

COURS ANALYSE NUMÉRIQUE

2020 - 2021

ESATIC UP MATHS

22 septembre 2021

Table des matières

1	Equations non linéaires	4
1.1	introduction	4
1.2	Définitions	4
1.3	Localisation des zéros	5
1.4	Méthode de dichotomie (ou bisection)	6
1.5	Méthode de Lagrange (ou Regula falsi)	8
1.6	méthode de la sécante	9
1.6.1	Critère d'arrêt	10
1.7	Méthode de Newton	10
1.8	Méthode de point fixe	11
1.9	Convergence des algorithmes	13
1.9.1	Méthodes de point fixe	13
1.9.2	Méthode de Newton	15
1.9.3	Méthode de la sécante	15
1.10	Résumés	16
1.10.1	Méthode de la dichotomie :	16
1.10.2	Méthode de la sécante	16
1.10.3	Méthode du point fixe / d'itération	17
1.10.4	Méthode de Newton(Raphson)/de la tangente	17
1.10.5	Arrêt de l'algorithme	18
2	Interpolation et approximation polynômiale	19
2.1	Interpolation de lagrange	19
2.2	Interpolation de Newton et différences divisées	21
2.3	Erreur dans l'interpolation de Lagrange	23
3	Dérivation et intégration numérique	25
3.1	Introduction	25
3.2	Dérivation numérique	26
3.2.1	Dérivée première	26
3.2.2	Dérivée d'ordre supérieur	28

TABLE DES MATIÈRES

3.3	Intégration numérique : méthodes composites	28
3.3.1	Principe	28
3.3.2	Méthode des rectangles	29
3.3.3	Méthode des trapèzes	30
3.3.4	Méthode de Simpson	30
3.4	Analyse de l'erreur dans les méthodes d'intégration	31
3.4.1	Théorie	31
3.4.2	Application aux méthodes usuelles	33
3.5	Exemple d'application	34
4	Résolution numérique des équations différentielles	36
4.1	Généralité	36
4.2	Méthodes de résolution numérique	37
4.2.1	Méthode à un pas	38
4.2.2	Méthode d'Euler	39
4.2.3	Méthodes de Runge-Kutta	44

Notations

Chapitre 1

Equations non linéaires

1.1 introduction

Un des problèmes classiques en mathématiques appliquées est celui de la recherche des valeurs pour lesquelles une fonction donnée s'annule. Dans certains cas bien particuliers, comme pour les fonctions $x \mapsto x + 1$, $x \mapsto \cos(2x)$, ou encore $x \mapsto x^2 - 2x + 1$, le problème est simple car il existe pour ces fonctions des formules qui donnent les zéros explicitement. Toutefois, pour la plupart des fonctions $f : \mathbb{R} \rightarrow \mathbb{R}$ il n'est pas possible de résoudre l'équation $f(x) = 0$ explicitement et il faut recourir à des méthodes numériques. Ainsi par exemple une brève étude de la fonction $f(x) = \cos(x) - x$ montre qu'elle possède un zéro à proximité de 0.7 mais ce zéro ne s'exprime pas au moyen de fonctions usuelles et pour en obtenir une valeur approchée il faut recourir à des méthodes numériques.

Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ une fonction continue donnée dont on veut évaluer numériquement un ou plusieurs zéros \hat{x} , c'est-à-dire qu'on cherche tous les \hat{x} tels que $f(\hat{x}) = 0$. Les méthodes numériques pour approcher \hat{x} consistent à :

1. localiser grossièrement le (ou les) zéro(s) de f en procédant à l'étude du graphe de f et/ou à des évaluations qui sont souvent de type graphique ; on note x_0 cette solution grossière ;
2. construire, à partir de x_0 une suite $x_0, x_1, x_2, \dots, x_n$ telle que $\lim_{x \rightarrow \infty} x_k = \hat{x}$ où $f(\hat{x}) = 0$. On dit alors que la méthode est convergente.

1.2 Définitions

Définition 1.1. Méthode itérative à deux niveaux

On appelle méthode itérative à deux niveaux un procédé de calcul de la forme

$$x_{k+1} = G(x_k), \quad k = 0, 1, 2, \dots$$

CHAPITRE 1. EQUATIONS NON LINÉAIRES

dans lequel on part d'une valeur donnée x_0 pour calculer x_1 , puis à l'aide de x_1 on calcul x_2 etc. La formule même est dite formule de récurrence. Le procédé est appelé convergent si x_k tend vers un nombre fini lorsque k tend vers $+\infty$. Il est bien évident qu'une méthode itérative n'est utile que s'il y a convergence vers les valeurs cherchées.

On peut parfaitement envisager des méthodes itératives multi-niveaux, comme par exemples les schémas à trois niveaux dans lesquels on part de deux valeurs données x_0 et x_1 pour calculer x_2 , puis à l'aide de x_1 et x_2 on calcule x_3 etc.

Définition 1.2. Ordre de convergence

Soit p un entier positif. On dit qu'une méthode (à deux niveaux) convergente est d'ordre p s'il existe une constante C telle que

$$|\hat{x} - x_{k+1}| \leq C|\hat{x} - x_k|^p$$

ou encore

$$\lim_{x \rightarrow \infty} \frac{x_{k+1} - \hat{x}}{(x_{k+1} - \hat{x})^p} = C$$

Si $p = 1$ (et $C < 1$) on parle de convergence linéaire, si $p = 2$ on parle de convergence quadratique.

1.3 Localisation des zéros

Définition 1.3. Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ une fonction. On dit que \hat{x} est un zéro de f si $f(\hat{x}) = 0$. Il est dit simple si $f'(\hat{x}) \neq 0$, multiple sinon. Si f est de classe C^p avec $p \in \mathbb{N}$, on dit que \hat{x} est un zéro de multiplicité p si

$$\begin{cases} f^{(i)}(\hat{x}) = 0 & \forall i = 0, \dots, p-1 \\ f^{(p)}(\hat{x}) \neq 0 \end{cases}$$

Pour localiser grossièrement le (ou les) zéro(s) de f on va d'abord étudier la fonction f , puis on va essayer d'utiliser un corollaire du théorème des valeurs intermédiaires et le théorème de la bijection afin de trouver un intervalle qui contient un et un seul zéro.

Théorème 1.1. Des valeurs intermédiaires

Soit f une fonction continue sur un intervalle $I = [a; b]$ de \mathbb{R} . Alors f atteint toutes les valeurs intermédiaires entre $f(a)$ et $f(b)$. Autrement dit :

si $f(a) \leq f(b)$ alors pour tout $d \in [f(a); f(b)]$ il existe $c \in [a; b]$ tel que $f(c) = d$;

si $f(a) \geq f(b)$ alors pour tout $d \in [f(b); f(a)]$ il existe $c \in [a; b]$ tel que $f(c) = d$.

Ce théorème donne alors le corollaire immédiat suivant.

Corollaire 1.1. *Soit une fonction continue $f : [a, b] \rightarrow \mathbb{R}$. Si $f(a) \times f(b) < 0$, alors il existe (au moins un) $\hat{x} \in]a; b[$ tel que $f(\hat{x}) = 0$*

Ce théorème garantit juste l'existence d'un zéro. Pour l'unicité on essaiera d'appliquer le théorème de la bijection dont l'énoncé est rappelé ci-dessous.

Théorème 1.2. De la bijection

Soit f une fonction continue et strictement monotone sur un intervalle I de \mathbb{R} , alors f induit une bijection de I dans $f(I)$. De plus, sa bijection réciproque est continue sur $f(I)$, monotone sur $f(I)$ et de même sens de variation que f .

Ayant encadré les zéros de f , la construction de suites qui convergent vers ces zéros peut se faire à l'aide de plusieurs méthodes numériques. Ci-dessous on décrit les méthodes les plus connues et on étudie leurs propriétés (convergence locale, globale, vitesse de convergence, etc.)

1.4 Méthode de dichotomie (ou bisection)

Dans les méthodes de dichotomie, à chaque pas d'itération on divise en deux un intervalle donné et on choisit le sous-intervalle où f change de signe.

Concrètement, soit $f : [a; b] \rightarrow \mathbb{R}$ une fonction strictement monotone sur un intervalle $[a; b]$. On suppose que l'équation $f(x) = 0$ n'a qu'une et une seule solution dans cet intervalle. On se propose de déterminer cette valeur avec une précision donnée. Soit $[a_0; b_0]$ un intervalle dans lequel $f(a_0)f(b_0) < 0$ et soit $c_0 \in]a_0; b_0[$. Si $f(a_0)f(c_0) < 0$, alors la racine appartient à l'intervalle $[a_0; c_0]$ et on reprend le procédé avec $a_1 = a_0$ et $b_1 = c_0$. Sinon, c'est-à-dire si $f(a_0)f(c_0) \geq 0$ on pose $a_1 = c_0$ et $b_1 = b_0$. On construit ainsi une suite d'intervalles emboîtés $[a_k; b_k]$. Les suites a_k et b_k sont adjacentes et convergent vers \hat{x} .

Définition 1.4. *Soit deux points a_0 et b_0 (avec $a_0 < b_0$) d'images par f de signe contraire (i.e. $f(a_0)f(b_0) < 0$). En partant de $I_0 = [a_0; b_0]$, la méthode de dichotomie produit une suite de sous-intervalles $I_k = [a_k; b_k]$, $k \geq 0$, avec $I_k \subset I_{k-1}$ pour $k \geq 1$ et tels que $f(a_k)f(b_k) < 0$.*

On découpe l'intervalle $[a_k; b_k]$ en deux intervalles de même longueur, i.e. on divise $[a_k; b_k]$ en $[a_k; c_k]$ et $[c_k; b_k]$ où $c_k = \frac{a_k + b_k}{2}$.

Pour l'itération suivante, on pose soit $[a_{k+1}, b_{k+1}] = [a_k; c_k]$ soit $[a_{k+1}, b_{k+1}] = [c_k; b_k]$ de sorte à ce que $f(a_{k+1})f(b_{k+1}) < 0$. La suite $(c_k)_{k \in \mathbb{N}}$ converge vers \hat{x} puisque la longueur de ces intervalles tend vers 0 quand k tend vers $+\infty$

Algorithme 1.4.1.

$$\left\{ \begin{array}{l} \text{Pour } k = 0, 1, 2, \dots, N, \text{ faire} \\ \quad c =: \frac{a_k + b_k}{2} \\ \quad \text{si } f(a_k) \times f(c) \leq 0, \quad a_{k+1} = a_k, \quad b_{k+1} := c \\ \quad \text{sinon } a_{k+1} := c, \quad b_{k+1} := b_k \end{array} \right.$$

Remarque 1.1. On peut choisir le temps d'arrêt N pour que :

$$\frac{1}{2^N}(b_0 - a_0) < \epsilon = \text{précision choisie.}$$

Remarque 1.2. Avec la méthode de la dichotomie, les itérations s'achèvent à la m -ème étape quand $|x_m - \hat{x}| \leq |I_m| < \epsilon$, où ϵ est une tolérance fixée et $|I_m|$ désigne la longueur de l'intervalle I_m . Clairement $I_k = \frac{b-a}{2^k}$, donc pour avoir une erreur $|x_m - \hat{x}| < \epsilon$, on doit prendre le plus petit m qui vérifie

$$m \geq \log_2 \left(\frac{b-a}{\epsilon} \right) = \frac{\ln(b-a) - \ln(\epsilon)}{\ln(2)}.$$

Notons que cette inégalité est générale, elle ne dépend pas du choix de la fonction f .

Exemple 1.1. Utiliser la méthode de la dichotomie pour calculer le zéro de la fonction

$$f(x) = x^3 + 2x - 1.$$

f est un polynôme donc dérivable sur \mathbb{R} . Pour tout $x \in \mathbb{R}$, $f'(x) = 3x^2 + 2 \geq 2 > 0$. Donc f est strictement croissante sur \mathbb{R} . On a aussi $\lim_{x \rightarrow +\infty} f(x) = \lim_{x \rightarrow +\infty} x^3 = +\infty$ et

$\lim_{x \rightarrow -\infty} f(x) = \lim_{x \rightarrow -\infty} x^3 = -\infty$. Le théorème de la bijection assure alors que l'équation $f(x) = 0$ admet une unique solution α dans \mathbb{R} .

De plus f est une fonction continue et strictement croissante sur l'intervalle $[0; 1]$, car sur \mathbb{R} , et $f(0) \times f(1) = -0,125 < 0$, le théorème de la bijection assure alors que l'équation $f(x) = 0$ admet une unique solution α dans $[0; 1]$.

Conséquence : L'équation $f(x) = 0$ admet une unique solution α dans \mathbb{R} , et $\alpha \in [0; 1]$.

Mettre en place sur une feuille Excel le calcul d'un encadrement à d'amplitude 10^{-3} de la solution de l'équation $f(x) = 0$, par la méthode de Dichotomie.

Déterminons N le nombre d'itérations nécessaires pour obtenir une valeur approchée de la solution à 10^{-3} près. N est le plus petit m qui vérifie

$$m \geq \frac{\ln(1-0) - \ln(10^{-3})}{\ln(2)} = 9,9658.$$

Donc $N = 10$.

n	a_n	b_n	$c_n = \frac{a_n + b_n}{2}$	$f(a_n)$	$f(c_n)$	$f(a_n) \times f(c_n)$
0	0	1	0,5	-1	0,125	-0,125
1	0	0,5	0,25	-1	-0,484375	0,484375
2	0,25	0,5	0,375	-0,484375	-0,19726563	0,09555054
3	0,375	0,5	0,4375	-0,19726563	-0,04125977	0,00813913
4	0,4375	0,5	0,46875	-0,04125977	0,04049683	-0,00167089
5	0,4375	0,46875	0,453125	-0,04125977	-0,00071335	2,9433.10 ⁻⁰⁵
6	0,453125	0,46875	0,4609375	-0,00071335	0,01980734	-1,413.10 ⁻⁰⁵
7	0,453125	0,4609375	0,45703125	-0,00071335	0,00952607	-6,7954.10 ⁻⁰⁶
8	0,453125	0,45703125	0,45507813	-0,00071335	0,00440115	-3,1396.10 ⁻⁰⁶
9	0,453125	0,45507813	0,45410156	-0,00071335	0,0018426	-1,3144.10 ⁻⁰⁶
10	0,453125	0,45410156	0,45361328	-0,00071335	0,0005643	-4,0254.10 ⁻⁰⁷

$$\alpha \in [0,4531250; 0,4541016],$$

$$amplitude = 0,0009766 < 0,001.$$

1.5 Méthode de Lagrange (ou Regula falsi)

Principe 1.1. Soit deux points a_0 et b_0 (avec $a_0 < b_0$) d'images par f de signe contraire (i.e. $f(a_0) \times f(b_0) < 0$). En partant de $I_0 = [a_0, b_0]$, la méthode de LAGRANGE produisent une suite de sous-intervalles $I_k = [a_k, b_k]$, $k \geq 0$, avec $I_k \subset I_{k-1}$ pour $k \geq 1$ et tels que $f(a_k) \times f(b_k) < 0$.

Dans la méthode de LAGRANGE, plutôt que de diviser l'intervalle $[a_k, b_k]$ en deux intervalles de même longueur, on découpe $[a_k, b_k]$ en $[a_k, c_k]$ et $[c_k, b_k]$ où c_k est l'abscisse du point d'intersection de la droite passant par $(a_k, f(a_k))$ et $(b_k, f(b_k))$ et l'axe des abscisses, autrement dit c_k est solution de l'équation

$$\frac{f(b_k) - f(a_k)}{b_k - a_k}(c - b_k) + f(b_k) = 0$$

qui est

$$c_k = b_k - \frac{b_k - a_k}{f(b_k) - f(a_k)} f(b_k) = \frac{a_k f(b_k) - b_k f(a_k)}{f(b_k) - f(a_k)}.$$

Dans les deux cas, pour l'itération suivante, on pose soit $[a_{k+1}, b_{k+1}] = [a_k, c_k]$ soit $[a_{k+1}, b_{k+1}] = [c_k, b_k]$ de sorte à ce que $f(a_{k+1}) \times f(b_{k+1}) < 0$. La suite $(c_k)_{k \in \mathbb{N}}$ converge vers \hat{x} puisque la longueur de ces intervalles tend vers 0 quand k tend vers $+\infty$.

Exemple 1.2. Utiliser la méthode de LAGRANGE pour calculer le zéro de la fonction

$$f(x) = x^3 - 4x - 8.95$$

dans l'intervalle $[2; 3]$ avec une précision de 10^{-2} .

k	a_k	x_k	b_k	signe de $f(a_k)$	signe de $f(x_k)$	signe de $f(b_k)$
0	2.000000	2.596666667	3.000000	-	-	+
1	2.596666667	2.690262642	3.000000	-	-	+
2	2.690262642	2.702092263	3.000000	-	-	+
3	2.702092263	2.703541518	3.000000	-	-	+
4	2.703541518	2.703718378	3.000000	-	-	+
5	2.703718378	2.703739951	3.000000	-	-	+
6	2.703739951	2.703742582	3.000000	-	-	+

1.6 méthode de la sécante

Soit f admettant un zéro dans l'intervalle $[x_{-1}; x_0]$. Pour obtenir une première approximation de ce zéro, l'idée est de remplacer f par son interpolé linéaire sur $[x_{-1}; x_0]$, soit par :

$$Y(x) = f(x_0) + (x - x_0) \frac{f(x_0) - f(x_{-1})}{x_0 - x_{-1}}$$

l'unique fonction linéaire dont les valeurs coïncident avec celles de f en x_{-1} et x_0 . Pour calculer x_1 on prend l'intersection de l'axe des abscisses avec la droite passant par les points $(x_0, f(x_0))$ et $(x_{-1}, f(x_{-1}))$. L'approximation x_1 est alors obtenue en résolvant

$$Y(x_1) = 0$$

soit

$$x_1 = x_0 - f(x_0) \frac{x_0 - x_{-1}}{f(x_0) - f(x_{-1})}$$

Définition 1.5. Méthode de la Sécante

Il s'agit d'une méthode à trois niveaux : approcher les zéros de f se ramène à calculer la limite de la suite récurrente

$$\begin{cases} x_0 \text{ donné} \\ x_1 \text{ donné} \\ x_{k+1} = x_k - f(x_k) \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} \end{cases}$$

Algorithme 1.6.1. x_0 et x_{-1} étant donnés,

$$\begin{cases} \text{pour } k = 0, 1, 2; \dots \\ x_{k+1} = x_k - f(x_k) \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} \end{cases}$$

Figure : Méthode de la sécante

1.6.1 Critère d'arrêt

Cet algorithme n'est évidemment pas complet tant qu'on n'a pas précisé un critère d'arrêt. Nous verrons plus loin que, généralement, x_n converge vers la solution \hat{x} cherchée ce qui signifie que pour n grand, x_n est voisin de \hat{x} .

Un critère d'arrêt souvent utilisé consiste à choisir a priori une tolérance ϵ et à terminer l'algorithme lorsque

$$|f(x_n) - f(x_{n-1})| < \epsilon.$$

Pour éviter à l'ordinateur de tourner sans s'arrêter lorsqu'il n'y a pas convergence, il est évidemment indispensable de toujours mettre un critère limitant le nombre total d'itérations.

Exemple 1.3. *Approximation de $\sqrt{10}$. Soit f définie par $f(x) = x^2 - 10$, f continue, strictement croissante et convexe sur $[0, +\infty[$.*

Intervalle $[3, 4]$: $f(3) \leq 0$ et $f(4) \geq 0$, donc $\sqrt{10} \in [3, 4]$

$a_0 = 3$	$\epsilon_0 \leq 0,1666 \dots$
$a_1 = 3,14285714285 \dots$	$\epsilon_1 \leq 0,02040 \dots$
$a_2 = 3,16000000000 \dots$	$\epsilon_2 \leq 0,00239 \dots$
$a_3 = 3,16201117318 \dots$	$\epsilon_3 \leq 0,00028 \dots$
$a_4 = 3,16224648985 \dots$	$\epsilon_4 \leq 3,28 \dots \cdot 10^{-5}$
$a_5 = 3,16227401437 \dots$	$\epsilon_5 \leq 3,84 \dots \cdot 10^{-6}$
$a_6 = 3,16227723374 \dots$	$\epsilon_6 \leq 4,49 \dots \cdot 10^{-7}$
$a_7 = 3,16227761029 \dots$	$\epsilon_7 \leq 5,25 \dots \cdot 10^{-8}$
$a_8 = 3,16227765433 \dots$	$\epsilon_8 \leq 6,14 \dots \cdot 10^{-9}$

1.7 Méthode de Newton

Ici, au lieu d'assimiler la courbe $y = f(x)$ à une sécante, on l'assimile à la tangente en un point $(x_n, f(x_n))$ soit, la droite d'équation :

$$Y = f(x_n) + f'(x_n)(x - x_n)$$

Son intersection avec l'axe des abscisses fournit une approximation de la solution \hat{x} . A nouveau, on renouvelle le procédé jusqu'à obtenir une approximation suffisante.

Algorithme 1.7.1. x_0 étant donné

$$\begin{cases} \text{pour } k = 0, 1, 2; \dots \\ x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \end{cases}$$

Remarque 1.3. *Cet algorithme est initié à l'aide d'un seul point. Il peut être vu comme une modification de l'algorithme précédent où le quotient $\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}$ a été remplacé par $\frac{1}{f'(x_k)}$.*

1.8 Méthode de point fixe

Définition 1.6. Point fixe

Soit $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ une fonction. Si \hat{x} est tel que $\varphi(\hat{x}) = \hat{x}$, on dit qu' \hat{x} est un point fixe de φ .

La méthode du point fixe consiste à d'abord remplacer l'équation

$$f(x) = 0 \tag{1.1}$$

par l'équation

$$\varphi(x) = x \tag{1.2}$$

ayant mêmes solutions. On est ainsi ramené à la recherche des points fixes de l'application φ . Le remplacement de (1.1) par (1.2) est toujours possible en posant, par exemple, $\varphi(x) = f(x) + x$. Ce n'est évidemment pas forcément le meilleur choix. L'équation étant sous la forme (1.2), on a alors l'algorithme suivant :

Algorithme 1.8.1. *On choisit x_0*

$$\begin{cases} \text{pour } k = 0, 1, 2; \dots \\ x_{k+1} = \varphi(x_k) \end{cases}$$

cette méthode est justifiée par

Proposition 1.1. *Soit $\varphi : [a; b] \rightarrow [a; b]$ continue et $x_0 \in [a; b]$. Si x_n converge vers \hat{x} alors $\varphi(\hat{x}) = \hat{x}$.*

La démonstration est immédiate en passant à la limite dans l'égalité $x_{n+1} = \varphi(x_n)$ et en utilisant la continuité de φ au point \hat{x} .

Exemple 1.4. *Commençons par un exemple simple. On cherche à résoudre l'équation du second degré $x^2 - 2x - 3 = 0$. Il n'est pas nécessaire de recourir aux méthodes numériques pour résoudre ce problème, dont les deux solutions sont $r_1 = 3$ et $r_2 = -1$. Cet exemple permet cependant de mieux comprendre ce qui se passe lorsqu'on utilise l'algorithme des points fixes. Puisqu'il y a une infinité de façons différentes de transformer cette équation sous la forme $x = \varphi(x)$, nous en choisissons trois au hasard. Vous pouvez bien sûr recourir à d'autres.*

CHAPITRE 1. EQUATIONS NON LINÉAIRES

$$x = \sqrt{2x + 3} = \varphi_1(x) \text{ (en isolant } (x^2))$$

$$x = \frac{3}{x - 2} = \varphi_2(x)$$

$$x = \frac{x^2 - 3}{2} = \varphi_3(x)$$

Si on applique l'algorithme des points fixes à chacune des fonctions $\varphi_i(x)$ en partant de $x_0 = 4$ on obtient pour $\varphi_1(x)$

$$\begin{array}{rcl} x_1 & = & \varphi_1(4) = 3,3166248 \\ x_2 & = & \varphi_1(3,3166248) = 3,1037477 \\ x_3 & = & \varphi_1(3,1037477) = 3,0343855 \\ x_4 & = & \varphi_1(3,0343855) = 3,0114402 \\ \vdots & & \vdots \\ x_{10} & = & \varphi_1(x_9) = 3,0000157 \end{array}$$

L'algorithme semble donc converger vers la racine $r_1 = 3$. Reprenons l'exercice avec $\varphi_2(x)$, toujours en partant de $x_0 = 4$:

$$\begin{array}{rcl} x_1 & = & \varphi_2(4) = 1,5 \\ x_2 & = & \varphi_2(1,5) = -6,0 \\ x_3 & = & \varphi_2(-6,0) = -0,375 \\ x_4 & = & \varphi_2(-0,375) = -1,2631579 \\ \vdots & & \vdots \\ x_{10} & = & \varphi_2(x_9) = -1,0003387 \end{array}$$

On remarque que contrairement au cas précédent, les itérations convergent vers la racine $r_2 = -1$. En dernier lieu essayons l'algorithme avec la fonction φ_3

$$\begin{array}{rcl} x_1 & = & \varphi_3(4) = 6,5 \\ x_2 & = & \varphi_3(6,5) = 19,625 \\ x_3 & = & \varphi_3(19,625) = 191,0703 \\ x_4 & = & \varphi_3(191,0703) = 18252,43 \\ \vdots & & \vdots \end{array}$$

Visiblement, les itérations tendent vers l'infini et aucune des deux solutions possibles ne sera atteinte.

Cet exemple montre clairement que l'algorithme des points fixes, selon le choix de la fonction itérative $\varphi(x)$, converge vers l'une ou l'autre des racines et peut même diverger complètement dans certains cas. Il faut donc une analyse plus fine afin de déterminer dans quelles conditions la méthode des points fixes est convergente.

1.9 Convergence des algorithmes

1.9.1 Méthodes de point fixe

Théorème 1.3. Convergence (globale) des itérations de point fixe

Considérons une fonction $\varphi : [a; b] \rightarrow \mathbb{R}$. On se donne $x_0 \in [a; b]$ et on considère la suite $x_{k+1} = \varphi(x_k)$ pour $k \geq 0$. Si les deux conditions suivantes sont satisfaites :

1. **condition de stabilité** $\varphi(x) \in [a; b]$ pour tout $x \in [a; b]$.
2. **Condition de contraction stricte** IL existe $K \in [0, 1]$ tel que $|\varphi(x) - \varphi(y)| \leq K|x - y|$, pour tout $x, y \in [a; b]$, alors :
 - φ est continue,
 - φ a un seul point fixe \hat{x} dans $[a; b]$,
 - la suite $x_{n+1} = \varphi(x_n)$ converge vers \hat{x} pour tout choix de x_0 dans $[a; b]$.

preuve. Continuité La condition de contraction stricte implique que φ est continue puisque, si on prend une suite $(y_k)_{k \in \mathbb{N}} \in [a; b]$ qui converge vers un élément x de $[a; b]$, alors nous avons $|\varphi(x) - \varphi(y_n)| \leq K|x - y_n|$ et par suite $\lim_{k \rightarrow \infty} \varphi(y_k) = \varphi(x)$.

Existence : Commençons par prouver l'existence d'un point fixe de φ . La fonction $g(x) = \varphi(x) - x$ est continue dans $[a; b]$ et, grâce à la condition de stabilité, on a $g(a) = \varphi(a) - a \geq 0$ et $g(b) = \varphi(b) - b \leq 0$. En appliquant le théorème des valeurs intermédiaires, on en déduit que g a au moins un zéro dans $[a; b]$, i.e. φ a au moins un point fixe dans $[a; b]$.

unicité : L'unicité du point fixe découle de la condition de contraction stricte. En effet, si on avait deux points fixes distincts \hat{x}_1 et \hat{x}_2 , alors

$$|\hat{x}_1 - \hat{x}_2| = |\varphi(\hat{x}_1) - \varphi(\hat{x}_2)| \leq K|\hat{x}_1 - \hat{x}_2| < |\hat{x}_1 - \hat{x}_2|$$

ce qui est impossible.

Convergence : Prouvons à présent que la suite x_k converge vers l'unique point fixe \hat{x} quand k tend vers $+\infty$ pour toute donnée initiale $x_0 \in [a; b]$. On a

$$|x_{k+1} - \hat{x}| = |\varphi(x_k) - \varphi(\hat{x})| \leq K|x_k - \hat{x}|$$

où $K < 1$ est la constante de contraction. En itérant $k + 1$ fois cette relation on obtient :

$$|x_{k+1} - \hat{x}| \leq K^{k+1}|x_0 - \hat{x}|$$

i.e., pour tout $k \geq 0$

$$\frac{|x_{k+1} - \hat{x}|}{|x_0 - \hat{x}|} \leq K^{k+1}.$$

En passant à la limite quand k tend vers $+\infty$ on obtient $|x_{k+1} - \hat{x}|$ tend vers zéro. □

Il est important de disposer d'un critère pratique assurant qu'une fonction φ est contractante stricte. Pour cela, rappelons quelques définitions.

CHAPITRE 1. EQUATIONS NON LINÉAIRES

Théorème 1.4. Si $\varphi(x) : [a; b] \rightarrow [a; b]$ est de classe $\mathcal{C}^1([a; b])$ et si $|\varphi'(x)| < 1$ pour tout $x \in [a; b]$ alors la condition de contraction stricte est satisfaite avec $K = \max_{[a; b]} |\varphi'(x)|$.

Proposition 1.2. Soit \hat{x} un point fixe d'une fonction $\varphi \in \mathcal{C}^{p+1}$ pour un entier $p \geq 1$ dans un intervalle $[a; b]$ contenant \hat{x} . Si $\varphi^i(x) = 0$ pour $1 \leq i \leq p$ et $\varphi^{p+1}(x) \neq 0$, alors la méthode de point fixe associée à la fonction d'itération φ est d'ordre $p + 1$ et

$$\lim_{k \rightarrow +\infty} \frac{|x_{k+1} - \hat{x}|}{|(x_{k+1} - \hat{x})^{p+1}|} = \frac{\varphi^{p+1}(x)}{(p+1)!}.$$

Remarque 1.4. Il est clair qu'une convergence est d'autant plus rapide que son ordre est grand. En effet si $|x_k - \hat{x}|$ est petit $|x_k - \hat{x}|^2$ est encore plus petit...

Remarque 1.5. En particulier on a :

$$\begin{aligned} |x_{n+p} - x_n| &= |x_{n+p} - x_{n+p-1} + x_{n+p-1} - x_{n+p-2} + \cdots + x_{n+1} - x_n| \\ &\leq |x_{n+p} - x_{n+p-1}| + |x_{n+p-1} - x_{n+p-2}| + \cdots + |x_{n+1} - x_n| \\ &\leq (K^{n+p-1} + K^{n+p-2} + \cdots + K^n) |x_1 - x_0| \\ &\leq K^n \frac{1 - K^p}{1 - K} |x_1 - x_0| \text{ or } 0 < 1 - K^p \leq 1 \\ &\leq \frac{K^n}{1 - K} |x_1 - x_0| \end{aligned}$$

faisons tendre p vers l'infini, on obtient :

$$|\hat{x} - x_n| \leq \frac{K^n}{1 - K} |x_1 - x_0|,$$

et que $\frac{K^n}{1 - K}$ est une estimation de l'erreur entre la solution \hat{x} et le n -ième itéré de la suite x_n . Donc plus K est proche de 0, plus \hat{x} est proche de x_n .

Remarque 1.6. La convergence d'une méthode de points fixes est également assujettie au choix de la valeur initiale x_0 . En effet, un mauvais choix de x_0 peut résulter en un algorithme divergent même si les conditions du théorème 1.3 sont respectées.

Cela nous amène à définir le bassin d'attraction d'une racine \hat{x} .

Définition 1.7. Le bassin d'attraction de la racine \hat{x} pour la méthode de points fixes $x_{k+1} = \varphi(x_k)$ est l'ensemble des valeurs initiales x_0 pour lesquelles x_k tend vers \hat{x} lorsque k tend vers l'infini.

Définition 1.8. Un point fixe \hat{x} de la fonction $\varphi(x)$ est dit attractif si

$$|\varphi'(\hat{x})| < 1$$

et répulsif si

$$|\varphi'(\hat{x})| > 1$$

le cas où $|\varphi'(\hat{x})| = 1$ est indéterminé.

1.9.2 Méthode de Newton

Revenons un instant à la méthode du point fixe. D'après la formule de Taylor à l'ordre 2, on a, en supposant φ suffisamment régulière :

$$x_{k+1} - \hat{x} = \varphi(x_k) - \varphi(\hat{x}) = (x_k - \hat{x})\varphi'(\hat{x}) + \frac{1}{2}(x_k - \hat{x})^2\varphi''(\hat{x})(\xi_n)$$

où $\xi_n \in [x_k; \hat{x}]$. Pour k grand on a donc

$$x_{k+1} - \hat{x} \approx (x_k - \hat{x})\varphi'(\hat{x})$$

et la vitesse de convergence sera d'autant plus grande que $\varphi'(\hat{x})$ est plus petit. Le cas le plus favorable est celui où $\varphi'(\hat{x}) = 0$.

Si M_2 est un majorant de $\varphi''(x)$ sur $[a; b]$, on a alors,

$$|x_{k+1} - \hat{x}| \leq \frac{M_2}{2}|x_k - \hat{x}|$$

La convergence est alors au moins d'ordre 2 !

Si on revient à l'algorithme de Newton, on voit qu'il s'agit en fait d'un algorithme de point fixe pour la fonction $\varphi(x) = x - \frac{f(x)}{f'(x)}$

Théorème 1.5. Soit f deux fois continument différentiable sur un intervalle ouvert de centre \hat{x} vérifiant

$$f(\hat{x}) = 0 \text{ et } f'(\hat{x}) \neq 0$$

Alors, si x_0 est choisi assez près de \hat{x} , la méthode de Newton :

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \forall k$$

produit une suite x_k convergeant au moins quadratiquement vers \hat{x} .

1.9.3 Méthode de la sécante

Théorème 1.6. Soit f deux fois continument différentiable sur un intervalle ouvert de centre \hat{x} vérifiant

$$f(\hat{x}) = 0 \text{ et } f'(\hat{x}) \neq 0$$

Alors, si x_0 et x_1 sont choisis assez près de \hat{x} , l'algorithme de la sécante :

$$x_{k+1} = x_k - f(x_k) \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}, \forall k$$

converge vers \hat{x} et la convergence est au moins d'ordre $p = 1,618\dots$.

1.10 Résumés

1.10.1 Méthode de la dichotomie :

[suite d'intervalles encadrant la racine]

1. $[a_0, b_0] = [a, b]$.
2. $[a_n, b_n](\text{connu}) \Rightarrow w = \frac{(a_n + b_n)}{2}$.
3. — $f(a_n) \cdot f(w) < 0 \Rightarrow [a_{n+1}, b_{n+1}] = [a_n, w]$
 — $f(a_n) \cdot f(w) > 0 \Rightarrow [a_{n+1}, b_{n+1}] = [w, b_n]$
 — $f(a_n) \cdot f(w) = 0 \Rightarrow w = \text{racine}$
4. **Avantages :**
 - Convergence certaine vers la racine (si elle est unique sur l'intervalle), si f est continue.
 - Pour une précision ε , il faut faire n itération avec $n > \ln \frac{\frac{b-a}{\varepsilon}}{\ln 2}$ (car l'intervalle après n itérations est de longueur $\frac{b-a}{2^n}$).
5. **Inconvénients :**
 - Convergence linéaire (lente).
 - Nécessité du changement de signe sur l'intervalle.

1.10.2 Méthode de la sécante

[suite de valeurs tendant vers racine]

1. Fixer x_0 et x_1
2. Calculer

$$x_{n+2} = x_{n+1} - f(x_{n+1}) \frac{x_{n+1} - x_n}{f(x_{n+1}) - f(x_n)} \quad (n = 0, 1, 2, \dots)$$

3. **Avantages :**
 - Pour une fonction à un seul zéro peu d'hypothèses de départ.
 - Convergence rapide (pour racines simples, avec bon choix de x_0 et x_1)
4. **Inconvénients :**
 - Lente convergence.
 - Manque de précision pour une racine multiple.
 - Risque d'une division par zéro.

1.10.3 Méthode du point fixe / d'itération

1. Choisir x_1
2. Calculer

$$x_{n+1} = g(x_n) \quad (n = 1, 2, \dots)$$

3. Choisir $g(x)$ tel que $x = g(x)$.
4. **Avantages :** Facile à mettre en oeuvre, $g(x)$ qui converge rapidement n'est souvent pas trop difficile trouver.
5. **Inconvénients :**
 - Bien choisir $g(x)$, tel que $g(x) = x$. Il faut parfois différents $g(x)$ pour trouver différentes racines d'une même équation.
 - Convergence si $g'(x_0) < 1$ (vérifier graphiquement si graphe de $g'(x) < \text{pente } y = x$). (divergence si $g(x)$ pas bonne)
 - Convergence lente (si $g(x)$ est valable).

1.10.4 Méthode de Newton(Raphson)/de la tangente

(remplacer la fonction par sa tangente : approximation. Zéro de la tangente est proche de celui de la fonction si la tangente est très semblable à la fonction) [génère une suite de tangentes au graphe de $F(x)$ et une suite de valeurs x_n]

1. Choisir x_1 proche du zéro cherché.
2. Calculer :

$$x_{n+1} = x_n - \frac{F(x_n)}{F'(x_n)} \quad (n = 1, 2, \dots)$$

[avec $F(x) = 0$ équation à résoudre].

Ou approximation : $x_{n+1} = x_n - h \frac{F(x_n)}{F(x_n + h) - F(x_n)}$ ($n = 1, 2, \dots$ et h très petit).

3. **Avantages :**
 - Converge quadratiquement si racine simple.
 - Méthode valable pour trouver zéros complexes.
4. **Inconvénients :**
 - Converge linéairement si racine multiple.
 - Si racine multiple : F et F' très petites \Rightarrow convergence très lente.
 - Besoin de calculer la dérivée de la fonction.

1.10.5 Arrêt de l'algorithme

Arrêter l'algorithme quand l'écart entre deux solution x_n et x_{n+1} est assez faible (de l'ordre de 10^{-k} , avec k que l'on veut).

Chapitre 2

Interpolation et approximation polynômiale

Dans ce chapitre, on dispose d'une fonction f , connue par exemple uniquement par ses valeurs en certains points, et on cherche à remplacer ou à approcher f par une fonction plus simple, le plus souvent par un polynôme. Nous verrons dans ce contexte, l'interpolation qui consiste à rechercher un polynôme qui passe exactement par les points donnés, l'interpolation par morceaux, en particulier les fonctions splines où l'expression du polynôme est différente sur chaque sous-intervalle, l'approximation uniforme ou l'approximation au sens des moindres carrés ou on cherche à approcher au mieux la fonction, soit pour la norme uniforme (ou norme infinie), soit pour la norme euclidienne.

2.1 Interpolation de lagrange

Soit $f : [a; b] \rightarrow \mathbb{R}$ continue en $n + 1$ points distincts x_0, x_1, \dots, x_n de l'intervalle $[a; b]$. Il s'agit de construire un polynôme P de degré inférieur ou égal à n tel que

$$\forall i = 0, 1, \dots, n, \quad P(x_i) = f(x_i) \quad (2.1)$$

Théorème 2.1. *Il existe un et un seul polynôme de degré inférieur ou égal à n solution de (2.1). Le polynôme s'écrit*

$$P_n(x) = \sum_{i=0}^n f(x_i) L_i(x) \quad (2.2)$$

où

$$L_i(x) = \prod_{k=0, k \neq i}^n \frac{x - x_k}{x_i - x_k} \quad (2.3)$$

Remarque 2.1. Le polynôme P_n est appelé polynôme d'interpolation de Lagrange de la fonction f aux points x_0, x_1, \dots, x_n . Les polynômes $L_i(x)$ sont appelés polynômes de base de Lagrange associés à ces points.

Démonstration. Preuve Existence On vérifie directement que le polynôme donné par (2.2) est solution de (2.1) (on utilise le fait que $L_i(x_j) = \delta_{ij}$).

Unicité Soit Q un autre polynôme solution. Alors $\forall i = 0, 1, \dots, n, Q(x_i) - P(x_i) = 0$. Ainsi $Q - P$ est un polynôme de degré inférieur ou égal à n s'annulant en $n + 1$ points. Il est donc identiquement nul. \square

Exemple 2.1. Interpolation linéaire On applique (2.2) avec $n = 1$ pour trouver

$$P_1(x) = f(x_0) \frac{x - x_1}{x_0 - x_1} + f(x_1) \frac{x - x_0}{x_1 - x_0}$$

ceci s'écrit encore

$$P_1(x) = f(x_0) \frac{f(x_1) - f(x_0)}{x_1 - x_0} (x - x_0) \quad (2.4)$$

Exemple 2.2. Pour exprimer le polynôme d'interpolation sous la forme de Lagrange, il faut définir les polynômes associés à chacun des points, $L_i(x)$. Par exemple pour la série de points suivante :

x	0	1	2
y	2,6	2,7	2,9

$$L_0(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} = \frac{(x - 1)(x - 2)}{(0 - 1)(0 - 2)} = \frac{1}{2}x^2 - \frac{1}{2}x + 1$$

$$L_1(x) = \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} = \frac{(x - 0)(x - 2)}{(1 - 0)(1 - 2)} = -x^2 + 2x$$

$$L_2(x) = \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} = \frac{(x - 0)(x - 1)}{(2 - 0)(2 - 1)} = \frac{1}{2}x^2 - \frac{1}{2}x$$

On obtient alors le polynôme d'interpolation :

$$\begin{aligned} P_n(x) &= \sum_{i=0}^2 y_i L_i(x) = y_0 L_0 + y_1 L_1 + y_2 L_2 \\ &= 2,6 \left(\frac{1}{2}x^2 - \frac{1}{2}x + 1 \right) + 2,7 (-x^2 + 2x) + 2,9 \left(\frac{1}{2}x^2 - \frac{1}{2}x \right) \\ P_n(x) &= 0,05x^2 + 0,05x + 2,6 \end{aligned}$$

Remarque 2.2. L'écriture (2.2) du polynôme d'interpolation est intéressante au point de vue théorique, mais peu du point de vue numérique : elle a un caractère peu algorithmique. De plus, son évaluation requiert trop d'opérations élémentaires (le calcul d'un des polynômes de Lagrange de degré n nécessite $4n^2$ opérations). On lui préfère la formule de Newton. On rappelle qu'il n'existe qu'un seul polynôme d'ordre $n - 1$ interpolant n points.

2.2 Interpolation de Newton et différences divisées

On veut maintenant déterminer une nouvelle forme du polynôme d'interpolation P qui ne nécessite pas de recalculer toutes les fonctions de base lors de l'ajout d'un nouveau point. Si l'on a qu'un point d'interpolation x_0 alors,

$$P(x) = f(x_0).$$

Si l'on ajoute un point x_1 alors

$$P(x) = f(x_0) + a_1(x - x_0)$$

et

$$P(x_1) = f(x_1)$$

implique

$$a_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0}.$$

Si l'on a $n + 1$ points x_0, x_1, \dots, x_n , on cherche

$$P(x) = a_0 + a_1(x - x_0) + \dots + a_n \prod_{k=0}^{n-1} (x - x_k)$$

Remarquer que si l'on a déterminé a_0, a_1, \dots, a_{i-1} alors de l'égalité,

$$P(x_i) = f(x_i) = a_0 + a_1(x_i - x_0) + \dots + a_i(x_i - x_0)(x_i - x_1) + \dots (x_i - x_{i-1})$$

on déduit que a_i est déterminé par la formule :

$$a_i = \frac{f(x_i) - a_0 - a_1(x_i - x_0) - \dots - a_{i-1}(x_i - x_0)(x_i - x_1) \dots (x_i - x_{i-2})}{(x_i - x_0)(x_i - x_1) \dots (x_i - x_{i-1})}.$$

Ce qui montre que a_i ne dépend que des points x_0, \dots, x_i . On pose alors pour tout $k \in \{1, 2, \dots, n\}$

$$a_k = f[x_0, \dots, x_k].$$

Où $f[.]$ désigne les différences divisées de f définies par :

$$i = 0, \dots, n \quad f[x_i] = f(x_i) \tag{2.5}$$

$$f[x_{i-1}, x_i] = \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}}$$

$$f[x_0, \dots, x_n] = \frac{f[x_1, \dots, x_n] - f[x_0, \dots, x_{n-1}]}{x_n - x_0} \tag{2.6}$$

CHAPITRE 2. INTERPOLATION ET APPROXIMATION POLYNÔMIALE

Avec cette notation, on a la formule d'interpolation de Newton

$$P(x) = f(x_0) + f[x_0, x_1](x - x_0) + \cdots + f[x_0, \dots, x_n] \prod_{k=0}^{n-1} (x - x_k) \quad (2.7)$$

$$= \sum_{i=0}^n f[x_0, \dots, x_i] \prod_{k=0}^{i-1} (x - x_k) \quad (2.8)$$

Remarque 2.3. Par convention $\prod_{k=0}^{i-1} (x - x_k) = 1$ si $i < 1$

Lemme 2.1. Soit $n \in \mathbb{N}$ et soient $x_0, x_1, \dots, x_n, n+1$ réels distincts, pour toute permutation σ sur $\{0, 1, \dots, n\}$, on a :

$$f[x_0, \dots, x_n] = f[x_{\sigma(0)}, \dots, x_{\sigma(n)}]$$

Pour expliciter le processus récursif, les différences divisées peuvent être calculées en les disposant de la manière suivante dans un tableau :

i	x_i	y_i	$f[x_{i-1}, x_i]$	$f[x_{i-2}, x_{i-1}, x_i]$	$f[x_{i-3}, x_{i-2}, x_{i-1}, x_i]$	$f[x_{i-4}, x_{i-3}, x_{i-2}, x_{i-1}, x_i]$	\cdots
0	x_0	y_0					
1	x_1	y_1	$f[x_0, x_1]$				
2	x_2	y_2	$f[x_1, x_2]$	$f[x_0, x_1, x_2]$			
3	x_3	y_3	$f[x_2, x_3]$	$f[x_1, x_2, x_3]$	$f[x_0, x_1, x_2, x_3]$		
4	x_4	y_4	$f[x_3, x_4]$	$f[x_2, x_3, x_4]$	$f[x_1, x_2, x_3, x_4]$	$f[x_0, x_1, x_2, x_3, x_4]$	
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\ddots

Si on reprend l'exemple de la section précédente :

Exemple 2.3.

x	0	1	2
y	2,6	2,7	2,9

on obtient :

$$\begin{aligned}
 f[x_0, x_1] &= \frac{f(x_0) - f(x_1)}{x_0 - x_1} = \frac{2,6 - 2,7}{0 - 1} = 0,1 \\
 f[x_1, x_2] &= \frac{f(x_1) - f(x_2)}{x_1 - x_2} = \frac{2,7 - 2,9}{1 - 2} = 0,2 \\
 f[x_0, x_1, x_2] &= \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0} = \frac{0,2 - 0,1}{2 - 0} = 0,05
 \end{aligned}$$

ou en utilisant le tableau

i	x_i	y_i	$f[x_{i-1}, x_i]$	$f[x_{i-2}, x_{i-1}, x_i]$	$f[x_{i-3}, x_{i-2}, x_{i-1}, x_i]$
0	0	2.6			
1	1	2.7	$f[x_0, x_1] = 0.1$		
2	2	2.9	$f[x_1, x_2] = 0.2$	$f[x_0, x_1, x_2] = 0.05$	

d'où

$$\begin{aligned}
 P(x) &= \sum_{i=0}^n f[x_0, \dots, x_i] \prod_{k=0}^{i-1} (x - x_k) \\
 &= f(x_0) + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1) \\
 &= 2,6 + 0,1(x - 0) + 0,05(x - 0)(x - 1) \\
 &= 2,6 + 0,05x + 0,05x^2
 \end{aligned}$$

Algorithme 2.2.1.

2.3 Erreur dans l'interpolation de Lagrange

Le but de l'interpolation étant de remplacer l'évaluation de $f(x)$ par celle de $P(x)$, il est important de connaître l'erreur. Soit P_n le polynôme d'interpolation de f aux points x_0, x_1, \dots, x_n . On pose

$$e_n(x) = f(x) - P_n(x)$$

Proposition 2.1. Pour tout réel $x \in \mathbb{R}$, on a

$$e_n(x) = \begin{cases} 0 & \text{si } x \in \{x_0, \dots, x_n\} \\ f[x_0, \dots, x_n, x] \prod_{k=0}^n (x - x_k) & \text{sinon} \end{cases}$$

Preuve 1. On suppose que $x \notin \{x_0, \dots, x_n\}$. Soit P_{n+1} le polynôme d'interpolation de f aux points $\{x_0, \dots, x_n\}$. Alors

$$e_n(x) = f(x) - P_n(x) = P_{n+1}(x) - P_n(x) = f[x_0, \dots, x_n, x] \prod_{k=0}^n (x - x_k)$$

Théorème 2.2. Soient f une fonction définie sur \mathbb{R} , de classe C^k et x_0, \dots, x_n , des réels distincts. On pose $a = \min\{x_0, \dots, x_n\}$ et $b = \max\{x_0, \dots, x_n\}$. Alors il existe $\xi \in]a; b[$ tel que

$$f[x_0, \dots, x_k] = \frac{f^{(k)}(\xi)}{k!}$$

Preuve 2. Dans le cas où $k = 1$, c'est une application du théorème des accroissements finis. Supposons $k > 1$ quelconque. Soit P_k le polynôme d'interpolation de f aux points x_0, \dots, x_k . Alors, la fonction :

$$e_k(x) = f(x) - P_k(x)$$

s'annule au moins en $k + 1$ points distincts : x_0, \dots, x_k . Donc, d'après le théorème de Rolle, la fonction e'_k s'annule au moins en k points dans $]a, b[$. La fonction e''_k s'annule au moins en $k - 1$ points dans $]a, b[$. De même, la fonction $(e_k)^{(k)}$ s'annule en au moins un point dans $]a, b[$. Donc :

$$\exists \xi \in]a, b[\text{ tel que } (e_k)^{(k)}(\xi) = f^{(k)}(\xi) - P_k^{(k)}(\xi) = 0.$$

Or

$$P_k^{(k)}(\xi) = k! f[x_0, \dots, x_k].$$

Donc

$$f[x_0, \dots, x_k] = \frac{f^{(k)}(\xi)}{k!}$$

Théorème 2.3. Soient f une fonction définie sur \mathbb{R} , de classe C^{n+1} sur $]a, b[$, x_0, \dots, x_n , des réels distincts de $[a, b]$ et P_n le polynôme d'interpolation de f sur le support x_0, \dots, x_n . Alors pour tout x dans $[a, b]$ il existe $\xi \in [a, b]$ tel que

$$e_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{k=0}^n (x - x_k)$$

Preuve 3. Elle découle de la proposition et du Théorème précédents.

Théorème 2.4. Soient f une fonction définie sur \mathbb{R} , de classe C^{n+1} sur $]a, b[$, x_0, \dots, x_n , des réels distincts de $[a, b]$ et P_n le polynôme d'interpolation de f sur le support x_0, \dots, x_n . Alors

$$|f(x) - P_n(x)| \leq \frac{M_{n+1}}{(n+1)!} |\pi_n(x)| \quad (2.9)$$

où

$$M_{n+1} = \max_{a \leq x \leq b} |f^{(n+1)}(x)| \quad (2.10)$$

et

$$\pi_n(x) = \prod_{i=0}^n (x - x_i) \quad (2.11)$$

Lemme 2.2. Sous les hypothèses du théorème, il existe $\xi \in [a, b]$ tel que

$$f[x_0, \dots, x_n] = \frac{f^{(n)}(\xi)}{n!}$$

Chapitre 3

Dérivation et intégration numérique

3.1 Introduction

Le but de ce chapitre est double : **indiquer comment calculer de façon approchée une dérivée et décrire les méthodes de base pour le calcul numérique d'intégrales.** Un point commun de ces deux thèmes qui vont nous occuper est l'utilisation de polynômes d'interpolation. L'idée sous-jacente est simple : la fonction avec laquelle on travaille est trop compliquée ou pas assez connue pour qu'on puisse faire des opérations simples comme la dérivation ou l'intégration, on choisit alors de la remplacer par un polynôme et on dérive ou on intègre celui-ci. On présentera assez rapidement les formules de dérivation approchée qui en résultent avec un aperçu sur les méthodes de différences finies. On passera plus de temps sur le calcul d'intégrales du type

$$\int_a^b f(x)w(x)dx \quad (3.1)$$

où w est une fonction poids (continue strictement positive sur $]a; b[$) et $f(x)w(x)$ est intégrable sur $[a; b]$. Les méthodes d'intégration numérique que nous décrirons consistent toutes à remplacer l'intégrale (3.1) par une expression le plus souvent de la forme :

$$\sum_{i=0}^N A_i f(x_i) \quad (3.2)$$

où les x_i sont des points distincts de $[a; b]$ et A_i des coefficients réels, le tout choisi pour que la différence

$$E = \int_a^b f(x)w(x)dx - \sum_{i=0}^N A_i f(x_i)$$

soit petite.

3.2 Dérivation numérique

Dans ce paragraphe, la fonction f n'est bien sûr pas connue par une formule explicite mais

- ou bien par ses valeurs sur un ensemble discret (en supposant les points assez proches pour que la notion de dérivée ait un sens)
- ou bien, le plus souvent, par un algorithme de calcul ou une formule compliquée qui permet, au moins en théorie, de la calculer en tout point. On suppose bien sûr que la dérivée n'est pas accessible par un procédé analogue.

Dans toute la suite, on supposera f connue ou calculable aux points $\dots, x_{i-2}, x_{i-1}, x_i, x_{i+1}, x_{i+2}, \dots$ qu'on supposera proches. On notera $h_i = x_{i+1} - x_i$.

3.2.1 Dérivée première

Supposons qu'on veuille calculer une valeur approchée de $f'(x_i)$. Une première idée, déjà évoquée dans le chapitre précédent, consiste à remplacer f par une fonction spline passant par les mêmes points et à prendre la dérivée de la fonction spline. D'un point de vue numérique, cette idée est très bonne, assez stable et donne de bons résultats en terme d'erreur. Malheureusement, elle est assez lourde à mettre en place, nécessite beaucoup de calculs. Elle a de plus un petit côté moralement désagréable : la dérivée est une notion purement locale, en ce sens que la dérivée de f au point x_i ne dépend que des valeurs prises par f au voisinage de x_i , alors que la fonction spline dépend globalement de f par l'intermédiaire d'un système linéaire qui fait intervenir toutes les valeurs de f .

On préfère donc le plus souvent utiliser une idée plus simple : on écrit un polynôme d'interpolation au voisinage du point x_i et on dérive celui-ci. Les formules vont varier en fonction du nombre de points qu'on choisit pour écrire le polynôme d'interpolation (en général 2 ou 3, plus rarement 4 ou 5).

Formules à deux points

Le polynôme d'interpolation sur les deux points x_i et x_{i+1} s'écrit :

$$P(x) = f(x_i) + f[x_i, x_{i+1}](x - x_i).$$

On a donc $P'(x) = f[x_i, x_{i+1}]$, ce qui fournit la

1 Formule décentrée à droite

$$f'(x_i) \simeq \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}. \quad (3.3)$$

On a aussi la

2 Formule décentrée à gauche

$$f'(x_i) \simeq \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}}. \quad (3.4)$$

Formules à trois points

On choisit d'interpoler sur les points x_{i-1}, x_i, x_{i+1} (ce qui est moralement plus satisfaisant). Dans ce cas on a :

$$P(x) = f(x_{i-1}) + f[x_{i-1}, x_i](x - x_{i-1}) + f[x_{i-1}, x_i, x_{i+1}](x - x_{i-1})(x - x_i).$$

On a donc

$$P'(x) = f[x_{i-1}, x_i] + f[x_{i-1}, x_i, x_{i+1}](x_i - x_{i-1}),$$

ce qui fournit, après simplification la

Formule centrée

$$f'(x_i) \simeq f(x_{i+1}) \frac{h_{i-1}}{h_i(h_{i-1} + h_i)} + f(x_i) \left(\frac{1}{h_{i-1}} - \frac{1}{h_i} \right) - f(x_{i-1}) \frac{h_i}{h_{i-1}(h_{i-1} + h_i)} \quad (3.5)$$

La formule (3.5) se simplifie notablement dans le cas de points équidistants ($h_{i-1} = h_i = h$) pour donner la

Formule centrée-points équidistants

$$f'(x_i) \simeq \frac{f(x_{i+1}) - f(x_{i-1}))}{2h}. \quad (3.6)$$

Remarquons que la formule ci-dessus n'est autre que la moyenne des deux formules décentrées dans le cas de points équidistants.

Erreur

Pour le calcul théorique de l'erreur commise quand on remplace $f'(x_i)$ par l'une des formules approchées ci-dessus, on revient à la démonstration effectuée dans le cadre de l'interpolation. On obtient :

$$\left| f'(x_i) - \frac{f(x_i) - f(x_{i-1}))}{x_i - x_{i-1}} \right| \leq \frac{M_2}{2} h_{i-1}. \quad (3.7)$$

On a de même

$$\left| f'(x_i) - \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} \right| \leq \frac{M_2}{2} h_i \quad (3.8)$$

et dans le cas centré, avec les points équidistants

$$\left| f'(x_i) - \frac{f(x_{i+1}) - f(x_{i-1}))}{2h} \right| \leq \frac{M_3}{6} h^2 \quad (3.9)$$

3.2.2 Dérivée d'ordre supérieur

Le principe est exactement le même, il faut simplement prendre garde que le degré du polynôme d'interpolation soit suffisant pour que sa dérivée n-ième soit non nulle ! Par exemple, pour la dérivée seconde, on choisit en général d'interpoler sur 3 points, ce qui donne (dans le cas de points équidistants)

Dérivée seconde - points équidistants

$$f''(x_i) \simeq \frac{f(x_{i+1}) + f(x_{i-1}) - 2f(x_i)}{h^2}. \quad (3.10)$$

avec une erreur majorée par

$$\left| f''(x_i) - \frac{f(x_{i+1}) + f(x_{i-1}) - 2f(x_i)}{h^2} \right| \leq \frac{M_4}{12} h^2 \quad (3.11)$$

Remarque 3.1. Application : les méthodes de différences finies. Les formules ci-dessus sont à la base des méthodes de différences finies pour résoudre de façon approchée des équations différentielles ou aux dérivées partielles.

3.3 Intégration numérique : méthodes composites

3.3.1 Principe

On veut calculer $\int_a^b f(x)dx$.

On décompose l'intervalle $[a; b]$ en $a = x_0 < x_1 < \dots < x_n = b$.

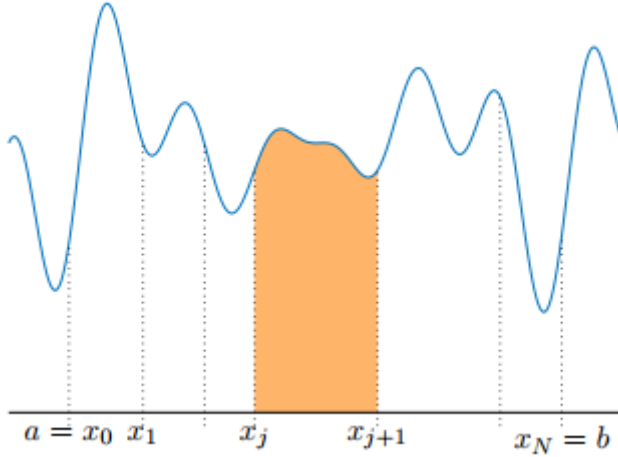


On a alors :

$$\int_a^b f(x)dx = \sum_{i=1}^n \int_{x_{i-1}}^{x_i} f(x)dx.$$

Sur chaque $[x_{i-1}; x_i]$, on applique une méthode d'intégration élémentaire consistant à remplacer $f(x)$ par

$$P_i = f(x_{i,0}) + f[x_{i,0}, x_{i,1}](x - x_{i,0}) + \dots + f[x_{i,0}, x_{i,1}, \dots, x_{i,l}](x - x_{i,0}) \dots (x - x_{i,l}) \quad (3.12)$$



le polynôme d'interpolation de f en des points $x_{i,0}, \dots, x_{i,l}$ de l'intervalle $[a; b]$ (qui peuvent être ou non dans $[x_{i-1}, x_i]$).

3.3.2 Méthode des rectangles

On prend $l = 0$: on approche f par une constante $f(x_{i,0})$ où $x_{i,0} \in [x_{i-1}; x_i]$. D'où la formule de quadrature (i.e. intégration numérique) élémentaire :

$$\int_{x_{i-1}}^{x_i} f(x)dx \simeq h_i f(\xi_i), \quad \xi_i \in [x_{i-1}; x_i], \quad h_i = x_i - x_{i-1} \quad (3.13)$$

et la formule de quadrature composée :

$$\int_a^b f(x)dx \simeq \sum_{i=1}^n h_i f(\xi_i), \quad (3.14)$$

On reconnait là une somme de Riemann dont on sait qu'elle converge vers $\int_a^b f(x)dx$ quand $\max_{1 \leq i \leq n} h_i \rightarrow 0$ si f est Riemann-intégrable. Les choix courants pour ξ_i sont :

— rectangles à gauche

$$\xi_i = x_{i-1} \Rightarrow \int_a^b f(x)dx \simeq \sum_{i=1}^n h_i f(x_{i-1})$$

— rectangles à droite

$$\xi_i = x_i \Rightarrow \int_a^b f(x)dx \simeq \sum_{i=1}^n h_i f(x_i)$$

— formule du point milieu

$$\xi_i = \frac{x_{i-1} + x_i}{2} = x_{i-\frac{1}{2}} \Rightarrow \int_a^b f(x)dx \simeq \sum_{i=1}^n h_i f(x_{i-\frac{1}{2}})$$

3.3.3 Méthode des trapèzes

On prend $l = 1$ et on remplace f par son interpolé linéaire aux points $[x_{i-1}; x_i]$

$$\int_{x_{i-1}}^{x_i} f(x)dx \simeq \frac{1}{2}(x_i - x_{i-1})(f(x_i) + f(x_{i-1}))$$

$$\int_a^b f(x)dx \simeq \sum_{i=1}^{n-1} \frac{h_i + h_{i+1}}{2} f(x_i) + \frac{1}{2}(h_1 f(x_0) + h_n f(x_n))$$

et dans le cas d'une subdivision régulière ($h_i = h, \forall i = 1, \dots, n$)

$$\int_a^b f(x)dx \simeq h \left[\sum_{i=1}^{n-1} f(x_i) + \frac{1}{2}(f(x_0) + f(x_n)) \right]$$

3.3.4 Méthode de Simpson

Cette fois on interpole f aux points x_{i-1}, x_i et $x_{i-\frac{1}{2}} = \frac{x_{i-1} + x_i}{2}$ soit, en utilisant la formule d'interpolation de Newton (2.8)

$$\begin{aligned} \int_{x_{i-1}}^{x_i} f(x)dx &\simeq \int_{x_{i-1}}^{x_i} f(x_{i-1}) + f[x_{i-1}, x_i](x - x_{i-1}) + f[x_{i-1}, x_i, x_{i-\frac{1}{2}}](x - x_{i-1})(x - x_i)dx \\ &= h_i f(x_{i-1}) + \frac{1}{2} h_i^2 f[x_{i-1}, x_i] + f[x_{i-1}, x_i, x_{i-\frac{1}{2}}] \int_{x_{i-1}}^{x_i} (x - x_{i-1})(x - x_i)dx \end{aligned}$$

or

$$\int_{x_{i-1}}^{x_i} (x - x_{i-1})(x - x_i)dx = \frac{1}{3} h_i^3 - \frac{1}{2} h_i^3$$

(en écrivant $x - x_i = x - x_{i-1} + x_{i-1} - x_i$)

et

$$\begin{aligned} h_i^2 f[x_{i-1}, x_i] &= h_i (f(x_i) - f(x_{i-1})) \\ h_i^3 f[x_{i-1}, x_{i-\frac{1}{2}}, x_i] &= h_i^2 \left(\frac{f(x_i) - f(x_{i-\frac{1}{2}})}{\frac{h_i}{2}} - \frac{f(x_{i-\frac{1}{2}}) - f(x_{i-1})}{\frac{h_i}{2}} \right). \end{aligned}$$

On obtient donc

$$\int_{x_{i-1}}^{x_i} f(x)dx = \frac{h_i}{6}(f(x_{i-1}) + 4f(x_{i-\frac{1}{2}}) + f(x_i))$$

La formule composée dans le cas de pas constants devient

$$\int_a^b f(x)dx = \frac{h}{6} \left[f_0 + f_n + 2 \sum_{i=1}^{n-1} f_i + 4 \sum_{i=1}^n f_{i-\frac{1}{2}} \right]$$

où on note

$$f_i = f(x_i), \quad f_{i-\frac{1}{2}} = f(x_{i-\frac{1}{2}}) = f\left(\frac{x_i + x_{i-1}}{2}\right)$$

Remarque 3.2. Pour cette méthode, il est nécessaire de connaître f aux points x_i et en leurs milieux, soit en un nombre impair de points. De manière générale, on appelle Formules de Newton-Cotes (fermées) les formules composées lorsque, sur chaque intervalle $[x_{i-1}, x_i]$, f est interpolée en l points équidistants. Les cas $l = 1$ et $l = 2$ correspondent donc respectivement aux formules des trapèzes et de Simpson, le cas $l = 4$ est connu dans la littérature sous le nom de Boole-Villarceau.

3.4 Analyse de l'erreur dans les méthodes d'intégration

3.4.1 Théorie

Nous évaluons ici l'erreur faite lorsqu'on remplace $\int_a^b f(x)dx$ par $\int_a^b P(x)dx$ où P est le polynôme d'interpolation de f aux points x_0, \dots, x_N .

On a vu au chapitre précédent que si $P_N(x)$ est le polynôme d'interpolation de f aux points x_0, \dots, x_N alors :

$$f(x) - P_N(x) = f[x_0, x_1, \dots, x_N, x] \Psi_N \quad (3.15)$$

avec $\Psi_N = (x - x_0)(x - x_1) \dots (x - x_N)$.

Ainsi

$$E(f) = \int_a^b f(x)dx - \int_a^b P_N(x)dx = \int_a^b f[x_0, x_1, \dots, x_N, x] \Psi_N dx \quad (3.16)$$

L'expression de $E(f)$ se simplifie dans deux cas particuliers.

Premier cas Ψ_N est de signe constant sur $[a, b]$, on peut alors utiliser le théorème de la moyenne pour les intégrales qui affirme que, si g et h sont continues sur $[a, b]$ et si $h(x) \geq 0$ sur $[a, b]$, il existe $\xi \in [a, b]$ tel que

$$\int_a^b g(x)h(x)dx = g(\xi) \int_a^b h(x)dx$$

CHAPITRE 3. DÉRIVATION ET INTÉGRATION NUMÉRIQUE

Appliquer à (3.16), ceci montre l'existence de $\xi \in [a, b]$ tel que

$$E(f) = f[x_0, x_1, \dots, x_N, \xi] \int_a^b \Psi_N(x) dx$$

D'après le lemme 2.2 du chapitre précédent, on en déduit encore qu'il existe $\eta \in [c, d]$ ($[c, d]$ est un intervalle contenant les x_i , a et b) tel que

$$E(f) = \frac{1}{(N+1)!} f^{N+1}(\eta) \int_a^b \Psi_N(x) dx \quad (3.17)$$

(On supposera ici et dans toute la suite que f est assez régulière).

Deuxième cas $\int_a^b \Psi_N(x) dx = 0$, dans ce cas, on utilise

$$f[x_0, x_1, \dots, x_N, x] = f[x_0, x_1, \dots, x_N, x_{N+1}] + (x_{N+1} - x)f[x_0, x_1, \dots, x_N, x_{N+1}, x] \quad (3.18)$$

valable pour x_{N+1} arbitraire. Alors (3.18) devient

$$E(f) = \int_a^b f[x_0, x_1, \dots, x_N, x_{N+1}] \Psi_N(x) dx + \int_a^b f[x_0, x_1, \dots, x_N, x_{N+1}, x] \Psi_{N+1}(x) dx$$

$$E(f) = \int_a^b f[x_0, x_1, \dots, x_N, x_{N+1}, x] \Psi_{N+1}(x) dx. \quad (3.19)$$

Si on peut choisir x_{N+1} de telle façon que $\Psi_{N+1}(x)$ reste de signe constant sur $[a, b]$, on aura d'après (3.17)

$$E(f) = \frac{1}{(N+2)!} f^{N+2}(\eta) \int_a^b \Psi_{N+1}(x) dx \quad (3.20)$$

pour un $\eta \in [c, d]$ ($[c, d]$ est un intervalle contenant les x_i , a et b)

Définition 3.1. On dit qu'une méthode d'intégration numérique est d'ordre k si elle est exacte pour tous les polynômes de degré inférieur ou égal à k .

Remarque 3.3. Lorsque $\Psi_N(x)$ est de signe constant sur $[a, b]$, d'après (3.17), on a une méthode d'ordre N puisque $P^{(N+1)}(\eta) = 0$ pour tout polynôme de degré inférieur ou égal à N . Dans la deuxième situation, si (3.17) est vraie, on aura même une méthode d'ordre $N+1$ (avec un même nombre de points).

3.4.2 Application aux méthodes usuelles

Méthode des rectangles à gauche ou à droite

$N = 0, x_0 = a, \Psi_0(x) = x - a$. On peut appliquer (3.17) qui donne

$$E(f) = f'(\eta) \int_a^b (x - a) dx = \frac{(b - a)^2}{2} f'(\eta)$$

La méthode est d'ordre 0 et exacte seulement pour les constantes.

$$\left| \int_{x_{i-1}}^{x_i} f(x) dx - h_i f(x_{i-1}) \right| \leq \frac{h_i^2}{2} f'(\eta_i) \leq h_i^2 2M_1$$

où $M_1 = \max_{a \leq x \leq b}$. Pour la formule composée, il faut ajouter les erreurs provenant de chaque intégration élémentaire. Dans le cas d'un pas constant, on obtient :

$$\left| \int_a^b f(x) dx - h \sum_{i=1}^N f(x_{i-1}) \right| \leq \frac{M_1}{2} (b - a) h$$

Formule du point milieu

$$N = 0, x_0 = \frac{a + b}{2}, \Psi_0(x) = x - \frac{a + b}{2}.$$

Ici, $\int_a^b \Psi_0(x) dx = 0$. On va donc obtenir une méthode d'ordre 1 (exacte pour toute fonction linéaire) bien qu'on interpole par une constante. En effet, d'après (3.20), appliqué

avec $x_{N+1} = x_0 = x_1 = \frac{a + b}{2}$

$$\left| \int_a^b f(x) dx - \frac{b - a}{2} (f(a) + f(b)) \right| \leq \frac{M_2}{2} \int_a^b \left(x - \frac{b - a}{2} \right)^2 dx = \frac{M_2}{24} (b - a)^3$$

avec $M_2 = \max_{a \leq x \leq b} |f''(x)|$. Pour la formule composée associée, on obtient

$$\left| \int_a^b f(x) dx - h \sum_{i=1}^N f(x_{i-\frac{1}{2}}) \right| \leq \frac{M_2}{24} h^2 (b - a)$$

Méthode des trapèzes

$$N = 1, x_0 = a, x_1 = b, \Psi_1(x) = (x - a)(x - b).$$

Comme Ψ_1 est de signe constant sur $[a; b]$ on peut appliquer (3.17) pour trouver

$$\left| \int_a^b f(x) dx - \frac{b - a}{2} (f(a) + f(b)) \right| \leq \frac{|f''(\eta)|}{2} \int_a^b (x - a)(x - b) dx$$

soit

$$\left| \int_a^b f(x) dx - \frac{b - a}{2} (f(a) + f(b)) \right| \leq \frac{M_2}{12} (b - a)^3$$

Pour la formule composée de la méthode des trapèzes, on en déduit que l'erreur est majorée par

$$E(f) \leq \frac{M_2}{12} h^3 (b - a)$$

Méthode de Simpson

$N = 2, x_0 = a, x_1 = \frac{a+b}{2}, x_2 = b, \Psi_2(x) = (x-a)(x-\frac{a+b}{2})(x-b).$

Par raison de symétrie, on vérifie que

$$\int_a^b \Psi_2(x) dx = 0$$

comme $\Psi_3(x) = (x-a)(x-\frac{a+b}{2})^2(x-b)$ est de signe constant sur $[a; b]$ on peut appliquer (3.20) pour obtenir

$$\left| \int_a^b f(x) dx - \frac{b-a}{6} (f(a) + 4f(\frac{b-a}{2}) + f(b)) \right| \leq \frac{M_4}{24} \int_a^b \Psi_3(x) dx$$

soit tous calculs faits

$$|E(f)| \leq \frac{M_4}{90} \left(\frac{b-a}{2}\right)^5$$

$$M_4 = \max_{a \leq x \leq b} |f^{(4)}(x)|.$$

La formule composée correspondante avec pas constant h donnera une erreur majorée par

$$|E(f)| \leq \frac{M_4}{2880} h^4 (b - a)$$

Par ailleurs, la méthode de Simpson (bien que reposant sur une interpolation à trois points) est exacte pour tout polynôme de degré inférieur ou égal à 3.

3.5 Exemple d'application

Exemple 3.1. Calcul de $\int_0^1 e^{-x^2} dx$.

Utilisons les méthodes élémentaires précédentes à l'aide des valeurs de $f(x) = e^{-x^2}$ aux points $0, \frac{1}{2}, 1$. soit

$$f(0) = 1, \quad f\left(\frac{1}{2}\right) = 0,77880, \quad f(1) = e^{-1} = 0,36788.$$

On a les résultats

CHAPITRE 3. DÉRIVATION ET INTÉGRATION NUMÉRIQUE

	<i>rectangle</i>	<i>trapèzes</i>	<i>Simpson</i>
$\int_0^1 e^{-x^2} dx$	1	0,68394	0,74718
<i>erreur</i>	0,25318	-0,06288	0,00036

La valeur exacte est 0,74682. Ces résultats sont tout à fait en accord avec les considérations théoriques précédentes. Noter comment la méthode de Simpson donne une approximation à $4 \cdot 10^{-4}$ avec seulement 3 valeurs de f . Ceci est bien sûr dû au fait que les dérivées d'ordre supérieur ne varient pas trop sur l'intervalle $[0, 1]$.

Exemple 3.2. Calcul d'une valeur approchée de $\ln 2 = \int_1^2 \frac{1}{t} dt$.

Si l'on pose, pour tout x de $[1, 2]$, $f(x) = \frac{1}{x}$, on a pour tout x de $[1, 2]$, $f'(x) = -\frac{1}{x^2}$ et $f''(x) = \frac{2}{x^3}$ et donc $\sup_{x \in [1,2]} |f'(x)| = 1$ et $\sup_{x \in [1,2]} |f''(x)| = 2$.

(1) Utilisation de la Méthode des rectangles

Avec $n = 10$, on obtient

$$\left| \ln 2 - \frac{1}{10} \left(1 + \frac{10}{11} + \frac{10}{12} + \frac{10}{13} + \frac{10}{14} + \frac{10}{15} + \frac{10}{16} + \frac{10}{17} + \frac{10}{18} + \frac{10}{19} \right) \right| \leq \frac{1}{20}.$$

Ce qui donne

$$\begin{aligned} \left| \ln 2 - \frac{33464927}{46558512} \right| &\leq \frac{1}{20}, \\ \text{puis } \frac{33464927}{46558512} - \frac{1}{20} &\leq \ln 2 \leq \frac{33464927}{46558512} + \frac{1}{20}. \end{aligned}$$

On en déduit que

$$0.66877 \leq \ln 2 \leq 0.76877.$$

(2) Utilisation de la Méthode des trapèzes

Avec $n = 10$, on obtient Posons

$$A = \frac{1}{10} \left(1 + \frac{10}{11} + \frac{10}{12} + \frac{10}{13} + \frac{10}{14} + \frac{10}{15} + \frac{10}{16} + \frac{10}{17} + \frac{10}{18} + \frac{10}{19} \right).$$

$$\left| \ln 2 - \frac{1}{20} \left(\frac{1}{2} - 1 \right) - A \right| \leq \frac{1}{600}.$$

Ce qui donne

$$\begin{aligned} \left| \ln 2 - \frac{161504821}{232792560} \right| &\leq \frac{1}{600}, \\ \text{puis } \frac{161504821}{232792560} - \frac{1}{600} &\leq \ln 2 \leq \frac{161504821}{232792560} + \frac{1}{600}. \end{aligned}$$

On en déduit que

$$0.6921 \leq \ln 2 \leq 0.69544.$$

Chapitre 4

Résolution numérique des équations différentielles

Les équations différentielles décrivent l'évolution de nombreux phénomènes dans des domaines variés. Une équation différentielle est une équation impliquant une ou plusieurs dérivées d'une fonction inconnue. Si toutes les dérivées sont prises par rapport à une seule variable, on parle d'équation différentielle ordinaire. Une équation mettant en jeu des dérivées partielles est appelée équation aux dérivées partielles.

4.1 Généralité

On appelle équation différentielle une équation reliant une fonction et ses dérivées successives. Si l'équation ne fait intervenir que la fonction et sa dérivée, on parle d'équation du premier ordre.

Soit Ω un ouvert de $\mathbb{R} \times \mathbb{R}^n$. Etant données $f : \Omega \subset \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ qui à tout couple $(t, y) \in \Omega$ associe $f(t, y)$ continu, un point $t_0 \in \mathbb{R}$, un point quelconque $Y_0 \in \mathbb{R}^n$, nous cherchons une fonction $y : I \rightarrow \mathbb{R}^n$, où I est un voisinage de t_0 dans \mathbb{R} , qui à associe $y(t)$, continûment dérivable, telle que y vérifie **le problème de Cauchy** suivant :

$$\begin{cases} y'(t) = f(t, y(t)) \\ y(t_0) = Y_0 \end{cases} \quad \text{condition initiale ou condition de Cauchy} \quad (4.1)$$

On dit qu'il y a unicité au problème de Cauchy en (t_0, Y_0) s'il existe au moins une solution à ce problème et si pour toutes solutions $\varphi : I \rightarrow \mathbb{R}^n$ et $\psi : J \rightarrow \mathbb{R}^n$, les fonctions φ et ψ coïncident sur $I \cap J$.

L'étude mathématique de l'existence et de l'unicité d'une solution y est délicate et constitue une branche entière des mathématiques. Nous nous contenterons de donner une condition suffisante à l'existence et l'unicité d'une solution au problème de Cauchy. Nous ne nous intéresserons ensuite qu'à la résolution numérique de ces équations différentielles.

CHAPITRE 4. RÉOLUTION NUMÉRIQUE DES ÉQUATIONS DIFFÉRENTIELLES

Théorème 4.1. *Théorème de Cauchy-Lipschitz (Admis)*

Si f est continue dans Ω et si f est localement lipschitzienne par rapport à sa deuxième variable, ie si pour tout $(t_0, Y_0) \in \Omega$ il existe un voisinage V de (t_0, Y_0) et $L > 0$ tels que

$$\forall (t, y), (t, z) \in V, |f(t, y) - f(t, z)| \leq |y - z|,$$

alors pour tout $(t_0, Y_0) \in \Omega$, il existe I voisinage de t_0 dans \mathbb{R} et une application y continûment dérivable de I dans \mathbb{R}^n solution unique de

$$\begin{cases} y'(t) = f(t, y(t)) \\ y(t_0) = Y_0 \end{cases} \quad (4.2)$$

Remarque 4.1. Si f est de classe C^1 alors f est localement lipschitzienne par rapport à sa deuxième variable, et le théorème de Cauchy-Lipschitz s'applique.

4.2 Méthodes de résolution numérique

Les méthodes numériques permettent de résoudre la majorité des équations différentielles indépendamment de leurs types, (la méthode d'Euler par exemple s'applique sur les équations linéaires et non linéaires). La méthode de résolution choisie, doit vérifier les critères suivants :

la consistance, la stabilité et la précision que nous verrons dans la suite du cours Une équation différentielle d'ordre n , peut être écrite sous forme d'un système de n équations différentielles d'ordre 1 :

$$\begin{cases} \frac{du}{dx} = f_1(u, x) \\ \frac{df_1}{dx} = f_2(f_1, x) \\ \vdots \\ \frac{df_{n-1}}{dx} = f_n(f_{n-1}, x) \end{cases} \quad (4.3)$$

Donc la résolution d'une équation différentielle d'ordre n , se ramène à la résolution combinée de n équations différentielles à valeurs initiales de type Cauchy :

$$\begin{cases} \frac{du}{dx} = f(u, x) \\ u(x_0) = u_0 \end{cases} \quad (4.4)$$

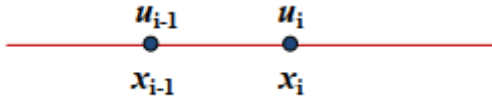
Le domaine de calcul $[x_0, x_n]$ est subdivisé en n sous intervalles $\{[x_{i-1}, x_i], i = 1, 2, \dots, n\}$.

Le pas de discrétisation h est supposé constant :

$$h = x_1 - x_0 = x_2 - x_1 = \dots = x_n - x_{n-1}$$

L'intégration de l'équation 4.4, entre deux noeuds successifs x_{i-1}, x_i :

CHAPITRE 4. RÉOLUTION NUMÉRIQUE DES ÉQUATIONS DIFFÉRENTIELLES



Permet d'aboutir à l'expression :

$$u_i = u_{i-1} + \int_{x_{i-1}}^{x_i} f(x, u) dx$$

Dans les méthodes à pas simple, seulement l'information u_{i-1} disponible au noeud x_{i-1} est utilisée pour calculer la valeur de u_i au noeud x_i . L'égalité ci-dessus, s'écrit sous la forme :

$$u_i = u_{i-1} + h * \Phi(x_{i-1}, u_{i-1})$$

Suivant le schéma numérique utilisé pour évaluer l'intégrale $\int_{x_{i-1}}^{x_i} f(x, u) dx$, (rectangles, Trapèzes, Simpson, ... etc); il existe plusieurs méthodes pour résoudre l'équation différentielle 4.4.

4.2.1 Méthode à un pas

Les méthodes à un pas s'écrivent de la manière suivante :

$$\begin{aligned} y_{j+1} &= y_j + h\phi(t_j, y_j, h), \\ y(0) &= Y_0 \end{aligned}$$

avec ϕ continue par rapport aux trois variables. Choisir une méthode revient à choisir la fonction ϕ . Quelles conditions imposer alors à ϕ pour que la méthode fonctionne ?

Exemple 4.1. Schéma d'Euler : $\phi(t, y, h) = f(t, y)$

On pose $e_j(h) = y(t_j) - y_j, \forall j = 0, 1, \dots, n$: l'erreur à l'étape j .

Objectif 1 : Quand h tend vers 0, $y(t_j)$ doit tendre vers y_j (i.e on a convergence de la solution approchée trouvée y_j vers la solution exacte au point $t_j, y(t_j)$..

Définition 4.1. Une méthode à un pas est dite **convergente** sur $[0, T]$ si quelle que soit Y_0 condition initiale

$$\begin{aligned} \lim_{h \rightarrow 0} \max_{0 \leq j \leq n} |e_j(h)| &= 0, \\ \lim_{h \rightarrow 0} \max_{0 \leq j \leq n} |y(t_j) - y_j| &= 0 \end{aligned}$$

Objectif 2 : Il faut que la méthode approche bien l'équation différentielle.

Définition 4.2. Une méthode à un pas est dite **consistante** si pour tout u solution de $y' = f(t, u)$

$$\lim_{h \rightarrow 0} \max_{0 \leq j \leq n} \left| \frac{1}{h} (y(t_j) - y_j) - \phi(t, y, h) \right| = 0$$

CHAPITRE 4. RÉOLUTION NUMÉRIQUE DES ÉQUATIONS DIFFÉRENTIELLES

Objectif 3 : Connaître et contrôler la répercussion des erreurs d'arrondi

Définition 4.3. Une méthode à un pas est dite **stable** si il existe K indépendant de h telle que, pour tous Y_0 et \tilde{Y}_0 satisfaisant

$$\begin{cases} y_{j+1} = y_j + h\phi(t_j, y(t_j), h) \\ \tilde{y}_{j+1} = \tilde{y}_j + h\phi(t_j, y(t_j), h) + \varepsilon_j \end{cases}$$

on a

$$|y_j - \tilde{y}_j| \leq K(|Y_0 - \tilde{Y}_0| + \sum_{i=0}^{j-1} |\varepsilon_i|), \quad \forall 0 \leq j \leq n.$$

Ces trois notions sont liées pour les schémas à un pas. En effet

Proposition 4.1. Pour une méthode à un pas, la convergence implique la consistance.

Théorème 4.2. Si une méthode à un pas est consistante et stable alors elle est convergente.

Proposition 4.2. Une condition nécessaire et suffisante pour qu'une méthode à un pas soit consistante est : $\phi(t, y, h) = f(t, y, 0)$, $\forall t \in [0, T], y \in \mathbb{R}^n$.

Justification

$$\lim_{h \rightarrow 0} \frac{y(t+h) - y(t)}{h} = y'(t) = f(t, y(t))$$

Comme ϕ est continue,

$$\lim_{h \rightarrow 0} \phi(t, y(t), h) = \phi(t, y(t), 0)$$

soit

$$\phi(t, y, h) = f(t, y, 0)$$

Proposition 4.3. Si ϕ est localement lipschitzienne par rapport à sa deuxième variable, alors la méthode est stable

Définition 4.4. On dit qu'une méthode est **d'ordre p** si

$$\max_j \left| \frac{1}{h} (y(t_{j+1}) - y(t_j)) - \phi(t_j, y(t_j), h) \right| = O(h^p)$$

4.2.2 Méthode d'Euler

On suppose que les hypothèses du théorème (4.1) sont vérifiées. Nous cherchons à approcher les valeurs de la solution y de l'équation (4.1) pour différentes valeurs de

CHAPITRE 4. RÉOLUTION NUMÉRIQUE DES ÉQUATIONS DIFFÉRENTIELLES

$t \in I = [0, T]$. Soit h la longueur du pas, on définit $t_j = jh, j = 0, 1, 2, \dots, n$:

$$\begin{aligned} \int_{t_j}^{t_{j+1}} y'(t) dt &= \int_{t_j}^{t_{j+1}} f(t, y(t)) dt \\ y(t_{j+1}) - y(t_j) &= \underbrace{\int_{t_j}^{t_{j+1}} f(t, y(t)) dt}_{\text{intégration numérique}} \end{aligned}$$

On pose $y_j = y(t_j)$ et $y_{j+1} = y(t_{j+1})$

En utilisant la formule des rectangles à gauche, on obtient :

$$\begin{aligned} y(t_{j+1}) - y(t_j) &= (t_{j+1} - t_j)(f(t_j, y_j)) \\ y(t_{j+1}) &= y(t_j) + (t_{j+1} - t_j)(f(t_j, y_j)) \end{aligned}$$

On obtient le schéma d'Euler explicite (explicite car on a directement y_{j+1} en fonction de y_j) :

$$\begin{aligned} y_{j+1} &= y_j + hf(t_j, y_j) \\ y(0) &= Y_0 \end{aligned}$$

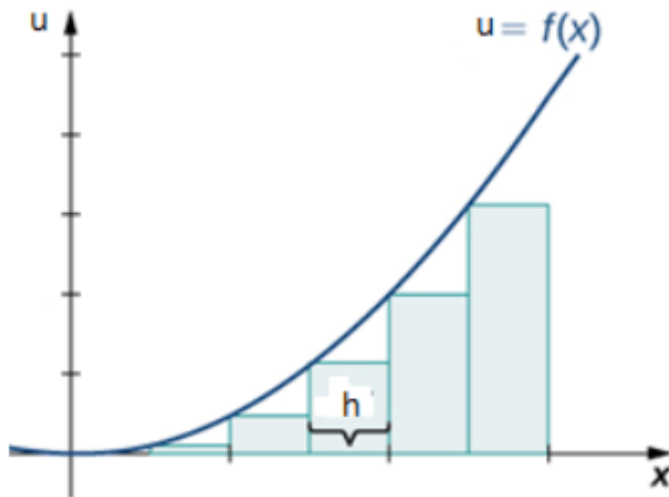


FIGURE 4.1 – Méthode des rectangles à gauche

En utilisant la formule des rectangles à droite, on obtient :

$$\begin{aligned} y(t_{j+1}) - y(t_j) &= (t_{j+1} - t_j)(f(t_{j+1}, y_{j+1})) \\ y(t_{j+1}) &= y(t_j) + (t_{j+1} - t_j)(f(t_{j+1}, y_{j+1})) \end{aligned}$$

CHAPITRE 4. RÉOLUTION NUMÉRIQUE DES ÉQUATIONS DIFFÉRENTIELLES

On obtient le schéma d'Euler implicite (implicite car y_{j+1} est des deux côtés de l'équation) :

$$\begin{aligned}y_{j+1} &= y_j + hf(t_j, y_{j+1}) \\ y(0) &= Y_0\end{aligned}$$

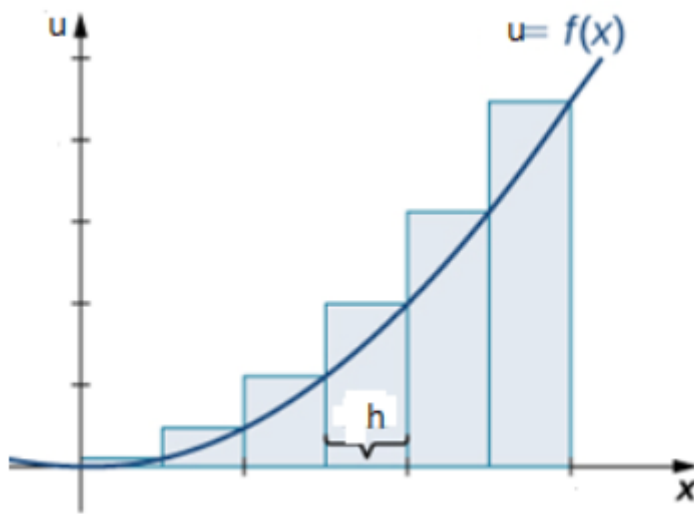


FIGURE 4.2 – Méthode des rectangles à droite

La valeur de y_j doit être calculée en fonction de y_j lui-même (implicite). Pour enlever ce problème d'implicite, on peut utiliser l'une des deux techniques suivantes :

Technique du prédicteur-correcteur :

Prédicteur : y_j^* est estimée par la méthode d'Euler explicite.

$$y_j^* = y_{j-1} + h * f(x_{j-1}, y_{j-1})$$

Correcteur : la valeur prédite y_j^* est utilisée pour corriger la solution implicite y_i .

$$y_j = y_{j-1} + h * f(x_j, y_j^*)$$

Cette équation est de la forme : $\Phi(y_j) = 0$, qu'on peut résoudre par l'une des méthodes étudiées au chapitre I. (Bissection, Newton, Point-fixe,).

CHAPITRE 4. RÉOLUTION NUMÉRIQUE DES ÉQUATIONS DIFFÉRENTIELLES

Remarque 4.2. *Le schéma d'Euler explicite est plus simple à utiliser mais moins précis que le schéma d'Euler implicite. De plus, le schéma d'Euler implicite est plus stable que le schéma d'Euler explicite.*

Ces méthodes sont peu précises, d'ordre 1 seulement et l'erreur est donc d'ordre 2 :

Exemple 4.2. Soit $\alpha > 0$ et l'équation

$$\begin{cases} y'(t) = -\alpha y(t) \\ y(0) = 1 \end{cases}$$

La fonction $y(t) = \exp(-\alpha t)$ est solution.

Par le schéma d'Euler explicite :

$$y_{j+1} = (1 - \alpha h)y_j \iff y_j = (1 - \alpha h)^n$$

Donc comme $y(t)$ tend vers 0 quand t tend vers l'infini, $|1 - \alpha h| < 1$, soit la condition de stabilité $h < \frac{2}{\alpha}$.

Par le schéma d'Euler implicite :

$$(1 + \alpha h)y_{j+1} = y_j \iff y_j = \frac{1}{(1 + \alpha h)^n}$$

Alors y_j tend vers 0 pour tout $h > 0$. Il n'y a pas de condition de stabilité.

Remarque 4.3. Le θ -schéma

En combinant les schémas d'Euler explicite et implicite, on obtient le θ -schéma :

$$\begin{aligned} y_{j+1} &= y_j + h[\theta f(t_{j+1}, y_{j+1}) + (1 - \theta)f(t_j, y_j)], 0 \leq \theta \leq 1 \\ y(0) &= Y_0 \end{aligned}$$

On retrouve le schéma d'Euler explicite pour $\theta = 0$, et l'implicite pour $\theta = 1$.

Pour les schémas implicites, la valeur de départ est en général donnée par un schéma explicite. Les schémas implicites sont à considérer comme des équations de point fixe.

Exemple Soit le problème à valeur initiale suivant :

$$\begin{cases} \frac{du}{dt} = u + e^{2t} \\ u(0) = 2 \end{cases}$$

Calculer $u(0.5)$, en utilisant un pas $\Delta t = 0.25$

Solution

Avec un pas $\Delta t = 0.25$, on doit effectuer 2 itérations de calcul, pour évaluer u à $t = 0.5$

CHAPITRE 4. RÉOLUTION NUMÉRIQUE DES ÉQUATIONS DIFFÉRENTIELLES

Méthode d'Euler explicite :

$$u_e(0.25) = u_e(0) + \Delta t * [u_e(0) + e^{2*t_0}] = 2 + 0.25 * (2 + e^{2*0}) = 2.7500$$

$$u_e(0.5) = u_e(0.25) + \Delta t * [u_e(0.25) + e^{2*t_1}] = 2.75 + 0.25 * (2.75 + e^{2*0.25}) = 3.8497$$

Méthode d'Euler implicite :

On utilise la technique du prédicteur-correcteur. La valeur prédite u_m^* est calculée par la méthode d'Euler explicite.

$$u_m^*(0.25) = u_e(0.25) = 2.7500$$

$$u_{im}(0.25) = u_{im}(0) + \Delta t * [u_m^*(0.25) + e^{2*t_1}] = 2 + 0.25 * (2.75 + e^{2*0.25}) = 3.0997$$

$$u_m^*(0.5) = u_e(0.5) = 3.8497$$

$$u_{im}(0.5) = u_{im}(0.25) + \Delta t * [u_m^*(0.5) + e^{2*t_2}] = 3.0997 + 0.25 * (3.8497 + e^{2*0.5}) = 4.7417$$

Comparaison avec la valeur exacte :

La solution analytique est : $u_{ex}(t) = e^t(e^t + 1)$

$$u_{ex}(0.5) = e^{0.5}(e^{0.5} + 1) = 4.3670$$

Pour comparer les résultats numériques des deux méthodes avec la valeur exacte, on calcule l'erreur relative : $\epsilon = \frac{|u - u_{ex}|}{u_{ex}}$

$u_e(0.5)$	$u_m(0.5)$	$u_{ex}(0.5)$	$\epsilon_{u_e(0.5)}$	$\epsilon_{u_m(0.5)}$
3.8497	4.7417	4.3670	11.84%	8.58%

FIGURE 4.3 – Comparaison des méthodes d'Euler explicite et implicite

Remarque 4.4. On peut augmenter la précision de ces méthodes en choisissant un pas de discrétisation h très petit. Mais cela augmentera énormément le temps et le coût de calcul.

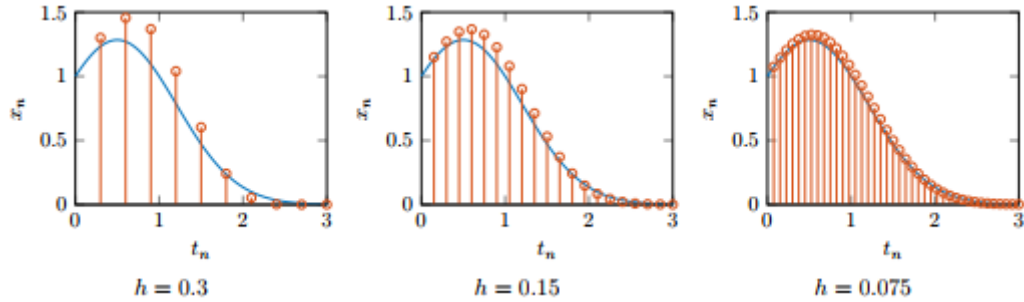


FIGURE 4.4 – Évolution par la méthode d'Euler

4.2.3 Méthodes de Runge-Kutta

Les méthodes de Runge-Kutta visent à étendre à la résolution d'équations différentielles ordinaires l'usage des techniques de calcul approché d'intégrales que sont les formules de quadrature interpolatoires. Elles font pour cela appel à de multiples évaluations de la fonction f , en des points obtenus par substitutions successives (pour les méthodes explicites) sur chaque sous-intervalle de la grille de discrétisation. Cet « échantillonnage » de la dérivée de la courbe intégrale recherchée permettant de réaliser l'intégration numérique approchée de cette dernière au moyen d'une somme pondérée des valeurs recueillies. Pour illustrer cette idée, donnons un premier exemple.

Exemple de méthode de Runge-Kutta explicite

On considère le problème

$$\begin{cases} y'(t) = f(t, y(t)), & t \in]t_0, t_f] \\ y(t_0) = Y_0, \end{cases}$$

et une subdivision $t_0 \leq t_1 \leq \dots \leq t_N = t_0 + T$.

Sur l'intervalle $[t_n, t_{n+1}]$, on a $y(t_{n+1}) = y(t_n) + \int_{t_n}^{t_{n+1}} f(s, y(s)) ds$ soit encore

$$y(t_{n+1}) = y(t_n) + h_n \int_0^1 f(t_n + \sigma h_n, y(t_n + \sigma h_n)) d\sigma$$

ou bien

$$y(t_{n+1}) = y(t_n) + h_n \int_0^1 g(\sigma) d\sigma, \quad g(\sigma) = f(t_n + \sigma h_n, y(t_n + \sigma h_n)).$$

On peut approcher le calcul de $\int_0^1 g(\sigma) d\sigma$ par la formule de quadrature (rectangle à gauche, rectangle à droite ou trapèze).

On veut maintenant utiliser une quadrature générale.

$$\int_0^1 g(\sigma) d\sigma \approx \sum_{i=1}^s b_i g(c_i), \quad \text{ou } g(c_i) = f(t_n + c_i h_n, y(t_n + c_i h_n))$$

CHAPITRE 4. RÉOLUTION NUMÉRIQUE DES ÉQUATIONS DIFFÉRENTIELLES

Le problème est que l'on ne connaît pas $y(t_n + c_i h_n) := y(t_{n,i})$. On évalue la fonction y au point $t_{n,i} = t_n + c_i h_n$ par une quadrature

$$y(t_{n,i}) = y(t_n) + h_n \int_0^{c_i} g(\sigma) d\sigma$$

Pour simplifier la présentation, on choisit ici des méthodes explicites où on approche $\int_0^{c_i} g(\sigma) d\sigma$ avec des valeurs de $g(c_j) = f(t_{n,j}, y(t_{n,j}))$, $j = 1, \dots, i-1$

$$\int_0^{c_i} g(\sigma) d\sigma \approx \sum_{j=1}^{i-1} a_{ij} g(c_j)$$

et on a donc

$$y(t_{n,i}) \approx y(t_n) + h_n \sum_{j=1}^{i-1} a_{ij} f(t_{n,j}, y(t_{n,j}))$$

$$y(t_{n+1}) \approx y(t_n) + h_n \sum_{i=1}^s b_i f(t_{n,i}, y(t_{n,i}))$$

avec $t_{n,i} = t_n + c_i h_n$.

La méthode de Runge-Kutta à s -étages explicite est donnée par

$$\begin{cases} k_i = f(t_{n,i}, y_n + h_n \sum_{j=1}^{i-1} a_{ij} k_j), & i = 1, \dots, s \\ y_{n+1} = y_n + h_n \sum_{i=1}^s b_i k_i \end{cases}$$

Une telle méthode est donc entièrement caractérisée par la donnée des coefficients $(a_{ij})_{1 \leq i, j \leq s}$, $(b_i)_{1 \leq i \leq s}$ et $(c_i)_{1 \leq i \leq s}$, que l'on a coutume de présenter dans le tableau de Butcher

c_1	a_{11}	a_{12}	\cdots	a_{1s}
c_2	a_{21}	a_{22}	\cdots	a_{2s}
\vdots	\vdots	\vdots		\vdots
c_s	a_{s1}	a_{ss}	\cdots	a_{ss}
	b_1	b_2	\cdots	b_s

Tableau de Butcher

Remarque 4.5. En se référant au tableau de Butcher, on voit qu'une méthode de Runge-Kutta est explicite si et seulement si la matrice A est strictement triangulaire inférieure. Si A est seulement triangulaire inférieure, c'est-à-dire si $a_{ij} = 0$ pour tout couple (i, j) de $1, \dots, s^2$ tel que $i < j$ mais qu'au moins l'un des coefficients a_{ii} , $1 \leq i \leq s$, est non nul, la méthode est semi-implicite.

CHAPITRE 4. RÉOLUTION NUMÉRIQUE DES ÉQUATIONS DIFFÉRENTIELLES

Remarque 4.6. Dans la méthode d'Euler l'ordre de précision est équivalent à la troncature de la série de Taylor à l'ordre 1 :

$$u_i = u_{i-1} + h * \underbrace{\left(\frac{du}{dx}\right)_{i-1}}_{f(x_{i-1}, u_{i-1})} + \theta^{(2)}$$

$\theta^{(2)}$: L'erreur de troncature d'ordre 2 :

$$\theta^{(2)} \simeq \frac{h^2}{2!} \underbrace{\left(\frac{d^2u}{dx^2}\right)_{i-1}}_{f'(x_{i-1}, u_{i-1})}$$

Pour augmenter la précision (ordre de la méthode), le calcul des dérivées f'' , $f^{(3)}$, ... $f^{(n)}$ augmente énormément le temps et le coût de calcul.

Carle Runge et Martin Kutta, ont développé des méthodes d'ordres plus élevés, par l'insertion des noeuds intermédiaires dans le pas de calcul.

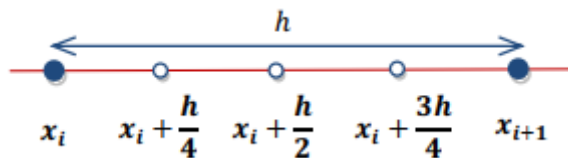


FIGURE 4.5 – Subdivision du pas de calcul dans les méthodes de Runge-Kutta

Méthode de Runge-Kutta d'ordre 2

Un noeud est inséré au milieu de l'intervalle $[x_i, x_{i+1}]$. La méthode est d'ordre 2 et l'erreur d'approximation est d'ordre 3 :

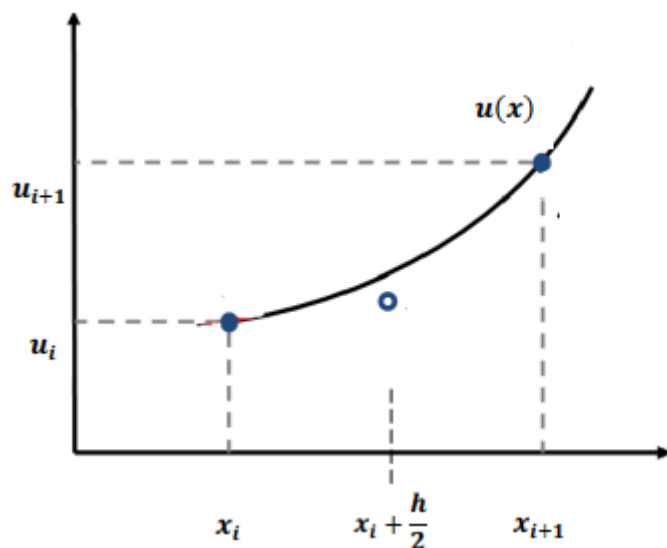


FIGURE 4.6 – Méthode de Runge-Kutta d'ordre 2

La solution de l'équation différentielle 4.4 est obtenue en deux étapes :

$$\begin{cases} k_1 = h * f(x_i, u_i) \\ k_2 = h * f(x_i + \frac{h}{2}, u_i + \frac{k_1}{2}) \\ u_{i+1} = u_i + k_2 \end{cases}$$

Méthode de Runge-Kutta d'ordre 4

Deux points intermédiaires sont insérés dans l'intervalle $[x_i, x_{i+1}]$

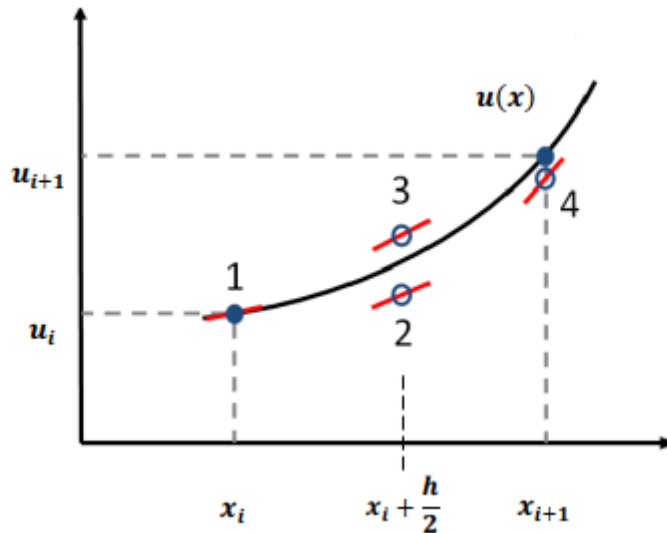


FIGURE 4.7 – Méthode de Runge-Kutta d'ordre 4

Méthode de Runge-Kutta d'ordre 4

La solution de l'équation différentielle 4.4 est obtenue en quatre étapes :

$$\begin{cases} k_1 = h * f(x_i, u_i) \\ k_2 = h * f(x_i + \frac{h}{2}, u_i + \frac{k_1}{2}) \\ k_3 = h * f(x_i + \frac{h}{2}, u_i + \frac{k_2}{2}) \\ k_4 = h * f(x_i + h, u_i + k_3) \\ u_{i+1} = u_i + \frac{k_1}{6} + \frac{k_2}{3} + \frac{k_3}{3} + \frac{k_4}{6} \end{cases}$$

Exemple

On considère le problème à valeur initiale, de l'exemple précédent :

$$\begin{cases} \frac{du}{dt} = u + e^{2t} \\ u(0) = 2 \end{cases}$$

Calculer $u(0.5)$, en utilisant un pas $t = 0.25$

Utiliser la méthode de Runge-Kutta d'ordre 2 et la méthode de Runge-Kutta d'ordre 4, pour calculer $u(0.5)$, en utilisant un pas $\Delta t = 0.25$

Comparer l'ordre de précision des résultats numériques des méthodes d'Euler et des méthodes de Runge-Kutta.

Solution

CHAPITRE 4. RÉOLUTION NUMÉRIQUE DES ÉQUATIONS DIFFÉRENTIELLES

Méthode de Runge-Kutta d'ordre 2 :

t=0.25

$$k_1 = \Delta t * [u_{r2}(0) + e^{2*t_0}] = 0.25 * [2 + e^{2*0}] = 0.7500$$

$$k_2 = \Delta t * [(u_{r2}(0) + \frac{k_1}{2}) + e^{2*(t_0 + \frac{\Delta t}{2})}] = 0.25 * [2 + \frac{0.7500}{2} + e^{2*(0 + \frac{0.25}{2})}] = 0.9147$$

$$u_{r2}(0.25) = u_{r2}(0) + k_2 = 2 + 0.9147 = 2.9147$$

t=0.5

$$k_1 = \Delta t * [u_{r2}(0.25) + e^{2*t_1}] = 0.25 * [2.9147 + e^{2*0.25}] = 1.1408$$

$$k_2 = \Delta t * [(u_{r2}(0.25) + \frac{k_1}{2}) + e^{2*(t_1 + \frac{\Delta t}{2})}] = 0.25 * [2.9147 + \frac{1.1408}{2} + e^{2*(0.25 + \frac{0.25}{2})}] = 1.4005$$

$$u_{r2}(0.5) = u_{r2}(0.25) + k_2 = 2.9147 + 1.4005 = 4.3152$$

Méthode de Runge-Kutta d'ordre 4 :

t=0.25 :

$$k_1 = \Delta t * [u_{r4}(0) + e^{2*t_0}] = 0.25 * [2 + e^{2*0}] = 0.7500$$

$$k_2 = \Delta t * [(u_{r4}(0) + \frac{k_1}{2}) + e^{2*(t_0 + \frac{\Delta t}{2})}] = 0.25 * [2 + \frac{0.7500}{2} + e^{2*(0 + \frac{0.25}{2})}] = 0.9147$$

$$k_3 = \Delta t * [(u_{r4}(0) + \frac{k_2}{2}) + e^{2*(t_0 + \frac{\Delta t}{2})}] = 0.25 * [2 + \frac{0.9147}{2} + e^{2*(0 + \frac{0.25}{2})}] = 0.9353$$

$$k_4 = \Delta t * [(u_{r4}(0) + k_3) + e^{2*(t_0 + \Delta t)}] = 0.25 * [2 + 0.9353 + e^{2*(0 + 0.25)}] = 1.1460$$

CHAPITRE 4. RÉOLUTION NUMÉRIQUE DES ÉQUATIONS DIFFÉRENTIELLES

$$u_{r4}(0.25) = u_{r4}(0) + \frac{k_1}{6} + \frac{k_2}{3} + \frac{k_3}{3} + \frac{k_4}{6} = 2 + \frac{0.7500}{6} + \frac{0.9147}{3} + \frac{0.9353}{3} + \frac{1.1460}{6} = 2.9327$$

t=0.5 :

$$k_1 = \Delta t * [u_{r4}(0.25) + e^{2*t_1}] = 0.25 * [2.9327 + e^{2*0.25}] = 1.1453$$

$$k_2 = \Delta t * [(u_{r4}(0.25) + \frac{k_1}{2}) + e^{2*(t_1 + \frac{\Delta t}{2})}] = 0.25 * [2.9327 + \frac{1.1453}{2} + e^{2*(0.25 + \frac{0.25}{2})}] = 1.4381$$

$$k_3 = \Delta t * [(u_{r4}(0.25) + \frac{k_2}{2}) + e^{2*(t_1 + \frac{\Delta t}{2})}] = 0.25 * [2.9327 + \frac{1.4056}{2} + e^{2*(0.25 + \frac{0.25}{2})}] = 1.4381$$

$$k_4 = \Delta t * [(u_{r4}(0.25) + k_3) + e^{2*(t_1 + \Delta t)}] = 0.25 * [2.9327 + 1.4381 + e^{2*(0.25 + 0.25)}] = 1.7723$$

$$u_{r4}(0.5) = u_{r4}(0.25) + \frac{k_1}{6} + \frac{k_2}{3} + \frac{k_3}{3} + \frac{k_4}{6} = 2.9327 + \frac{1.1453}{6} + \frac{1.4056}{3} + \frac{1.4381}{3} + \frac{1.7723}{6} = 4.3669$$

Comparaison des valeurs numériques avec la valeur exacte :

$u_e(0.5)$	$u_m(0.5)$	$u_{r2}(0.5)$	$u_{r4}(0.5)$	$u_{ex}(0.5)$	$\epsilon_e(0.5)$	$\epsilon_m(0.5)$	$\epsilon_{r2}(0.5)$	$\epsilon_{r4}(0.5)$
3.8497	4.7417	4.3152	4.3669	4.3670	11.84%	8.58%	1.19%	0.0023%

FIGURE 4.8 – Comparaison de la précision des méthodes numériques par rapport à la valeur exacte