$F(X + H) - F(X) = (X + H)^2 - X^2 = XH + HX + H^2.$$

On a donc

$$F(X + H) - F(X) = DF(X)(H) + H^2 \text{ avec } DF(X)(H) = XH + HX.$$

L'application $DF(X)$ ainsi définie est bien une application linéaire de $\mathcal{M}_n(\mathbb{R})$ dans $\mathcal{M}_n(\mathbb{R})$, ce qui prouve que F est différentiable.

Attention, on ne peut pas écrire que $DF(X)(H) = 2XH$, car les matrices H et X ne commutent pas en général. On rappelle que deux matrices diagonalisables commutent si et seulement si elles sont *diagonalisables dans la même base*.

5.1 L'algorithme de Newton s'écrit :

$$\begin{aligned}
&X^{(k)} \text{ connu },\\
&X^{(k+1)} = X^{(k)} + Y\\
&\text{avec } Y \text{ solution de } X^{(k)}Y + YX^{(k)} = -(X^{(k)})^2 + A
\end{aligned}$$

5.2 (i) Notons d'abord que $X^{(0)} = Id$ et s'écrit donc bien $X^{(0)} = P\text{diag}(\mu_1^{(0)}, \cdots, \mu_n^{(0)})P^{-1}$ puisque $\mu_i^{(0)} = 1$ pour tout i. Supposons l'hypothèse de récurrence

(HR) $X^{(k)}$ est une matrice symétrique définie positive de valeurs propres $(\mu_i^{(k)})$ associées aux vecteurs propres $\boldsymbol{u}_i$ (vecteurs colonnes de la matrice P).

Montrons que $\boldsymbol{u}_i$ est vecteur propre de la matrice $X^{(k+1)}$; par définition,

$$X^{(k+1)}\boldsymbol{u}_i = X^{(k)}\boldsymbol{u}_i + Y\boldsymbol{u}_i = \mu_i^{(k)}\boldsymbol{u}_i + Y\boldsymbol{u}_i,$$

et par définition de Y, on a :

$$X^{(k)}Y\boldsymbol{u}_i + YX^{(k)}\boldsymbol{u}_i = -(X^{(k)})^2\boldsymbol{u}_i + A\boldsymbol{u}_i.$$

Donc, comme $\boldsymbol{u}_i$ est vecteur propre de A associé à λ_i, on a par hypothèse de récurrence :

$$X^{(k)}Y\boldsymbol{u}_i + \mu_i^{(k)}Y\boldsymbol{u}_i = -(\mu^{(k)})^2\boldsymbol{u}_i + \lambda_i\boldsymbol{u}_i. \tag{2.53}$$

Ecrivons alors la décomposition de $Y\boldsymbol{u}_i$ sur la base $(\boldsymbol{u}_j)_{j=1,\ldots,n}$:

$$Y\boldsymbol{u}_i = \sum_{j=1}^{n} \alpha_j\boldsymbol{u}_j.$$

En reportant dans (2.53) et en utilisant à nouveau l'hypothèse de récurrence, on obtient :

$$\sum_{j=1}^{n} \alpha_j\mu_j^{(k)}\boldsymbol{u}_j + \mu_i^{(k)}\sum_{j=1}^{n} \alpha_j\boldsymbol{u}_j = -(\mu_i^{(k)})^2\boldsymbol{u}_i + \lambda_i\boldsymbol{u}_i. \tag{2.54}$$

En prenant le produit scalaire de cette équation avec $\boldsymbol{u}_\ell$, $\ell \neq i$, on obtient que :

$$(\mu_\ell^{(k)} + \mu_i^{(k)})\alpha_\ell = 0.$$

Comme la matrice $X^{(k)}$ est sdp, on a $\mu_\ell^{(k)} + \mu_i^{(k)} > 0$ et donc $\alpha_\ell = 0$ pour tout $\ell \neq i$. En prenant ensuite le produit scalaire de l'équation (2.54) avec $\boldsymbol{u}_i$, on obtient que :

$$2\mu_i^{(k)}\alpha_i = -(\mu_i^{(k)})^2 + \lambda_i.$$

On a donc finalement
$$Y\boldsymbol{u}_i = \frac{1}{2\mu_i^{(k)}}(-(\mu_i^{(k)})^2 + \lambda_i)\boldsymbol{u}_i$$
ce qui détermine Y de manière unique, et montre également que Y est symétrique, puisque Y s'écrit sous la forme
$$Y = P\Delta^{(k)}P^{-1} \text{ avec } \Delta^{(k)} = \text{ diag } \left(\frac{1}{2\mu_1^{(k)}}(-(\mu_1^{(k)})^2 + \lambda_1), \ldots, \frac{1}{2\mu_n^{(k)}}(-(\mu_n^{(k)})^2 + \lambda_n)\right).$$
On peut enfin calculer $X^{(k+1)} = X^{(k)} + Y = P(D^{(k)} + \Delta^{(k)})P^{-1}$. Comme
$$(D^{(k)} + \Delta^{(k)})_i = \mu_i^{(k)} + \frac{1}{2\mu_i^{(k)}}(-(\mu_i^{(k)})^2 + \lambda_i),$$
on a donc bien
$$X^{(k+1)} = P^{-1}D^{(k+1)}P \text{ avec } D^{(k+1)} = \text{diag}(\mu_1^{(k+1)}, \ldots, \mu_n^{(k+1)})\ ;$$
de plus $X^{(k+1)}$ est s.d.p. puisqu'on a montré à la question 1 que $\mu_i^{(k+1)} \geq \sqrt{\lambda_i}$.
5.2 (ii) On a vu à la question 1 que $\mu_i^{(k)}$ converge vers $\sqrt{\lambda_i}$ lorsque $k \to +\infty$. On en déduit que $X^{(k)}$ converge vers B quand $k \to +\infty$.

Corrigé de l'exercice 102 page 178 (Autre démonstration de la convergence locale de Newton) Remarquons d'abord que
$$Dg(\boldsymbol{x}) = Dg(\bar{\boldsymbol{x}}) - Dg(\bar{\boldsymbol{x}}) + Dg(\boldsymbol{x}) = Dg(\bar{\boldsymbol{x}})(Id + S)$$
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 $\|Dg(x) - Dg(\bar{x})\| \leq \varepsilon$. 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
$$x, y \in B(\bar{x}, a) \Longrightarrow \|g(y) - g(x) - Dg(x)(y - x)\| \leq a_2\|y - x\|^2.$$

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 dont on veut majorer la 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

$$\begin{aligned} \|\varphi(1) - \varphi(0)| &= \|g(y) - g(x) - Dg(x)(y-x)\| \\ &\leq \int_0^1 \|Dg(x + t(y-x))(y-x) - Dg(x)(y-x)\|dt \\ &\leq \|y-x\| \int_0^1 \|Dg(x + t(y-x)) - Dg(x)\|dt. \end{aligned} \tag{2.55}$$

Pour majorer $\|Dg(x + t(y-x)) - Dg(x))\|$, on utilise alors le théorème des accroissements finis (parfois aussi appelé "théorème de la moyenne") appliqué à Dg ; de l'inégalité (2.55), on tire donc que pour $x, y \in B(\bar{x}, a)$ et $t \in]0, 1[$:

$$\|Dg(x + t(y-x)) - Dg(x)\| \leq t\|y-x\| \sup_{c \in B(\bar{x}, a)} \|D(Dg)(c)\|_{\mathcal{L}(\mathbb{R}^n, \mathcal{M}_n(\mathbb{R}))}. \tag{2.56}$$

Comme $D(Dg) = D^2 g$ 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 (2.56) que :

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

et de l'inégalité (2.55) on déduit ensuite que

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

ce qui termine la preuve. On peut alors appliquer le théorème 2.20 pour obtenir le résultat de convergence locale du théorème 2.19.

Corrigé de l'exercice 99 page 178 (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 :

$$z \cdot A\bar{x} - \bar{\lambda} z \cdot \bar{x} - \nu \bar{x} \cdot \bar{x} = 0, \tag{2.57}$$

et comme $A\bar{x} = \bar{\lambda}\bar{x}$ et $\bar{x} \cdot \bar{x} = 1$, ceci entraîne que $\nu = 0$. En revenant à (2.57) 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 100 page 178 (Modification de la méthode de Newton)
1. Si $A \in \mathcal{M}_n(\mathbb{R})$ et $\lambda > 0$, alors $A^tA + \lambda Id$ est symétrique définie positive. En prenant $A = Df(x^{(k)})$, on obtient que la matrice $Df(x^{(k)})^tDf(x^{(k)}) + \lambda Id$ est symétrique définie positive donc inversible, ce qui montre que la suite $(x^{(k)})_{n\in\mathbb{N}}$ est bien définie.

2. Soit φ la fonction φ définie par $\varphi(t) = f(tx^{(k)} + (1-t)\overline{x})$, alors $\varphi(0) = f(\overline{x}) = 0$ et $\varphi(1) = f(x^{(k)})$. Donc

$$f(x^{(k)}) = \varphi(1) - \varphi(0) = \int_0^1 \varphi'(t)dt = (x^{(k)} - \overline{x})\int_0^1 f'(tx^{(k)} + (1-t)\overline{x})dt.$$

On a donc

$$f(x^{(k)}) = (x^{(k)} - \overline{x})g(x^{(k)}), \tag{2.58}$$

où $g(x) = \int_0^1 f'(tx+(1-t)\overline{x})dt$. La fonction g est continue car $f \in C^1(\mathbb{R},\mathbb{R})$, et $g(x) \to f'(\bar{x})$ lorsque $x \to \bar{x}$.

La suite $(x^{(k)})_{n\in\mathbb{N}}$ est définie par

$$x^{(k+1)} = x^{(k)} - \frac{f'(x^{(k)})}{f'(x^{(k)})^2 + \lambda}f(x^{(k)}).$$

En utilisant (2.58), on a donc :

$$x^{(k+1)} - \overline{x} = a_k(x^{(k)} - \overline{x}),\ \text{où } a_k = 1 - \frac{f'(x^{(k)})g(x^{(k)})}{f'(x^{(k)})^2 + \lambda}.$$

Soit a la fonction définie par :

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

On a $a(\bar{x}) = 1 - \frac{f'(\bar{x})^2}{f'(x)^2+\lambda} \in]0,1[$.
En posant $\eta = \min\{a(\bar{x})/2, (1-a(\bar{x}))/2\}$, on a donc $\eta \in]0,1/2[$ et $a(\bar{x}) \in]2\eta, 1-2\eta[$.
Comme g est continue, il existe $\alpha \in \mathbb{R}_+^*$ t.q. si $x \in]\overline{x}-\alpha, \overline{x}+\alpha[$, alors $a(x) \in]\eta, 1-\eta[$. Donc si $x^{(0)} \in]\overline{x}-\alpha, \overline{x}+\alpha[$, on a $a_0 \in]\eta, 1-\eta[$ et $x^{(1)} - \overline{x} = a_0(x^{(0)} - \overline{x})$, et par récurrence sur k, $a_k \in]\eta, 1-\eta[$ et $x^{(k)} - \overline{x} = \prod_{i=0}^{k-1} a_i(x^{(0)} - \overline{x}) \to 0$ lorsque $k \to +\infty$, ce qui prouve la convergence locale de la méthode.

3. Par définition de la méthode, on a :

$$x^{(k+1)} - \overline{x} = x^{(k)} - \overline{x} - (Df(x^{(k)})^tDf(x^{(k)}) + \lambda Id)^{-1}Df(x^{(k)})^tf(x^{(k)})$$

En posant, pour $t \in \mathbb{R}$, $\varphi(t) = f(tx^{(k)} + (1-t)\overline{x})$, on a :

$$\begin{aligned} f(x^{(k)}) - f(\overline{x}) &= \int_0^1 \varphi'(t)dt \\ &= G(x^{(k)})(x^{(k)} - \overline{x}), \end{aligned}$$

où $G \in C(\mathbb{R}^n, \mathcal{M}_n(\mathbb{R}))$ est définie par

$$G(x) = \int_0^1 Df(tx + (1-t)\overline{x})dt.$$

On a donc :

$$x^{(k+1)} - \overline{x} = H(x^{(k)})(x^{(k)} - \overline{x}), \tag{2.59}$$

où $H \in \mathcal{C}(\mathbb{R}^n, \mathcal{M}_n(\mathbb{R}))$ est définie par :

$$H(x) = Id - (Df(x)^t Df(x) + \lambda Id)^{-1} Df(x)^t G(x).$$

On veut montrer que $\|H(x^{(k)})\| < 1$ si $x^{(k)}$ est suffisamment proche de $\overline{x}$. On va pour cela utiliser la continuité de H autour de $\overline{x}$. On a :

$$H(\overline{x}) = Id - (Df(\overline{x})^t Df(\overline{x}) + \lambda Id)^{-1} Df(\overline{x})^t Df(\overline{x}).$$

La matrice $B = Df(\overline{x})^t Df(\overline{x})$ est évidemment symétrique. On a donc :

$$\begin{aligned} H(\overline{x})^t &= (Id - (B + \lambda Id)^{-1} B)^t \\ &= Id - B(B + \lambda Id)^{-1} \end{aligned}$$

Pour montrer que $H(\overline{x})$ est symétrique, il reste à montrer que B et $(B + \lambda Id)^{-1}$ commutent.

$$\text{Or } (B + \lambda Id)(B + \lambda Id)^{-1} = Id.$$

$$\begin{aligned} \text{Donc } B(B + \lambda Id)^{-1} &= Id - \lambda(B + \lambda Id)^{-1} \\ &= (B + \lambda Id)^{-1}(B + \lambda Id) - \lambda(B + \lambda Id)^{-1} \\ &= (B + \lambda Id)^{-1} B. \end{aligned}$$

On en déduit que $H(\overline{x})$ est symétrique. On a donc $\|H(\overline{x})\|_2 = \rho(H(\overline{x}))$. Calculons le rayon spectral de $H(\overline{x})$. Comme $Df(\overline{x})^t Df(\overline{x})$ est diagonalisable dans $\mathbb{R}$, il existe $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ et $(f_1, \ldots, f_n)$ base orthonormée telle que :

$$Df(\overline{x})^t Df(\overline{x}) \ f_i = \lambda_i f_i, \ i = 1, \ldots, n.$$

De plus $\lambda_i > 0$, $i = 1, \ldots, n$. En effet $\lambda_i = Df(\bar{x}) f_i \cdot Df(\bar{x}) f_i > 0$ car le rang de $Df(\bar{x})$ est supposé être égal à n (et son noyau est donc réduit à $\{0\}$).
On a

$$H(\overline{x}) f_i = f_i - (Df(\overline{x})^t Df(\overline{x}) + \lambda Id)^{-1} \lambda_i f_i$$

$$\text{Or } (Df(\overline{x})^t Df(\overline{x}) + \lambda Id) f_i = (\lambda_i + \lambda) f_i, \text{ donc}$$

$Df(\overline{x})^t Df(\overline{x}) + \lambda Id)^{-1} f_i = \dfrac{1}{\lambda_i + \lambda} f_i$. On en déduit que

$$H(\overline{x}) f_i = \mu_i f_i, \ i = 1, \ldots, n, \text{ où } \mu_i = 1 - \frac{\lambda_i}{\lambda_i + \lambda}.$$

On a donc $0 < \mu_i < 1$ et donc $\rho(H(\bar{x})) < 1$. On en déduit que $\|H(\overline{x})\|_2 < 1$, et par continuité il existe $\alpha > 0$ tel que si $x \in B(\overline{x}, \alpha)$ alors $\|H(x)\|_2 < 1$.
On déduit alors de (2.59) que la méthode est localement convergente.

Corrigé de l'exercice 103 page 179 (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 :

$$f'(x_n)(x_{n+1} - x_n) = -f(x_n) \tag{2.60}$$

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 (2.60) que

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

On a 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 104 page 179 (Point fixe et Newton) 1) Pour que la suite $(x_n)_{n\in\mathbb{N}}$ soit bien définie, il faut que $g' \circ f(x_n) \neq 0, \forall n \in \mathbb{N}$. On remarque tout d'abord que $g'(f(\bar{x})) = g'(\bar{x}) \neq 0$, par hypothèse. Or $g \in C^3(\mathbb{R}, \mathbb{R})$ et $f \in C^1(\mathbb{R}, \mathbb{R})$, donc $g' \circ f$ est continue ; on en déduit qu'il existe $\eta > 0$ tel que $|g' \circ f(x)| > \dfrac{|g'(\bar{x})|}{2} > 0, \forall x \in]\bar{x} - \eta, \bar{x} + \eta[$.
Pour montrer que la suite est bien définie, il reste à montrer que la suite $(x_n)_{n\in\mathbb{N}}$ est incluse dans cet intervalle. Pour ce faire, on va montrer que h est contractante sur un intervalle $]\bar{x} - \alpha, \bar{x} + \alpha[$. En effet, on a

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

Donc

$$h'(\bar{x}) = 1 - \frac{1}{(g'(\bar{x}))^2}((g'(\bar{x}))^2 = 0.$$

Comme $g \in C^3(\mathbb{R}, \mathbb{R})$ et $f \in C^1(\mathbb{R}, \mathbb{R})$, on a $h \in C^1(\mathbb{R}, \mathbb{R})$ et donc h' est continue. On en déduit qu'il existe $\beta > 0$ t.q. $h'(x) < 1, \forall x \in]\bar{x} - \beta, \bar{x} + \beta[$. Soit $\alpha = \min(\eta, \beta)$. Sur $I_\alpha =]\bar{x} - \alpha, \bar{x} + \alpha[$, on a donc $g' \circ f(x) \neq 0$ et $h'(x) < 1$. En particulier, h' est donc contractante sur I_α. Donc si $x_0 \in I_\alpha$, on a

$$|x_1 - \bar{x}| = |h(x_0) - h(\bar{x})| < |x_0 - \bar{x}|$$

et donc $x_1 \in I_\alpha$. On en déduit par récurrence que $(x_n)_{n\in\mathbb{N}} \subset I_\alpha$, et donc que la suite est bien définie.

Par le théorème du point fixe, on en déduit également que $(x_n)_{n\in\mathbb{N}}$ converge vers l'unique point fixe de h sur I_α, c'est à dire $\bar{x}$, lorsque $n \to +\infty$.

2) Montrons que $\dfrac{|x_{n+1}-\bar{x}|}{|x_n-\bar{x}|^2}$ est borné indépendamment de n. En effet, on a :

$$\begin{aligned} x_{n+1}-\bar{x} &= h(x_n)-\bar{x} \\ &= x_n-\bar{x}-\frac{g(x_n)}{g'\circ f(x_n)}. \end{aligned}$$

Comme $g(\bar{x})=0$, on a donc :

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

Par le théorème des accroissements finis, il existe $\xi_n \in I_\alpha$ tel que

$$\frac{g(x_n)-g(\bar{x})}{x_n-\bar{x}} = g'(\xi_n).$$

On a donc

$$x_{n+1}-\bar{x} = \frac{x_n-\bar{x}}{g'\circ f(x_n)}\left[g'(f(x_n))-g'(\xi_n)\right]$$

Comme $g \in C^3$, on peut appliquer à nouveau le théorème des accroissements finis sur g' : il existe $\zeta_n \in I_\alpha$ tel que

$$g'(f(x_n))-g'(\xi_n) = g''(\zeta_n)(f(x_n)-\xi_n).$$

On a donc :

$$\begin{aligned} |x_{n+1}-\bar{x}| &= \frac{|x_n-\bar{x}|}{|g'\circ f(x_n)|}|g''(\zeta_n)||f(x_n)-\bar{x}+\bar{x}-\xi_n| \\ &\leq 2\frac{|x_n-\bar{x}|}{|g'(\bar{x})|} \quad |g''(\zeta_n)|2|x_n-\bar{x}| \end{aligned}$$

On a donc finalement

$$|x_{n+1}-\bar{x}| \leq \frac{4}{|g'(\bar{x})|}\sup_{I_\alpha}|g''||x_n-\bar{x}|^2$$

Ce qui montre que la convergence est d'ordre 2.

3)Allons-y pour les développements limités, dans la joie et l'allégresse... On pose $\alpha = g'(\bar{x})$, $\beta = g''(\bar{x})$, et $\gamma = f'(\bar{x})$. on notera dans la suite a, b, et c des fonctions bornées de $\mathbb{R}$ dans $\mathbb{R}$, telles que :
$g(x) = (x-\bar{x})\alpha + (x-\bar{x})^2\beta + (x-\bar{x})^3 a(x)$
$g'(x) = \alpha + 2\beta(x-\bar{x}) + (x-\bar{x})^2 b(x)$
$f(x) = \bar{x} + \gamma(x-\bar{x}) + (x-\bar{x})^2 c(x)$.
On a donc :

$$\begin{aligned} g'(f(x)) &= \alpha + 2\beta(f(x)-\bar{x}) + (f(x)-\bar{x})^2 b(x) \\ &= \alpha + 2\beta\gamma(x-\bar{x}) + (x-\bar{x})^2 d(x), \end{aligned}$$

où d est aussi une fonction bornée de $\mathbb{R}$ dans $\mathbb{R}$. On en déduit que

$$\begin{aligned} h(x) &= x - \frac{(x-\bar{x})\alpha + (x\bar{x})^2\beta + (x-\bar{x})^3 a(x)}{\alpha + 2\beta\gamma(x-\bar{x}) + (x-\bar{x})^2 d(x)} \\ &= x - \left[(x-\bar{x}) + (x-\bar{x})\frac{\beta}{\alpha} + (x-\bar{x})^3\widetilde{a}(x)\right]\left[1 - 2\frac{\beta\gamma}{\alpha}(x-\bar{x}) + (x-\bar{x})^2 d(x).\right] \end{aligned}$$

On en déduit que $h(x) = \bar{x} + \dfrac{\beta}{\alpha}(2\gamma - 1)(x - \bar{x})^2 + (x - \bar{x})^3 \widetilde{d}(x)$, où $\widetilde{d}$ est encore une fonction bornée.
Cette formule donne :

$$x_{n+1} - \bar{x} = \frac{\beta}{\alpha}(2\gamma - 1)(x_n - \bar{x})^2 + (x_n - \bar{x})^3 \widetilde{d}(x_n),$$

ce qui redonne l'ordre 2 trouvé à la question 2 ; dans le cas où $\gamma = \dfrac{1}{2}$, i.e. $f'(\bar{x}) = \dfrac{1}{2}$, on obtient bien une convergence cubique.

4) Comme g' ne s'annule pas sur I_β, la fonction f est de classe C^2 sur I_β, et $f'(\bar{x}) = \dfrac{1}{2}$.
Soit $\gamma = \min(\alpha, \beta)$, où α est défini à la question 1.
Les hypothèses des questions 1 et 3 sont alors bien vérifiées, et l'algorithme converge de manière au moins cubique.

Corrigé de l'exercice 105 page 179 (Variante de la méthode de Newton)

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

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

Donc

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

On en déduit que

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

Ceci entraîne :

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

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

2. On a :

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

Par hypothèse, on a donc

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

On en déduit par récurrence que

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

Ceci entraîne en particulier que

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

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

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

$$\text{Donc} \quad \frac{|x^{(k+1)}-\bar{x}|}{|x^{(k)}-\bar{x}|} \underset{n\to+\infty}{\to} |1-\frac{f'(\bar{x})}{f'(y)}| = \beta \geq 0$$

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

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

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. Pour $n \geq 1$, l'algorithme se généralise en :

$$\begin{cases} x^{(0)} = x_0 \\ x^{(k+1)} = x^{(k)} - (DF(y))^{-1}f(x^{(k)}). \end{cases}$$

On a donc

$$x^{(k+1)}-x_0 = x^{(k)}-x_0-(DF(y))^{-1}(f(x^{(k)})-f(x_0))-(DF(y))^{-1}f(x_0). \tag{2.61}$$

On définit $\varphi : \mathbb{R} \to \mathbb{R}^n$ par $\varphi(t) = f(tx^{(k)}+(1-t)x^{(0)})$. On a

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

et donc

$$\begin{aligned} f(x^{(k)})-f(x^{(0)}) &= \varphi(1)-\varphi(0)\\ &= \int_0^1 \varphi'(t)dt\\ &= \int_0^1 Df(tx^{(k)}+(1-t)x^{(0)})(x^{(k)}-x^{(0)})dt. \end{aligned}$$

L'égalité (2.61) s'écrit donc

$$\begin{aligned} x^{(k+1)}-x^{(0)} &= \left(Id-(Df(y))^{-1}\int_0^1 Df(tx^{(k)}+(1-t)x^{(0)})dt\right)(x^{(k)}-x^{(0)})-(Df(y))^{-1}f(x_0)\\ &= (Df(y))^{-1}\left(\int_0^1 \left(Df(y)-Df(tx^{(k)}+(1-t)x^{(0)}\right)dt\right)(x^{(k)}-x^{(0)})-(Df(y))^{-1}f(x_0). \end{aligned}$$

On en déduit que :

$$\begin{aligned}\|x^{(k+1)}-x^{(0)}\| \leq \|(Df(y))^{-1}\| \int_0^1 \|Df(y)-Df(tx^{(k)}+(1-t)x^{(0)})\|dt\,\|x^{(k)}-x^{(0)}\|\\ +\|(Df(y))^{-1}\|\|f(x_0)\|. \end{aligned} \tag{2.62}$$

Si on suppose que $x^{(k)} \in I$, alors $tx^{(k)} + (1-t)x^{(0)} \in I$. 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 2.62 donne que

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

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

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

Corrigé de l'exercice 101 page 178 (Méthode de Newton)

1. Soient u et v solutions du système non linéaire considéré. On a alors :

$$(A(u-v))_i + \alpha_i(f(u_i) - f(v_i)) = 0 \text{ pour tout } i = 1, \ldots, n.$$

Donc $\displaystyle\sum_{i=1}^{n}(A(u-v))_i(u-v)_i + \sum_{i=1}^{n}\alpha_i(f(u_i) - f(v_i))(u_i - v_i) = 0.$

Comme f est croissante, on a $f(u_i) - f(v_i)(u_i - v_i) \geq 0 \ \forall i = 1, \ldots, n$.

On en déduit que $A(u-v)\cdot(u-v) = 0$. Comme A est symétrique définie positive, ceci entraîne $u = v$.

2. Soit F la fonction de $\mathbb{R}^n$ dans $\mathbb{R}^n$ définie par :

$$(F(u))_i = (Au)_i + \alpha_i f(u_i), \quad i = 1, \ldots, n.$$

Comme $f \in \mathcal{C}^2(\mathbb{R}, \mathbb{R})$, on a $F \in \mathcal{C}^2(\mathbb{R}^n, \mathbb{R}^n)$. La méthode de Newton de recherche d'un zéro de F s'écrit

$$u^{(k+1)} = u^{(k)} - (DF(u^{(k)}))^{-1}F(u^{(k)}).$$

D'après un théorème du cours, la méthode de Newton converge localement (avec convergence d'ordre 2) si la matrice jacobienne $DF(\bar{u})$ est inversible, où $\bar{u}$ est l'unique solution de $F(\bar{u}) = 0$.

Calculons $DF(\bar{u})$:

on a

$$(F(u))_i = (Au)_i + \alpha_i f(u_i), \text{ pour } i = 1, \ldots, n, \text{ et donc}$$

$$\begin{aligned} (DF(u)\cdot v)_i &= (Av)_i + \alpha_i f'(u_i)v_i \\ &= ((A+D)v)_i. \end{aligned}$$

où $D = \mathrm{diag}(\alpha_1 f'(u_1), \ldots, \alpha_n f'(u_n))$. Comme $\alpha_i > 0$ et $f'(u_i) \geq 0$ pour tout $i = 1, n$, la matrice $A + D$ est symétrique définie positive donc $DF(\bar{u})$ est inversible.

On en déduit que la méthode de Newton converge localement.

Corrigé de l'exercice 106 page 180 (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 :

$$x_{n+1} - \overline{x} = x_n - \overline{x} - \frac{(f(x_n))^2}{f(x_n + f(x_r)) - f(x_n)}. \tag{2.63}$$

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 $\xi_n(t) = \displaystyle\int_0^t \xi_n'(s)ds + \xi_n(0)$.
On obtient alors :

$$f(x_n + f(x_n)) - f(x_n) = f(x_n)\left[f(x_n)\int_0^1\int_0^t f''(x_n + sf(x_n))ds + f'(x_n)\right]. \tag{2.64}$$

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 $b(x) \to \dfrac{1}{2}f''(\overline{x})$ lorsque $x \to \overline{x}$
L'égalité (2.64) s'écrit alors :

$$f(x_n + f(x_n)) - f(x_n) = (f(x_n))^2 b(x_n) + f(x_n)f'(x_n). \tag{2.65}$$

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 (2.63) et (2.65) :

$$x_{n+1} - \overline{x} = x_n - \overline{x} - \frac{f(x_n)}{f(x_n)b(x_n) + f'(x_n))} \tag{2.66}$$

On a maintenant $-f(x_n) = f(\overline{x}) - f(x_n) = \int_0^1 \varphi'(t)dt$ 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 :

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

De cette égalité et de (2.66), on obtient :

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

On en déduit :

$$(x_{n+1} - \overline{x}) = (x_n - \overline{x})^2 a(x_n) \tag{2.67}$$

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

$$|a(x) - a(\overline{x})| \leq \varepsilon. \tag{2.68}$$

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 $\gamma = \min(\eta_1, \dfrac{1}{2(\beta+1)})$; si $x \in B(\overline{x}, \gamma)$, alors $|a(x)| \leq \beta + 1$ grâce à (2.68), et $|x - \overline{x}| \leq \dfrac{1}{2(\beta+1)}$.
On déduit alors de (2.67) 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 (2.67), 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 107 page 180 (Méthode de Newton-Tchebychev)
1.
  (a) PF comme Point Fixe... parce que c'est effectivement un algorithme de point fixe.
  (b) Comme $f'(x) = 0$ la continuité de f' donne que pour tout $\varepsilon > 0$, il existe $\alpha(\varepsilon)$ tel que si $|y - \overline{x}| \leq \alpha(\varepsilon)$ alors $|f'(y)| \leq \varepsilon$. On obtient donc le résultat souhaité en prenant $\varepsilon = \frac{1}{2}$ et $a = \alpha(\frac{1}{2})$.
  (c) Itération $n = 0$ On a bien $x_0 \in B(\bar{x}, a)$. Itération $n + 1$ On suppose $x_n \in B(\bar{x}, \frac{a}{2^n})$. Comme x_n et $\bar{x} \in B(\bar{x}, a)$, on a par le théorème des acroissements finis :

  $$x_{n+1} - \bar{x} = f(x_n) - f(\bar{x}) = f'(c)(x_n - \bar{x}) \text{ avec } c \in B(\bar{x}, a),$$

  et donc $|x_{n+1} - \bar{x}| \leq \frac{1}{2}|x_n - \bar{x}| < \frac{a}{2^{n+1}}$ et $x_{n+1} \in B(\bar{x}, \frac{a}{2^{n+1}})$.
  (d) On vient de montrer que si $x_0 \in V$ alors $|x_n - \bar{x}| < \frac{a}{2^{n+1}} \to 0$ lorsque $n \to +\infty$, ce qui montre le résultat.
  (e) Comme $x_n \in B(\bar{x}, a)$, par Taylor Lagrange, il existe $c_n \in B(\bar{x}, a)$ tel que

  $$f(x_n) = f(\bar{x}) + (x_n - \bar{x})f'(\bar{x}) + \frac{1}{2}(x_n - \bar{x})^2 f''(\bar{x}) + \frac{1}{6}(x_n - \bar{x})^3 f'''(c_n)$$

  Donc $|x_{n+1} - \bar{x}| \leq \beta |x_n - \bar{x}|^3$, avec $\beta = \frac{1}{6} \sup_{x \in B(\bar{x},a)} |f'''(x)|$

2. En posant $f(x) = x + h(x)g(x)$, on a

$$f'(x) = 1 + h'(x)g(x) + h(x)g'(x)$$
$$f''(x) = h''(x)g(x) + 2h'(x)g'(x) + h(x)g''(x)$$

On en déduit que $f'(\bar{x}) = 1 + h(\bar{x})g'(\bar{x})$ et $f''(\bar{x}) = 2h'(\bar{x})g'(\bar{x}) + h(\bar{x})g''(\bar{x})$. En appliquant la question 1, on remarque que pour avoir la convergence cubique, il suffit que $f'(\bar{x}) = f''(\bar{x}) = 0$, c.à.d.

$$h(\bar{x}) = -\frac{1}{g'(\bar{x})} \text{ et } h'(\bar{x}) = \frac{g''(\bar{x})}{2(g'(\bar{x}))^2}.$$

3. Si g' ne s'annule pas, on peut définir la fonction h de $\mathbb{R}$ dans $\mathbb{R}$ par

$$h(x) = -\frac{1}{g'(x)} - \frac{g''(x)[g(x)]^2}{2[g'(x)]^3}$$

On vérifie alors facilement que $h \in \mathcal{C}^3(\mathbb{R}, \mathbb{R})$, et que :

$$h(\bar{x}) = -\frac{1}{g'(\bar{x})} \text{ et } h'(\bar{x}) = \frac{g''(\bar{x})}{2(g'(\bar{x}))^2}.$$

Les conditions des questions 1 et 2 sont donc satisfaites et la suite construite par Tchebychev converge localement avec une vitesse cubique.

Montrons maintenant que la suite converge localement même si g' s'annule. Comme $g'(\bar{x}) \neq 0$ et que g' est continue, il existe $b > 0$ tel que $g'(y) \neq 0$ pour tout $y \in B(\bar{x}, b)$. On prend $x_0 \in B(\bar{x}, b)$. On a vu à la question 1 que $|x_n - \bar{x}| \leq \frac{1}{2^n}|x_0 - \bar{x}|$ si $x_0 \in B(\bar{x}, a)$. Donc si $x_0 \in B(\bar{x}, \min(a, b))$, la suite est bien définie (car elle reste dans une boule où g' ne s'annule pas) et converge localement avec une vitesse cubique.

Corrigé de l'exercice 108 page 181 (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 :

   $$\begin{aligned} (f(x_n) - f(x_{n-1})e_{n+1} &= e_n f(x_n) - e_n f(x_{n-1}) - f(x_n)e_n + f(x_n)e_{n-1} &\quad (2.69)\\ &= -e_n f(x_{n-1}) + f(x_n)e_{n-1} &\quad (2.70)\\ &= e_n e_{n-1}(\frac{f(x_n)}{e_n} - \frac{f(x_{n-1})}{e_{n-1}}). &\quad (2.71) \end{aligned}$$

   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

$$\mu'(x) = \frac{f'(x) - \mu(x)}{x - \overline{x}},\ \forall x > \overline{x}. \tag{2.72}$$

Or

$$\mu(x) = \frac{1}{x - \overline{x}}(f(x) - f(\overline{x})) \tag{2.73}$$

$$= \frac{1}{x - \overline{x}}(f(x) - (f(x) + (\overline{x} - x)f'(x) + \frac{1}{2}(\overline{x} - x)^2 f''(x)(\overline{x} - x)^3\varepsilon(x). \tag{2.74}$$

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 (2.72) :

$$\mu'(x) = \frac{1}{2}f''(x) + (x - \overline{x})\varepsilon(x),\ \forall x > \overline{x}. \tag{2.75}$$

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 (2.75) et on obtient :

$$\lim_{x \to \overline{x}} \mu'(x) = \frac{1}{2}f''(\overline{x}). \tag{2.76}$$

(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 (2.76), 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 : $a_{n+1} = a_n a_{n-1} \geq M e_n M e_{n-1}$. Or par la question 2a, $e_n e_{n-1} = M_n e_{n+1} \leq M e_{n+1}$. On en déduit que la relation est encore vérifiée au rang $n + 1$.

(b) Par définition, $a_i = M e_i = M(x_i - \overline{x})$, 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

$$b_{n+1} = b_n + b_{n-1},\ \forall n \geq 1 \tag{2.77}$$

L'ensemble de suites $(b_n)_{n\in\mathbb{N}}$ vérifiant (2.77) 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 (2.77) 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 (2.77), 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 $d = \frac{1+\sqrt{5}}{2}$, $\alpha = e^{|D|}$ 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 faible que l'ordre de convergence de la méthode de Newton.

Université de Marseille
Licence de Mathématiques, 3eme année, analyse numérique et optimisation
Téléenseignement 2017-18, corrigé du 1er devoir

Exercice 1 (Déterminant d'une matrice sous forme de blocs, exercice 16)**.**
Soient $A \in \mathcal{M}_n(\mathbb{R})$ $(n > 1)$, $b, c \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$. On s'intéresse à la matrice $\bar{A} \in \mathcal{M}_{n+1}(\mathbb{R})$ définie sous forme de blocs de la manière suivante :

$$\bar{A} = \begin{bmatrix} A & b \\ c^t & \lambda \end{bmatrix} \tag{1}$$

On montre dans cet exercice que les deux assertions suivantes sont, sauf cas particuliers, fausses :

A1 $\det(\bar{A}) = \lambda \det(A) - \det(bc^t)$,

A2 $\det(\bar{A}) = \lambda \det(A) - c^t\, b$,

1. Dans cette question, on prend $n \geq 2$, $A = 0$, $b = c$ et on suppose que $b \neq 0$.
   (a) Montrer que $\mathrm{rg}(\bar{\mathrm{A}}) \leq 2$ et en déduire que $\bar{A}$ n'est pas inversible.

   Corrigé – *Les n premières colonnes de $\bar{A}$ sont colinéaires au dernier vecteur de la base canonique de $\mathbb{R}^{n+1}$. L'image de $\bar{A}$ est donc incluse dans l'espace vectoriel engendré par le dernier vecteur de la base canonique de $\mathbb{R}^{n+1}$ et la dernière colonne de $\bar{A}$. On a donc $\mathrm{rg}(\bar{\mathrm{A}}) = \dim(\mathrm{Im}(\bar{\mathrm{A}})) \leq 2$.*
   Le théorème de rang donne alors $\dim(\mathrm{Ker}(\bar{A})) = n + 1 - \dim(\mathrm{Im}(\bar{A})) \geq n - 1 > 0$ et donc $\bar{A}$ n'est pas inversible.

   (b) En déduire que l'assertion A2 est fausse pour cet exemple.

   Corrigé – *Pour cet exemple, on a $\det(\bar{A}) = \lambda \det(A) = 0$ mais $c^t\, b = b \cdot b \neq 0$. L'assertion A2 est donc fausse pour cet exemple.*

2. Dans cette question, on suppose que A est symétrique définie positive, $\lambda = 0$, $b = c$ et que $b \neq 0$.
   (a) Montrer que $\bar{A}$ est inversible et que $\mathrm{rg}(\mathrm{bb^t}) = 1$.

   Corrigé – *Soient $x \in \mathbb{R}^n$ et $y \in \mathbb{R}$ tel que $\begin{bmatrix} A & b \\ b^t & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Ceci donne $Ax + yb = 0$ et $b \cdot x = 0$. En prenant le produit scalaire de la première équation avec x, on obtient, avec la deuxième équation,*

   $$0 = (Ax + yb) \cdot x = Ax \cdot x.$$

   Comme A est s.d.p., ceci donne $x = 0$. On a alors aussi $yb = 0$ et donc $y = 0$ (car $b \neq 0$). On a ainsi prouvé que $\bar{A}$ était inversible (et donc $\det(\bar{A}) \neq 0$).
   Pour tout $x \in \mathbb{R}^n$, on a $bb^t x = (b \cdot x)b$. Ceci montre que $\mathrm{Im}(bb^t) = \{\alpha b, \alpha \in \mathbb{R}\}$, et donc $\mathrm{rg}(\mathrm{bb^t}) = 1$ (car $b \neq 0$).

   (b) En déduire que l'assertion A1 est fausse pour cet exemple.

   Corrigé – *Comme $n > 1$, le théorème du rang donne $\dim(\mathrm{Ker}(bb^t)) = n - \mathrm{rg}(\mathrm{bb^t}) = \mathrm{n} - 1 > 0$ et donc $\det(bb^t) = 0$. Comme $\lambda = 0$, on a donc $\lambda \det(A) - \det(bc^t) = 0$. Mais $\det(\bar{A}) \neq 0$, L'assertion A1 est donc fausse pour cet exemple.*

Exercice 2 (Décompositions LU et de Choleski, exercice 28)**.**
Soit $M = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 8 & 10 \\ 1 & 10 & 18 \end{bmatrix}$.

1. Calculer les mineurs principaux de M. En déduire que M admet des décompositions LU et de Choleski.

Corrigé – *Les mineurs principaux de M valent 1, 4 et 4. Comme ils sont non nuls, la matrice M admet une décomposition LU sans pivot.*
La matrice M est symétrique. Comme tous les coefficients diagonaux de la matrice U (donnée par la décomposition LU sans pivot de M) sont strictement positifs, la matrice M est s.d.p. (proposition 1.23 du polycopié). Elle admet donc une décomposition de Choleski.

2. Donner la décomposition LU de M.

Corrigé – *La décomposition LU de M est* $L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix}$ *et* $U = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 4 & 8 \\ 0 & 0 & 1 \end{bmatrix}$.

3. Donner la décomposition de Choleski de M.

Corrigé – *La décomposition de Choleski de M est* $L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ 1 & 4 & 1 \end{bmatrix}$.

Exercice 3 (Vitesse de convergence pour la méthode de Jacobi, exercice 72)**.**
Soient A une matrice carrée d'ordre n, inversible, et $b \in \mathbb{R}^n$, $n > 1$. On pose $\bar{x} = A^{-1}b$. On note D la partie diagonale de A, $-E$ la partie triangulaire inférieure stricte de A et $-F$ la partie triangulaire supérieure stricte de A. On suppose que D est inversible et on note B_J la matrice des itérations de la méthode de Jacobi, c'est-à-dire $B_J = D^{-1}(E+F)$. On rappelle que la méthode de Jacobi s'écrit
Initialisation : $x^{(0)} \in \mathbb{R}^n$,
Itérations : pour tout $k \geq 0$, $Dx^{(k+1)} = (E+F)x^{(k)} + b$.
On munit $\mathbb{R}^n$ d'une norme notée $\|\cdot\|$. On note ρ le rayon spectral de B_J.

1. On suppose que B_J est diagonalisable dans $\mathbb{R}$ (c'est-à-dire qu'il existe une base de $\mathbb{R}^n$ formée de vecteurs propres de B_J). Montrer qu'il existe $C > 0$, dépendant de A, b, $x^{(0)}$ et de la norme choisie sur $\mathbb{R}^n$, mais indépendant de k, telle que
$$\|x^{(k)} - \overline{x}\| \leq C\rho^k \text{ pour tout } k \geq 0.$$

Corrigé – *Soit $f_1, \ldots, f_n$ une base de $\mathbb{R}^n$ formée de vecteurs propres de B_J. On a donc, pour tout $i \in \{1, \ldots, n\}$, $B_J f_i = \lambda_i f_i$ avec $\lambda_i \in \mathbb{R}$ et $\rho = \max\{|\lambda_i|, i \in \{1, \ldots, n\}\}$.*
La suite $(x^{(k)})_{k \in \mathbb{N}}$ donnée par la méthode de Jacobi vérifie, pour tout $k \geq 0$, $x^{(k+1)} - \overline{x} = B_J(x^{(k)} - \overline{x})$. En écrivant $x^{(0)} - \overline{x}$ dans la base $f_1, \ldots, f_n$ on a $x^{(0)} - \overline{x} = \sum_{i=1}^n \alpha_i f_i$. Par récurrence sur k, on en déduit, pour tout $k \geq 0$,
$$x^{(k)} - \overline{x} = \sum_{i=1}^n \lambda_i^k \alpha_i f_i.$$
Comme $|\lambda_i| \leq \rho$, on en déduit $\|x^{(k)} - \overline{x}\| \leq \rho^k \sum_{i=1}^n |\alpha_i| \|f_i\|$, En prenant $C = \sum_{i=1}^n |\alpha_i| \|f_i\|$, on obtient le résultat demandé .
N.B. Ce résultat reste vrai si B_J est diagonalisable dans $\mathbb{C}$.

2. On ne suppose plus que B_J est diagonalisable. Montrer que pour tout $\varepsilon > 0$, il existe $C_\varepsilon > 0$, dépendant de A, b, $x^{(0)}$, ε et de la norme choisie sur $\mathbb{R}^n$, mais indépendant de k, telle que
$$\|x^{(k)} - \overline{x}\| \leq C_\varepsilon(\rho + \varepsilon)^k \text{ pour tout } k \geq 0.$$

Corrigé – *Soit $\varepsilon > 0$. On sait qu'il existe une norme sur $\mathbb{R}^n$, notée $\|\cdot\|_\varepsilon$, telle que la norme induite sur $\mathcal{M}_n(\mathbb{R})$ (encore notée $\|\cdot\|_\varepsilon$) vérifie $\|B_J\|_\varepsilon \leq \rho + \varepsilon$. (On rappelle que cette norme dépend de ε et B_J.)*
Avec cette norme, on a donc
$$\|x^{(k)} - \overline{x}\|_\varepsilon \leq \|B_J\|_\varepsilon^k \|x^{(0)} - \overline{x}\|_\varepsilon \leq \|x^{(0)} - \overline{x}\|_\varepsilon (\rho + \varepsilon)^k.$$
D'autre part, comme sur $\mathbb{R}^n$ toutes les normes sont équivalentes, il existe $D \in \mathbb{R}$ tel que $\|\cdot\| \leq D\|\cdot\|_\varepsilon$. On a donc
$$\|x^{(k)} - \overline{x}\|_\varepsilon \leq C_\varepsilon(\rho + \varepsilon)^k \text{ avec } C_\varepsilon = D\|x^{(0)} - \overline{x}\|_\varepsilon.$$

Dans la suite de cet exercice on prend $n = 2$ et $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$.

3. Dans cette question, on choisit, pour norme dans $\mathbb{R}^2$, la norme euclidienne, c'est-à-dire $\|x\|^2 = x_1^2 + x_2^2$ si $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Montrer qu'il existe C (dépendant de b et $x^{(0)}$, mais non de k) telle que

$$\|x^{(k)} - \overline{x}\| = C\rho^k \text{ pour tout } k \geq 0. \tag{2}$$

Corrigé – $B_J = D^{-1}(E+F) = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}$. *On a donc* $\rho = \frac{1}{2}$.

Soit $x^{(0)} \in \mathbb{R}^2$, $x^{(0)} - \overline{x} = \alpha_1 e_1 + \alpha_2 e_2$. *On a alors*

$$x^{(k)} - \overline{x} = (\frac{1}{2})^k(\alpha_1 e_1 + \alpha_2 e_2) \text{ si } k \text{ est pair,}$$

$$x^{(k)} - \overline{x} = (\frac{1}{2})^k(\alpha_1 e_2 + \alpha_2 e_1) \text{ si } k \text{ est impair.}$$

On en déduit que $|x^{(k)} - \overline{x}| = \rho^k|x^{(0)} - \overline{x}|$ *pour tout* $k \geq 0$.

4. Montrer qu'il existe des normes dans $\mathbb{R}^2$ pour lesquelles la conclusion de la question 3 est fausse (c'est-à-dire pour lesquelles la suite $(\|x^{(k)} - \overline{x}\|/\rho^k)_{k\in\mathbb{N}}$ n'est pas une suite constante sauf éventuellement pour des valeurs particulières de $x^{(0)}$).

Corrigé – *On prend, par exemple,* $\|x\| = \sqrt{2x_1^2 + x_2^2}$ *pour* $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

Avec les notations de la question précédente, on a $\|x^{(k)} - \overline{x}\| = \rho^k\sqrt{2\alpha_1^2 + \alpha_2^2}$ *si* k *est pair et* $\|x^{(k)} - \overline{x}\| = \rho^k\sqrt{2\alpha_2^2 + \alpha_1^2}$ *si* k *est impair.*

La suite $(\|x^{(k)} - \overline{x}\|/\rho^k)_{k\in\mathbb{N}}$ *est une suite constante seulement si* $\alpha_1^2 = \alpha_2^2$.