

I.1. Introduction

Le module « méthodes numériques appliquées », rassemble toutes les techniques de calcul qui permettent de résoudre de manière exacte ou, le plus souvent, de manière approchée un problème physique représenté par un modèle mathématique.

I.2. Résolution des systèmes d'équations linéaires et non linéaires

Soit $M_n(\mathbb{R})$, l'espace vectoriel des matrices carrées d'ordre n et à coefficients réels. Soit $A \in M_n(\mathbb{R})$ et $B \in \mathbb{R}^n$, on cherche le vecteur $X \in \mathbb{R}^n$, solution du système linéaire $AX = B$. Ce système admet une solution unique lorsque le déterminant de A est non nul, ce que nous supposons dans la suite.

Remarquons que la résolution de ce système à l'aide des formules de Cramer est impraticable lorsque n est 'grand', car ces formules nécessitent approximativement ' $n!$ ' opérations arithmétiques élémentaires ; par exemple si $n=15$, cela représente de l'ordre de 4 mois de calcul pour un ordinateur moyen effectuant opérations à la seconde. Ce temps de calcul évolue de manière exponentielle avec n .

On sépare généralement les problèmes en deux classes suivant les caractéristiques de la matrice A :

- La matrice A est de taille réduite et est pleine. Par taille réduite, on comprend les matrices d'ordre inférieur à 100, par matrice pleine, on signifie que A comporte peu d'éléments nuls ;
- La matrice A est éparse (creuse) et de grande taille. De grande taille nomme des matrices d'ordres égaux à plusieurs centaines ou plusieurs milliers. La matrice éparse possède peu d'éléments non nuls.

Il est à remarquer qu'il n'existe pas de règle définitive pour le choix entre les méthodes directes et les méthodes itératives.

Une méthode directe conduit à une solution en un nombre fini d'étapes, et sans les erreurs arrondie.

- Une méthode itérative fait passer d'un estimé $X^{(k)}$ de la solution à un autre estimé $X^{(k+1)}$ de cette solution. S'il ya convergence, la solution ne pourrait donc être atteinte qu'après un nombre infini d'itérations.

Ce premier chapitre est consacré aux méthodes employées pour résoudre les problèmes les plus fréquemment rencontrées pour les systèmes linéaires. Par définition, on peut écrire celles-ci comme : $AX = b$, où A est une matrice de l'ordre de $M \times N$:

Un système de m équations à n inconnues $x_1, x_2, \dots, x_n$ s'écrit sous forme matricielle : $AX = B$ où A est une matrice comportant m lignes et n colonnes,

X est le vecteur colonne dont les composantes sont les x_i et B , le second membre, est aussi un vecteur colonne avec n composantes. Le vecteur X est appelé solution du système

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \dots & \dots & \dots & \dots \\ a_{M1} & a_{M2} & \dots & a_{MN} \end{bmatrix}; \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_N \end{bmatrix}; \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_M \end{bmatrix}$$

Si $N = M$, il y a autant d'équations que d'inconnues et si aucune des équations n'est une combinaison linéaire des $N-1$ autres, la solution existe et est unique

I.2. 1. Résolution des systèmes d'équations linéaires

I.2.1.a. Méthodes de Jacobi

Pour les systèmes linéaires, cette méthode illustre bien les techniques de point fixe : si on suppose connues toutes les inconnues sauf l'inconnue i , cette dernière peut être déterminée à l'aide de l'équation i . Procédant ainsi sur toutes les inconnues, on obtient un nouvel itéré x^k à partir de l'ancien x^{k-1} . Cependant, comme la méthode de Gauss-Seidel, elle peut présenter un intérêt dans le cas de systèmes non-linéaires. Sous forme matricielle, cela se traduit par :

$x^k = (I - D^{-1}A) x^{k-1} + D^{-1}b$ où D est la matrice diagonale, contenant la diagonale de A . La matrice d'itération est donc ici

$B = I - D^{-1}A$. On peut noter l'utilisation d'un vecteur de stockage $\bar{x}$ supplémentaire.

Algorithme :

```

- Vecteur de départ  $x^{(0)}$ 
Tant que  $R > \varepsilon$ 
    itération sur  $k$ 
        boucle sur  $i = 1, n$ 
             $\bar{x}_i \leftarrow A_{i,1:n} x_{1:n}^{k-1}$ 
             $\bar{x}_i \leftarrow (b_i - \bar{x}_i) / A_{i,i}$ 
        fin
         $x^{(k)} \leftarrow \bar{x}$ 
Fin

```

Convergence :

Si A est symétrique, définie positive, alors :

$$\rho_J = \rho(B) < 1,$$

I.2.1.b. Méthodes de Gauss-Seidel

Le concept est semblable à celui de la méthode de Jacobi, à ceci près que l'on utilise les nouvelles valeurs estimées dès qu'elles sont disponibles, et non pas à la fin de l'itération. Sous forme matricielle, cela se traduit par:

$$x^k = (D - E)^{-1}(Fx^{k-1} + b)$$

Où « $-E$ » est la partie triangulaire inférieure de A , et « $-F$ » sa partie triangulaire supérieure. La matrice d'itération est ici :

$$B = (D - E)^{-1}F = (I - D^{-1}E)^{-1}D^{-1}F$$

Algorithme :

Vecteur de départ $x^{(0)}$

Tant que $R > \varepsilon$

 itération sur k

 boucle sur $i = 1, 2, \dots, n$

$$x_i^{k-1} \leftarrow 0$$

$$\beta \leftarrow A_{i,1:n} x^{(k-1)}$$

$$x_i^k \leftarrow (b_i - \beta) / A_{i,i}$$

fin

Fin

Convergence :

Quand A est symétrique, définie positive, on est assuré de la convergence puisque $\rho_{GS} < 1$.

Mais l'estimation de la valeur du rayon spectral est difficile car il dépend en particulier de la numérotation des degrés de liberté employée.

I.2.1.c. Méthode de Newton–Raphson

En supposant la fonction f est de classe C^1 et que le zéro ξ est simple, la méthode de Newton–Raphson fait le choix

$$\alpha(x) = \frac{1}{f'(x)}$$

La relation de récurrence définissant cette méthode est alors

$$\forall k \in \mathbb{N}, x^{(k+1)} = x^{(k)} - \frac{f(x^{(k)})}{f'(x^{(k)})},$$

L'initialisation $x(0)$ étant donnée. Dans cette méthode, toute nouvelle approximation du zéro est construite au moyen d'une linéarisation de l'équation $f(x) = 0$ autour de l'approximation précédente. En effet, si l'on remplace $f(x)$ au voisinage du point $x^{(k)}$ par l'approximation affine obtenue en tronquant au premier ordre le développement de Taylor de f en $x^{(k)}$ et qu'on résout l'équation linéaire résultante :

$$f(x^{(k)}) + (x - x^{(k)})f'(x^{(k)}) = 0,$$

en notant sa solution $x^{(k+1)}$, et il en résulte que, géométriquement parlant, le point $x^{(k+1)}$ est l'abscisse du point d'intersection entre la tangente à la courbe de f au point $(x^{(k)}, f(x^{(k)}))$ et l'axe des abscisses

Par rapport à toutes les méthodes introduites jusqu'à présent, on pourra remarquer que la méthode de Newton nécessite à chaque itération l'évaluation des deux fonctions f et f' au point courant $x^{(k)}$. Cet effort est compensé par une vitesse de convergence accrue, puisque cette méthode est d'ordre deux si le zéro recherché est simple.

I.2. 2. Résolution des systèmes d'équations non linéaires

La méthode Newton peut s'appliquer à la résolution d'un système de plusieurs équations non linéaires :

$$\begin{cases} f(x, y) = 0 \\ g(x, y) = 0 \end{cases}$$

A partir d'un couple de valeurs approchées (x_1, y_1) d'une solution du système, on peut déterminer deux accroissements h et k à donner à x_1 et y_1 de manière à ce que :

$$\begin{cases} f(x_1 + h, y_1 + k) = 0 \\ g(x_1 + h, y_1 + k) = 0 \end{cases}$$

En développant en premier ordre, on obtient :

$$\begin{cases} f(x_1 + h, y_1 + k) = f(x_1, y_1) + hf'(x_1) + kf'(y_1) = 0 \\ g(x_1 + h, y_1 + k) = g(x_1, y_1) + hg'(x_1) + kg'(y_1) = 0 \end{cases}$$

Où l'on a posé :

$$\begin{cases} f'(x_1) = \frac{\partial f(x_1, y_1)}{\partial x} \\ f'(y_1) = \frac{\partial f(x_1, y_1)}{\partial y} \end{cases} \text{ et } \begin{cases} g'(x_1) = \frac{\partial g(x_1, y_1)}{\partial x} \\ g'(y_1) = \frac{\partial g(x_1, y_1)}{\partial y} \end{cases}$$

Les quantités h et k s'obtiennent donc, en résolvant le système linéaire suivant :

$$\begin{cases} hf'(x_1) + kf'(y_1) = -f(x_1, y_1) \\ hg'(x_1) + kg'(y_1) = -g(x_1, y_1) \end{cases} \text{ avec } \begin{cases} h = x_{n+1} - x_n \\ k = y_{n+1} - y_n \end{cases}$$

Le calcul est alors relancé jusqu'à ce que h et k deviennent inférieurs à une valeur ε que l'on se donne (selon la précision voulue pour le calcul).

Ainsi, l'algorithme correspondant est :

$$\begin{cases} x_{n+1} = x_n - \left(\frac{f(x_n, y_n) \cdot g'(y_k) - g(x_n, y_n) \cdot f'(y_k)}{\Delta} \right) \\ y_{n+1} = y_n - \left(\frac{g(x_n, y_n) \cdot f'(y_k) - f(x_n, y_n) \cdot g'(y_k)}{\Delta} \right) \end{cases}$$

Avec :

$$\Delta = f'(x_n) \cdot g'(y_n) - f'(y_n) \cdot g'(x_n)$$

I.3. Interpolation et approximation

L'intérêt de remplacer une fonction quelconque par un polynôme l'approchant aussi précisément que voulu sur un intervalle donné est évident d'un point de vue numérique et informatique, puisqu'il est très aisé de stocker et de manipuler, c'est-à-dire additionner, multiplier, dériver ou intégrer, des polynômes dans un calculateur. Pour ce faire, il semble naturel de chercher à utiliser un polynôme d'interpolation de Lagrange associé aux valeurs prises par la fonction en des nœuds choisis.

I.3.1 Rappel et définitions

Soit $P_n(x)$ l'espace des polynômes de degré inférieur ou égal à n .

On rappelle que

$\{1, x, x^2, \dots, x^n\}$ est une base de $P_n(x)$ ($\dim P_n(x) = n + 1$)

On note δ_{ij} le symbole de Kronecker : $\delta_{ij} = 1$ si $i = j$; $\delta_{ij} = 0$ si $i \neq j$.

Soient f une fonction continue sur $[a, b]$, $x_1, \dots, x_n$ n points de $[a, b]$ et $g_1, \dots, g_n$ des réels de même signe.

Alors on a :

$$\sum_{i=1}^{i=n} f(x_i)g_i = f(c) \sum_{i=1}^{i=n} g_i \quad \text{où } c \in [a, b]$$

I.3.1.a. Définitions

Soit f une fonction réelle définie sur un intervalle $[a, b]$ contenant $n + 1$ points distincts

$x_0, x_1, \dots, x_n$. Soit P_n un polynôme de degré inférieur ou égal à n .

On dit que P_n est un interpolant de f ou interpole f en $x_0, x_1, \dots, x_n$ si :

$$P_n(x_i) = f(x_i) \text{ pour } 0 \leq i \leq n$$

I.3.1.b. Méthode de Lagrange

Soient $x_0, x_1, \dots, x_n$ $(n + 1)$ points deux à deux distincts d'un intervalle $[a, b]$ de $\mathbb{R}$

On appelle interpolant de Lagrange les polynômes L_i définis pour $i = 0, \dots, n$ par :

$$L_i(x) = \prod_{j=0, j \neq i}^{j=n} \frac{(x - x_j)}{(x_i - x_j)} = \frac{(x - x_0)(x - x_1) \dots (x - x_{i-1})(x - x_{i+1}) \dots (x - x_n)}{(x_i - x_0)(x_i - x_1) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_n)}$$

On a en particulier :

$$L_0(x) = \prod_{j=1}^{j=n} \frac{(x - x_j)}{(x_0 - x_j)} = \frac{(x - x_1)(x - x_2) \dots (x - x_i) \dots (x - x_n)}{(x_0 - x_1)(x_0 - x_2) \dots (x_0 - x_i) \dots (x_0 - x_n)}$$

$$L_n(x) = \prod_{j=0}^{j=n-1} \frac{(x - x_j)}{(x_n - x_j)} = \frac{(x - x_0)(x - x_1) \dots (x - x_i) \dots (x - x_{n-1})}{(x_n - x_0)(x_n - x_1) \dots (x_n - x_i) \dots (x_n - x_{n-1})}$$

Si on prend :

$$P_n(x) = L_0(x)f(x_0) + L_1(x)f(x_1) + \dots + L_n(x)f(x_n)$$

Alors :

$$P_n(x_i) = f(x_i) \text{ pour } 0 \leq i \leq n$$

Exemple :

Si $x_0 = -1$, $x_1 = 0$ et $x_2 = 1$,

$f(x_0) = 2$, $f(x_1) = 1$, $f(x_2) = -1$, on obtient :

$$\begin{aligned} L_0(x) &= \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} = \frac{x(x - 1)}{(-1)(-1 - 1)} = \frac{x(x - 1)}{2} \\ L_1(x) &= \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} = \frac{(x - (-1))(x - 1)}{-(-1)(-1)} = \frac{(x + 1)(x - 1)}{-1} \\ L_2(x) &= \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} = \frac{(x + 1)x}{(1 - (-1))(1 - 0)} = \frac{(x + 1)x}{2} \end{aligned}$$

Alors :

$$\begin{aligned}
P_2(x) &= L_0(x)f(x_0) + L_1(x)f(x_1) + L_2(x)f(x_2) \\
&= \frac{x(x-1)}{2}f(x_0) + \frac{(x+1)(x-1)}{-1}f(x_1) + \frac{(x+1)x}{2}f(x_2) \\
&= 2\frac{x(x-1)}{2} + \frac{(x+1)(x-1)}{-1} - \frac{(x+1)x}{2} \\
&= -\frac{1}{2}x^2 - \frac{3}{2}x + 1
\end{aligned}$$

On peut facilement vérifier que :

$$\begin{aligned}
P_2(x_0) &= P_2(-1) = \frac{(-)(-1-1)}{2} = 2 = f(x_0) \\
P_2(x_1) &= P_2(0) = \frac{(1)(-1)}{-1} = 1 = f(x_1) \\
P_2(x_2) &= P_2(1) = -\frac{(1+1)1}{2} = -1 = f(x_2)
\end{aligned}$$

En déduisant donc, que : Les polynômes de Lagrange ont les propriétés suivantes :

- ✓ $L_j(x)$ est un polynôme de degré n ; $\forall j = 0, \dots, n$
- ✓ $L_j(x_j) = 1 \quad \forall j = 0, \dots, n$ et $L_j(x_i) = 0$ pour tout $j \neq i$.
- ✓ la famille $\{L_0(x), L_1(x), \dots, L_n(x)\}$ est une base de $P_n(x)$.

I.3.1.c. Méthode des différences divisées

Cette méthode consiste d'abord à établir des propriétés de continuité et de dérivabilité pour la fonction de la variable réelle x définie par $[x_0; x_1, \dots, x_n; x]f$, où les points $x_0, \dots, x_n$ sont distincts et contenus dans un intervalle $[a; b]$ borné de $\mathbb{R}$ et x appartient à $[a; b]$.

On définit les différences divisées d'ordre i de f aux points (x_i) comme suit :

$$\begin{aligned}
[f(x_0)] &= f(x_0) \\
[f(x_0), f(x_1)] &= \frac{f(x_1) - f(x_0)}{x_1 - x_0} \\
[f(x_0), f(x_1), \dots, f(x_i)] &= \frac{[f(x_1), f(x_2), \dots, f(x_i)] - [f(x_0), f(x_1), \dots, f(x_{i-1})]}{x_i - x_0}
\end{aligned}$$

Pour $i \geq 2$

Exemple :

Si $x_0 = -1$, $x_1 = 0$ et $x_2 = 1$, $f(x_0) = 2$, $f(x_1) = 1$, $f(x_2) = -1$, on obtient :

$$[f(-1)] = 2$$

$$[f(-1), f(0)] = \frac{1-2}{0-(-1)} = -1$$

$$[f(0)] = 1$$

$$[f(-1), f(0), f(1)] = \frac{-2-(-1)}{1-(-1)} = \frac{-1}{2}$$

$$[f(0), f(1)] = \frac{-1-1}{1-0} = -2$$

$$[f(1)] = 1$$

En déduisant donc, que La valeur d'une différence divisée est indépendante de l'ordre des x_i

On a ainsi :

$$[f(x_0), f(x_1)] = \frac{f(x_1) - f(x_0)}{x_1 - x_0} = \frac{f(x_1)}{x_1 - x_0} + \frac{f(x_0)}{x_0 - x_1} = [f(x_1), f(x_0)]$$

$$\begin{aligned} [f(x_0), f(x_1), f(x_2)] &= \frac{f(x_2)}{(x_2 - x_1)(x_2 - x_0)} + \frac{f(x_1)}{(x_1 - x_2)(x_1 - x_0)} + \frac{f(x_0)}{(x_0 - x_1)(x_0 - x_2)} \\ &= [f(x_2), f(x_1), f(x_0)] = [f(x_1), f(x_0), f(x_2)] \end{aligned}$$

D'une manière générale on a :

$$[f(x_0), f(x_1), \dots, f(x_k)] = \sum_{i=0}^{i=k} \frac{f(x_i)}{(x_i - x_0) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_k)}$$

I.4. Intégration numérique

I.4.1. Méthode des Trapèzes :

Soit $x_0 < x_1 < \dots < x_i < x_{i+1} < \dots < x_n$ une subdivision uniforme et $P_1(x)$ un polynôme de degré 1 interpolant f aux points x_i et x_{i+1} de chaque intervalle $[x_i, x_{i+1}]$.

$$[P_1(x_i) = f(x_i) \text{ et } P_1(x_{i+1}) = f(x_{i+1})]$$

En approchant sur chaque sous intervalle $[x_i, x_{i+1}]$, $f(x)$ par $P_1(x)$ on obtient :

$$\begin{aligned} f(x) &\simeq P_1(x) = f(x_i) + [f(x_i), f(x_{i+1})](x - x_i) \\ f(x) &\simeq P_1(x) = f(x_i) + \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}(x - x_i) \end{aligned}$$

Et en conséquence :

$$I(f) = \int_a^b f(x) dx \simeq \int_a^b P_1(x) dx = \frac{h}{2} [f(x_i) + f(x_{i+1})]$$

I.4.2. Méthode de Simpson

Soit $P_2(x)$ un polynôme de degré 2 vérifiant :

$$P_2(x_i) = f(x_i), P_2(x_{i+1}) = f(x_{i+1}) \text{ et } P_2(x_{i+2}) = f(x_{i+2})$$

En approchant sur chaque sous intervalle $[x_i, x_{i+2}]$, $f(x)$ par $P_2(x)$ on obtient :

$$\begin{aligned} f(x) &\simeq P_2(x) \\ &\simeq f(x_0) + [f(x_i), f(x_{i+1})](x - x_i) + [f(x_i), f(x_{i+1}), f(x_{i+2})](x - x_i)(x - x_{i+1}) \end{aligned}$$

Et en déduisant :

$$I(f) = \int_{x_i}^{x_{i+2}} f(x) dx \simeq \int_{x_i}^{x_{i+2}} P_2(x) dx = \frac{h}{3} [f(x_i) + 4f(x_{i+1}) + f(x_{i+2})]$$

On a

$$\int_{x_i}^{x_{i+2}} P_2(x) dx = I(P_2) + J(P_2) + K(P_2)$$

Où :

$$I(P_2) = \int_{x_i}^{x_{i+2}} f(x_i) dx$$

$$J(P_2) = \int_{x_i}^{x_{i+2}} \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} (x - x_i) dx$$

$$K(P_2) = \int_{x_i}^{x_{i+2}} [f(x_i), f(x_{i+1}), f(x_{i+2})](x - x_i)(x - x_{i+1}) dx$$

On fait le changement de variables suivant :

$$(x - x_i) = ht \rightarrow dx = hdt$$

Quand :

$$x = x_i \rightarrow t = 0 \text{ et si } x = x_{i+2} \rightarrow t = 2$$

On obtient donc :

$$I(P_2) = \int_0^2 f(x_i) h dt = 2hf(x_i)$$

$$J(P_2) = \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} \int_0^2 h^2 t dt$$

$$= h[f(x_{i+1}) - f(x_i)] \left[\frac{t^2}{2} \right]_0^2$$

$$= 2h[f(x_{i+1}) - f(x_i)]$$

$$K(P_2) = [f(x_i), f(x_{i+1}), f(x_{i+2})] \int_0^2 (x - x_i)(x - x_{i+1}) dx$$

$$= [f(x_i), f(x_{i+1}), f(x_{i+2})] \int_0^2 h^3 t(t-1) dt$$

$$= \frac{f(x_{i+2}) - 2f(x_{i+1}) + f(x_i)}{2h^2} h^3 \left[\frac{t^3}{3} - \frac{t^2}{2} \right]_0^2$$

$$= [f(x_{i+2}) - 2f(x_{i+1}) + f(x_i)] \frac{h}{3}$$

Et enfin :

$$\int_{x_i}^{x_{i+2}} P_2(x) dx = I(P_2) + J(P_2) + K(P_2) = \frac{h}{3} [f(x_i) + 4f(x_{i+1}) + f(x_{i+2})]$$

I.4.3. Méthode composée des Trapèzes :

Pour chercher une approximation de l'intégrale sur tout l'intervalle $[a, b]$, il

Suffit d'écrire :

$$\begin{aligned} \int_{a=x_0}^{b=x_n} f(x) dx &= \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} f(x) dx \\ &= \sum_{i=0}^{n-1} \left[\frac{h}{2} [f(x_i) + f(x_{i+1})] + \sum_{i=0}^{n-1} -\frac{h^3}{12} f''(\eta_i) \right] \\ &= \frac{h}{2} [f(x_0) + 2f(x_1) + \dots + 2f(x_{n-1}) + f(x_n)] - \frac{h^3}{12} \sum_{i=0}^{n-1} f''(\eta_i) \\ &= \frac{h}{2} [f(x_0) + 2f(x_1) + \dots + 2f(x_{n-1}) + f(x_n)] - \frac{(b-a)}{12} h^2 f''(\eta) \end{aligned}$$

Avec

$$\eta \in [a, b]$$

I.5. Résolution des équations différentielles ordinaires

La résolution numérique des équations différentielles est éventuellement le domaine de l'analyse numérique où les applications sont les plus nombreuses. Que ce soit en mécanique ou en génie électrique, on aboutit souvent à la résolution d'équations différentielles. On ne sait cependant pas toujours donner une expression exacte de la solution, et il faut alors l'approximer numériquement.

Les équations différentielles nous servent pour modéliser le processus physique (sous forme d'une équation devinette), puis de trouver les expressions mathématiques (les solutions) qui vont décrire le comportement futur de forme d'une équation différentielle (la devinette) et ensuite à trouver l'expression de la réponse et analyser le comportement (résultats de simulation), évitant un effondrement ultérieur du système qui porterait conséquences sur les vies humaines. Très peu d'équations différentielles sont solubles analytiquement. De plus, chaque type d'équation requiert une méthode particulière de résolution. Par suite, la résolution de la plus part des équations différentielles nécessite l'utilisation de méthodes numériques (exp : Euler, Runge-Kutta, Adams Moulton, ...).

Une équation est dite différentielle si elle comporte une ou plusieurs dérivées de la variable à résoudre.

Une équation différentielle est dite ordinaire lorsque la variable à résoudre ne dépend que d'une seule autre variable.

L'ordre d'une équation différentielle est déterminé selon le plus haut degré de dérivée apparaissant dans l'équation. Par exemple une équation différentielle du premier ordre ne contient que la dérivée première et possiblement la fonction elle-même. Une équation du deuxième ordre contient obligatoirement la dérivée seconde et éventuellement la dérivée première et/ou la fonction elle-même.

Une équation différentielle est dite linéaire si elle est formée par une combinaison linéaire de la variable et de ses dérivées.

La forme implicite d'une équation différentielle est: $F(x, y, y') = 0$, Exemple: $2x + y' + 2y = 0$.

La forme explicite d'une équation différentielle est : $F(x, y) = y'$, Exemple: $-2x - 2y = y'$.

La solution à une équation différentielle est une fonction $h(x)$, c'est à dire une expression explicite de la combinaison de la variable x (dans laquelle la dépendance envers la variable y a disparu), et qui substitué à y dans l'équation, satisfait cette équation dans une plage de valeurs de la variable : $a < x < b$.

Exemple1 : Solution explicite

La devinette (l'équation différentielle) est: $xy' = 2y$

Devinez alors une fonction $h(x)$ qui une fois substituée à y vérifie bien cette équation.

Solution :

Bien sûr, pour l'instant, on essaie de deviner un peu au hasard, mais nous verrons que nous sommes ici pour apprendre une méthode systématique de résolution de ce genre d'équations.

Donc si on essaie de substituer : $y = h(x) = x^2$

Nous allons voir que : $y' = h'(x) = 2x$

Et alors : $x(2x) = 2x^2$: Donc, nous dirons que $h(x) = x^2$ est une solution explicite de l'équation.

Exemple2: Solution implicite

Soit à trouver une solution pour l'équation suivante : $yy' = -x$

Solution :

Nous pouvons voir que l'expression suivante est solution implicite de l'équation ci-dessus :

$$h(x, y) = x^2 + y^2 - 1 = 0.$$

En effet, il suffit de réarranger sous la forme $y^2 = 1 - x^2$ et de dériver les deux membres de cette égalité par rapport à la variable x pour voir qu'elle vérifie bien l'équation de départ :

$$2y'y = 0 - 2x$$

I.5.1. Classification suivant les conditions

Les équations différentielles peuvent être classées en deux catégories suivant le types des conditions:

- les équations différentielles aux conditions initiales.
- les équations différentielles aux conditions aux limites.

I.5.1.a. Equations différentielles aux conditions initiales

Soit l'équation différentielle suivante :

$$A(t) \cdot \frac{d^2 y}{dt^2} + B(t) \cdot \frac{dy}{dt} + C(t) \cdot y = g(t)$$

Où A, B, C et g sont des fonctions connues de t. On donne en plus de cette équation différentielle, les conditions initiales suivantes:

$$y(0) = y_0 \text{ et } \left(\frac{dy}{dt} \right)_0 = V_0$$

Toute équation différentielle d'ordre n avec des conditions initiales peut être remplacée par un système de n équations différentielles couplées du premier ordre. En posant : $dy / dt = z$, l'équation précédente devient :

$$\frac{dz}{dt} = -\frac{B(t)}{A(t)} \cdot z - \frac{C(t)}{A(t)} \cdot y + \frac{g(t)}{A(t)}$$

Avec :

$$\begin{cases} y(0) = y_0 \\ z(0) = V_0 \end{cases}$$

Ainsi, une équation d'ordre n peut être ramenée à un système du premier ordre:

$$\begin{cases} \frac{d\vec{y}}{dt} = \vec{f}(\vec{y}, t) ; \vec{y} = \{y_1, y_2, \dots, y_n\} ; \vec{f} = \{f_1, f_2, \dots, f_n\} \\ \vec{y}(0) = \vec{y}_0 \end{cases}$$

Soit par exemple l'équation: $y''' + y'' + y' + y = f(t)$ avec les Conditions initiales:

$$y(0) = y_0, y'(0) = v_0 \text{ et } y''(0) = g_0$$

Pour simplifier la suite, écrivons plutôt: $y_1''' + y_1'' + y_1' + y_1 = f(t)$(*)

Et posons: $y_2 = y_1'$

L'équation (*) s'écrit maintenant: $y_2'' + y_2' + y_2 + y_1 = f(t)$(**)

Qui est encore du deuxième ordre, on pose alors : $y_3 = y_2'$

L'équation (**) devient : $y_3' + y_3 + y_2 + y_1 = f(t)$

Au final, (*) s'écrit comme le système de trois équations du premier degré:

$$\begin{cases} y_1' = y_2 \\ y_2' = y_3 \\ y_3' = -y_3 - y_2 - y_1 + f(t) \end{cases} \quad \text{CI:} \begin{cases} y_1(0) = y_0 \\ y_2(0) = v_0 \\ y_3(0) = g_0 \end{cases}$$

I.5.1.b. Equations différentielles aux conditions aux limites

A titre d'exemple on cite le cas de l'équation de la propagation de la chaleur le long d'une barre de longueur L définie comme suit:

$$\begin{cases} \frac{d^2 y}{dx^2} + D \cdot \frac{dy}{dx} + E \cdot y = h(x) \\ y(0) = y_0 \\ y(L) = y_L \end{cases}$$

Ce problème est à conditions aux limites.

Si une condition aux limites est donnée en terme de gradient, par exemple $v_0 = dy(0)/dx$ plutôt que $y(0)$ lui-même, on peut adopter l'une ou l'autre des techniques suivantes :

- On écrit y_0 en fonction de y_1 à partir de $(y_1 - y_0)/h = v_0$ et on réarrange les termes de la première équation pour se ramener comme précédemment à une équation matricielle.
- On introduit un point « fantôme » y_{-1} tel que $(y_{-1} - y_1)/2h = v_0$ et on écrit une nouvelle équation correspondante à y_{-1} , y_0 et y_1 en y remplaçant y_{-1} par $y_{-1} = y_1 + 2hv_0$. Après réarrangement des termes, on obtient une matrice avec une colonne et ligne supplémentaire qu'il ne reste qu'à inverser pour résoudre le système.

I.5.2. Méthode d'Euler

La plus simple et la plus connue des méthodes d'approximation des solutions des équations différentielles ordinaires est la méthode d'Euler donnée par :

$\Phi(x, y, h) = f(x, y)$, de l'ordre 1, stabilité acquise (ϕ et f lipchitzienne) avec :

$$f(x_n, y_n) = \frac{y_{n+1} - y_n}{h}$$

La formule s'écrit donc :

$$y_{n+1} = y_n + hf(x_n, y_n)$$

I.5.2.a. Algorithme d'Euler général

- Etant donné un pas de temps h , une condition initiale (x_0, y_0) et un nombre maximal d'itérations N ;
- Pour $0 \leq n \leq N$

$$Y(n+1) = y(n) + h \cdot f(t(n), y(n))$$

$$t(n+1) = t(n) + h$$

Ecrire $t(n+1)$ et $y(n+1)$
- Arrêt

I.5.2.b. Algorithme d'Euler modifiée

- Etant donné un pas de temps h , une condition initiale (x_0, y_0) et un nombre maximal d'itérations N ;
- Pour $0 \leq n \leq N$

$$\hat{y} = y(n) + h \cdot f(t(n), y(n))$$

$$y(n+1) = y(n) + \frac{h}{2} (f(t(n), y(n)) + f(t(n+1), \hat{y}))$$

$$t(n+1) = t(n) + h$$

Ecrire $t(n+1)$ et $y(n+1)$
- Arrêt.

I.5.3. Méthode de Runge-Kutta**I.5.3.a. Méthode de Runge-Kutta d'ordre 2**

$$\Phi(x, y, h) = \beta f(x, y) + \alpha f(x + \lambda h, y + \mu h f(x, y)).$$

- L'ordre 1 impose $\alpha + \beta = 1$
- L'ordre 2 impose $\lambda = \mu$ et $\alpha \lambda = \alpha \mu = 1/2$.

Si $\alpha = 1$, alors $\beta = 0$ et $\lambda = \mu = 1/2$, d'où la formule :

$$y_{n+1} = y_n + hf\left(x_n + \frac{h}{2}, y_n + \frac{h}{2} f(x_n, y_n)\right)$$

Soit le prédicteur $y_{n+1/2}$, la méthode s'écrit :

$$\begin{cases} y_{n+1/2} = y_n + \frac{h}{2} f(x_n, y_n) \\ y_{n+1} = y_n + hf(x_n + h/2, y_{n+1/2}) \end{cases}$$

La méthode est d'ordre 2, et correspond à une amélioration de la méthode d'Euler : le taux d'accroissement est calculé pour le point milieu : $(x_n + h/2, y_{n+1/2})$.

I.5.3.b. Méthode de Runge-Kutta d'ordre 4

- Etant donné un pas de temps h , une condition initiale (x_0, y_0) et un nombre maximal d'itérations N ;
- Pour $0 \leq n \leq N$

$$k_1 = f(x, y)$$

$$k_2 = f(x + h/2, y + h/2 \cdot k_1)$$

$$k_3 = f(x + h/2, y + h/2 \cdot k_2)$$

$$k_4 = f(x + h, y + h \cdot k_3)$$

$$y(n+1) = y(n) + \frac{1}{6} (k_1 + 2k_2 + 2k_3 + k_4)$$

$$t(n+1) = t(n) + h$$

Ecrire $t(n+1)$ et $y(n+1)$

- Arrêt

I.5.4. Méthode d'Adams

On déduit ces méthodes de la forme intégrale, en évaluant de manière approchée l'intégrale de f entre t_n et t_{n+1} . On suppose les nœuds de discrétisation équirépartis, c'est-à-dire $t_j = t_0 + jh$, avec $h > 0$ et $j \geq 1$. On intègre alors, au lieu de f , son polynôme d'interpolation aux $\tilde{p} + \vartheta$ nœuds distincts,

Où $\vartheta = 1$ quand les méthodes sont explicites (dans ce cas $\tilde{p} \geq 0$) et $\vartheta = 2$ quand les méthodes sont implicites (dans ce cas $\tilde{p} \geq -1$). Les schémas obtenus ont la forme suivante :

$$u_{n+1} = u_n + h \sum_{j=-1}^{\tilde{p}+\vartheta} b_j f_{n-j}$$

Les nœuds d'interpolation peuvent être ou bien :

- $t_n, t_{n-1}, \dots, t_{n-\tilde{p}}$ (dans ce cas $b_{-1} = 0$ et la méthode est explicite) ;

Ou bien

- $t_{n+1}, t_n, \dots, t_{n-\tilde{p}}$ (dans ce cas $b_{-1} \neq 0$ et le schéma est implicite).

Ces schémas implicites sont appelés méthodes d'Adams-Moulton, et les explicites sont appelés

I.5.4.a. Méthodes d'Adams - Bashforth.

En prenant $\tilde{p} = 0$, on retrouve la méthode d'Euler progressive, puisque le polynôme d'interpolation de degré zéro au nœud t_n est simplement $\Pi_0 f = f_n$. Pour $\tilde{p} = 1$, le polynôme d'interpolation linéaire aux nœuds t_{n-1} et t_n est

$$\Pi_1 f(t) = f_n + (t - t_n) \frac{f_{n-1} - f_n}{t_{n-1} - t_n}$$

Comme :

$$\Pi_1 f(t_n) = f_n \text{ et } \Pi_1 f(t_{n+1}) = 2f_n - f_{n-1}$$

On obtient :

$$\int_{t_n}^{t_{n+1}} \Pi_1 f(t) dt = \frac{h}{2} [\Pi_1 f(t_n) + \Pi_1 f(t_{n+1})] = \frac{h}{2} [3f_n - f_{n-1}]$$

Le schéma d'Adams – Bashforth à deux pas est donc :

$$u_{n+1} = u_n + \frac{h}{2} [3f_n - f_{n-1}]$$

Si $\tilde{p} = 2$, on trouve de façon analogue le schéma d'Adams-Bashforth à trois pas.

$$u_{n+1} = u_n + \frac{h}{12} [23f_n - 16f_{n-1} + 5f_{n-2}]$$

Et pour $\tilde{p} = 3$, on a le schéma d'Adams-Bashforth à quatre pas.

$$u_{n+1} = u_n + \frac{h}{24} (55f_n - 59f_{n-1} + 37f_{n-2} - 9f_{n-3})$$

Remarquer que les schémas d'Adams-Bashforth utilisent $\tilde{p}+1$ nœuds et sont des méthodes à $\tilde{p}+1$ pas (avec $\tilde{p} \geq 0$). De plus, les schémas d'Adams-Bashforth à q pas sont d'ordre q .

I.5.4.a. Méthodes d'Adams-Moulton.

Si $\tilde{p} = -1$, on retrouve le schéma d'Euler rétrograde. Si $\tilde{p} = 0$, on construit le polynôme d'interpolation de degré un de f aux nœuds t_n et t_{n+1} , et on retrouve le schéma de Crank-Nicolson. Pour la méthode à deux pas ($\tilde{p} = 1$), on construit le polynôme d'interpolation de degré 2 de f aux nœuds t_{n-1} , t_n , t_{n+1} , et on obtient le schéma suivant :

$$u_{n+1} = u_n + \frac{h}{12} [5f_{n+1} + 8f_n - f_{n-1}]$$

Les schémas correspondant à $\tilde{p} = 2$ et 3 sont respectivement donnés par :

$$u_{n+1} = u_n + \frac{h}{24} (9f_{n+1} + 19f_n - 5f_{n-1} + f_{n-2})$$

Et

$$u_{n+1} = u_n + \frac{h}{720} (251f_{n+1} + 646f_n - 264f_{n-1} + 106f_{n-2} - 19f_{n-3})$$

Les schémas d'Adams-Moulton utilisent $\tilde{p} + 2$ nœuds et sont à $\tilde{p} + 1$ pas si $\tilde{p} \geq 0$, la seule exception étant le schéma d'Euler rétrograde ($\tilde{p} = -1$) qui est à un pas et utilise un nœud. Les schémas d'Adams-Moulton à q pas sont d'ordre $q + 1$, excepté à nouveau le schéma d'Euler rétrograde qui est une méthode à un pas d'ordre un.

I.6. Conclusion

Le rappel que nous avons cité sur les méthodes numériques consiste à entamer une branche des mathématiques appliquées s'intéressant au développement d'outils et de méthodes numériques pour le calcul d'approximations de solutions de problèmes de mathématiques qu'il serait difficile, voire impossible, d'obtenir par des moyens analytiques. Son objectif est notamment d'introduire des procédures calculatoires détaillées susceptibles d'être mises en œuvre par des calculateurs (électroniques, électrotechniques, mécaniques ou humains) et d'analyser leurs caractéristiques et leurs performances.