

Le 27 janvier 2012

**Mathématiques appliquées
et numériques**
Licence 3/ TAO (2011-2012)
Présentation synthétique du cours

Janvier 2012

Cours donné en 3^e année
de Licence de Sciences de la planète Terre
par Michael Ghil et Jean Roux
École normale supérieure, Paris

Premier cours

Résolution équations non-linéaires

Il s'agit dans ce premier cours de calculer x telle que $f(x) = 0$. Les notions de dérivée, de norme vectorielle (que l'on rappelle) et de matrice sont nécessaires.

1.1 Prérequis

Notions de dérivée et de matrice.

Normes vectorielles

Définition 1.1.1. : Soit X un espace vectoriel de dimension finie sur le corps K des scalaires. On appelle norme vectorielle toute application de X sur \mathbb{R}_+ , notée $x \rightarrow \|x\|$, possédant les propriétés suivantes :

- 1. $\|x\| > 0 \Leftrightarrow x \neq 0$ et $\|0\| = 0$
- 2. $\|\alpha x\| = |\alpha|\|x\|$, $\forall \alpha \in K$
- 3. $\|x + y\| \leq \|x\| + \|y\|$, $\forall x, y \in X$

1.2 Résolution explicite

C'est rarement possible.

Pour les polynômes du premier et du second degré nous avons :

$$\begin{aligned} P_1(x) \equiv ax + b = 0 & \quad \rightarrow \quad x = -b/a; \\ P_2(x) \equiv ax^2 + bx + c = 0 & \quad \rightarrow \quad x = (-b \pm \sqrt{b^2 - 4ac})/2a. \end{aligned}$$

Pour le troisième degré,

$$P_3(x) \equiv ax^3 + bx^2 + cx + d = 0,$$

nous avons (depuis 1535) les formules de J. Cardan pour une équation du type $z^3 + pz + q = 0$, toute équation générale du type $P_3(x) = 0$ pouvant s'écrire sous cette forme par le changement de variable $x = z - b/3a$.

Les équations polynomiales de degré quatre sont résolubles par radicaux par les méthodes de L. Ferrari ou de R. Descartes.

Par contre il a été rigoureusement démontré par N. H. Abel (en 1824) que, pour les degrés cinq et plus, il n'existe plus de telles solutions explicites.

En pratique nul n'utilise aujourd'hui les solutions analytiques au-delà du second degré. D'où l'examen de méthodes numériques.

Commençons donc par le cas le plus simple, celui de la recherche des racines d'une équation scalaire.

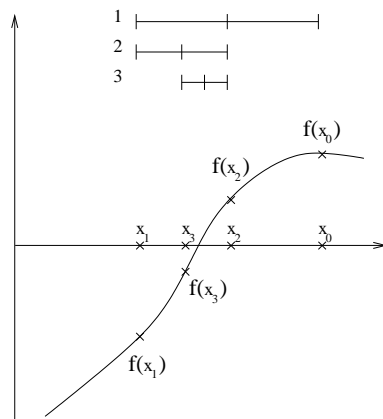


Figure 1.1: Méthode de la dichotomie

1.3 Méthodes numériques de résolution d'équations scalaires

On cherche à résoudre numériquement

$$f(x) = 0, x \in \mathbb{R},$$

f étant à valeurs réelles et continue.

Supposons que $f \not\equiv 0$ et qu'il existe $x_0, x_1 \in \mathbb{R}$ tels que

$$f(x_0) < 0 \text{ et } f(x_1) > 0 \quad ;$$

alors le théorème des valeurs intermédiaires nous assure de l'existence d'une solution x^* , située entre x_0 et x_1 .

Remarquons que tout polynôme réel de degré impair possède cette propriété et possèdera donc une solution réelle pouvant être calculée par les algorithmes qui suivent. Évidemment cela ne veut pas dire que nous pourrions calculer ainsi toutes les racines, en particulier celles qui sont complexes. Par exemple l'équation $x^2 + 1 = 0$ à coefficients réels n'a pas de solution réelle, les deux solutions complexes $+i$ et $-i$ ne pourront pas être estimées par les méthodes numériques ci-après.

1.3.1 Méthode de la dichotomie

Cette méthode est aussi appelée méthode de la *bissection*. Pour l'appliquer, il faut partir de deux points, $x_0 \neq x_1$, tels que $f(x_0)f(x_1) < 0$.

De façon générale, on notera $f_k \equiv f(x_k)$, x_k étant l'abscisse du $k^{\text{ième}}$ point obtenu par une méthode donnée de recherche de x^* .

Dans cette méthode, on utilise comme x_2 le point situé au milieu du segment $[x_0, x_1]$, $x_2 = (x_0 + x_1)/2$. Pour continuer, on vérifie lequel des deux

points initiaux a l'image de signe opposé à f_2 – c'est-à-dire $f_0 f_2 < 0$ ou $f_1 f_2 < 0$ –, et on retient celui-là.

La même méthode est itérée avec les deux nouveaux points. A chaque itération le milieu est utilisé comme nouveau point, ainsi que le point dont l'image est de signe opposé à celle du point milieu. Avec cette méthode il est certain que la distance entre deux points diminue de moitié à chaque itération.

Cet algorithme revient à sélectionner à chaque étape un intervalle qui contient toujours au moins un zéro de f .

On vérifie que $s_{k+1} - r_{k+1} = (s_k - r_k)/2 = (s_0 - r_0)/2^{k+1}$, la distance entre r_k et s_k tend rapidement vers zéro, c'est-à-dire que

$$\lim_{n \rightarrow \infty} (s_k - r_k) = 0. \quad (1.3.1)$$

On démontre la convergence des suites (r_k) et (s_k) vers un même point x^* qui est une racine de f . La méthode est illustrée à la Figure 1.1.

1.3.2 Méthode de la “fausse position”

On suppose toujours $f \in \mathcal{C}^0$ et on se donne x_0 et x_1 tels que $x_1 < x_0$ et $f_1 f_0 < 0$. La méthode de la “fausse position” est semblable à la méthode de la bisection sauf que, à chaque pas, on ne détermine pas z_k comme le point milieu de l'intervalle $[r_k, s_k]$ mais par la formule suivante

$$z_k = \frac{r_k f(s_k) - s_k f(r_k)}{f(s_k) - f(r_k)}. \quad (1.3.2)$$

Cette formule est basée sur l'idée simple suivante. Soit $f_1 f_0 < 0$, on détermine la fonction linéaire $F(x)$ telle que $F(x_0) = f_0$ et $F(x_1) = f_1$. Le point c tel que $F(c) = 0$ est donné par

$$c = \frac{x_1 f(x_0) - x_0 f(x_1)}{f(x_0) - f(x_1)}; \quad (1.3.3)$$

notons que le dénominateur est non nul car $f_1 f_0 < 0$. En posant $r_k = x_1$, $s_k = x_0$ et $z_k = c$ on obtient (1.3.2).

1.3.3 Méthode de la sécante

La méthode de la sécante est très proche de la méthode de la fausse position, sauf qu'elle s'applique même si $f_1 f_0 > 0$. Elle retient systématiquement les deux derniers points calculés. Dans (1.3.3) en prenant les itérés x_{k-1} , x_k et les valeurs associées f_{k-1} et f_k de la fonction f , en posant $x_{k+1} = c$, il vient pour $n = 1, 2, \dots$

$$x_{k+1} = \frac{x_k f_{k-1} - x_{k-1} f_k}{f_{k-1} - f_k}. \quad (1.3.4)$$

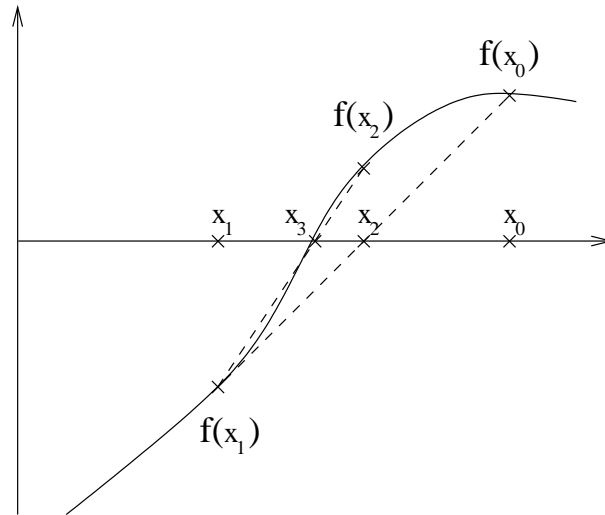


Figure 1.2: Méthode de la sécante

Évidemment la méthode de la sécante avorte si $f_k = f_{k-1}$. Si par ailleurs la méthode converge alors f_k et f_{k-1} convergent vers zéro et le dénominateur de (1.3.4) tend vers zéro.

Pour éviter cette difficulté, et en supposant que $f' \in \mathcal{C}^0$, autrement dit que $f \in \mathcal{C}^1$, on écrit

$$x_{k+1} = x_k - \frac{x_k - x_{k-1}}{f_k - f_{k-1}} f_k. \quad (1.3.5)$$

La méthode est illustrée à la Figure 1.2. Elle ne donne clairement pas satisfaction quand $f'(x^*)$ devient trop petit et il faut donc surveiller cet aspect dans toute programmation de la méthode.

La méthode de la sécante converge usuellement plus vite que la méthode de bisection ou de la fausse position, mais elle peut ne pas converger même si ces dernières convergent.

1.3.4 Méthode itérative de Picard ou des approximations successives

C'est une méthode itérative de type *point fixe*. Soient $f : \mathbb{R} \rightarrow \mathbb{R}$ et la racine ξ de l'équation $f(x) = 0$ unique dans l'intervalle $I = [a, b]$.

L'idée de base est de remarquer que trouver ξ comme racine de f est équivalent à la trouver comme solution de $\xi = F(\xi)$, en définissant

$$F(x) \equiv x + Mf(x). \quad (1.3.6)$$

On résout (1.3.6) par la méthode du point fixe qui consiste à choisir x_0

arbitrairement dans I et à calculer x_k récursivement par

$$x_k = F(x_{k-1}), \quad k = 1, 2, \dots \quad (1.3.7)$$

Cette méthode est très générale. Dans le cas unidimensionnel, $n = 1$, par lequel nous avons commencé ici, un exemple de convergence de la méthode est illustré dans la Figure 1.3. Il est facile de vérifier que, dans cet exemple, $|F'(\xi)| < 1$ et de se convaincre, par le même type de croquis, qu'il y a divergence de l'itération (1.3.7) si $|F'(\xi)| > 1$.

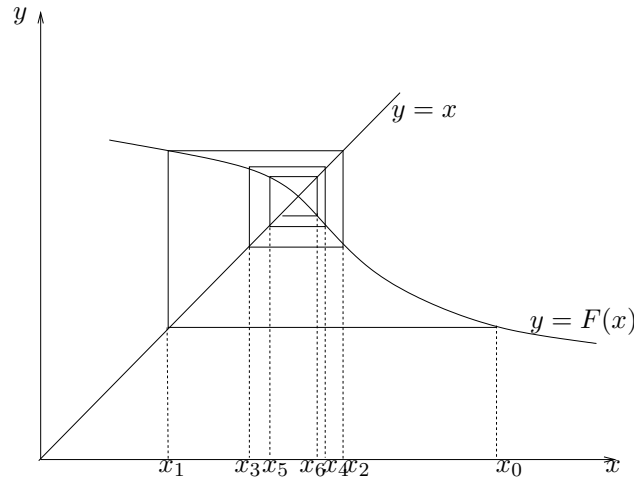


Figure 1.3: Itérés de l'algorithme du point fixe

En effet, il n'est pas nécessaire que f , et donc F , soient dérivables : il suffit que F soit une contraction lipschitzienne pour que le point fixe existe, soit unique et puisse être calculé par l'itération de point fixe. La situation est donc résumée par le théorème suivant :

Théorème 1.3.1. *Soient $I = [a, b]$ un intervalle de \mathbb{R} et F une fonction satisfaisant aux deux conditions suivantes :*

- i) $F(x) \in I$ pour tout $x \in I$,*
- ii) F vérifie la condition de Lipschitz*

$$|F(x_1) - F(x_2)| \leq L|x_1 - x_2| \quad (1.3.8)$$

pour tous $x_1, x_2 \in I$ avec une constante de Lipschitz $L < 1$.

Alors, pour tout choix de x_0 dans I , la suite définie par $x_k = F(x_{k-1})$, $k = 1, 2, \dots$, converge vers l'unique solution de l'équation $x = F(x)$.

Les deux conditions (i) et (ii) ci-dessus définissent une *application contractante* sur I .

La Figure 1.4 montre trois itérations de Picard, avec deux points attracteurs séparés par un point répulsif.

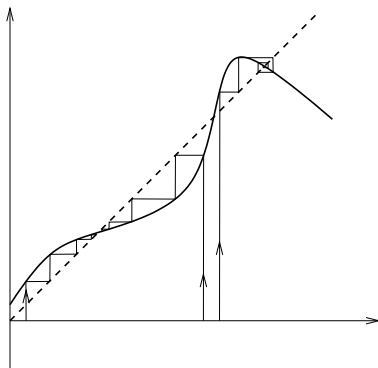


Figure 1.4: Trois itérations de Picard

1.3.5 Méthode de Newton

Soit toujours le calcul du zéro, supposé pour l'instant unique, de la fonction f dans $I = [a, b]$. Nous essayons ici de nous affranchir de la condition, assez contraignante, de l'existence d'une condition contractante de Lipschitz ($L < 1$). Ceci même si l'on perd la convergence "globale" de la méthode.

Pouvons-nous déterminer une fonction $h \in \mathcal{C}^1$ – dépendante de f , mais aisément calculable –, de sorte que le processus précédent appliqué à la fonction F ,

$$F(x) \equiv x + h(x)f(x),$$

converge vers ξ ? Ce qui revient à remplacer la constante M de (1.3.6) par la fonction $h(x)$.

Une façon de faire est de choisir $h(x) = -1/f'(x)$. Nous sommes alors conduits à l'algorithme itératif suivant, de type (1.3.7) :

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \quad k = 0, 1, 2, \dots \quad (1.3.9)$$

C'est la *méthode de Newton*. La convergence est *locale* mais elle est très rapide.

Géométriquement cette méthode est équivalente au remplacement d'un petit arc de la courbe $y = f(x)$ par la tangente menée par un certain point de la courbe. Voir la Figure 1.5 où l'on suppose, pour fixer les idées que $f'' > 0$ pour $x \in [a, b]$ et $f(b) > 0$.

1.4 Méthodes multidimensionnelles

On étend à la dimension n ce que l'on vu aux paragraphes 1.3.4 et 1.3.5, respectivement. On considère donc l'application $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, pour $n \geq 2$, qui généralise le cas scalaire de $n = 1$, traité jusqu'ici.

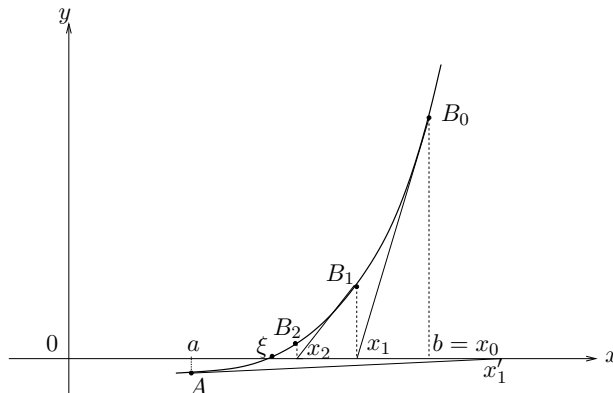


Figure 1.5: Itérés de l'algorithme de Newton

Nous mentionnerons enfin le cas – très intéressant en pratique –, où la dérivée de cette application f n'est pas connue analytiquement, ou qu'elle est plutôt coûteuse à calculer à partir de son expression analytique. Cette dernière situation survient surtout pour n très grand, comme c'est le cas dans la plupart des applications aux sciences de la planète et de la vie.

1.4.1 Applications contractantes et points fixes

Supposons que l'on ait $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ telle que

$$\exists L < 1, \text{ tel que } \forall (x, y) \in \mathbb{R}^{2n} \quad \|f(x) - f(y)\| \leq L\|x - y\|. \quad (1.4.1)$$

Il s'agit encore d'une application contractante et le théorème suivant s'applique:

Théorème 1.4.1. *Soit f une application contractante, alors*

1. *Il existe une solution unique x^* de*

$$f(x) = x.$$

2. *$\forall x_0 \in \mathbb{R}^n$, la suite*

$$x_{k+1} = f(x_k)$$

converge vers le point fixe x^ .*

1.4.2 Méthode de Newton multidimensionnelle

La méthode (1.3.9) suggère, pour tout indice $k \geq 0$, de définir la méthode itérative

$$x_{k+1} = x_k - (f'(x_k))^{-1}f(x_k), \quad (1.4.2)$$

où $f'(x_k)$ désigne la matrice jacobienne de f au point itéré x_k , soit

$$f'(x_k) \equiv \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}, \quad (1.4.3)$$

toutes les dérivées partielles étant estimées au point x_k . La formule (1.4.2) suppose évidemment que $f'(x_k)$ est inversible, ce qui doit être vérifié et justifié.

Il est judicieux, plutôt que de calculer l'inverse de la matrice jacobienne, de résoudre le système linéaire $f'(x_k)(x_{k+1} - x_k) = -f(x_k)$ à chaque étape.

Par contre, comme on peut s'attendre par ailleurs à ce que, moyennant certaines hypothèses, les matrices $(f'(x_k))^{-1}$ varient "peu" (surtout au voisinage du zéro), il est naturel de songer à remplacer cette suite de matrices par une matrice approximante, calculée une fois pour toutes, par exemple $(f'(x_0))^{-1}$; ou encore on peut imaginer de garder le même élément pendant un nombre fixé d'itérations successives.

C'est pour prendre en compte l'éventualité de telles variantes que l'on définit généralement la méthode de Newton sous la forme suivante. On considère l'ensemble des applications linéaires continues de \mathbb{R}^n dans \mathbb{R}^n et on se donne une suite (A_k) d'éléments de cet ensemble supposés continuellement inversibles. On définit alors la suite (x_k) par

$$x_{k+1} = x_k - A_k^{-1}f(x_k), \quad k \geq 0, \quad (1.4.4)$$

x_0 étant un point arbitraire de Ω fermé borné de \mathbb{R}^n .

Naturellement, en l'absence d'hypothèses, rien ne garantit que la suite (x_k) reste dans l'ouvert Ω , et qu'elle converge vers x^* racine de $f(x) = 0$.

La preuve de la convergence de la suite (1.4.4) vers x^* est non triviale, même si elle utilise aussi une technique de point fixe.

Comme dans le cas $n = 1$ la convergence est locale, c'est le propre des méthodes de Newton.

1.4.3 Méthodes quasi-Newton

La méthode de Newton exige l'évaluation de $n^2 + n$ fonctions scalaires – n^2 pour $f'(x_k)$ et n pour $f(x_k)$ –, et la résolution d'un système linéaire en $O(n^3)$ opérations élémentaires par itération. C'est pour obvier à ces inconvénients que les méthodes appelées *quasi-Newton* ont été construites.

L'idée est d'approcher les matrices jacobienes $f'(x_k)$ par des opérateurs matriciels B_k , tels qu'il ne soit jamais nécessaire de résoudre un système linéaire en $O(n^3)$ opérations élémentaires et de calculer des dérivées.

On calcule B_k^{-1} en fonction de B_{k-1}^{-1} . L'intérêt est, de plus, que cette actualisation de la matrice inverse se fait en seulement $O(n^2)$ opérations élémentaires : pour $n \geq 100$ grand, le gain est déjà considérable !