Analyse numérique

1. Interpolation
2. Lissage
3. Intégration numérique
4. Racines d'une équation non linéaire
5. Résolution de systèmes linéaires
6. Équations différentielles

Interpolation (approximation de fonctions)

Idée : déterminer une fonction f à partir de données et telle que f possède certaines propriétés par rapport à ces données.

Le problème

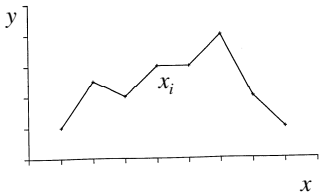
Une fonction f : [a, b] $\rightarrow$ R est connue par les valeurs qu'elle prend en $n + 1$ points $x_i \in$ [a, b], 1 = 0,1,... , n

On voudrait connaître $f(x)$, $\forall x \in$ [a, b]

Interpolation linéaire:

Inconvénients:

- la fonction n'est pas dérivable
- précision



On cherche d'autres solutions au problème en imposant certaines propriétés à l'approximation :

- passer par les points $(x_i, f(x_i))$;
- posséder certaines propriétés de régularité comme la dérivabilité;
- avoir une certaine forme, etc.

Une façon courante de procéder consiste à rechercher l'approximation sous une forme polynomiale.

Interpolation polynomiale

Données : $n+1$ points $(x_i, f(x_i))$, $i = 1, 2, 3, \ldots, n+1$.

On cherche un polynôme $P_n(x)$ définie par

$$P_n(x) = a_o + a_1 x + a_2 x^2 + \ldots + a_n x^n$$

tel que $P_n(x_i) = f(x_i)$ pour $i = 0,1,\ldots,n.$

Les coefficients inconnus a_i sont donc solution du système

$$\begin{pmatrix} 1 & x_0 & \ldots & x_0^n \\ 1 & x_1 & \ldots & x_1^n \\ \vdots & & & \vdots \\ 1 & x_n & \ldots & x_n^n \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} f(x_0) \\ f(x_1) \\ \\ f(x_n) \end{pmatrix}$$

On a donc un système linéaire à résoudre. Celui-ci admet une solution unique si les x_i sont distincts.

Interpolation à l'aide de formule de Lagrange

Données : $n+1$ de points $(x_i, f(x_i))$, $i = 0, 1, 2, 3 \ldots n$

On cherche un polynôme P(x) (de degré le plus petit possible) tel que $P(x_i) = f(x_i)$

- On peut écrire le polynôme P recherché sous la forme (**formule de Lagrange**):

$$P(x) = \sum_{j=0}^{n} f(x_j) L_j(x) \quad \text{avec} \quad L_j(x_i) = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \quad (*)$$

En effet, dans ce cas $P(x_i) = \sum_{j=0}^{n} f(x_j) L_j(x_i) = f(x_i)$

- Déterminons les polynômes $L_j(x)$ (j=0,1,…,n) qui satisfont la condition (*):

$$L_j(x) = \alpha (x - x_o)(x - x_1)\ldots(x - x_{j-1})(x - x_{j+1})\ldots(x - x_n)$$

$$L_j(x_j) = \alpha (x_j - x_o)(x_j - x_1)\ldots(x_j - x_{j-1})(x_j - x_{j+1})\ldots(x_j - x_n) = 1$$

donc $$\alpha = \frac{1}{(x_j - x_o)(x_j - x_1)\ldots(x_j - x_{j-1})(x_j - x_{j+1})\ldots(x_j - x_n)}$$

$$\mathrm{L}_j(x) = \frac{(x - x_o)(x - x_1)\ldots(x - x_{j-1})(x - x_{j+1})\ldots(x - x_n)}{(x_j - x_o)(x_j - x_1)\ldots(x_j - x_{j-1})(x_j - x_{j+1})\ldots(x_j - x_n)} = \prod_{\substack{i=0 \\ i \neq j}}^{n} \frac{x - x_i}{x_j - x_i}$$

Lissage

L'idée: déterminer la "meilleure" courbe, de forme choisi,qui représente le phénomène étudié; la courbe ne passe pas obligatoirement par les points donnés.

Le problème

Données : n points (x_i, y_i), $i = 1, 2, 3, \ldots, n$.

On choisit le type de fonction $f(x, \Lambda)$ et on doit uniquement déterminer les paramètres Λ ($\Lambda \in R^p$) de cette fonction.

On cherche Λ^*, la meilleure valeur de Λ, telle que :

$$d(f(x,\Lambda^*), y(x)) \leq d(f(x,\Lambda), y(x)) \quad \forall \Lambda \in R^p$$

où $d(f(x,\Lambda^*), y(x))$ représente la distance entre la fonction f (avec les paramètres Λ^*) et les points $(x_i, f(x_i))$; elle doit être minimale, c'est-à-dire la plus petite pour tous les $\Lambda \in R^p$

Remarques

- Les fonctions f choisies couramment sont les fonctions linéaires, les polynômes, les exponentielles et les fonctions trigonométriques.
- Les résultats d'un lissage dépendent du choix d'une fonction et d'une distance.

Droite des moindres carrés (droite de régression linéaire)

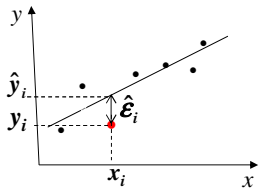
Données : n de points (x_i, y_i), $i = 1, 2, 3, \ldots n$

On cherche à déterminer les meilleurs paramètres a,b de la droite $y(x) = a\,x + b$

tels que $S(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2$ soit minimale.

Valeur ajustée de y_i : $\hat{y}_i = ax_i + b$

Le résidu en i (erreur) : $\hat{\varepsilon}_i = y_i - \hat{y}_i$



Remarques

- Les x_i sont donc implicitement supposés connus exactement
- Le choix de la somme des carrés des écarts repose sur deux considérations :

1° En prenant les carrés, il ne peut y avoir de phénomène de compensation comme ce pourrait être le cas en faisant la somme des distances algébriques;

2° L'introduction d'un carré permet la dérivation. Une fonction valeur absolue répondrait au 1° mais ne permettrait pas de dériver le critère.

$$S(a,b)=\sum_{i=1}^{n}(y_i-ax_i-b)^2 \quad \text{minimale}$$

$$\Updownarrow$$

$$\begin{cases}\dfrac{\partial S}{\partial a}(a,b)=0\\ \dfrac{\partial S}{\partial b}(a,b)=0\end{cases} \Leftrightarrow \begin{cases}-2\sum_{i=1}^{n}(y_i-ax_i-b)x_i=0\\ -2\sum_{i=1}^{n}(y_i-ax_i-b)=0\end{cases} \Leftrightarrow \begin{cases}a\sum_{i=1}^{n}x_i^2+b\sum_{i=1}^{n}x_i=\sum_{i=1}^{n}x_iy_i\\ a\sum_{i=1}^{n}x_i+nb=\sum_{i=1}^{n}y_i\end{cases} \quad (1)$$

Posons: $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$, $\bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i$, $\overline{x^2}=\frac{1}{n}\sum_{i=1}^{n}x_i^2$, $\overline{y^2}=\frac{1}{n}\sum_{i=1}^{n}y_i^2$, $\overline{xy}=\frac{1}{n}\sum_{i=1}^{n}x_iy_i$,

moyennes simples respectivement sur $x_i, y_i, x_i^2, y_i^2, x_iy_i$

Avec la notation ci-dessus (1) devient: $\begin{cases}a\overline{x^2}+b\bar{x}=\overline{xy}\\ a\bar{x}+b=\bar{y}\end{cases}$ (2)

Solution du système (2): $a=\dfrac{\overline{xy}-\bar{x}\,\bar{y}}{\overline{x^2}-\bar{x}^2}$ $b=\bar{y}-a\bar{x}$

Généralisation: lissage par un polynôme de degré *p*

Données : n de points (x_i, y_i), $i=1, 2, 3,\ldots n$

$$P(x)=a_px^p+a_{p-1}x^{p-1}+\ldots a_1x+a_0\,, \qquad a=\begin{bmatrix}a_0\\ a_1\\ \vdots\\ a_p\end{bmatrix}$$

On cherche a optimal tel que

$$S(a)=\sum_{i=1}^{n}\left(y_i-\sum_{k=0}^{p}a_kx_i^k\right)^2 \quad \text{soit minimale.}$$

La condition nécessaire d'optimalité:

$$\frac{\partial S}{\partial a_j}(a)=0 \qquad j=0,1,\ldots,p$$

$$\frac{\partial S}{\partial a_j}(a)=0 \Leftrightarrow \quad 2\sum_{i=1}^{n}\left(y_i-\sum_{k=0}^{p}a_k x_i^k\right)(-x_i^j)=0 \qquad j=0,1,...,p$$

$$\Updownarrow$$

$$\sum_{i=1}^{n}\sum_{k=0}^{p}a_k x_i^k x_i^j=\sum_{i=1}^{n}x_i^j y_i \qquad j=0,1,...,p$$

$$\Updownarrow$$

$$a_0\sum_{i=1}^{n}x_i^j+a_1\sum_{i=1}^{n}x_i^{j+1}+...+a_p\sum_{i=1}^{n}x_i^{j+p}=\sum_{i=1}^{n}x_i^j y_i \qquad j=0,1,...,p$$

$$\Updownarrow$$

$$\begin{cases} a_0\sum_{i=1}^{n}(x_i)^0+a_1\sum_{i=1}^{n}x_i+...+a_p\sum_{i=1}^{n}x_i^p=\sum_{i=1}^{n}(x_i)^0 y_i \\ a_0\sum_{i=1}^{n}x_i+a_1\sum_{i=1}^{n}x_i^2+...+a_p\sum_{i=1}^{n}x_i^{p+1}=\sum_{i=1}^{n}x_i y_i \\ \vdots \\ a_0\sum_{i=1}^{n}x_i^p+a_1\sum_{i=1}^{n}x_i^{p+1}+...+a_p\sum_{i=1}^{n}x_i^{2p}=\sum_{i=1}^{n}(x_i)^p y_i \end{cases} \Leftrightarrow \begin{bmatrix} 1 & \overline{x} & \overline{x^2} & ... & \overline{x^p} \\ \overline{x} & \overline{x^2} & \overline{x^3} & ... & \overline{x^{p+1}} \\ \overline{x^2} & \overline{x^3} & ... & ... & \overline{x^{p+2}} \\ ... & ... & ... & ... & ... \\ \overline{x^p} & \overline{x^{p+1}} & ... & ... & \overline{x^{2p}} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ ... \\ a_p \end{bmatrix} = \begin{bmatrix} \overline{y} \\ \overline{xy} \\ \overline{x^2 y} \\ ... \\ \overline{x^p y} \end{bmatrix}$$

où $\overline{x^q}=\dfrac{\sum_{i=1}^{n}(x_i)^q}{n}, \quad \overline{x^q y}=\dfrac{\sum_{i=1}^{n}(x_i)^q y_i}{n}$

q=0,1..,p

Résolution de ce système de p+1 équations donne les coefficients du polynôme de lissage

Lissage par une fonction exponentielle

Données : n de points (x_i, y_i), $i=1, 2, 3,\ldots n$

On cherche à approche ces points par une fonction f qui s'écrit:

$$f(x)=y_0\exp(-\alpha x)$$

Soit $S(y_0,\alpha)=\sum_{i=1}^{n}\left(y_i-y_0 e^{-\alpha x_i}\right)^2$

La condition d'optimalité s'écrit:

$$\begin{cases} \dfrac{\partial S}{\partial y_0}(y_0,\alpha)=0 \\ \dfrac{\partial S}{\partial \alpha}(y_0,\alpha)=0 \end{cases} \Leftrightarrow \begin{cases} -2\sum_{i=1}^{n}(y_i-y_0 e^{-\alpha x_i})e^{-\alpha x_i}=0 \\ -2\sum_{i=1}^{n}(y_i-y_0 e^{-\alpha x_i})y_0 e^{-\alpha x_i}x_i=0 \end{cases} \Rightarrow \begin{cases} y_0=\dfrac{\sum_{i=1}^{n}y_i e^{-\alpha x_i}}{\sum_{i=1}^{n}e^{-2\alpha x_i}} & (1) \\ \sum_{i=1}^{n}x_i y_i e^{-\alpha x_i}-y_0\sum_{i=1}^{n}x_i e^{-2\alpha x_i}=0 & (2) \end{cases}$$

En remplaçant y_0 dans (2) par sa valeur (1), on obtient:

$$\left(\sum_{i=1}^{n}e^{-2\alpha x_i}\right)\left(\sum_{i=1}^{n}x_i y_i e^{-\alpha x_i}\right)-\left(\sum_{i=1}^{n}y_i e^{-\alpha x_i}\right)\left(\sum_{i=1}^{n}x_i e^{-2\alpha x_i}\right)=0$$

Cette équation en α peut être résolue par la méthode de la sécante (cf. chapitre 4)

Un autre approche de lissage exponentiel

$$f(x) = y_0 \exp(-\alpha x)$$
$$\Downarrow$$
$$\ln(f(x)) = \ln(y_0) - \alpha x$$

On cherche $g(x) = ax + b$ à partir des couples x_i , $\ln(y_i)$.

On obtient alors, par une régression linéaire : $a = \dfrac{\bar{x}\overline{\ln y} - \overline{x \ln y}}{(\bar{x})^2 - \overline{x^2}}$, $\quad b = \overline{\ln y} - a\bar{x}$

où $\quad \overline{\ln y} = \sum_{i=1}^{n} \dfrac{\ln y_i}{n}$ et $\overline{x \ln y} = \sum_{i=1}^{n} \dfrac{x_i \ln y_i}{n}$

$a = -\alpha$ et $b = \ln(y_0)$ $\quad \Rightarrow \quad \alpha = -a$ et $y_0 = \exp(b)$

Remarques

- On évite ainsi la résolution d'une équation algébrique non linéaire.
- Les résultats obtenus par les deux méthodes, avec le même ensemble de données, sont légèrement différents. Ceci est dû au fait qu'on a minimisé des critères différents car la fonction logarithme introduit un poids différent pour chaque x_i.

Intégration numérique

Le problème

$$J = \int_a^b f(x)dx$$

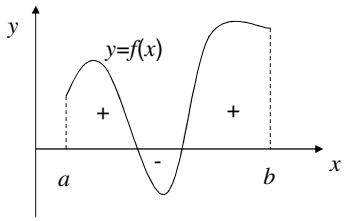
où les bornes d'intégration a et b sont supposées finies.

Le calcul numérique d'une intégrale est nécessaire :

- Quand on ne connaît pas de primitive de la fonction à intégrer;
- Quand on ne connaît la fonction qu'en un nombre fini de points;
- Quand une primitive n'est connue que sous une forme trop compliquée pour être facilement calculable.

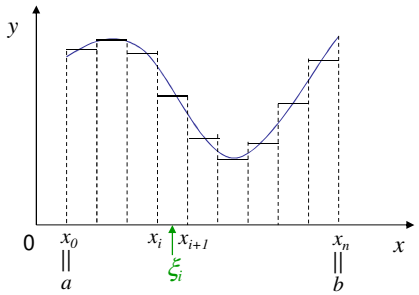
Interprétation géométrique

La valeur de l'intégrale correspond à l'aire sous la courbe (pour les intervalles où f(x)>0); il est pris avec le signe (-) si f(x)<0.



Méthode des rectangles

Idée: Approcher la fonction par des fonctions constantes par morceaux et évaluer l'aire sous la courbe en faisant la somme des aires des rectangles élémentaires.



- On divise l'intervalle d'intégration la, b] en n intervalles égaux de la longueur h:

$$h = \frac{b-a}{n}$$

h est appelé pas d'intégration ou pas de discrétisation

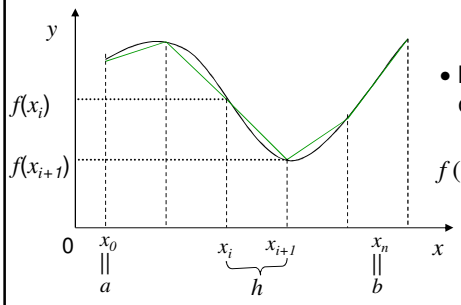
$$J = \int_a^b f(x)dx \cong \frac{b-a}{n}\left(f(\xi_0) + f(\xi_1) + \ldots + f(\xi_{n-1})\right)$$

Où $\xi_i \in [x_i, x_{i+1}]$

D'habitude on prend $\xi_i = \frac{x_i + x_{i+1}}{2}$.

Méthode des trapèzes

Idée: Approcher la courbe de f par des fonctions affines par morceaux, et évaluer l'aire sous la courbe en faisant la somme des aires des trapèzes élémentaires.



- Pour chaque pas d'intégration, on évalue l'aire du trapèze $(x_i,0)$, $(x_i, f(x_i))$, $(x_{i+1}, f(x_{i+1}))$, $(x_{i+1},0)$:

$$f(x_{i+1})\cdot h + \tfrac{1}{2}h\left(f(x_i) - f(x_{i+1})\right) = h\frac{f(x_i) + f(x_{i+1})}{2}$$

En supposant le pas h constant et on obtient:

$$J = \int_a^b f(x)dx \cong h\frac{f(a)+f(x_1)}{2} + h\frac{f(x_1)+f(x_2)}{2} + \ldots + h\frac{f(x_{n-1})+f(b)}{2} \quad \text{avec} \quad h = \frac{b-a}{n}$$

$$J = \int_a^b f(x)dx \cong \frac{b-a}{n}\left(\frac{f(a)}{2} + f(x_1) + \ldots + f(x_{n-1}) + \frac{f(b)}{2}\right)$$

Remarque 1

La méthode des trapèzes donne des résultats plus précis que la méthode des rectangles pour le même nombre de pas.

Remarque 2

Supposons que f est de classe C^2 sur $[a,b]$ et $\underline{M} = \min_{x\in[a,b]} f''(x)$, $\overline{M} = \max_{x\in[a,b]} f''(x)$.

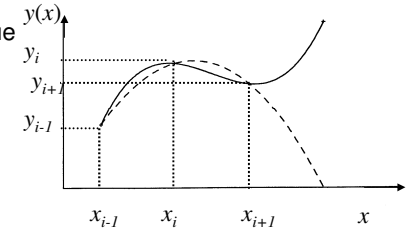
On peut alors majorer l'erreur de calcul de $J = \int_a^b f(x)dx$ à l'aide de la méthode des trapèzes par:

$$\frac{(b-a)^3}{12n^2}\underline{M} \le \varepsilon \le \frac{(b-a)^3}{12n^2}\overline{M}$$

Méthode de Simpson

La méthode consiste à diviser l'intervalle d'intégration $[a, b]$ en un nombre pair de pas égaux: $x_0 = a, x_1, x_2, x_3, ..., x_{2n-1}, x_{2n} = b$

et dans chaque intervalle $[x_{i-1}, x_{i+1}]$ on applique l'interpolation parabolique.



$$\int_{x_0}^{x_2} f(x)dx \cong \frac{b-a}{6n}(y_0 + 4y_1 + y_2)$$

$$\int_{x_2}^{x_4} f(x)dx \cong \frac{b-a}{6n}(y_2 + 4y_3 + y_4)$$

$$\vdots$$

$$\int_{x_{2n-2}}^{x_{2n}} f(x)dx \cong \frac{b-a}{6n}(y_{2n-2} + 4y_{2n-1} + y_{2n})$$

Soit $y_i = f(x_i)$.

Chaque parabole passe par trois points successifs, comme l'illustre la figure où $f(x)$ est représentée en trait plein et la parabole en trait pointillé.

$$\int_a^b f(x)dx \cong \frac{b-a}{6n}\left((y_0 + y_{2n}) + 2(y_2 + y_4 + ... + y_{2n-2}) + 4(y_1 + y_3 + .. + y_{2n-1})\right)$$

Remarque 1

Supposons que f est de classe C^4 sur $[a,b]$ et $\underline{M} = \min_{x\in[a,b]} f^{(4)}(x)$, $\overline{M} = \max_{x\in[a,b]} f^{(4)}(x)$.

On peut alors majorer l'erreur de calcul de $J = \int_a^b f(x)dx$ à l'aide de la méthode de Simpson par:

$$\frac{(b-a)^5}{180(2n)^4}\underline{M} \le \varepsilon \le \frac{(b-a)^5}{180(2n)^4}\overline{M}$$

L'encadrement de l'erreur varie comme $1/n^4$ pour la méthode de Simpson, alors qu'il varie comme $1/n^2$ pour la méthode des trapèzes.

Remarque 2

La méthode de Simpson est exacte pour un polynôme de degré inférieur ou égal à 3. Cette propriété résulte de la formule d'encadrement de l'erreur dans laquelle intervient la dérivée quatrième de la fonction.

Racines d'une équation non linéaire

Le problème

$$f : D \to R, \quad D \subset R$$

On cherche, s'il existe, $x \in D$ tel que: $f(x) = 0$.

Remarques

- Si des solutions analytiques sont connues, on utilise ces solutions, sinon il faut recourir à des méthodes itératives qui donnent une approximation de la solution cherchée en un nombre fini d'itérations.
- Si une fonction F est dérivable et si F' = f, alors résolution de $f(x) = 0$ est lié à la recherche des extremums de la fonction F.

Méthode du point fixe

Elle consiste à transformer le problème $f(x) = 0$ en l'écrivant sous la forme :

$$g(x) = x$$

et à utiliser une méthode itérative qui s'écrit:

x_0 initialisé à une valeur donnée

$$x_1 = g(x_0)$$
$$x_2 = g(x_1)$$
$$\vdots$$
$$x_n = g(x_{n-1})$$
$$x_{n+1} = g(x_n)$$

jusqu'à ce que la méthode converge, si elle converge.

Théorème du point fixe.

Soit g une fonction définie de $[a, b]$ dans $[a, b]$, dérivable sur $]a, b[$ et telle qu'il existe $L < 1$ tel que

$$\forall x \in]a,b[\quad |g'(x)| \leq L$$

alors la suite définie par $x_0 \in [a,b]$

$$x_{n+1} = g(x_n) \qquad \forall n \geq 0$$

converge et sa limite s est l'unique solution de l'équation du point fixe

$$g(x) = x$$

Remarque 1 On arrête les itérations lorsque deux valeurs successives de x_k sont « suffisamment » voisines. On peut utiliser par exemple l'un de deux critères:

- Convergence absolue: $|x_{k+1} - x_k| < \varepsilon$
- Convergence relative: $\left|\frac{x_{k+1} - x_k}{x_{k+1}}\right| < \varepsilon$

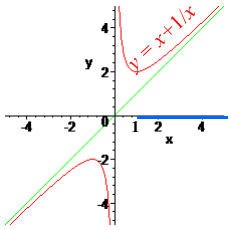
Remarque 2

Il ne suffit pas de vérifier que $|g'(x)|< 1$ (au lieu de $|g'(x)| \leq L< 1$) dans l'application du théorème du point fixe . C'est une erreur comme le montre l'exemple suivant :

Soit $g : [1, \infty[\to R, \quad g(x) = x + \frac{1}{x}$

$$\Downarrow$$

$$g'(x) = 1 - \frac{1}{x^2}$$



On a bien $|g'(x)| < 1$ mais $\sup_{x \geq 1} |g'(x)| = 1$

et l'équation $x = g(x)$ n'admet pas de point fixe dans [1, ∞[

Conclusion:

Il faut toujours déterminer une constante L, indépendante de x, pour appliquer le théorème de point fixe.

Remarque 3

La relation d'inclusion $g([a,b]) \subset [a,b]$ doit également être vérifiée, sous peine de sortir éventuellement du domaine de convergence.

Méthode de Newton

La méthode de Newton peut être considérée comme un cas particulier de la méthode du point fixe. Pour résoudre $f(x) = 0$, elle consiste à définir les itérations par :

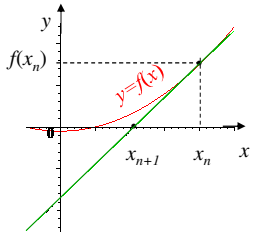
x_0 initialisation donnée

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

Ceci suppose que f est dérivable.

L'interprétation géométrique de la méthode de Newton

À partir d'un point x_n, on construit le point x_{n+1} en prenant la tangente à la courbe en x_n, et en considérant l'intersection de la tangente avec l'axe des abscisses.



La tangente à $f(x)$ en x_n: $y = f'(x_n)x + b$

$$f(x_n) = f'(x_n)x_n + b \Rightarrow b = f(x_n) - f'(x_n)x_n$$

$$(y = 0 \text{ pour } x_{n+1}) \Rightarrow 0 = f'(x_n)x_{n+1} + f(x_n) - f'(x_n)x_n$$

$$\Downarrow$$

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

Remarque 1

On arrête le calcul lorsque deux valeurs successives de x_n sont «suffisamment» voisines. On peut utiliser, par exemple, l'un de deux critères:

- Convergence absolue: $|x_{n+1} - x_n| < \varepsilon$
- Convergence relative: $\left|\frac{x_{n+1} - x_n}{x_n}\right| < \varepsilon$

Remarque 2

Si au cours des itérations on ait $f'(x_n) = 0$, il faut alors recommencer les calculs en changeant l'initialisation x_o.

Ordre de convergence d'une méthode itérative

On veut savoir si la suite définie par les itération $x_{n+1}=g(x_n)$ converge «rapidement » vers la sa limite s, c'est-à-dire comment diminue l'erreur $e_n=x_n-s$ d'une itération à la suivante. C'est pour quantifier ceci qu'on introduit la notion d'ordre de convergence.

Définition. La méthode définie par $x_{n+1}=g(x_n)$ est dite d'ordre de convergence p,

si $\frac{|x_{n+1} - s|}{|x_n - s|^p}$ a une limite finie quand n tend vers +∞.

Explication : D'après la formule de Taylor:

$$g(x) = g(s) + g'(s)(x-s) + \frac{g''(s)}{2!}(x-s)^2 + \frac{g'''(s)}{3!}(x-s)^3 + \ldots + \frac{g^{(p)}(s)}{p!}(x-s)^p + r_{p,s}(x)$$

où $r_{p,s}(x) = \frac{g^{(p+1)}(\xi)}{(p+1)!}(x-s)^{p+1}$ avec $\xi \in]x,s[$ et $\lim_{x\to s} \frac{r_{p,s}(x)}{(x-s)^p} = 0$

Alors $g(x_n) - g(s) = g'(s)(x_n - s) + \frac{g''(s)}{2!}(x_n - s)^2 + \ldots + \frac{g^{(p)}(s)}{p!}(x_n - s)^p + r_{p,s}(x_n)$

Donc $\lim_{n\to\infty} \frac{|x_{n+1} - s|}{|x_n - s|^p}$ existe ssi $g'(s) = \ldots = g^{(p-1)}(s) = 0$ et $g^{(p)}(s) \neq 0$

Conclusion L'ordre de convergence est le premier nombre p tel que $g^{(p)}(s) \neq 0$.
En particulier, si $g'(s) \neq 0$, la méthode est d'ordre 1.

Ordre de convergence de la méthode de Newton

Pour la méthode de Newton, la fonction du théorème du théorème du point fixe, est définie par:

$$g(x) = x - \frac{f(x)}{f'(x)}$$

D'où $g'(x) = 1 - \frac{f'(x)\cdot f'(x) - f(x)f''(x)}{(f'(x))^2} \quad \Rightarrow \quad g'(x) = \frac{f(x)f''(x)}{(f'(x))^2}$

Comme $g(s)=s \quad \Rightarrow f(s)=0 \quad \Rightarrow \; g'(s)=0$

$\Downarrow$

La méthode de Newton est au moins du deuxième ordre.

$$g''(x) = \frac{(f'(x))^2 f''(x) + f(x)f'(x)f^{(3)}(x) - 2f(x)(f''(x))^2}{(f'(x))^3}$$

$\Downarrow$

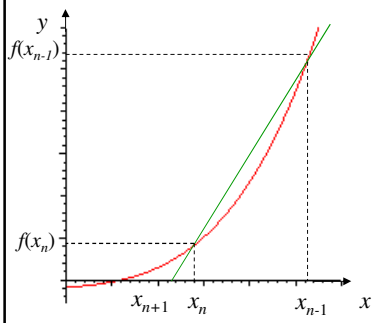
$$g''(s) = \frac{f''(s)}{f'(s)} \quad (\text{car } f(s) = 0)$$

Si $f'(s)\neq 0$ et $f''(s)\neq 0$ alors la méthode de Newton est du deuxième ordre.

Méthode de la sécante

Cette méthode est une variante de la méthode de Newton. Elle consiste à remplacer l'expression de la dérivée par une approximation de celle-ci.

Le principe de la méthode est illustré sur la figure:



On détermine la position de x_{n+1} comme intersection de la sécante (qui passe par les points $(x_n, f(x_n))$, $(x_{n-1}, f(x_{n+1}))$ avec l'axe d'abscisses.

La sécante: $y = ax+b, \quad a = \frac{f(x_{n-1}) - f(x_n)}{x_{n-1} - x_n}$

$$f(x_n) = ax_n + b \;\Rightarrow\; b = f(x_n) - ax_n$$

$$(y = 0 \;\text{ pour }\; x_{n+1}) \Rightarrow 0 = ax_{n+1} + b \;\Rightarrow\; x_{n+1} = -\frac{b}{a}$$

$$\Rightarrow x_{n+1} = \frac{ax_n - f(x_n)}{a} = x_n - f(x_n)\frac{1}{a} \quad \Rightarrow \quad \boxed{x_{n+1} = x_n - f(x_n)\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}}$$

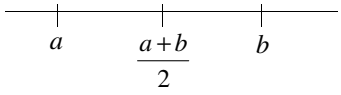
Le principal intérêt de la méthode de la sécante est donc de ne pas nécessiter l'expression de la dérivée et c'est pour cela qu'elle est utilisée dans les calculatrices.

Le procédé de dichotomie

$f : D \to R, \quad D \subset R$

On cherche des solutions de l'équation $f(x) = 0$ avec f continue.

- S'il existe $a,b \in D$ tel que $f(a)\cdot f(b) < 0$, alors il existe au moins une racine dans l'intervalle fermé [a,b].



On peut alors considérer le point $\frac{a+b}{2}$ et calculer $f\left(\frac{a+b}{2}\right)$.

- La solution s se trouve dans le nouvel intervalle ayant pour l'extremités $(a+b)/2$ et le point a ou b selon que $f(a)$ ou $f(b)$ est de signe contraire à $f((a+b)/2)$.
- A chaque itération on diminue de moitié la longueur de l'intervalle. Au bout de n itérations, la longueur de l'intervalle sera $(b-a)/2^n$.

Au début de processus de recherche de racine on emplois souvent le procédé de dichotomie, puis dès que l'intervalle contient s est suffisamment petit, on utilise les méthodes plus rapides, par exemple la méthode de Newton.

Résolution de systèmes linéaires

Le problème:

On cherche à résoudre un système :

$$\begin{cases} a_{11}x_1 + a_{12}x_2 \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 \ldots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 \ldots + a_{nn}x_n = b_n \end{cases}$$

Forme matricielle:

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

$$A \cdot x = b$$

où A est une matrice à n lignes et n colonnes;

b est un vecteur donné dans R^n;

x est le vecteur cherché dans R^n.

- On suppose que le problème admet une solution, c'est-à-dire que $\det(A) \neq 0$

Remarque 1: **Inversion de la matrice**

$$x = A^{-1} \cdot b$$

Pratiquement, ou n'inverse jamais une matrice pour résoudre un système linéaire car le nombre d'opérations à effectuer est très largement supérieur à celui nécessité par d'autres méthodes.

Remarque 2: **Méthode des déterminants (de Cramer)**

La méthode de Cramer donne : $\boxed{x_j = \frac{D_j}{D}}$

où D est le déterminant de la matrice A: $D = \det A$

D_j est le déterminant de la matrice déduite de la matrice A en y remplaçant la j[ème] colonne par la colonne second membre

$$D_j = \det \begin{bmatrix} a_{11} & \dots & a_{1,j-1} & b_1 & a_{1,j+1} & \dots & a_{1n} \\ a_{21} & \dots & a_{2,j-1} & b_2 & a_{2,j+1} & \dots & a_{2n} \\ \vdots & & & & & & \vdots \\ \vdots & & & & & & \vdots \\ \vdots & & & & & & \vdots \\ a_{n-1,1} & \dots & a_{n-1,j-1} & b_{n-1} & a_{n,j+1} & \dots & a_{n-1,n} \\ a_{n,1} & \dots & a_{n,j-1} & b_n & a_{n,j+1} & \dots & a_{n,n} \end{bmatrix}$$

- Le nombre d'opérations élémentaires nécessitées par la méthode de Cramer est de l'ordre de (n+2)!.

Exemple Supposons qu'une machine effectue 10^6 opérations par seconde.

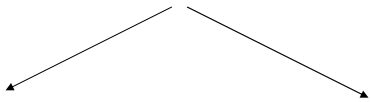
n = 5 $\Rightarrow$ 5040 opérations $\Rightarrow 5 \cdot 10^{-3}$ s

n = 10 $\Rightarrow 4{,}79 \cdot 10^8$ opérations $\Rightarrow$ 476 s $\cong$ 8 min

n = 15 $\Rightarrow 3{,}56 \cdot 10^{14}$ opérations $\Rightarrow 3{,}56 \cdot 10^8 \cong 5{,}9 \cdot 10^6$ min $\cong$ 4097 jours

La méthode de Cramer, est inadaptée à la résolution effective d'un système linéaire si n est élevé.

Méthodes de résolution numérique



Méthodes de résolution directes

Résolution par élimination

Résolution à l'aide de décomposition LU

Méthode de Gauss-Jordan

Méthodes itératives

Méthode de Jacobi

Méthode de Gauss-Seidel

Résolution par élimination (méthode de Gauss)

L'idée : transformer le système en système triangulaire supérieur (ou inférieure) et le résoudre par substitutions successives.

Exemple: Soit $n = 3$ et $a_{11} \neq 0$

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases}$$

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 & L_1 \to L_1 \\ a_{22}'x_2 + a_{23}'x_3 = b_2' & L_2 \to L_2 - \frac{a_{21}}{a_{11}}L_1 \\ a_{32}'x_2 + a_{33}'x_3 = b_3' & L_3 \to L_3 - \frac{a_{31}}{a_{11}}L_1 \end{cases}$$

a_{11} est appelé premier pivot;
Si a_{11} =0 il faut permuter les équations de façon à amener à la place de a_{11} un terme non nul; si la matrice A est inversible c'est toujours possible.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 & L_1 \to L_1 \\ a_{22}'x_2 + a_{23}'x_3 = b_2' & L_2 \to L_2 \\ a_{33}''x_3 = b_3'' & L_3 \to L_3 - \frac{a_{32}'}{a_{22}'}L_2 \end{cases}$$

Résolution d'un système triangulaire est facile:

$$\begin{cases} x_3 = \frac{b_3''}{a_{33}''} \\ x_2 = \frac{b_2' - a_{23}'x_3}{a_{22}'} \\ x_1 = \frac{b_1 - a_{12}x_2 - a_{13}x_3}{a_{11}} \end{cases}$$

Généralisation: procédure de la méthode de Gauss

1er étape: triangularisation de la matrice A

Notons: $A^{(0)} = A$ la matrice initiale
$b^{(0)} = b$ la colonne second membre (initiale)
$A^{(k)}$ la matrice du système à l'issue de k-ème étape k=1,…,n-1
$b^{(k)}$ la colonne second membre à l'issue de k-ème étape

Calcul des $a_{ij}^{(k)}$ et des $b_i^{(k)}$:

- Si $a_{kk}^{(k-1)} = 0$, trouver une ligne i, $i>k$ avec $a_{ik}^{(k-1)} \neq 0$ et permuter les lignes k et i. On note la matrice résultante $A^{(k-1)}$. Faire de même pour le vecteur $b^{(k-1)}$.

- Pour i=k+1, k+2,…,n et j=k+1, k+2,…,n calculer :

$$m_{ik} = \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}$$
$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - m_{ik} a_{kj}^{(k-1)}$$
$$b_i^{(k)} = b_i^{(k-1)} - m_{ik} b_k^{(k-1)}$$

$$A^{(0)} = A \Rightarrow A^{(1)} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & & a_{1n}^{(1)} \\ 0 & a_{22}^{(1)} & & a_{2n}^{(1)} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n1}^{(1)} & .. & a_{nn}^{(1)} \end{bmatrix} \Rightarrow A^{(2)} = \begin{bmatrix} a_{11}^{(2)} & a_{12}^{(2)} & a_{13}^{(2)} & \dots & a_{1n}^{(2)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & & a_{2n}^{(2)} \\ 0 & 0 & a_{33}^{(2)} & & a_{3n}^{(2)} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & a_{n3}^{(2)} & \dots & a_{nn}^{(2)} \end{bmatrix}$$

$$A^{(k-1)} = \begin{bmatrix} a_{11}^{(k-1)} & \dots & a_{1,k-1}^{(k-1)} & a_{1k}^{(k-1)} & a_{1k+1}^{(k-1)} & \dots & a_{1n}^{(k-1)} \\ 0 & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & a_{k-1,k-1}^{(k-1)} & & & & \\ \vdots & \vdots & 0 & a_{kk}^{(k-1)} & & & a_{kn}^{(k-1)} \\ \vdots & \vdots & \vdots & a_{k+1,k}^{(k-1)} & & & a_{k+1,n}^{(k-1)} \\ 0 & \vdots & 0 & \vdots & & & \\ 0 & \dots & 0 & a_{n,k}^{(k-1)} & \dots & \dots & a_{nn}^{(k-1)} \end{bmatrix}$$

- Pour i=k+1, k+2,…,n et j=k+1, k+2,…,n :

$$m_{ik} = \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}$$
$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - m_{ik} a_{kj}^{(k-1)}$$
$$b_i^{(k)} = b_i^{(k-1)} - m_{ik} b_k^{(k-1)}$$

$$\Rightarrow A^{(k)} = \begin{bmatrix} a_{11}^{(k)} & \dots & a_{1,k-1}^{(k)} & a_{1k}^{(k)} & a_{1,k+1}^{(k)} & & a_{1n}^{(k)} \\ 0 & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & a_{k-1,k-1}^{(k)} & & & & \\ \vdots & \vdots & 0 & a_{kk}^{(k)} & a_{k,k+1}^{(k)} & & a_{kn}^{(k)} \\ \vdots & \vdots & \vdots & 0 & a_{k+1,k+1}^{(k)} & & a_{k+1,n}^{(k)} \\ 0 & 0 & 0 & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & a_{n,k+1}^{(k)} & \dots & a_{nn}^{(k)} \end{bmatrix}$$

2ème étape: résolution du système triangulaire

$$\begin{cases} a_{11}^{(n-1)}x_1 + \ldots + a_{1k}^{(n-1)}x_k + a_{1,k+1}^{(n-1)}x_{k+1} + \ldots + a_{1n}^{(n-1)}x_n = b_1^{(n-1)} \\ \vdots \\ a_{kk}^{(n-1)}x_k + a_{k,k+1}^{(n-1)}x_{k+1} + \ldots + a_{kn}^{(n-1)}x_n = b_k^{(n-1)} \\ \vdots \\ a_{nn}^{(n-1)}x_n = b_n^{(n-1)} \end{cases}$$

D'où

$$\begin{cases} x_n = \dfrac{b_n^{(n-1)}}{a_{nn}^{(n-1)}} \\ \vdots \\ x_k = \dfrac{b_k^{(n-1)} - \sum_{i=k+1}^{n} a_{ki}^{(n-1)}x_i}{a_{kk}^{(n-1)}} \\ \vdots \\ x_1 = \dfrac{b_1^{(n-1)} - \sum_{i=2}^{n} a_{1i}^{(n-1)}x_i}{a_{11}^{(n-1)}} \end{cases}$$

Remarque

Le nombre d'opérations élémentaires nécessitées par la méthode de Gauss est de l'ordre de $\frac{2}{3}n^3$.

si $n = 15 \Rightarrow$ 2250 opérations $\Rightarrow 2 \cdot 10^{-3}$ s

Méthode de Gauss avec le choix de pivots partiels

Dans le cas où les pivots choisis dans la méthode de Gauss sont très petits, les résultats obtenus peuvent être erronés.

Exemple $\begin{cases} \varepsilon x_1 + x_2 = 1 \\ x_1 + x_2 = 2 \end{cases}$ avec $\varepsilon \neq 0$ mais $|\varepsilon| << 1$

Solution: $x_1 = \dfrac{1}{1-\varepsilon}$, $x_2 = \dfrac{1-2\varepsilon}{1-\varepsilon}$; si $|\varepsilon| << 1 \Rightarrow x_1 \approx 1$ et $x_2 \approx 1$

Par la méthode de Gauss: $\begin{cases} \varepsilon x_1 + x_2 = 1 \\ \left(1 - \frac{1}{\varepsilon}\right)x_2 = 2 - \frac{1}{\varepsilon} \end{cases}$

Prenons par exemple $\varepsilon = 10^{-9}$.

Avec 8 chiffres significatifs $1 - 10^9 = -10^9$ et $2 - 10^9 = -10^9$.

Alors $\begin{cases} 10^{-9}x_1 + x_2 = 1 \\ -10^9 x_2 = -10^9 \end{cases}$ d'où $\begin{cases} x_2 = 1 \\ x_1 = 0 \end{cases}$ contradiction avec le résultat précédent.

Dans la méthode de Gauss des difficultés surviennent lorsque l'élément pivot est petit , car les erreur d'arrondi sont relativement importants et affectent toute la suite des opérations.

Choix de pivot

Il existe une manière simple de procéder pour que la méthode soit stable:

on choisis comme pivot l'élément $a_{ik}^{(k-1)}$ $(i \geq k)$ de plus grande valeur absolue et on permute l'équation k avec l'équation i

$$A^{(k-1)} = \begin{bmatrix} a_{11}^{(k-1)} & \dots & a_{1k}^{(k-1)} & a_{1,k+1}^{(k-1)} & a_{1,k+2}^{(k-1)} & \dots & a_{1n}^{(k-1)} \\ 0 & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & a_{kk}^{(k-1)} & a_{k,k+1}^{(k-1)} & & & a_{kn}^{(k-1)} \\ \vdots & \vdots & \vdots & \vdots & & & \vdots \\ \vdots & \vdots & a_{ik}^{(k-1)} & a_{i,k+1}^{(k-1)} & & & a_{in}^{(k-1)} \\ 0 & \vdots & \vdots & \vdots & & & \vdots \\ 0 & \dots & a_{nk}^{(k-1)} & a_{n,k+1}^{(k-1)} & \dots & \dots & a_{nn}^{(k-1)} \end{bmatrix}$$

ligne k

ligne i

L'élément de plus grande valeur absolue dans la colonne k de $A^{(k-1)}$ et tel que $i \geq k$

Décomposition LU de la matrice A

Soit $A \cdot x = b$ et $\det(A) \neq 0$

On pourrait résoudre le système facilement si $\boldsymbol{A = LU}$

où L est une matrice triangulaire inférieure dont les éléments diagonaux sont égaux à 1:

$$L = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ l_{21} & 1 & 0 & \dots & 0 \\ l_{31} & l_{32} & 1 & \dots & 0 \\ \vdots & \vdots & & & \vdots \\ l_{n1} & l_{n2} & \dots & l_{n2} & 1 \end{bmatrix}$$

et U est une matrice triangulaire supérieure: $U = \begin{bmatrix} u_{11} & u_{12} & u_{13} & \dots & u_{1n} \\ 0 & u_{22} & u_{23} & \dots & u_{2n} \\ 0 & 0 & u_{33} & \dots & u_{3n} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & u_{nn} \end{bmatrix}$

Dans ce cas le système $Ax=b$ s'écrit $LUx=b$

d'où $\begin{cases} Ly = b \\ y = Ux \end{cases}$

La résolution du système comporte alors 3 étapes:

1) Calcul de L et U
2) Résolution du système $Ly=b$, ce qui donne y
3) Résolution du système $Ux=y$, ce qui donne x

Étapes faciles car L et U sont triangulaires

Condition nécessaire et suffisantes d'existence de la décomposition LU

Déf. A matrice régulière ssi detA≠0

Matrices principales de la matrice A:

$$A_1=[a_{11}],\quad A_2=\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix},\quad A_3=\begin{bmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{bmatrix},\quad \dots\quad A_n=\begin{bmatrix}a_{11}&a_{12}&\dots&a_{1n}\\a_{21}&a_{22}&\dots&a_{2n}\\\vdots&\vdots&\vdots&\vdots\\a_{n1}&a_{n2}&\dots&a_{nn}\end{bmatrix}$$

Théorème

Une matrice régulière d'ordre n possède une factorisation $A=LU$ où L est une matrice triangulaire inférieure à diagonale unitaire et U est une matrice triangulaire supérieure, si et seulement si toutes les matrices principales de A sont régulières. Dans ce cas la factorisation est unique.

$$\left(\det A\neq 0 \text{ et } \exists! L\ \exists! U \quad A=LU\right)\Leftrightarrow \text{matrices principales de A sont réglières}$$

Calcul de L et U

On peut obtenir la décomposition LU d'une matrice en écrivant de façon matricielle la méthode de Gauss.

En effet, l'application de la méthode de Gauss au système linéaire $Ax=b$ donne:

$$T_{n-1}T_{n-2}\dots T_2T_1Ax=b'$$

Où T_k $(k=1,..,n-1)$ sont des matrices de transformations effectués: $T_kA^{(k-1)}=A^{(k)}$

et $T_{n-1}T_{n-2}\dots T_2T_1A$ est une matrice triangulaire supérieure

Notons $U=T_{n-1}T_{n-2}\dots T_2T_1A$ Alors $A=\left(T_{n-1}T_{n-2}\dots T_2T_1\right)^{-1}U$

Chaque matrice T_i est une matrice triangulaire inférieure à diagonale unitaire

$$T_k=\begin{bmatrix}1&0&\dots&0&0&\dots&0\\0&1&\dots&0&0&\dots&0\\0&0&\ddots&0&\vdots&&\vdots\\\vdots&\vdots&\dots&1&&&0\\\vdots&\vdots&\vdots&-m_{k+1,k}&\vdots&0&0\\0&\vdots&0&\vdots&0&1&0\\0&\dots&0&-m_{n,k}&0&0&1\end{bmatrix}$$

$$m_{ik}=\frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}$$

D'où $L=\left(T_{n-1}T_{n-2}\dots T_2T_1\right)^{-1}=$

$$\begin{bmatrix}1&0&\dots&0&0&\dots&0\\m_{21}&1&\dots&0&0&\dots&0\\m_{31}&m_{32}&\dots&0&\vdots&&\vdots\\\vdots&\vdots&\dots&1&&&0\\\vdots&\vdots&\vdots&m_{k+1,k}&\vdots&0&0\\m_{n-1,1}&m_{n-1,2}&\vdots&\vdots&\vdots&1&0\\m_{n,1}&m_{n2}&\dots&m_{n,k}&\dots&m_{n,n-1}&1\end{bmatrix}$$

une matrice triangulaire inférieure à diagonale unitaire

alors $A=LU$

Méthode de Gauss-Jordan

C'est une variante de la méthode de Gauss, mais au lieu de trianguliser la matrice on la diagonalise.

Exemple: Soit $n = 3$ et $a_{11} \neq 0$

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases}$$

$$\begin{cases} x_1 + a_{12}'x_2 + a_{13}'x_3 = b_1' & L_1 \to L_1 \cdot \frac{1}{a_{11}} \\ a_{22}'x_2 + a_{23}'x_3 = b_2' & L_2 \to L_2 - \frac{a_{21}}{a_{11}} L_1 \\ a_{32}'x_2 + a_{33}'x_3 = b_3' & L_3 \to L_3 - \frac{a_{31}}{a_{11}} L_1 \end{cases}$$

Si a_{11} =0 il faut permuter les équations de façon à amener à la place de a_{11} un terme non nul; si la matrice A est inversible c'est toujours possible.

$$\begin{cases} x_1 + a_{13}''x_3 = b_1'' & L_1 \to L_1 - \frac{a_{12}'}{a_{22}'} L_2 \\ x_2 + a_{23}''x_3 = b_2'' & L_2 \to L_2 \cdot \frac{1}{a_{22}'} \\ a_{33}''x_3 = b_3'' & L_3 \to L_3 - \frac{a_{32}'}{a_{22}'} L_2 \end{cases} \Rightarrow \begin{cases} x_1 = b_1''' & L_1 \to L_1 - \frac{a_{13}''}{a_{33}''} L_3 \\ x_2 = b_2''' & L_2 \to L_2 - \frac{a_{23}''}{a_{33}''} L_3 \\ x_3 = b_3''' & L_3 \to L_3 \cdot \frac{1}{a_{33}''} \end{cases}$$

Généralisation: procédure de la méthode de Gauss-Jordan

Notons: $A^{(0)} = A$ la matrice initiale

$b^{(0)} = b$ la colonne second membre (initiale)

$A^{(k)}$ la matrice du système à l'issue de k-ème étape k=1,…,n-1

$b^{(k)}$ la colonne second membre à l'issue de k-ème étape

Calcul des $a_{ij}^{(k)}$ et des $b_i^{(k)}$:

- Si $a_{kk}^{(k-1)} = 0$, trouver une ligne i, $i>k$ avec $a_{ik}^{(k-1)} \neq 0$ et permuter les lignes k et i.

  On note la matrice résultante $A^{(k-1)}$; Faire de même pour le vecteur $b^{(k-1)}$.

$$A^{(k-1)} = \begin{bmatrix} 1 & 0 & 0 & \dots & a_{1k}^{(k-1)} & \dots & a_{1n}^{(k-1)} \\ 0 & 1 & 0 & \vdots & \vdots & & \vdots \\ 0 & 0 & 1 & & a_{ik}^{(k-1)} & & a_{kn}^{(k-1)} \\ \vdots & 0 & 0 & \vdots & \vdots & & \vdots \\ \vdots & \vdots & 0 & 0 & a_{kk}^{(k-1)} & & a_{in}^{(k-1)} \\ 0 & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & a_{nk}^{(k-1)} & \dots & a_{nn}^{(k-1)} \end{bmatrix}$$

- Pour i=1,..,n et j=k,…,n

Si $i=k$ $\quad a_{kj}^{(k)} = \frac{a_{kj}^{(k-1)}}{a_{kk}^{(k-1)}}, \quad b_k^{(k)} = \frac{b_k^{(k-1)}}{a_{kk}^{(k-1)}}$

Si $i \neq k$ $\quad a_{ij}^{(k)} = a_{ij}^{(k-1)} - a_{ik}^{(k-1)} a_{kj}^{(k)}$

$\qquad\qquad b_i^{(k)} = b_i^{(k-1)} - a_{ik}^{(k-1)} b_k^{(k)}$

Méthode de Gauss-Jordan trouve des applications lorsqu'on a plusieurs systèmes à résoudre avec la même matrice A et lorsqu'on veut inverser une matrice.

Résolutions itératives

- Ces méthodes ne donnent généralement qu'un résultat approché en un nombre fini d'itérations.

- Elles sont utilisées lorsqu'on a un système de grande dimension à résoudre ou lorsque le système présente des propriétés particulières comme la symétrie par exemple.

Méthode de Jacobi (méthode des déplacements simultanés)

L'idée: construire une suite de vecteurs x^k qui s'approchent (au sens de la norme dans R^n) de la solution x quand k augmente.

Exemple: Soit $n = 3$ $\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases}$ (*)

On résout la 1ère équation par rapport à x_1, la second par rapport à x_2, etc…, ce qui donne:

$$\begin{cases} a_{11}x_1 = b_1 - a_{12}x_2 - a_{13}x_3 \\ a_{22}x_2 = b_2 - a_{21}x_1 - a_{23}x_3 \\ a_{33}x_3 = b_3 - a_{31}x_1 - a_{32}x_2 \end{cases} \qquad \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} - \begin{bmatrix} 0 & a_{12} & a_{13} \\ a_{21} & 0 & a_{23} \\ a_{31} & a_{32} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Donnons les valeurs initiales arbitraires $x_1^{(0)}, x_2^{(0)}, x_3^{(0)}$; en les substituant dans le second membre du système (*) on obtient de nouvelles valeurs $x_1^{(1)}, x_2^{(1)}, x_3^{(1)}$:

$$\begin{cases} x_1^{(1)} = (b_1 - a_{12}x_2^{(0)} - a_{13}x_3^{(0)}) / a_{11} \\ x_2^{(1)} = (b_2 - a_{21}x_1^{(0)} - a_{23}x_3^{(0)}) / a_{22} \\ x_3^{(1)} = (b_3 - a_{31}x_1^{(0)} - a_{32}x_2^{(0)}) / a_{33} \end{cases}$$

Ce nouvel ensemble porté dans le 2ème membre de (*) donne un autre ensemble $x_1^{(2)}, x_2^{(2)}, x_3^{(2)}$ et ainsi de suite.

Généralisation: procédure de la méthode de Jacobi

$A \cdot x = b$ A est une matrice à $n \times n$; b est un vecteur dans R^n;

Le processus itératif de Jacobi s'écrit: $a_{ii} x_i^{(k+1)} = b_i - \sum_{\substack{j=i \\ j \neq i}}^{n} a_{ij} x_j^{(k)}$

ou sous forme matricielle: $Dx^{(k+1)} = b - (A - D)x^{(k)}$

où D est la matrice diagonale de A: $D = \begin{bmatrix} a_{11} & 0 & 0 & \dots & 0 \\ 0 & a_{22} & 0 & \dots & 0 \\ 0 & 0 & a_{33} & \dots & \vdots \\ \vdots & \vdots & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & a_{nn} \end{bmatrix}$

$x^{(0)}$ est le vecteur initiale arbitraire

$x^{(k)}$ est le vecteur obtenu à la k-ème itération

$$Dx^{(k+1)} = b - (A - D)x^{(k)} \Rightarrow x^{(k+1)} = D^{-1}b - (D^{-1}A - I)x^{(k)}$$

$$\Downarrow$$

Relation d'itération de Jacobi: $\boxed{x^{(k+1)} = (I - D^{-1}A)x^{(k)} + D^{-1}b}$ (1)

D'autre part $Ax = b \Leftrightarrow (D + A - D)x = b \Leftrightarrow Dx = b - (A - D)x$

$$\Updownarrow$$

$$\boxed{x = (I - D^{-1}A)x + D^{-1}b} \quad (2)$$

Alors la solution du système $Ax=b$ est le point fixe associé à la relation d'itération

On soustrait (1) et (2) et on obtient: $x^{(k+1)} - x = (I - D^{-1}A)(x^{(k)} - x)$

D'où $x^{(k+1)} - x = (I - D^{-1}A)^2(x^{(k-1)} - x) = \dots = (I - D^{-1}A)^{k+1}(x^{(0)} - x)$

$\varepsilon^{(k+1)} = x^{(k+1)} - x$ L'erreur à la k^{ème} itération

$\varepsilon^{(k+1)} = (I - D^{-1}A)^{k+1}\varepsilon^{(k)} = B\varepsilon^{(k)}$ $B = I - D^{-1}A$

Matrice d'itération de Jacobi

Rappel de quelques notions d'algèbre

Soit B une matrice $n \times n$

(1) $\lambda \in C$ est une valeur propre de B ssi $\exists u \in C^n$ tel que $Bu = \lambda u$

(2) Les racines du polynôme caractéristique $P_B(\lambda)=\det(B- \lambda I)$ sont les valeurs propres de B.

(3) Le rayon spectral de B est le nombre $\rho(B) = \max_{1 \le i \le n} |\lambda_i|$

Théorème (convergence de la méthode de Jacobi)

La suite définie par $x^{(k+1)} = (I - D^{-1}A)x^{(k)} + D^{-1}b$ converge vers la solution du système $Ax=b$ quelque soit $x^{(0)}$ si et seulement si $\rho(I-D^{-1}A) < 1$.

Remarque 1 Plus $\rho(B)$ est petit plus vite la méthode converge. $(B = I - D^{-1}A)$

Remarque 2 On arrête le calcul lorsque deux valeurs successives de $x^{(k)}$ sont « suffisamment » voisines. On peut utiliser par exemple l'un de deux critères:

- Convergence absolue: $\sum_{j=1}^{n} \left| x_j^{(k+1)} - x_j^{(k)} \right| < \varepsilon$
- Convergence relative: $\sum_{j=1}^{n} \left| \frac{x_j^{(k+1)} - x_j^{(k)}}{x_j^{(k+1)}} \right| < \varepsilon$

Méthode de Gauss-Seidel (méthode des déplacements successifs)

L'idée: des itérations semblable à celles de la méthode de Jacobi mais dès qu'une composante est calculée, elle est utilisée.

Exemple: Soit $n = 3$ $\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases}$

On calcule d'abord $x_1^{(1)}$ à l'aide de valeurs arbitraires $x_2^{(0)}, x_3^{(0)}$, ce qui donne:

$$x_1^{(1)} = (b_1 - a_{12}x_2^{(0)} - a_{13}x_3^{(0)}) / a_{11}$$

C'est cette nouvelle valeur de x_1 et non $x_1^{(0)}$ qui est portée dans la 2ème équation:

$$x_2^{(1)} = (b_2 - a_{21}x_1^{(1)} - a_{23}x_3^{(0)}) / a_{22}$$

De même, dans la troisième équation on porte $x_1^{(1)}, x_2^{(1)}$ au lieu de $x_1^{(0)}, x_2^{(0)}$:

$$x_3^{(1)} = (b_3 - a_{31}x_1^{(1)} - a_{32}x_2^{(1)}) / a_{33}$$

On utilise toujours la plus récente des valeurs calculées.

Généralisation: procédure de la méthode de Gauss-Seidel

$A \cdot x = b$ A est une matrice à $n \times n$; b est un vecteur dans R^n;

Le processus itératif de Gauss-Seidel s'écrit:

$$a_{ii}x_i^{(k+1)} = b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}$$

Notons: $A = E + D + F$

$$E = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ a_{21} & 0 & 0 & \dots & 0 \\ a_{31} & a_{32} & 0 & \dots & \vdots \\ \vdots & \vdots & \dots & \dots & 0 \\ a_{n1} & a_{n2} & \dots & a_{nn-1} & 0 \end{bmatrix}, \quad D = \begin{bmatrix} a_{11} & 0 & 0 & \dots & 0 \\ 0 & a_{22} & 0 & \dots & 0 \\ 0 & 0 & a_{33} & \dots & \vdots \\ \vdots & \vdots & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & a_{nn} \end{bmatrix}, \quad F = \begin{bmatrix} 0 & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & 0 & a_{23} & \dots & a_{2n} \\ 0 & 0 & 0 & \dots & \vdots \\ \vdots & \vdots & 0 & \dots & a_{n-1,n} \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

Forme matricielle:

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

$$\Downarrow$$

$$x^{(k+1)} = -(D+E)^{-1}Fx^{(k)} + (D+E)^{-1}b \quad \text{où } F = A - D - E$$

$$\Downarrow$$

$$\boxed{x^{(k+1)} = \left(I - (D+E)^{-1}A\right)x^{(k)} + (D+E)^{-1}b}$$

Relation d'itération de Gauss-Seidel

Théorème (convergence de la méthode de Gauss-Seidel)

La suite définie par $x^{(k+1)} = \left(I - (D+E)^{-1}A\right)x^{(k)} + (D+E)^{-1}b$ converge vers la solution du système $Ax=b$ quelque soit $x^{(0)}$ si et seulement si

$$\rho\,(I-(D+E)^{-1}A) < 1.$$

Matrice d'itération de Gauss-Seidel: $B = I - (D+E)^{-1}A$

Remarque 1 Plus $\rho(B)$ est petit plus vite la méthode converge.

Remarque 2 On arrête le calcul lorsque deux valeurs successives de $x^{(k)}$ sont « suffisamment » voisines. On peut utiliser, par exemple, l'un de deux critères:

- Convergence absolue:

$$\sum_{j=1}^{n} \left| x_j^{(k+1)} - x_j^{(k)} \right| < \varepsilon$$

- Convergence relative:

$$\sum_{j=1}^{n} \left| \frac{x_j^{(k+1)} - x_j^{(k)}}{x_j^{(k+1)}} \right| < \varepsilon$$

Équations différentielles

La résolution numérique d'une équation différentielle est appliquée si on ne connaît pas de solution analytique ou lorsque celle-ci est connue sous une forme difficilement exploitable.

Le problème

On cherche une (ou des) fonction(s) $y(x)$ définie sur intervalle $I \subset$ R telle(s) que:

$$y'(x) = f(x, y(x)) \quad \forall x \in I$$

Condition initiale : $y(x_0) = y_0$

L'idée générale:

Chercher une solution par une méthode pas-à-pas; c'est-à-dire que la solution (approchée) n'est connue qu'en un nombre fini n de points (ce nombre n est choisi par l'utilisateur).

On utilise développement de Taylor : $y(x+h) = y(x) + hy'(x) + \frac{h^2}{2} y''(x) + \ldots$

$$= y(x) + hf(x, y(x)) + \frac{h^2}{2} f'(x, y(x)) + \ldots$$

avec $y'' = \frac{df}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial x}$

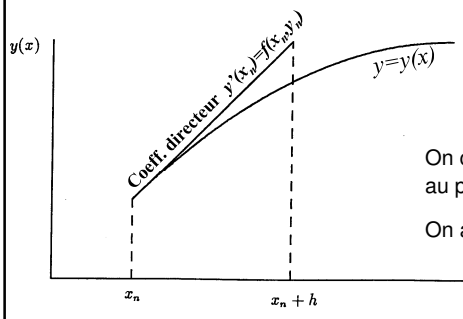
Première méthode d'Euler

Cette méthode consiste à tronquer le développement ci-dessus au premier ordre et à écrire :

$$y(x+h) \cong y(x) + hf(x, y(x))$$ où h est le pas de la discrétisation.

donc $y_{n+1} = y_n + hf(x_n, y_n)$ n = 0,1,2,...

Interprétation géométrique de la méthode d'Euler.



On considère la tangente à la courbe solution $y=y(x)$ au point d'abscisse x.

On assimile donc localement la courbe à sa tangente.

L'intérêt pratique de cette méthode est très limité et on doit la considérer comme une première approche des méthodes pas-à-pas.

Exemple

Considérons l'équation et la condition initiale suivantes : $y'(x) = x + y(x)$

$y(0) = 0$

dont la solution analytique s'écrit : $y(x) = \exp(x) - x - 1$

Développement de 1er ordre: $y(x+h) \cong y(x) + hf(x, y(x)) = y(x) + h(x + y(x))$

En vue de la programmation ou des calculs avec une calculatrice, on écrit :

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

En prenant un pas h = 0,2 on trouve les résultats consignés dans le tableau :

n	x_n	y_n	$0{,}2(x_n+y_n)$	*exact*	$\lvert erreur \rvert$
0	0,0	0,000	0,000	0,000	0,000
1	0,2	0,000	0,040	0,021	0,021
2	0,4	0,040	0,088	0,092	0,052
3	0,6	0,128	0,146	0,222	0,094
4	0,8	0,274	0,215	0,426	0,152
5	1,0	0,489		0,718	0,229

Si on diminue le pas h pour améliorer la précision des résultats, on augmente le volume des calculs et la précision reste médiocre.

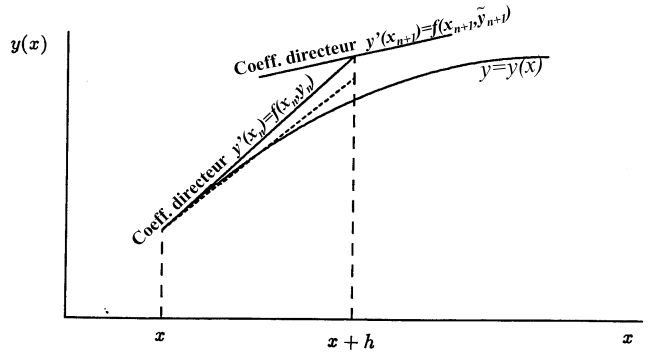
Deuxième méthode d'Euler (améliorée)

Le principe de la méthode consiste à prédire une première valeur $\tilde{y}$ de la solution et à prendre une valeur moyenne de la pente de la tangente entre la valeur courante et la valeur prédite; c'est à dire

on calcule : $\tilde{y}_{n+1} = y_n + hf(x_n, y_n)$

et on obtient : $y_{n+1} = y_n + \frac{1}{2}h\left[f(x_n, y_n) + f(x_{n+1}, \tilde{y}_{n+1})\right]$

Une interprétation géométrique de la méthode :



Généralités sur les équations différentielles ordinaires (EDO)

On cherche une solution $y(x), x \in [a, b]$ d'équation: $y'(x) = f(x, y(x))$

Avec la condition initiale : $y(x_0) = y_0$

Une condition suffisante d'existence et d'unicité de la solution de ce système est donnée par le théorème suivant :

Théorème de Cauchy-Lipschitz

Si $f(x, y)$ est continue sur $[a, b] \times R$ et lipschitzienne indépendamment de x par rapport à y, c'est-à-dire

$$\exists L > 0 \ \ \forall (x, y, z) \in [a,b] \times R^2 \quad |f(x, y) - f(x, z)| \leq L|y - z|$$

et si $(x_0, y_0) \in [a,b] \times R,$

alors $y'(x) = f(x, y(x))$ avec $y(x_0) = y_0$

admet une solution unique sur $[a, b]$.

Remarque

Le point le plus délicat dans l'application du théorème est la détermination d'une constante L (L ne dépend ni de x, ni de y).

D'une manière générale, les méthodes à pas séparés s'écrivent :

$$y_{n+1} = y_n + h \cdot \varphi(x_n, y_n, h)$$

Le choix de φ permet de caractériser ces méthodes et entraîne les propriétés de :

- consistance;
- stabilité;
- convergence.

Consistance

Une méthode définie par $y_{n+1} = y_n + h \cdot \varphi(x_n, y_n, h)$

est consistante pour l'équation $y'(x) = f(x, y(x)) \quad \forall x \in [a, b]$

si φ est continue par rapport à (x, y, h) et si

$$\forall (x, y) \in [a,b] \times R \qquad \varphi(x, y, 0) = f(x, y)$$

Exemples : Les méthodes d'Euler et d'Euler améliorée sont consistantes.

Euler:	Euler améliorée:
$\varphi(x, y, h) = f(x, y)$	$\varphi(x, y, h) = \frac{1}{2}\left[f(x, y) + f(x + h, y + hf(x, y)\right]$
$\varphi(x, y, 0) = f(x, y)$	$\varphi(x, y, 0) = f(x, y)$

Stabilité

Une méthode définie par $y_{n+1} = y_n + h \cdot \varphi(x_n, y_n, h)$

est stable pour l'équation $y'(x) = f(x, y(x)) \quad \forall x \in [a,b]$

s'il existe L >0 tel que $\forall (x, y, z, h) \in [a,b] \times R^3 \quad |\varphi(x, y, h) - \varphi(x, z, h)| \le L|y - z|$

Remarque 1

Si f est lipschitzienne, les méthodes d'Euler et d'Euler améliorée sont stables.

Remarque 2

Si φ est dérivable par rapport à y et si φ_y' est bornée, alors la méthode définie par $y_{n+1} = y_n + h \cdot \varphi(x_n, y_n, h)$ est stable.

Convergence

Une méthode à pas séparés définie par $y_{n+1} = y_n + h \cdot \varphi(x_n, y_n, h)$

est convergente pour l'équation $y'(x) = f(x, y(x)) \quad \forall x \in [a,b]$

avec f vérifiant les hypothèses du théorème de Cauchy-Lipschitz, si

$$\lim_{h \to 0} \max_{x_k \in [a,b]} |y_k - y(x_k)| = 0$$

où $y(x)$ est l'unique solution.

Théorème

Une méthode à pas séparés qui est consistante et stable est convergente.

Ordre d'une méthode

On dit qu'une méthode est d'ordre $p > 0$ *si*

$$\sup_n \left| \frac{y(x_{n+1}) - y(x_n)}{h} - \varphi(x_n, y(x_n), 0) \right| \leq L \cdot h^p$$

où L est indépendant de h.

Exemple : Considérons la méthode d'Euler et un développement de Taylor tronqué à l'ordre 2; on a :

$$y(x_{n+1}) - y(x_n) - hf(x_n, y(x_n)) = \frac{h^2}{2} y''(\xi_n) + o(h^3) \ , \quad \xi_n \in]x_n, x_{n+1}[$$

d'où $$\left| \frac{y(x_{n+1}) - y(x_n)}{h} - \varphi(x_n, y(x_n), 0) \right| \leq \frac{h}{2} M$$

si y'' est borné par M sur $[a, b]$.

Donc la méthode d'Euler est d'ordre 1.

Méthode de Runge-Kutta d'ordre 2

Soit à résoudre : $y'(x) = f(x, y(x))$

$y(x_0) = y_0$

►La méthode consiste à poser: $y_{n+1} = y_n + ak_1 + bk_2$

avec $k_1 = hf(x_n, y_n)$

$k_2 = hf(x_n + \alpha h, y_n + \beta k_1)$

où a, b, α *et* β sont des constantes à déterminer.

►Le calcul des coefficients est basé sur le développement de Taylor.

Notation:

$$f_x = \frac{\partial f}{\partial x}, \quad f_y = \frac{\partial f}{\partial y}.$$

$$y'' = \frac{df}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial x} = f_x + f_y \cdot y' = f_x + f_y \cdot f$$

Soit $y_{n+1} = y_n + ak_1 + bk_2$ (1)

avec $k_1 = hf(x_n, y_n)$ (2)

$k_2 = hf(x_n + \alpha h, y_n + \beta k_1)$ (3)

où a, b, α et β sont des constantes à déterminer.

Le calcul des coefficients est basé sur le développement de Taylor:

$$y(x_{n+1}) = y(x_n) + hy'(x_n) + \frac{h^2}{2} y''(x_n) + o(h^3)$$

Sachant que $y'' = f_x + f_y \cdot f$ on obtient :

$y(x_{n+1}) = y(x_n) + hf(x_n, y_n) + \frac{h^2}{2}\left(f_x + f_y f\right) + o(h^3)$ où f et les dérivées sont prises en (x_n, y_n).

Or le développement de $f(x_n + \alpha h, y_n + \beta k_1)$ en (x_n, y_n) est $f(x_n, y_n) + \alpha h f_x + \beta k_1 f_y + o(h^2)$

Donc, d'après (3): $k_2 = h\left(f(x_n, y_n) + \alpha h f_x + \beta k_1 f_y + o(h^2)\right)$

$= h\left(f(x_n, y_n) + \alpha h f_x + \beta h f(x_n, y_n) f_y + o(h^2)\right)$

En reportant cette expression dans (1) et en tenant compte de (2) on a:

$y_{n+1} = y_n + ahf(x_n, y_n) + bh\left(f(x_n, y_n) + \alpha h f_x + \beta h f(x_n, y_n) f_y + o(h^2)\right)$

$= y_n + (a+b)hf + bh^2\left(\alpha f_x + \beta f f_y\right) + o(h^3)$ (5)

En comparant (4) et (5), on obtient: $\begin{cases} a + b = 1 \\ b\alpha = \frac{1}{2} \\ b\beta = \frac{1}{2} \end{cases}$

$$\begin{cases} a+b=1 \\ b\alpha=\frac{1}{2} \\ b\beta=\frac{1}{2} \end{cases}$$

On a trois équations et quatre inconnues donc un degré de liberté.

On choisit, par exemple: $a = b = 1/2$ et $\alpha = \beta = 1$,

d'où les relations du schéma de Runge-Kutta d'ordre 2 :

$$k_1 = hf(x_n, y_n)$$

$$k_2 = hf(x_n + h, y_n + k_1)$$

$$y_{n+1} = y_n + \tfrac{1}{2}(k_1 + k_2)$$

Ces relations sont identiques à celles du schéma de la méthode d'Euler améliorée.

Méthode de Runge-Kutta d'ordre 4

Les calculs développés jusqu'à l'ordre 2 peuvent être poursuivis à des ordres supérieurs. L'établissement des formules devient de plus en plus complexe. La question est alors de savoir si à l'augmentation de la complexité de la formule à programmer correspond une augmentation significative de la précision des résultats. L'ordre 4 est souvent utilisé.

Les relations du schéma de Runge-Kutta d'ordre 4 :

$$k_1 = hf(x_n, y_n)$$

$$k_2 = hf(x_n + \tfrac{h}{2}, y_n + \tfrac{k_1}{2})$$

$$k_3 = hf(x_n + \tfrac{h}{2}, y_n + \tfrac{k_2}{2})$$

$$k_4 = hf(x_n + h, y_n + k_3)$$

$$y_{n+1} = y_n + \tfrac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

(*)

Les formules sont simples et la précision suffisante dans de nombreux cas.

Exemple

On reprend l'exemple précédent: $y'(x) = x + y(x)$
$y(0) = 0$

Solution analytique: $y(x) = \exp(x) - x - 1$

En prenant un pas h = 0, 2, et on appliquant les formules (*) on obtient:

n	x_n	y_n	*exact*	$\|erreur\|$
0	0,0	0,000000	0,000000	0,000000
1	0,2	0,021400	0,021403	3×10^{-6}
2	0,4	0,091818	0,091825	7×10^{-6}
3	0,6	0,222107	0,222119	11×10^{-6}
4	0,8	0,425521	0,425541	20×10^{-6}
5	1,0	0,7182251	0,718282	31×10^{-6}

La méthode de Runge-Kutta d'ordre 4 donne relativement bons résultats.