Résolution d'équations non linéaires: point fixe, Newton

Part 1. Cas scalaire

1. contenu

Le **plan** de cette session, partie I, est le suivant:
- Position du problème,
- Méthode de dichotomie,
- Méthode de point fixe,
- Méthode de Newton,
- Méthode de la sécante.

2. Position du problème

Soit F une fonction à valeurs réelles, $F : I \subset \mathbb{R} \to \mathbb{R}$, F a-priori *non-linéaire* (ex. $\sin(x)$, $\sqrt{x}$ etc).
NB. F linéaire impliquerait que F puisse s'écrire sous la forme $F(x) = a \cdot x$ (ou $ax + b$ pour F affine).
On s'intéresse alors à résoudre l'équation (non linéaire) suivante:

$$\boxed{F(x) = 0}$$

par exemple $\sin(x) = 0$.

Pour *approcher numériquement* la solution de cette équation, on devra construire une suite $(x_k)_{k\in\mathbb{N}}$ qui converge vers *une* solution du problème.

A noter qu'il n'y généralement pas unicité de la solution, et il arrive que l'on en cherche plusieurs voire qu'on les cherche toutes.

Exemple. Comment fait votre calculatrice pour vous donner "la" valeur de $\sqrt{2}$?
La valeur approchée qu'elle indique est en fait le résultat d'un algorithme itératif de la résolution de l'équation non linèaire $x^2 - a = 0$.

Une première méthode de résolution de cette équation consiste à *localiser* une solution "pas à pas" (et à partir d'un point initial choisi).

Avant de présenter cette méthode dite de dichotomie, rappelons tout d'abord le résultat suivant.

Theorem 1. *(Théorème des valeurs intermédiaires). Soit a,b deux réels tels que $a < b$ et $F : [a, b] \to R$ une fonction continue sur $[a, b]$. On suppose que$F(a)F(b) < 0$, alors il existe au moins un $x^* \in]a, b[$ tel que $F(x^*) = 0$.*
Par ailleurs, un tel x^* est unique si F est de plus monotone (soit croissante, soit décroissante).

Exercice. Donner trois exemples illustrant ces propos (avec graphes à l'appui).

3. Méthode de dichotomie

On suppose dorénavant que $F : [a, b] \to R$ est strictement croissante et continue sur $[a, b]$ et est telle que $F(a)F(b) < 0$.

Le théorème précédent est alors à la base de la *méthode de dichotomie.*

Notons $a_0 = a$ et $b_0 = b$ et introduisons le point milieu $c = \frac{a_0+b_0}{2}$.

- Si $F(a_0)F(c) < 0$ alors nécessairement la solution $x^* \in [a_0, c]$. On pose alors $a_1 = a_0$ et$b_1 = c$ et on ré-itère le raisonnement.

- Sinon, on a nécessairement $x^* \in [c, b_0]$ et on pose $a_1 = c$ et $b_1 = b_0$.

Dans les deux cas, on obtient$x^* \in [a_1, b_1]$.

On itère ensuite ce même procédé par récurrence, ce qui donne l'algorithme de dichotomie.

> *Algorithme de dichotomie.* Supposons avoir construit a_n et b_n tels que $x^* \in [a_n, b_n]$.
> On pose $c = \frac{a_n+b_n}{2}$.
> Si $f(a_n)f(c) < 0$, on pose $a_{n+1} = a_n$ et $b_{n+1} = c$;
> Sinon, on pose $a_{n+1} = c$ et $b_{n+1} = b_n$.

Par récurrence, on montre que:
– La suite $(a_n)_{n\in N}$ est croissante, la suite $(b_n)_{n\in N}$ est décroissante,
– $|b_n - a_n| = \frac{b-a}{2^n}, \forall n \in \mathbb{N}$
– $x^* \in [a_n, b_n], \forall n \in \mathbb{N}$

Les suites $(a_n)_{n\in N}$ et $(b_n)_{n\in N}$ sont donc adjacentes, et convergent donc vers une même limite $l = x^*$.

De plus, si on se sert de $(a_n)_{n\in N}$ comme de la suite approchant x^*, on a la majoration suivante:

$$|a_n - x^*| \leq \frac{b-a}{2^n}, \forall n \in \mathbb{N}$$

Cet algorithme de dichotomie est robuste mais converge très lentement (de manière logarithmique).

4. Méthode de point fixe

Une alternative pour résoudre cette équation non linéaire $F(x) = 0$ est de reformuler l'equation sous forme de point fixe: $x = G(x)$ avec $G : J \subset \mathbb{R} \to \mathbb{R}$. Il suffit pour cela de poser $G(x) = x + F(x)$, ou encore:

$$\boxed{x = G_\lambda(x) \text{ avec } G_\lambda(x) = x + \lambda F(x)}$$

et $\lambda \in \mathbb{R}^*$.

L'intérêt de cette reformulation réside dans le *théorème de point fixe* que nous rappelons:

Theorem. *Soit $I = \mathbb{R}$ ou $I = [a, b]$ avec $a < b$ deux réels. Soit G une application de I dans $\mathbb{R}$ telle que:*
i) $G(I) \subset I$,
ii) il existe une constante k, $k \in [0, 1[$, telle que G est k-contractante, c'est-à-dire:

$$\forall (x, y) \in I^2, \;\; |G(x) - G(y)| \leq k|x - y|$$

Alors la fonction G admet un unique point fixe $x^* \in I$ tel que $G(x^*) = x^*$.
De plus, toute suite récurrente définie par

$$x_0 \in I; \;\; x_{n+1} = G(x_n), n \in \mathbb{N}$$

converge vers x^*.

D'où *l'algorithme de point fixe* appliqué à G_λ donne:

$$\boxed{x_0 \text{ donné}, \; x_{n+1} = x_n + \lambda F(x_n) \;, n \geq 0}$$

5. Méthode de point fixe (suite)

La méthode du point fixe s'illustre ainsi.

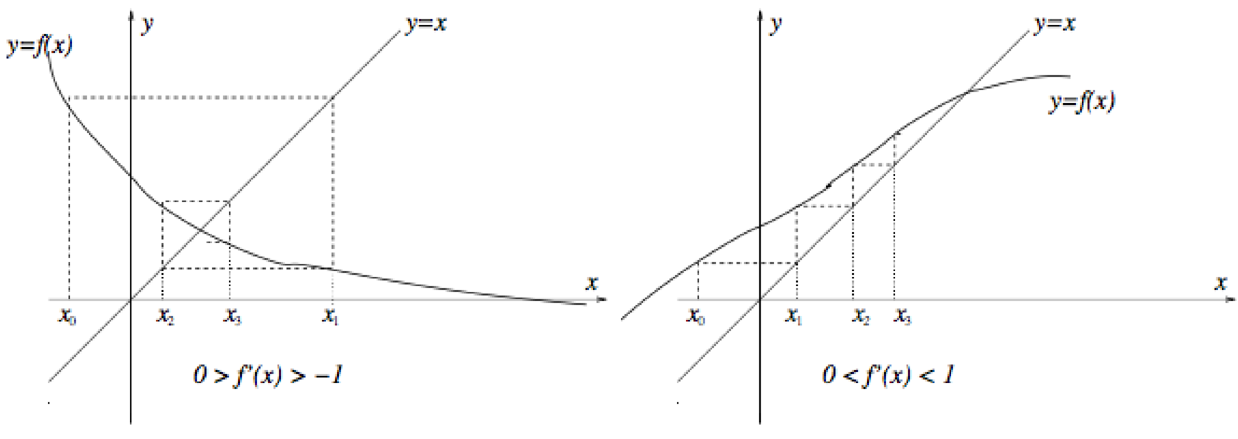


FIGURE 3.4 – Premiers itérés d'une méthode de point fixe.

Vitesse de convergence.

En choisissant λ tel que $G_\lambda(x)$ strictement contractante (de constante de Lipschitz $k(\lambda) < 1$), on montre à l'aide du théorème des accroissements finis que:

$$|x_n - x^*| \leq \frac{k(\lambda)^n}{(1 - k(\lambda))}|x_1 - x^*|$$

La vitesse de convergence de cet algorithme est alors d'ordre 1; elle est d'autant plus rapide que $k(\lambda)$ est petit.

On a alors un "degré de liberté" dans le choix du λ pour tenter d'accélérer l'algorithme.

Aussi, on pourrait tout aussi bien considérer une version généralisée de l'algorithme comme suit :$x_{n+1} = x_n + \lambda_n F(x_n)$, du moment que la fonction reste contractante, à savoir $|1 + \lambda_n F'(x_n) \leq k < 1, \; \forall n \in \mathbb{N}$.

6. Méthode de Newton

Le principe de la méthode de Newton est non plus de chercher à "résoudre" $F(x) = 0$ mais plutôt de résoudre l'équation *linéarisée.*

Pour cela on écrit un Développement Limité (D.L.) de Taylor à l'ordre 1 au point x_n:

$$F(x_n + h) = F(x_n) + h\,F'(x_n) + h\,\varepsilon(h)$$

Pour h petit (D.L. au voisinage du point x_n), $h\varepsilon(h)$ est encore plus petit (terme d'ordre deux et plus), on néglige alors le reste du D.L.$h\,\varepsilon(h)$ et on cherche l'"incrément" h tel que: $F(x_n + h) = 0$.

Ce qui nous conduit alors à résoudre à chaque itération l'équation $F(x_n) + hF'(x_n) = 0$.

Soit en posant $x_{n+1} = x_n + h$,

$$\boxed{x_{n+1} = x_n - \frac{F(x_n)}{F'(x_n)}\ , n \geq 0; x_0 \text{ donné}}$$

pour $F'(x_n) \neq 0$.

La force de la méthode de Newton est sa convergence rapide, si convergence...

Son point faible est le caractère *local* de sa convergence i.e. il faut bien choisir le point initial x_0 pour espérer une convergence ou encore une convergence vers telle ou telle solution.

7. Méthode de Newton: interprétation graphique

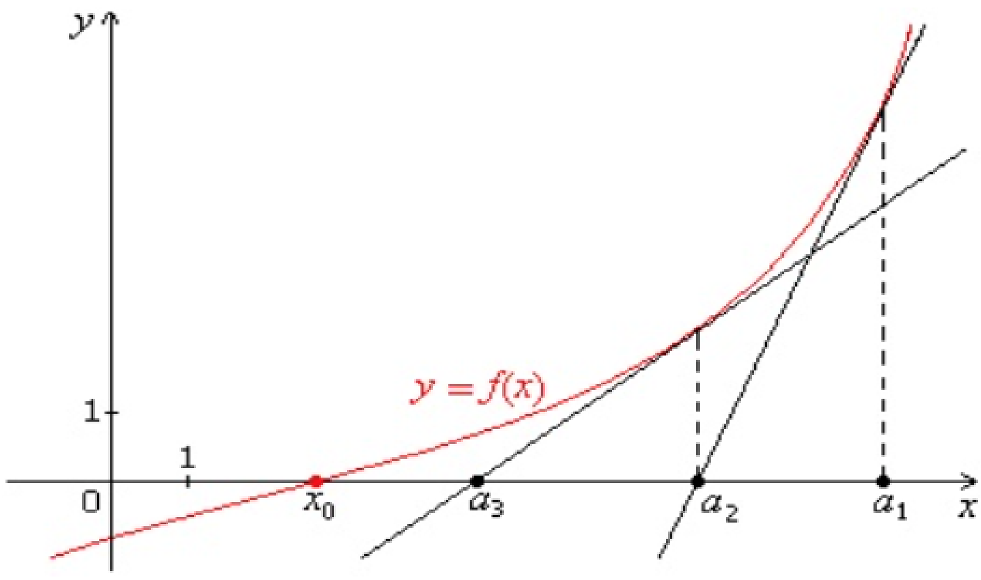
Avant d'étudier la convergence de cette méthode, nous en présentons une interprétation graphique.

L'algorithme de Newton peut s'écrire sous la forme suivante:

$$F(x_n) + F'(x_n)(x_{n+1} - x_n) = 0$$

Autrement dit, x_{n+1} est le point d'intersection entre l'axe des y et l'équation de droite $y = F(x_n) + F'(x_n)(x - x_n)$, cf Fig.

Cette droite étant la tangente à la courbe $y = f(x)$ au point x_n .



8. Méthode de Newton: convergence

Theorem. *Supposons que F soit de classe C^2; on suppose de plus que x^* est une racine simple de $F(x) = 0$, ce qui signifie que $F(x^*) = 0$ et $F'(x^*) \neq 0$.*

Alors l'algorithme de Newton converge localement, et de manière quadratique (i.e. d'ordre 2).

La convergence locale signifie qu'il existe $\eta > 0$ tel que si $|x^* - x_0| < \eta$, alors la suite définie par l'algorithme ($x_{n+1} = x_n - \frac{F(x_n)}{F'(x_n)}$; x_0 donné) converge vers x^*.

Une vitesse de convergence d'ordre 2 signifie qu'il existe $C > 0$ tel que:

$$|x_n - x^*| \leq C|x_{n-1} - x^*|^2$$

En pratique, cela signifie qu'à chaque itération le nombre de digits exact est doublé (en effet si $|x_{n-1} - x^*| = 10^{-p}$ alors $|x_n - x^*| \approx 10^{-2p}$), soit une convergence extrêmement rapide (4-5 itérations pour un résidu très faible); si convergence...

Nous vous invitons à étudier la démonstration de ce théorème dans le polycopié de cours.

9. Variante de Newton: méthode de la sécante

La méthode de Newton se généralise aux *systèmes* d'équations non linéaires (soit m équations non linéaires à m inconnues), c'est la partie 2 à venir de cette session.

Aussi le calcul de $F'(x_n)$ (soit la matrice jacobienne $DF(x_n)$ dans le cas de systèmes) peut être délicat à mener.

Une alternative simple consiste alors à a*pprocher la dérivée* par une différence divisée, ce qui donne l'algorithme suivant.

$$\boxed{x_{n+1} = x_n - F(x_n)\frac{x_n - x_{n-1}}{F(x_n) - F(x_{n-1})}\ , n \geq 0; x_0 \text{ donné}}$$

Il s'agit de la *méthode de la sécante.*

On peut montrer que comme pour la méthode de Newton, la convergence de cette variante est uniquement locale.

Par ailleurs, sa vitesse de convergence est égal au nombre d'or $(1 + \sqrt{5})/2$, soit mieux que celle de la méthode de point fixe qui est linéaire mais moins bien que celle de Newton qui est quadratique.

Part 2. Cas des systèmes