Analyse Numérique

Ian Schindler

27 janvier 2023

1 Introduction

Ce document s'appuie essentiallement sur [1].

2 Calculs Numériques Approchés

2.1 Définitions

Pour $x \in \mathbb{R}^*$ on a

$$x = \pm b^p m = \pm b^p \sum_{k=1}^{\infty} a_k b^{-k}, \tag{1}$$

ou $b \in \mathbb{N}^*$, $a_k \in \{0, \ldots, b-1\}$, $m := \sum_{k=1}^{\infty} a_k b^{-k}$ et $p \in \mathbb{Z}$. Puisque la capacité de la mémoire de l'ordinateur est finie, il retient

$$x \approx x' = \pm b^p m' = \pm b^p \sum_{k=1}^{N} a_k b^{-k},$$

ou $m' := \sum_{k=1}^{N} a_k b^{-k}$ et $0 < m \leq 1$. Pour ne pas perdre de précision on prend $a_1 \neq 0$ ce qui fixe p de façon unique. Définitions :

- b est la *base*, pour l'ordinateur, en général, $b = 2$ ou 16. On pourra utiliser $b = 10$ pour des exercices.
- m' est la *mantisse*, N est la *longueur* de la mantisse.

L'erreur en x est :

$$\begin{aligned}
\Delta x &:= |x - x'| \qquad (2)\\
&= b^p \sum_{k=N+1}^{\infty} a_k b^{-k} \\
&\leq b^p \sum_{k=N+1}^{\infty} (b-1) b^{-k} \\
&= b^p b^{-N}. \qquad (3)
\end{aligned}$$

Donc

$$\begin{aligned}
\frac{\Delta x}{|x|} &= \frac{\Delta m b^p}{m b^p} \\
&\leq \frac{b^{-N}}{b^{-1}} \\
\frac{\Delta x}{|x|} &\leq b^{1-N} = \varepsilon.
\end{aligned} \tag{4}$$

On appelle ε est la *précision relative.*

2.2 Arrondie

Pour réduire ε dans (4) on utilise l'arrondie. Pour x comme dans (1) on prend

$$x' = \begin{cases} b^p \sum_{k=1}^{N} a_k b^{-k} & \text{si } a_{N+1} < b/2 \\ b^p((\sum_{k=1}^{N-1} a_k b^{-k}) + (a_N + 1)b^{-N}) & \text{si } a_{N+1} \geq b/2. \end{cases} \tag{5}$$

2.3 Non associativité

Exemple : on prend $b = 10$ et $N = 3$. Alors, si $x = 8,22$, $y = 0,00317$ et $z = 0,00432$, on a

$$(x+y)' = 8,22317' = 8,22$$

et donc $((x + y)' + z)' = 8,22$. Par contre,

$$(y+z)' = 0,749 \times 10^{-2}$$

et donc en utilisant la régle d'arrondie, $(x+(y+z)')' = 8,23 \neq ((x+y)'+z)'$.

2.4 Erreur d'une Somme

L'équation (4) implique que

$$\begin{aligned}
\Delta(x' + y') &= |(x' + y')' - (x' + y')| \\
&\leq \varepsilon |x' + y'| \\
&\leq \varepsilon(|x'| + |y'|) \\
&\leq \varepsilon(|x| + \Delta x + |y| + \Delta y) \\
&\approx \varepsilon(|x| + |y|).
\end{aligned}$$

Donc

$$\begin{aligned}
\Delta(x + y) &= |(x + y) - (x' + y')'| \\
&= |(x + y) - (x' + y') + (x' + y') - (x' + y')'| \\
&\leq |x - x'| + |y - y'| + |(x' + y') - (x' + y')'| \\
&\leq \Delta x + \Delta y + \varepsilon(|x| + |y|) \\
&\leq \Delta x + \Delta y + \varepsilon(|x| + |y|)
\end{aligned} \tag{6}$$

Exemple :

$$S_n := \sum_{i=1}^{n} u_i \tag{7}$$

On suppose $\Delta u_i = 0$ et on note $S_k := \sum_{i=1}^{k} u_k$, $k \in \{1, \ldots, n\}$. Donc

$$\begin{aligned} \Delta S_k &\leq \Delta S_{k-1} + \varepsilon(S_{k-1} + u_k) && (8) \\ &= \Delta S_{k-1} + \varepsilon S_k. && (9) \end{aligned}$$

Par récurrence, on a

$$\begin{aligned} \Delta S &\leq \varepsilon(S_2 + \cdots + S_n) \\ &= \varepsilon(u_n + 2u_{n-1} + \cdots (n-1)u_2 + (n-1)u_1). \end{aligned}$$

Petites valeurs d'abord !

2.5 Erreur Produit

$$\begin{aligned} \Delta(x'y') &= |x'y' - (x'y')'| \\ &\leq \varepsilon|x'y'| \\ &\leq \varepsilon|(x + \Delta x)(y + \Delta y)| \\ &\approx \varepsilon|xy|. \end{aligned}$$

et

$$\begin{aligned} |x'y' - xy| &= |x'y' - x'y + x'y - xy| \\ &= |x'\Delta y + y\Delta x| \\ &\leq |x'|\Delta y + |y|\Delta x \\ &\leq (x + \Delta x)\Delta y + |y|\Delta x \\ &\approx |x|\Delta y + |y|\Delta x. \end{aligned}$$

D'ou

$$\begin{aligned} \Delta(xy) &= |xy - (x'y')'| \\ &= |xy - xy' + xy' - x'y' + x'y' - (x'y')'| \\ &\leq |xy - x'y'| + |xy' - x'y'| + |x'y' - (x'y')'| \\ &\leq |x|\Delta y + |y'|\Delta x + \varepsilon|xy| \\ &\approx |x|\Delta y + |y|\Delta x + \varepsilon|xy| && (10) \end{aligned}$$

2.6 Cumulation d'Erreurs Aléatoires

$$\begin{aligned} S &:= \sum_{i=0}^{\infty} u_i,\ u_i \geq 0, \\ S_k &:= \sum_{i=0}^{k} u_i = S_{k-1} + u_k. \end{aligned}$$

On suppose $\Delta u_i = 0$. On pose $\Delta S_k = \Delta S_{k-1} + \alpha_k$. L'inégalité (8) (et sa négative) implique :

$$\begin{aligned} |\alpha_k| &\leq \varepsilon S_k \\ &\leq \varepsilon S. \end{aligned}$$

On a

$$\begin{aligned} \Delta S_n &= \sum_{i=1}^{n} \alpha_i \\ &\leq n\varepsilon S. \end{aligned}$$

Hypothèse :

1. α_i sont des V.A. globalement indépendantes.
2. $E(\alpha_i) = 0$.

d'ou $\mathrm{var}(\Delta S_n) = \sum_{i=1}^{n} \mathrm{var}(\alpha_i)$. Mais $\mathrm{var}(\alpha_i) \leq \varepsilon^2 S^2$, et Bienaymé- Tchebychev :

$$\P(X \geq \beta) \leq \frac{\sigma^2}{\beta^2}\ \beta > 0$$

implique

$$\P(|\Delta S_n| \geq \alpha\sigma(\Delta S_n)) \leq \alpha^{-2}.$$

Exemple :

$$\P(|\Delta S_n| \geq 10\sqrt{n}\varepsilon S) \leq 0,01.$$

2.7 Phénomène de Compensation

Exemple : $b = 10$, $N = 10$, $x^2 - 1634x + 2 = 0$. $\Delta = 667487$. $\sqrt{\Delta} \approx 816,9987760$.

$$\begin{aligned} x_1 &= 817 + \sqrt{\Delta} = 1633,998776 \\ x_2 &= 817 - \sqrt{\Delta} = 0,0012240. \end{aligned}$$

Ici $x_1 x_2 = 2$, $x_2 = 2/x_1 \approx 1.223991125 \times 10^{-3}$.

3 Approximation Polynomiale

Notations : $\mathbb{R}_n[x] := \{$polynômes à coefficients réels de degré $\leq n\}$, dim $\mathbb{R}_n[x] = n+1$.
$f : [a,b] \mapsto \mathbb{R}$, alors

$$\|f\|_{[a,b]} := \sup_{x\in[a,b]} |f(x)|. \tag{11}$$

Propriétés :

1. $\|f\| \geq 0$, $\|f\| = 0 \implies f = 0$.
2. $\|\lambda f\| = |\lambda|\|f\|$.
3. $\|f+g\| \leq \|f\| + \|g\|$.
4. $\|fg\| \leq \|f\|\|g\|$.

$C([a,b]) := \{f : [a,b] \mapsto \mathbb{R}/f \text{ continue}\}$
$C^n([a,b]) := \{f : [a,b] \mapsto \mathbb{R}/f^{(n)} \text{ continue}\}$.

3.1 Méthode d'Interpolation de Lagrange

Pour la suite on pose :

$$\pi_{n+1}(x) := \prod_{i=0}^{n}(x - x_i) \in \mathbb{R}_{n+1}[x]. \tag{12}$$

3.1.1 Existence et unicité

Soit $f \in C([a,b])$ et soient $x_0, x_1, \dots, x_n \in [a,b]$, $x_i \neq x_j$ si $i \neq j$.

Problème : trouver $p_n \in \mathbb{R}_n[x]$ tel que $p_n(x_i) = f(x_i)\, \forall i \in \{0,1,\dots,n\}$.

On dit que p_n *interpole* f aux points x_i, ou bien p_n est le *polynôme d'interpolation* de f aux points x_i.

Posons

$$l_i := \prod_{j\neq i} \frac{(x - x_j)}{(x_i - x_j)}\ i = 0, 1, \dots, n \tag{13}$$

$l_i(x_j) = 0$ si $j \neq i$. $l_i(x_i) = 1$.

$$p_n(x) := \sum_{i=0}^{n} f(x_i) l_i(x) \in \mathbb{R}_n[x]. \tag{14}$$

Théorème 1
Soit $f \in C([a,b])$ et $x_i \in [a,b]$, $i = 0, 1, \dots, n$. Alors il existe un polynôme unique dans $\mathbb{R}_n[x]$ qui interpole f au points x_i donné par (14).

Preuve. L'existence est fait. Soit $q_n \in \mathbb{R}_n[x]$ une autre solution, donc $p_n(x_i) = q_n(x_i) = f(x_i)$ pour $i = 0, 1, \dots, n$. Donc x_i est racine de $p_n(x) - q_n(x)$ pour $i = 0, \dots, n$. Par conséquence, $\pi_{n+1}(x)$ divise $p_n(x) - q_n(x)$. Puisque π_{n+1} est de degré $n+1$ et $p_n - q_n$ est de degré n, $p_n(x) - q_n(x) = 0$. ∎

Remarque 1
Autre démonstration : resoudre le système

$$\sum_{j=0}^{n} a_j x_i^j = f(x_i)\ 0 \leq i \leq n. \tag{15}$$

Il y a une solution unique car le déterminant du système, dit le déterminant de Van der Monde

$$\begin{vmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{vmatrix} \neq 0. \tag{16}$$

Cette méthode n'est pas utilisé car le déterminant de Van der Monde est numériquement instable.

3.2 Méthode de Différence Divisée

Soit p_n le polynôme d'interpolation de f aux points $x_0, \cdots, x_n$. On note $f[x_0, \cdots, x_k]$ le coefficient directeur de p_k. Le polynôme $p_k - p_{k-1}$ s'annule aux points $x_0, \dots, x_{k-1}$ d'où :

$$p_k - p_{k-1} = f[x_0, \cdots, x_k](x - x_0) \cdots (x - x_{k-1}). \tag{17}$$

$p_0 = f(x_0) = f[x_0]$.

$$\begin{aligned} p_n(x) &= p_0 + \sum_{k=1}^{n} (p_k - p_{k-1}) \\ &= f(x_0) + \sum_{k=1}^{n} f[x_0, \dots, x_k](x - x_0) \cdots (x - x_{k-1}). \end{aligned} \tag{18}$$

Formule de récurrence pour la différence divisée d'ordre $k \geq 1$:

$$f[x_0, \dots, x_k] = \frac{f[x_1, \dots, x_k] - f[x_0, \dots, x_{k-1}]}{x_k - x_0}. \tag{19}$$

Vérification. Soit $q_{k-1} \in \mathbb{R}_{k-1}[x]$ le polynôme d'interpolation de f aux points $x_1, \dots, x_k$. Posons

$$\tilde{p}_k(x) = \frac{(x - x_0) q_{k-1}(x) - (x - x_k) p_{k-1}(x)}{x_k - x_0}$$

$\tilde{p}_k \in \mathbb{R}_k[x]$. De plus, $\tilde{p}_k(x_0) = f(x_0)$, $\tilde{p}_k(x_i) = f(x_i)$ et pour $0 < i < k$, $\tilde{p}_k(x_i) = f(x_i)$. Donc $\tilde{p}_k = p_k$ ce qui implique (19). En utilisant (18), on peut calculer p_n.

3.3 Formule d'Erreur

Théorème 2

Soient $f \in C^{n+1}([a,b])$, $x_i \in [a,b]$ et $p_n \in \mathbb{R}_n[x]$ le polynôme d'interpolation de f aux points x_i. Alors pour $x \in [a,b]$, il existe $\xi_x \in]a,b[$ tel que

$$|f(x) - p_n(x)| \leq |\frac{1}{(n+1)!}\pi_{n+1}(x) f^{(n+1)}(\xi_x)|. \tag{20}$$

Lemme 1

Soit $g \in C([a,b])$ et $c_0 < c_1 < \cdots < c_p \in [a,b]$ tels que $g(c_i) = 0\ \forall i$. Alors il existe $\xi \in]c_0, c_p[$ tel que $g^{(p)}(\xi) = 0$.

Démonstration. Par récurrence. Pour $p = 1$ c'est le théorème de Rolle. Supposons vraie pour $p-1$ et supposons on a $p+1$ points $c_0 < c_1 < \cdots < c_p$ avec $g(c_i) = 0$. Par le théorème de Rolle, il existe $p-1$ points c_i' ou $g'(c_i') = 0$. Par hypothèse il existe un point ξ tel que $g'^{(p-1)}(\xi) = 0$. On a $g^{(p)}(\xi) = 0$. ■

Démonstration du Théorème 2. Si $x = x_i$, $\pi_{n+1}(x) = 0$, tout point $\xi_x \in]a,b[$ convient. Supposons $x \neq x_i$ pour tout i. Soit $p_{n+1}(t) \in \mathbb{R}_{n+1}[x]$ le polynôme d'interpolation aux points $x, x_0, \ldots, x_n$. Posons

$$e_{n+1} := p_{n+1}(t) - p_n(t) = c\pi_{n+1}(t). \tag{21}$$

Soit

$$\begin{aligned} g(t) &:= f(t) - p_n(t) + (p_n(t) - p_{n+1}(t)) \\ &= f(t) - p_n(t) - c\pi_{n+1}(t). \end{aligned}$$

g s'annule en $x_0, \cdots, x_n$ et en x. On en déduit par le Lemme 1 qu'il existe $\xi_x \in]a,b[$, tel que $g^{(n+1)}(\xi_x) = 0$. On a $p_n^{(n+1)}(x) = 0$, $\pi_{n+1}^{(n+1)}(x) = n!$. On en déduit que $c = f^{(n+1)}(\xi_x)/(n+1)!$. Puisque

$$f(x) - p_n(x) = p_{n+1}(x) - p_n(x),$$

on obtient (20). ■

Corollaire 1

$$\|f - p_n\| \leq \frac{1}{(n+1)!}\|\pi_{n+1}\|\|f^{(n+1)}\|. \tag{22}$$

3.4 Cas des Points Équidistants

$[a,b]$. $x_i = a + ih = a + i\frac{b-a}{n}$, $0 \leq i \leq n$. Pour utiliser (22), on voudrait estimer $\|\pi_{n+1}\|$.

$$\begin{aligned} \pi_{n+1}(x) &= (x - x_0)\cdots(x - x_n) \\ &= h^{n+1}s(s-1)\cdots(s-n)\ s \in [0,n]. \end{aligned}$$

Soit $\phi(s) := |s(s-1)\cdots(s-n)|$ pour $s \in [0,n]$. $\phi(n-s) = \phi(s)$ donc

$$\max_{s\in[0,n]} \phi(s) = \max_{s\in[0,n/2]} \phi(s).$$

On a

$$\frac{\phi(s-1)}{\phi(s)} = \frac{n+1-s}{s} > 1 \; s \in [1, n/2].$$

Donc

$$\max_{s\in[0,n]} \phi(s) = \max_{s\in[0,1]} \phi(s) \leq n!.$$

Donc

$$|f(x) - p_n(x)| \leq \frac{h^{n+1}}{n+1} \max_{x\in]a,b[} |f^{(n+1)}(x)|. \tag{23}$$

Exemple. Pour avoir une erreur inférieur à 10^{-8}, si $h = 10^{-2}$ et $\|f^{(4)}\| \leq 4$

$$|f(x) - p_3(x)| \leq \frac{(10^{-2})^4}{4} \|f^{(4)}\| = 10^{-8}.$$

$$|f(x) - p_4(x)| \leq \frac{(10^{-2})^5}{5} \|f^{(5)}\|$$

erreur plus petite si $\|f^{(5)}\| < 500$, plus grande si $\|f^{(5)}\| > 500$.

Remarque 2
Pour les points equidistants on peut montrer que $\|\pi_{n+1}\| \geq \frac{1}{n^{3/2}}(\frac{b-a}{e})^{n+1}$.

3.5 Les Points de Tchebychev

Les polynômes de Tchebychev sont définis par

$$t_n(x) = \cos(n \arccos x). \tag{24}$$

Posons $\theta := \arccos x$. Donc

$$\begin{aligned} t_{n+1} + t_{n-1} &= \cos((n+1)\theta) + \cos((n-1)\theta) \\ &= \cos n\theta \cos\theta - \sin n\theta \sin\theta + \cos n\theta \cos(-\theta) - \sin n\theta \sin(-\theta) \\ &= 2\cos n\theta \cos\theta \\ &= 2x t_n(x). \end{aligned}$$

Donc on a $t_0(x) = 1$, $t_1(x) = x$ et les autres sont obtenus par la formule de récurrence $t_{n+1}(x) = 2xt_n(x) - t_{n-1}(x)$. On voit que $t_n \in \mathbb{R}_n[x]$ et que son coefficient directeur est 2^{n-1}. Puisque $x = \cos\theta \in [-1,1]$, les polynôme de Tchebychev sont définis sur $[-1,1]$. Si $t_n(x) = \cos n\theta = 0$, alors $n\theta = \frac{\pi}{2} + i\pi$, ou bien $\theta = \frac{2i+1}{2n}\pi$. À priori, $i \in \mathbb{Z}$, mais puisque $t_n \in \mathbb{R}_n[x]$ il n'y a que n racines differentes, et donc on peut prendre $i \in \{0, \ldots, n-1\}$.

Définition 1
Les *points d'interpolation de Tchebychev* d'ordre n sont les points $x_i = \cos(\frac{2i+1}{2n+2}\pi)$, $i \in \{0, \ldots, n\}$ racines de $t_{n+1}(x)$.

En prenant $u \in [-1, 1]$ et $u_i = \cos(\frac{2i+1}{2n+2}\pi)$, on peut écrire

$$\begin{aligned} t_{n+1}(u) &= 2^n \prod_{i=0}^{n} (u - u_i) && (25) \\ &= 2^n \pi_{n+1}(u). && (26) \end{aligned}$$

Pour se ramener à l'intervalle $[a, b]$,

$$\begin{aligned} [-1, 1] &\mapsto [a, b] \\ u &\mapsto x = \frac{a+b}{2} + \frac{b-a}{2} u. \end{aligned}$$

Donc

$$\begin{aligned} x_i &= \frac{a+b}{2} + \frac{b-a}{2} \cos(\frac{2i+1}{2n+2}\pi) \\ x - x_i &= \frac{b-a}{2}(u - u_i) \\ \pi_{n+1}(x) &= \prod_{i=0}^{n}(x - x_i) = \left(\frac{b-a}{2}\right)^{n+1} \prod_{i=0}^{n}(u - u_i), \end{aligned}$$

où $\prod_{i=0}^{n}(u - u_i) = \frac{1}{2^n} t_{n+1}(u)$. D'où

$$\pi_{n+1}(x) = \frac{(b-a)^{n+1}}{2^{2n+1}} t_{n+1}(u).$$

Puisque $\|t_{n+1}\| = 1$,

$$\begin{aligned} \|\pi_{n+1}\| &\leq \frac{(b-a)^{n+1}}{2^{2n+1}} \\ &= 2\left(\frac{b-a}{4}\right)^{n+1} \\ &< \left(\frac{b-a}{e}\right)^{n+1}. \end{aligned}$$

Exemple : $n = 30$,

$$\left(\frac{e}{4}\right)^{31} < 7 \cdot 10^{-6}.$$

4 Integration Numérique

4.1 Methodes de quadrature élémentaires et composées

4.1.1 Principe des méthodes numériques

On cherche une valeur approchée de $\int_\alpha^\beta f(x)\mathrm{d}x$. On partitionne $\alpha = \alpha_0 < \alpha_1 < \cdots < \alpha_k = \beta$. La formule de Chasles donne

$$\int_\alpha^\beta f(x)\mathrm{d}x = \sum_{i=0}^{k} \int_{\alpha_i}^{\alpha_{i+1}} f(x)\mathrm{d}x.$$

Méthodes de quadrature élémentaires

$$\int_{\alpha_i}^{\alpha_{i+1}} f(x)\mathrm{d}x = (\alpha_{i+1} - \alpha_i) \sum_{j=0}^{l_i} w_{i,j} f(\xi_{i,j}),$$

où $\sum_{j=0}^{l_i} w_{i,j} = 1$, $\xi_{i,j} \in [\alpha_i, \alpha_{i+1}]$, $0 \le j \le l_i$.

Méthodes de quadrature composée

$$\int_\alpha^\beta f(x)\mathrm{d}x = \sum_{i=0}^{k-1} (\alpha_{i+1} - \alpha_i) \sum_{j=0}^{l_i} w_{i,j} f(\xi_{i,j}).$$

Définition 2

Une méthode de quadrature est d'*ordre* N si la formule est exacte pour $f \in \mathbb{R}_N[x]$ et inéxacte pour la fonction $x \mapsto x^{N+1}$.

4.1.2 Exemples

(a) $l_i = 0$ pour tout i.

$$\begin{aligned} \int_{\alpha_i}^{\alpha_{i+1}} f(x)\mathrm{d}x &= (\alpha_{i+1} - \alpha_i) f(\xi_i) \\ \int_\alpha^\beta f(x)\mathrm{d}x &= \sum_{i=0}^{k-1} (\alpha_{i+1} - \alpha_i) f(\xi_i). \end{aligned}$$

C'est une somme de Riemann si

1. $\xi_i = \alpha_i$ rectangles à droite.
2. $\xi_i = \alpha_{i+1}$ rectangles à gauche.

La méthode est d'ordre 0.

Si $\xi_i = \frac{\alpha_i + \alpha_{i+1}}{2}$, c'est la méthode du point milieu, d'ordre 1.

(b) Interpolation linéaire ou la méthode des trapèze. $l_i = 1$ $\forall i$, $\xi_{i,0} = \alpha_i$, $\xi_{i,1} = \alpha_{i+1}$.

$$p_1(x) = \frac{(x - \alpha_i) f(\alpha_{i+1}) - (x - \alpha_{i+1}) f(\alpha_i)}{\alpha_{i+1} - \alpha_i}.$$

$$\begin{aligned}\int_{\alpha_i}^{\alpha_{i+1}} f(x)\mathrm{d}x &= \int_{\alpha_i}^{\alpha_{i+1}} p_1(x)\mathrm{d}x \\ &= (\alpha_{i+1}-\alpha_i)(\frac{1}{2}f(\alpha_i)+\frac{1}{2}f(\alpha_{i+1})).\end{aligned}$$

La méthode est d'ordre 1.
(c) Méthodes de Newton-Cotes : NC_l. $l_i = l\ \forall i$.

$$\begin{aligned}\xi_{i,j} &= \alpha_i + j\frac{\alpha_{i+1}-\alpha_i}{l}\ 0 \le j \le l \\ p_l(x) &= \sum_{j=0}^{l} f(\tau_j)L_j(x) \\ L_j(x) &= \prod_{k\neq j}\frac{x-\tau_k}{\tau_j-\tau_k}.\end{aligned}$$

Si $[\alpha_i, \alpha_{i+1}] = [-1, 1]$, $\tau_j = -1 + j\frac{2}{l}$. Ceci donne

$$\begin{aligned}\int_{-1}^{1} f(x) \quad \mathrm{d}x &\approx \int_{-1}^{1} p_l(x)\mathrm{d}x \\ &= 2\sum_{j=0}^{l} w_j f(\tau_j),\end{aligned}$$

avec $w_j = \frac{1}{2}\int_{-1}^{1} L_j(x)\mathrm{d}x$. Le calcul des w_j est facilité en remarquant que $\tau_{l-j} = -\tau_j$, $L_{l-j}(x) = L_j(-x)$ et $w_{l-j} = w_j$. Pour un domaine d'integration quelquonque, on remarque que les w_j sont invariants par le changement de variable

$$\begin{aligned}[-1,1] &\mapsto [a,b] \\ u &\mapsto x = \frac{a+b}{2} + \frac{b-a}{2}u.\end{aligned}$$

Proposition 1
Si l est paire, les méthodes NC_l sont d'ordre $l+1$. Si l est impaire, les méthodes NC_l sont d'ordre l.

Exemples :
— $l = 1$, c'est la méthode des trapèzes : $w_0 = w_1 = \frac{1}{2}$.
— $l = 2$, la méthode de Simpson : $w_0 = w_2 = \frac{1}{6}$, $w_1 = 2/3$.

4.2 Évaluation de l'Erreur

4.2.1 Rappels d'Analyse

Si $f \in C^{N+1}([\alpha, \beta])$ et $x, \in [\alpha, \beta]$, la formule de Taylor avec reste integrale est :

$$f(x) = \sum_{k=0}^{N} \frac{1}{k!} f^{(k)}(\alpha)(x-\alpha)^k + \int_{\alpha}^{x} \frac{1}{N!}(x-t)^N f^{(N+1)}(t)\mathrm{d}t. \tag{27}$$

Vérification. Par récurrence sur N. Pour $N = 0$:

$$f(x) = f(\alpha) + \int_\alpha^x f'(t)\mathrm{d}t.$$

La formule est vérifiée. Supposons la formule vérifiée à l'ordre $N-1$. Faisons une intégration par partie sur l'intégrale dans cette formule :

$$\int_\alpha^x \frac{1}{(N-1)!}(x-t)^{N-1} f^{(N)}(t)\mathrm{d}t = -\frac{1}{N!}(x-t)^N f^{(N+1)}(t)\Big]_\alpha^x + \int_\alpha^x \frac{1}{N!}(x-t)^N f^{(N+1)}(t)dt.$$

Ceci donne (27). ■

On définie $x_+ := \max\{x, 0\}$, et $x_- := (-x)_+$. (27) peux s'écrire :

$$f(x) = \sum_{k=0}^{N} \frac{1}{k!} f^{(k)}(\alpha)(x-\alpha)^k + \int_\alpha^\beta \frac{1}{N!}((x-t)_+)^N f^{(N+1)}(t)\mathrm{d}t. \tag{28}$$

La Formule de la Moyenne. Soient $0 \leq w(x)$ intégrable et $f \in C([a, b])$, alors il existe $\xi \in [a, b]$ tel que

$$\int_\alpha^\beta f(x)w(x)\mathrm{d}x = f(\xi)\int_\alpha^\beta w(x)\mathrm{d}x. \tag{29}$$

Vérification. Soient $m := \min_{x\in[a,b]} f(x)$ et $M := \max_{x\in[a,b]} f(x)$. Alors

$$c = m\int_\alpha^\beta w(x)\mathrm{d}x \leq \int_\alpha^\beta f(x)w(x)\mathrm{d}x \leq M\int_\alpha^\beta w(x)\mathrm{d}x = C.$$

Soit $G(t) := f(t)\int_\alpha^\beta w(x)\mathrm{d}x$. Puisqu'il existe $\xi_i \in [a, b]$, $i = 1, 2$, tel que $G(\xi_1) = c$ et $G(\xi_2) = C$, par le théorème des valeurs intermédiares, il existe $\xi \in [a, b]$ qui vérifie (29). ■

4.3 Le Noyau de Peano

Si notre méthode donne

$$\int_\alpha^\beta f(x)\mathrm{d}x \approx \sum_{j=0}^{l} \lambda_j f(x_j), \tag{30}$$

on définie

$$E(f) := \int_\alpha^\beta f(x)\mathrm{d}x - \sum_{j=0}^{l} \lambda_j f(x_j). \tag{31}$$

Définition 3
On définie

$$K_N(t) := E(x \mapsto ((x-t)_+)^N)\ t \in [\alpha, \beta]. \tag{32}$$

Théorème 3
Si la méthode est d'ordre N et $f \in C^{N+1}([\alpha,\beta])$,

$$E(f) = \frac{1}{N!}\int_\alpha^\beta K_N(t) f^{(N+1)}(t)\mathrm{d}t. \tag{33}$$

Démonstration. L'application $f \mapsto E(f)$ est une forme linéaire. Si $g : (x,t) \mapsto g(x,t)$ est intégrable sur $I := [\alpha,\beta]$, le théorème de Fubini implique que

$$E\left(x \mapsto \int_{t\in I} g(x,t)\mathrm{d}t\right) = \int_{t\in I} E(x \mapsto g(x,t))\mathrm{d}t.$$

Par (28) on peut écrire

$$f(x) = p_N(x) + \int_\alpha^\beta \frac{1}{N!}((x-t)_+)^N f^{(N+1)}(t)\mathrm{d}t.$$

On a $p_N \in \mathbb{R}_N[x]$, donc $E(p_N) = 0$ d'où

$$\begin{aligned}
E(f) &= E\left(x \mapsto \int_\alpha^\beta \frac{1}{N}((x-t)_+)^N f^{(N+1)}(t)\mathrm{d}t\right)\\
&= \int_\alpha^\beta E\left(x \mapsto \frac{1}{N!}((x-t)+)^N f^{(N+1)}(t)\right)\mathrm{d}t\\
&= \int_\alpha^\beta \frac{1}{N!} f^{(N+1)}(t) E(x \mapsto (x-t)_+)^N)\mathrm{d}t\\
&= \frac{1}{N!}\int_\alpha^\beta K_N(t) f^{(N+1)}(t)\mathrm{d}t.
\end{aligned}$$

■

Corollaire 2

$$|E(f)| \leq \frac{1}{N!}\|f^{(N+1)}\|_\infty \int_\alpha^\beta |K_N(t)|\mathrm{d}t. \tag{34}$$

Corollaire 3
Si K_N est de signe constant et $f \in C^{N+1}(I)$ alors $\exists \xi \in I$ tel que

$$E(f) = \frac{1}{(N+1)!} f^{(N+1)}(\xi) E(x \mapsto x^{N+1}). \tag{35}$$

Démonstration. Par le théorème de la moyenne, on a

$$E(f) = \frac{1}{N!} f^{N+1}(\xi) \int_\alpha^\beta K_N(t)\mathrm{d}t. \tag{36}$$

En prenant $f : x \mapsto x^{N+1}$ et en remplacant dans (36), on obtient :

$$\int_\alpha^\beta K_N(t)\mathrm{d}t = \frac{1}{N+1} E(x \mapsto x^{N+1}).$$

Remplacant cette expression pour $\int_\alpha^\beta K_N(t)\mathrm{d}t$ dans (36) donne (35). ■

4.4 Exemples

On prend l'intervalle $I = [-1, 1]$.

Méthode du point milieu. Méthode d'ordre 1.

$$E(f) = \int_{-1}^{1} f(x)\mathrm{d}x - 2f(0).$$

$$\begin{aligned}
K_1(t) &= E\left(x \mapsto (x-t)_+\right) \\
&= \int_{-1}^{1} (x-t)_+ \mathrm{d}x - 2(-t)_+ \\
&= \int_{t}^{1} (x-t)\mathrm{d}x - 2t_- \\
&= \left[\frac{1}{2}(x-t)^2\right]_t^1 - 2t_- \\
&= \frac{1}{2}(1-t)^2 - 2t_-.
\end{aligned}$$

Donc

$$K_1(t) = \begin{cases} \frac{1}{2}(1-t)^2 & t \geq 0 \\ \frac{1}{2}(1-t)^2 + 2t & t < 0. \end{cases}$$

D'où

$$\int_{-1}^{1} K_1(t)\mathrm{d}t = \int_{0}^{1} (1-t)^2 \mathrm{d}t = 1/3,$$

et

$$E(f) = \frac{1}{3} f''(\xi) \ \xi \in]-1, 1[.$$

Méthode du trapèze. Méthode d'ordre 1.

$$E(f) = \int_{-1}^{1} f(x)\mathrm{d}x - (f(-1) + f(1)).$$

$$\begin{aligned}
K_1(t) &= \int_{-1}^{1} (x-t)_+ \mathrm{d}x - ((-1-t)_+ + (1-t)_+) \\
&= \int_{t}^{1} (x-t)\mathrm{d}x - (1-t) \\
&= -\frac{1}{2}(1-t^2) \leq 0.
\end{aligned}$$

$$\int_{-1}^{1} K(t)\mathrm{d}t = -\frac{2}{3}.$$

$$E(f) = -\frac{2}{3} f''(\xi).$$

5 Méthodes Itératives Pour la Résolution d'Équations

5.1 Les suites de Cauchy

Un exemple : soit $S_N \stackrel{\text{def}}{=} \sum_{k=1}^{N} 1/k$. Alors $S_{N+1} - S_N = 1/(N+1)$, donc $\lim_{N\to\infty} |S_{N+1} - S_N| = 0$, mais $\lim_{N\to\infty} S_N = \infty$. Nous voudrions caracteriser les suites convergentes.

Définition 4
Le *supremum* (*infimum*) d'un ensemble $A \subset \mathbb{R}$ est la plus petite (plus grande) borne supérieure (inférieure) de l'ensemble A.

Le supremum (parfois $+\infty$) d'un ensemble réel A existe toujours par la construction des nombres réels par les coupures de Dedekind. On écrit $\sup A$ ou $\inf A$.

Soit $(x_p)_p \subset \mathbb{R}$. On définit

$$\limsup x_p \stackrel{\text{def}}{=} \lim_{N\to\infty} \sup_{p\geq N} x_p.$$

Si $\infty > M = \limsup x_p$, pour tout $\epsilon > 0$ et $N > 0$,il exist $l > N$ tel que $x_l > M - \epsilon$.

Définition 5
Une suite $(x_p)_{p\geq 0} \subset \mathbb{R}^N$ est *de Cauchy* si pour tout $\epsilon > 0$ il existe $N_\epsilon \in \mathbb{N}$ tel que pour tout p, $q \geq N_\epsilon$, $\|x_p - x_q\| < \epsilon$.

Théorème 4
Une suite $(x_p)_p \subset \mathbb{R}^N$ converge si et seulement si elle est de Cauchy.

Démonstration : On commence par le cas $N = 1$.
$\Longrightarrow$: Supposons $\lim_{p\to\infty} x_p = L$. Soit $\epsilon > 0$. Alors il exist $N_{\epsilon/2} \in \mathbb{N}$ tel que pour $p \geq N_{\epsilon/2}$, $\|x_p - L\| < \epsilon/2$. Donc pour tout p, $q \geq N_{\epsilon/2}$,

$$\begin{aligned}
\|x_p - x_q\| &= \|x_p - L + L - x_q\| \\
&\leq \|x_p - L\| + \|L - x_q\| \\
&\leq \epsilon/2 + \epsilon/2 = \epsilon.
\end{aligned}$$

Puisque ϵ est quelconque, la suite $(x_p)_p$ est de Cauchy.
$\Longleftarrow$: Supposons $(x_p)_p$ est de Cauchy et soient $L = \limsup x_p$ et $\epsilon > 0$. Alors il exist $N_{\epsilon/2} \in \mathbb{N}$ tel que pour p et $q \geq N_{\epsilon/2}$, $\|x_p - x_q\| < \epsilon/2$. Par la définition de $\limsup$ il y a $l > N_{\epsilon/2}$ tel que $\|x_l - L\| < \epsilon/2$. Il vient que pour $p > N_{\epsilon/2}$,

$$\begin{aligned}
\|x_p - L\| &= \|x_p - x_l + x_l - L\| \\
&\leq \|x_p - x_l\| + \|x_l - L\| \\
&= \epsilon/2 + \epsilon/2 = \epsilon.
\end{aligned}$$

Donc $\lim_{p\to\infty} x_p = L$.

Le cas $N > 1$, suit par le fait qu'une suite $(x_p)_p \subset \mathbb{R}^N$ est de Cauchy si et seulement si chaque coordonné de $(x_p)_p$ est de Cauchy. ■

5.2 Principe

Soit $E \subset \mathbb{R}^N$ un ensemble compact (fermé et borné) et $\phi : E \mapsto E$. Un *point fixe* $a \in E$ est un point tel que $\phi(a) = a$.

Définition 6
La fonction ϕ est *lipschitzienne* de rapport k dans E si

$$\|\phi(x) - \phi(y)\| \leq k\|x - y\| \;\forall\, x, y \in E. \tag{37}$$

Définition 7
La fonction ϕ est *contractante* dans E si ϕ est lipschitzienne de rapport $k < 1$ dans E.

Théorème 5
Soit $\phi : E \mapsto E$ contractante. Alors ϕ possède un point fixe a, unique dans E. De plus la suite $(x_p)_p$ définie par

$$x_{p+1} = \phi(x_p)\; x_0 \in E \tag{38}$$

converge vers a pour tout point $x_0 \in E$.

Démonstration. Unicité. Par l'absurde : supposons qu'il existe 2 points fixes a_1 et a_2. Alors par (37) on obtient

$$\|\phi(a_1) - \phi(a_2)\| < \|a_1 - a_2\|,$$

mais par la définition d'un point fixe on a

$$\|\phi(a_1) - \phi(a_2)\| = \|a_1 - a_2\|.$$

■

Éxistence. Soit $x_0 \in E$ et x_p définie par (38). On a

$$\|x_p - x_{p+1}\| = \|\phi(x_{p-1} - \phi(x_p)\| \leq k\|x_{p-1} - x_p\|. \tag{39}$$

Par récurrence, on obtient donc

$$\|x_p - x_{p+1}\| \leq k^p\|x_1 - x_0\|. \tag{40}$$

Pour $q > p$,

$$\begin{aligned}\|x_p - x_q\| &\leq \sum_{l=p}^{q-1} \|x_l - x_{l+1}\| \\ &\leq \|x_1 - x_0\| \sum_{l=p}^{q-1} k^l.\end{aligned}$$

Puisque

$$\begin{aligned}\lim_{p\to\infty}\sum_{1=p}^{q-1}k^l &\le \lim_{p\to\infty}\sum_{l=p}^{\infty}k^l\\ &= \lim_{p\to\infty}\frac{k^p}{1-k}=0,\end{aligned}$$

la suite $(x_p)_p$ est de Cauchy. Puisque E est fermé Il existe $a\in E$ tel que $\lim_{p\to\infty} x_p = a$. Puisque ϕ est continue, (38) implique que $\phi(a)=a$. ■

Vitesse de convergence.

$$\begin{aligned}\|x_p-a\| &= \|\phi(x_{p-1}-\phi(a)\|\\ &\le k\|x_{p-1}-a\|.\end{aligned}$$

Par récurrence, on obtient

$$\|x_p-a\|\le k^p\|x_0-a\|. \tag{41}$$

Remarque 3

Quand on écrit un algorithme utilisant la méthode du point fixe, on arrête souvent l'algorithme quand $\|x_p-x_{p+1}\|$ est petit. La justification est :

$$\begin{aligned}\|x_p-a\| &= \|x_p-x_{p+1}+x_{p+1}-a\|\\ &\le \|x_p-x_{p+1}\|+\|a-x_{p+1}|\\ &= \|x_p-x_{p+1}\|+\|\phi(a)-\phi(x_p)\|\\ &\le \|x_p-x_{p+1}\|+k\|a-x_p\|\end{aligned}$$

d'où

$$\|x_p-a\|\le\frac{1}{1-k}\|x_p-x_{p+1}\|. \tag{42}$$

5.3 Application

Soit $f:\mathbb{R}^N\mapsto\mathbb{R}^N$. On cherche a tel que $f(a)=0$. On cherche donc une fonction ϕ tel que $\phi(a)=a$ si et seulement si $f(a)=0$.

Pour le cas $N=1$, supposons $\phi\in C^1(I\mapsto I)$ et ϕ possède un point fixe dans I.

1. $|\phi'(a)|<1$. Soit k tel que $|\phi'(a)|<k<1$. Par continuité, il existe h tel que pour $x\in E=[a-h,a+h]$, $|\phi'(x)|\le k$. Pour $x\in E$, $|\phi(x)-\phi(a)|=|\phi(x)-a|\le|x-a|$, donc $\phi(E)\subset E$. Le point fixe a est un point fixe *attractif*. La suite $(x_p)_p$ définie par (38) converge vers a.
2. $|\phi'(a)|>1$. Soit k tel que $|\phi'(a)|>k>1$. Par continuité, il existe h tel que pour $x\in E=[a-h,a+h]$, $|\phi'(x)|\ge k$. Pour $x\in E$, $|\phi(x)-\phi(a)|=|\phi(x)-a|>|x-a|$. a est un point fixe *répulsif*. Dans ce cas on utilise $\phi^{-1}:E\mapsto E$. Car a est aussi un point fixe pour ϕ^{-1} et $|(\phi^{-1})'(a)|=|1/\phi'(a)|<1$ et pour $x\in E$, $|(\phi^{-1})'(x)|\le 1/k<1$.

3. $|\phi(a)| = 1$. Exemple (1) $\phi(x) = \sin x$, $x \in [0, \pi/2]$. $\sin x < x$, $x \in [0, \pi/2]$, 0 est un point fixe attractif.

   Exemple (2) $\phi(x) = \operatorname{sh} x$, $x \in [0, M]$. $\operatorname{sh} x > x$, $x \in [0, M]$. 0 est un point fixe répulsif.

   Conclusion : on ne peut rien conclure si $\phi'(a) = 1$.

Cas particulier : $\phi \in C^2(I)$ et $\phi'(a) = 0$. Soit M tel que $|\phi''(x)| \leq M$. On a pour $x \in I$,

$$\phi(x) = \phi(a) + \phi'(a)(x-a) + \frac{(x-a)^2}{2!}\phi''(c).$$

Donc

$$\begin{aligned} |\phi(x) - a| &= |\phi(x) - \phi(a)| \\ &\leq \frac{1}{2}M|x-a|^2. \end{aligned}$$

D'où

$$\frac{1}{2}M|\phi(x) - a| \leq (\frac{1}{2}M|x-a|)^2.$$

Par récurrence, on obtient

$$|x_p - a| \leq \frac{2}{M}|\frac{1}{2}M(x_0 - a)|^{2^p}. \tag{43}$$

La convergence est *quadratique*.

5.4 Exemple

Soit $f : x \mapsto x^3 - 4x + 1$. On cherche a tel que $f(a) = 0$. $f'(x) = 3x^2 - 4$. $f(a) = 0 \iff \frac{1}{4}(a^3 + 1) = a$. Donc a est un point fixe de $\phi : x \mapsto \frac{1}{4}(x^3 + 1)$. $\phi'(x) = \frac{3x^2}{4}$. L'analyse des variations de f donne 3 racines $a_1 < a_2 < a_3$. Par le théorème des valeurs intermediares on obtient

$$\begin{aligned} a_1 &\in [-2,5; -2] \\ a_2 &\in [0; 0,5] \\ a_3 &\in [1,5; 2] \end{aligned}$$

Pour

$$\begin{aligned} x &\in [-2,5; -2]\ \phi'(x) \geq \phi'(-2) = 3 > 1 \\ x &\in [0; 0,5]\ \phi'(x) \leq \phi'(0,5) = 3/16 < 1 \\ x &\in [1,5; 2]\ \phi'(x) \geq \phi'(1,5) = 27/16 > 1. \end{aligned}$$

Conclusion : a_1 et a_3 sont des points fixes répulsifs, a_2 est un point fixe attractif.

Pour a_2, on a

$$\begin{aligned} |x_p - a_2| &\leq (3/16)^p |x_0 - a_2| \\ &\leq (3/16)^p \frac{1}{2}. \end{aligned}$$

Pour a_1 et a_3, on utilise $\phi^{-1} : y \mapsto (4y-1)^{1/3}$ dans les intervalles $[\phi(-2,5); \phi(-2)]$ et $[\phi(1,5); \phi(2)] = [35/32; 9/4]$ respectivement. Par exemple pour a_3,

$$\begin{aligned} |y_p - a_3| &\leq (1/3)^p |x_0 - a_3| \\ &\leq (1/3)^p 37/72 \\ &\leq (1/3)^p \end{aligned}$$

5.5 La Méthode de Newton

On cherche une racine d'une fonction $f : \mathbb{R} \mapsto \mathbb{R}$. L'idée de la méthode de Newton est d'utiliser l'approximation affine de la fonction pour trouver la racine. On commence avec x_0 pas trop loin de la racine, et on calcul x_1, le point ou l'approximation affine est 0 :

$$0 = f(x_0) + f'(x_0)(x_1 - x_0),$$

ce qui donne

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

ou plus généralement :

$$x_{p+1} = x_p - \frac{f(x_p)}{f'(x_p)}. \tag{44}$$

On vérifie que $\phi : \mathbb{R} \mapsto \mathbb{R}$

$$\phi_N(x) \stackrel{\text{def}}{=} x - \frac{f(x)}{f'(x)} \tag{45}$$

est contractante. Si $f \in C^2(I)$,

$$\phi_N'(x) = \frac{f''(x) f(x)}{f'(x)^2}. \tag{46}$$

Si a est racine de f et $f'(a) \neq 0$, $\phi'(a) = 0$, donc a est un point fixe super attractif de ϕ_N. Plus généralement si $f'(x) \neq 0$ pour $x \in I = [a - r, a + r]$ on a :

Théorème 6
Soit $f \in C^2(I)$ et soit $M = \sup_{x \in I} \frac{\|f''\|}{|f'(x)|}$, alors pour $x \in I$,

$$\frac{M}{2} |\phi(x) - a| \leq \left(\frac{M}{2} |x - a| \right)^2. \tag{47}$$

Démonstration. Soit $x \in I$. Alors la formule de Taylor en x done :

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

En divisant par $f'(x)$ et en prenant le module, on trouve :

$$|a - x + f(x)/f'(x)| = \frac{1}{2} \frac{|f''(c)|}{|f'(x)|} (a-x)^2,$$

qui donne (47). ∎

Par récurrence on a donc :

Corollaire 4
Soit $f \in C^2(I)$ et M comme dans le Théorème 6 et la suite (x_p) définie par $x_{p+1} = \phi_N(x_p)$, alors

$$\frac{M}{2}|x_p - a| \leq |\frac{M}{2}(x_0 - a)|^{2^p}. \tag{48}$$

Références

[1] Jean-Pierre Demailly. *Analyse numérique et équations différentielles.* Broché, 2006.