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 monotonie en dimension $n = 1$. On va maintenant étudier cette méthode dans le cas n quelconque. Soient $g \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ et $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ 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, $g \in C^2$)**.** *Soient $g \in C^2(\mathbb{R}^n, \mathbb{R}^n)$ 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 $\boldsymbol{x}^{(0)} \in B(\bar{\boldsymbol{x}}, b) = \{\boldsymbol{x} \in \mathbb{R}^n, \|\boldsymbol{x} - \bar{\boldsymbol{x}}\| < b\}$ 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 de $\mathbb{R}^n$ dans $\mathbb{R}^n$ définie 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
$$\|\boldsymbol{y} - \bar{\boldsymbol{x}}\| \le \frac{1}{2}\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|.$$
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 par hypothèse, $Df \in C^1(\mathbb{R}^n, \mathbb{R}^n)$, on a
$$\|Df(\boldsymbol{z})\| = \|Df(\boldsymbol{z}) - Df(\bar{x})\|$$
$$\le \sup_{\xi\in B}\|Df(\xi)\|\|\boldsymbol{z} - \bar{\boldsymbol{x}}\| \tag{2.22}$$
$$\le \beta\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|. \tag{2.23}$$
En reportant cette majoration de $\|Df(\boldsymbol{z})\|$ dans (2.21), on obtient alors (avec $\beta = \sup_{\xi\in B}\|Df(\xi)\|$) :
$$\|y - \bar{\boldsymbol{x}}\| \le \beta\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|^2$$
ce qui donne la convergence locale au moins quadratique. ■

La condition $g \in C^2(\mathbb{R}^n, \mathbb{R}^n)$ 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 $\|Dg(\boldsymbol{x}))^{-1}\| \le a_1$;*
2. *si $\boldsymbol{x}, \boldsymbol{y} \in B(\bar{\boldsymbol{x}}, a)$ alors $\|g(\boldsymbol{y}) - g(\boldsymbol{x}) - Dg(\boldsymbol{x})(\boldsymbol{y} - \boldsymbol{x})\| \le a_2\|\boldsymbol{y} - \boldsymbol{x}\|^2$.*

Alors, si on pose : $b = \min\left(a, \dfrac{1}{a_1a_2}\right) > 0$, $\beta = a_1a_2$ 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. *$\|\boldsymbol{x}^{(k+1)} - \bar{\boldsymbol{x}}\| \le \beta\|\boldsymbol{x}^{(k)} - \bar{\boldsymbol{x}}\|^2 \;\; \forall k \in \mathbb{N}$.*

DÉMONSTRATION – Soit $\boldsymbol{x}^{(0)} \in B(\bar{\boldsymbol{x}}, b) \subset B(\bar{\boldsymbol{x}}, a)$ où $b \le a$. 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 $b \le a$, 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 $b \le \dfrac{1}{a_1a_2}$. Par hypothèse, on sait que si $\boldsymbol{x}, \boldsymbol{y} \in B(\bar{\boldsymbol{x}}, a)$, on a
$$\|g(\boldsymbol{y}) - g(\boldsymbol{x}) - Dg(\boldsymbol{x})(\boldsymbol{y} - \boldsymbol{x})\| \le a_2\|\boldsymbol{y} - \boldsymbol{x}\|^2.$$
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 :
$$\|g(\bar{\boldsymbol{x}}) - g(\boldsymbol{x}^{(k)}) - Dg(\boldsymbol{x}^{(k)})(\bar{\boldsymbol{x}} - \boldsymbol{x}^{(k)})\| \le a_2\|\bar{\boldsymbol{x}} - \boldsymbol{x}^{(k)}\|^2.$$
Comme $g(\bar{\boldsymbol{x}}) = 0$ et par définition de $\boldsymbol{x}^{(k+1)}$, on a donc :
$$\|Dg(\boldsymbol{x}^{(k)})(\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}) - Dg(\boldsymbol{x}^{(k)})(\bar{\boldsymbol{x}} - \boldsymbol{x}^{(k)})\| \le a_2\|\bar{\boldsymbol{x}} - \boldsymbol{x}^{(k)}\|^2,$$
et donc
$$\|Dg(\boldsymbol{x}^{(k)})(\boldsymbol{x}^{(k+1)} - \bar{\boldsymbol{x}})\| \le a_2\|\bar{\boldsymbol{x}} - \boldsymbol{x}^{(k)}\|^2. \tag{2.24}$$
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
$$\|\boldsymbol{x}^{(k+1)} - \bar{\boldsymbol{x}}\| \le \|Dg(\boldsymbol{x}^{(k)})^{-1}\| \; \|Dg(\boldsymbol{x}^{(k)})(\boldsymbol{x}^{(k+1)} - \bar{\boldsymbol{x}})\|.$$
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
$$\|\boldsymbol{x}^{(k+1)} - \bar{\boldsymbol{x}}\| \le a_1a_2\|\boldsymbol{x}^{(k)} - \bar{\boldsymbol{x}}\|^2 < a_1a_2b^2. \tag{2.25}$$

Or $a_1 a_2 b^2 < b$ car $b \leq \dfrac{1}{a_1 a_2}$. 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) :

$$a_1 a_2 \|\boldsymbol{x}^{(k+1)} - \bar{\boldsymbol{x}}\| \leq (a_1 a_2)^2 \|\bar{\boldsymbol{x}} - \boldsymbol{x}^{(k)}\|^2 = (a_1 a_2 \|\boldsymbol{x}^{(k)} - \bar{\boldsymbol{x}}\|)^2, \ \forall k \in \mathbb{N},$$

et donc par récurrence

$$a_1 a_2 \|\boldsymbol{x}^{(k)} - \bar{\boldsymbol{x}}\| \leq (a_1 a_2 \|\boldsymbol{x}^{(0)} - \bar{\boldsymbol{x}}\|)^{2^k}, \quad \forall k \in \mathbb{N}$$

Comme $\boldsymbol{x}^{(0)} \in B(\bar{\boldsymbol{x}}, b)$ et $b \leq \frac{1}{a_1 a_2}$, on a $a_1 a_2 \|\boldsymbol{x}^{(0)} - \bar{\boldsymbol{x}}\| < 1$ 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 :

$$\|\boldsymbol{x}^{(k+1)} - \bar{\boldsymbol{x}}\| \leq \beta \|\boldsymbol{x}^{(k)} - \boldsymbol{x}\|^2 \text{ avec } \beta = a_1 a_2.$$

■

Le théorème 2.19 peut aussi se démontrer comme corollaire du théorème 2.20. En effet, sous les hypothèses du théorème 2.19, 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 95 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.*