

Université Larbi Ben MHIDI Oum El Bouaghi
Faculté des Sciences et Sciences Appliquées pole Ain Beida
Département de Génie Mécanique

Premier cours en :

Méthodes Numériques

Conforme au programme de la 2^{eme} année ST.

Par : Dr. MAMERI Abdelbaki

Programme Méthodes numériques

Chapitre 1 : Résolution des équations non linéaires $f(x)=0$

1. Introduction générale,
2. Méthode de bissection,
3. Méthode des approximations successives (point fixe),
4. Méthode de Newton-Raphson.

Chapitre 2 : Interpolation polynomiale

1. Introduction générale,
2. Polynôme de Lagrange,
3. Polynômes de Newton.

Chapitre 3 : Intégration numérique

1. Introduction générale,
2. Méthode du trapèze,
3. Méthode de Simpson,
4. Formules de quadrature.

Chapitre 4 : Résolution des équations différentielles ordinaires (problème de la condition initiale ou de Cauchy).

1. Introduction générale,
2. Méthode d'Euler,
3. Méthode d'Euler améliorée,
4. Méthode de Runge-Kutta d'ordre quatre.

Chapitre 5 : Méthode de résolution directe des systèmes d'équations linéaires

1. Introduction et définitions,
2. Méthode de Gauss et pivotation,
3. Méthode de factorisation LU,
4. Méthode de factorisation de Choleski MMT,
5. Algorithme de Thomas (TDMA) pour les systèmes tri diagonales.

Chapitre 6 : Méthode de résolution approximative des systèmes d'équations linéaires

1. Introduction et définitions,
2. Méthode de Jacobi,
3. Méthode de Gauss-Seidel,
4. Utilisation de la relaxation.

Chapitre I : Résolution numérique des équations non linéaires à une seule variable

Dans ce chapitre on va étudier trois méthodes pour la résolution des équations non linéaires à une variable aussi dites équations à variable transcendante. Comme exemple de ces équations, on peut citer $f(x) = \sin(x) + x = 0$, $f(x) = \ln(x) - 2x + 3 = 0$. Ces équations ne possèdent pas une ou des racines exactes qui peuvent être calculées directement, c'est pourquoi on fait recours aux méthodes numériques pour trouver les solutions approchées de ces équations. Les racines calculées sont autant précises que l'on veut surtout lorsqu'on dispose des moyens de calcul.

Ces méthodes numériques permettent seulement le calcul d'une seule racine sur un intervalle bien choisi. Donc si l'équation possède plus d'une racine, il est nécessaire de les localiser dans des intervalles choisis soigneusement et de faire le calcul pour chaque racine à part.

1 Localisation des racines d'une équation $f(x)=0$.

Soit une équation $f(x)=0$ dont on cherche la solution sur un intervalle $[a,b]$, on commence par un tracé grossier de la fonction sur l'intervalle donnée puis on isole chaque racine dans un sous intervalle le plus étroit possible. La Fig. 1 montre le tracé d'une fonction f qui coupe l'axe des x en trois points, c'est-à-dire que l'équation $f(x)=0$ possède trois racines, on note les racines exactes par \bar{x}_1, \bar{x}_2 et \bar{x}_3 .

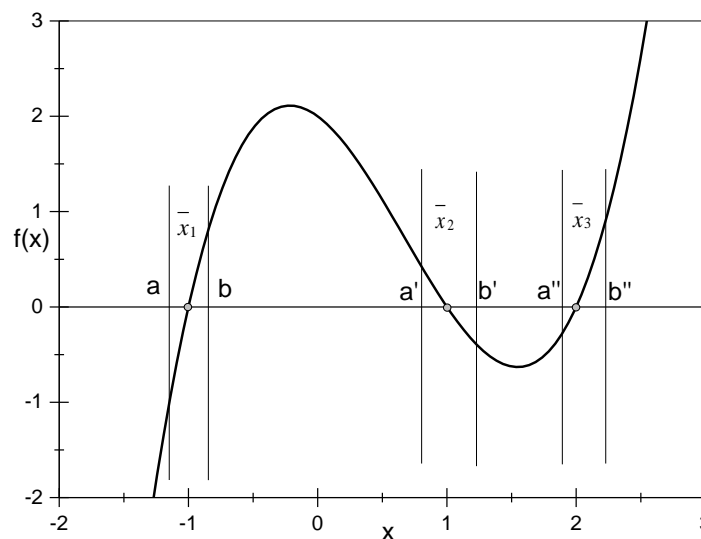


Fig. 1.1. Localisation des racines

On remarque que la fonction est continue sur chaque sous intervalle, aussi chaque sous intervalle :

- Englobe une seule racine tel que $\bar{x}_1 \in [a, b]$, $\bar{x}_2 \in [a', b']$ et $\bar{x}_3 \in [a'', b'']$.
- Vérifie la condition $f(a)f(b) < 0$, $f(a')f(b') < 0$ et $f(a'')f(b'') < 0$.

La forme de l'équation $f(x)=0$ peut être compliquée, dans ce cas s'il est possible on peut la décomposer en deux parties simples $g(x)=h(x)$.

Par exemple $f(x) = \ln(x) - x^2 + 2 = 0$, qui est assez compliqué pour le tracé, peut être décomposée en :

$$g(x) = h(x) \text{ avec } g(x) = \ln(x) \text{ et } h(x) = x^2 - 2$$

Le tracé de g et h est très simple, la solution de $f(x)=0$ se situe à l'intersection de g et h . Ensuite, on projette les points des intersection sur l'axe des x et on localise les racines. On peut facilement vérifier les intervalles trouvés en calculant $f(a_i)f(b_i) < 0$.

Exemple :

Prenons l'équation : $\ln(x) - x^2 + 2 = 0$ localisant ses racines, à ce point on ne connaît pas le nombre de racines de cette équation. On trace la courbe de la fonction :

$f(x) = \ln(x) - x^2 + 2 = 0$, les intersections de la courbe avec l'axe des x représentent les racines de cette fonction.

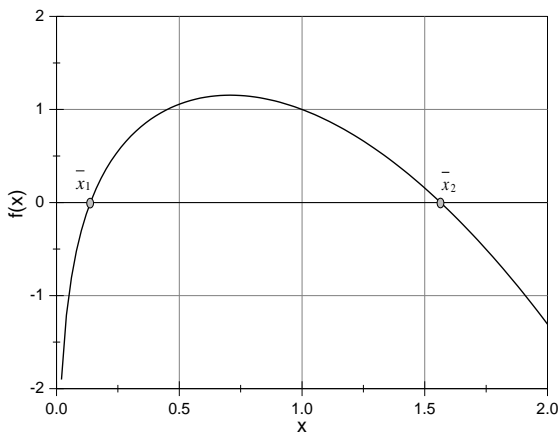


Fig. 1.2. Tracé de la fonction f

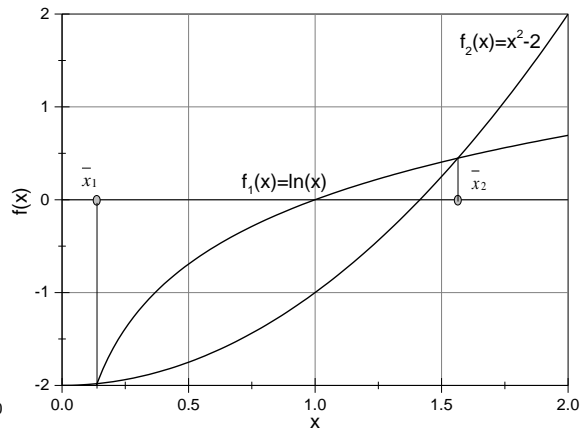


Fig. 1.3. Tracés des fonction f_1 et f_2

On note que cette équation a deux racines (Fig. 2) qui appartiennent par exemple aux intervalles $[0.1, 0.5]$ et $[1, 2]$. On peut aussi réécrire la fonction $f(x)=0$ sous une forme plus simple, par exemple : $\ln(x) - x^2 + 2 = 0 \Leftrightarrow \ln(x) = x^2 - 2$ ou bien

$f_1(x) = f_2(x)$; avec $f_1(x) = \ln(x)$ et $f_2(x) = x^2 - 2$. Ensuite on trace ces deux fonctions (Fig. 3), qui sont faciles sur les même axes, leurs intersections représentent les racines de $f(x)=0$.

2 Méthode de bisection où de dichotomie

C'est la méthode la plus simple et qui nécessite le plus de calculs, elle est basée sur le cas particulier (Théorème de Bolzano) du théorème des valeurs intermédiaires qui dit que :

1. Si $f(x)$ est continue sur l'intervalle $[a, b]$,
2. Si $f(a)$ et $f(b)$ ne sont pas de même signe, il existe au moins un réel c compris entre a et b tel que $f(c) = 0$ (car 0 est compris entre $f(a)$ et $f(b)$).

2.1 Principe de la méthode

Une fois les racines sont localisées chacune dans un intervalle, pour simplifier l'écriture soit par exemple $[a, b]$:

1. On divise l'intervalle en deux parties égales tel que $x_0 = \frac{a+b}{2}$.
2. On obtient deux sous intervalles $[a, x_0]$ et $[x_0, b]$, la racine \bar{x} doit obligatoirement appartenir à l'un d'eux. Pour vérifier, on calcule $f(a)f(x_0)$ et $f(x_0)f(b)$ le produit qui est négatif c'est celui qui correspond à l'intervalle qui contient la solution.

Notons le nouvel intervalle par $[a_1, b_1]$ tel que (Fig. 4):

$$a_1 = \begin{cases} a & \text{si } \bar{x} \in [a, x_0] \\ x_0 & \text{si } \bar{x} \in [x_0, b] \end{cases} \quad \text{et} \quad b_1 = \begin{cases} x_0 & \text{si } \bar{x} \in [a, x_0] \\ b & \text{si } \bar{x} \in [x_0, b] \end{cases}$$

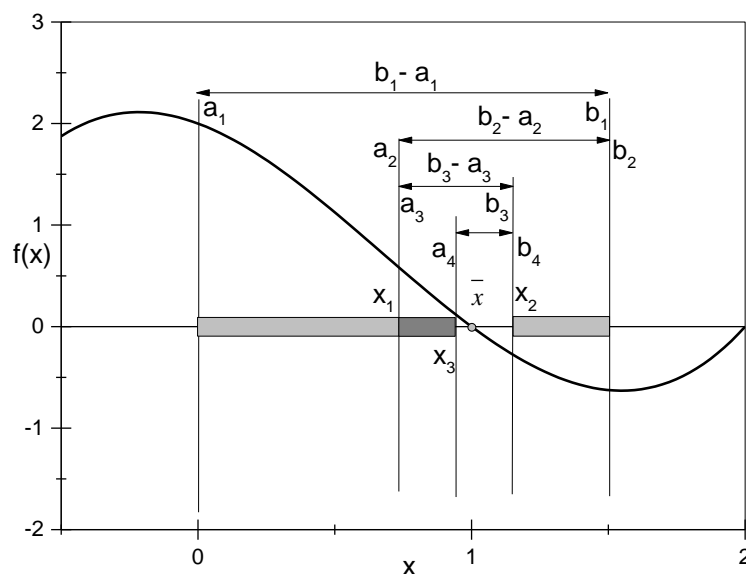


Fig. 1.4. Illustration de la méthode de bisection

En répétant (itérant) la même méthode pour l'intervalle obtenu on aura les valeurs :

$$x_1 = \frac{a_1+b_1}{2}, x_2 = \frac{a_2+b_2}{2}, \dots, \dots, x_n = \frac{a_n+b_n}{2}$$

La suite $\{x_n\}_{n=0,\infty}$ converge vers la solution \bar{x} de $f(x)=0$ lorsque $n \rightarrow \infty$.

2.2 Nombre de divisions pour avoir une précision ε donné.

Puisque chaque fois on divise l'intervalle en deux parties égales, on a :

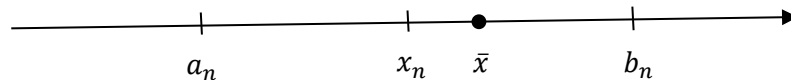
$$b_1 - a_1 = \frac{b - a}{2};$$

$$b_2 - a_2 = \frac{b_1 - a_1}{2} = \frac{1}{2} \frac{b - a}{2} = \frac{b - a}{2^2};$$

$$b_3 - a_3 = \frac{b_2 - a_2}{2} = \frac{1}{2} \frac{b - a}{2^2} = \frac{b - a}{2^3};$$

.....

$$b_n - a_n = \frac{b_{n-1} - a_{n-1}}{2} = \frac{1}{2} \frac{b - a}{2^{n-1}} = \frac{b - a}{2^n}.$$



Puisque $\bar{x} \in [a_n, b_n] = [a_n, x_n] \cup [x_n, b_n]$ on a $|x_n - \bar{x}| \leq \frac{1}{2} \frac{b-a}{2^n} = \frac{b-a}{2^{n+1}}$

Il faut que la différence $|x_n - \bar{x}|$ qui est l'erreur du calcul soit inférieure à une précision donnée ε , c'est-à-dire :

$$|x_n - \bar{x}| \leq \varepsilon$$

Alors, il suffit que

$$\frac{b-a}{2^{n+1}} \leq \varepsilon$$

Cela donne

$$n \geq \frac{\ln(\frac{b-a}{2\varepsilon})}{\ln(2)}$$

Le nombre de divisions ne dépend que de la longueur de l'intervalle est de la précision.

Cette méthode est inconditionnellement convergente, son problème c'est qu'elle est lente c'est pourquoi elle est utilisée pour démarrer d'autres méthodes plus élaborées.

Exemple : Calculons la première racine de l'équation $\ln(x) - x^2 + 2 = 0$ qui appartient à $[0.1, 0.5]$ avec une précision de **0.01**.

Calculons le nombre de divisions à faire :

$$n \geq \frac{\ln\left(\frac{b-a}{2\varepsilon}\right)}{\ln(2)} = \frac{\ln\left(\frac{0.5-0.1}{2*0.01}\right)}{\ln(2)} = 4.32 \text{ on prend } n=5 \text{ puisque } n \text{ est entier et supérieur à } 4.32.$$

$$f(a_1) = f(0.1) = -0.313 \text{ et } f(b_1) = f(0.5) = 1.057$$

$$x_1 = \frac{a_1 + b_1}{2} = \frac{0.1 + 0.5}{2} = 0.30; \quad f(0.3) = 0.706 > 0 \text{ donc } a_2 = 0.1 \text{ et } b_2 = 0.3$$

$$x_2 = \frac{a_2 + b_2}{2} = \frac{0.1 + 0.3}{2} = 0.20; \quad f(0.2) = 0.351 > 0 \text{ donc } a_3 = 0.1 \text{ et } b_3 = 0.2$$

$$x_3 = \frac{a_3 + b_3}{2} = \frac{0.1 + 0.2}{2} = 0.15; \quad f(0.15) = 0.080 > 0 \text{ donc } a_4 = 0.1 \text{ et } b_4 = 0.15$$

$$x_4 = \frac{a_4 + b_4}{2} = \frac{0.1 + 0.15}{2} = 0.125; \quad f(0.125) = -0.095 < 0 \text{ donc } a_5 = 0.125 \text{ et } b_4 = 0.125$$

$$x_5 = \frac{a_5 + b_5}{2} = \frac{0.125 + 0.125}{2} = 0.125; \quad f(0.125) = -0.003 \text{ La solution est } x_5 = 0.125$$

3 Méthode des approximations successives où du point fixe.

Soit g une fonction définie sur un intervalle $[a, b]$, le point \bar{x} qui vérifie $\bar{x} = g(\bar{x})$ avec $\bar{x} \in [a, b]$ est dit point fixe de la fonction g .

Cette méthode est basée sur le principe du point fixe d'une fonction, on écrit l'équation $f(x)=0$ sous la forme $x=g(x)$, ensuite on cherche le point fixe \bar{x} de la fonction g . Pour cela on crée la suite $x_{n+1} = g(x_n)$ ($n=0,1,2,\dots$) avec x_0 donnée par dichotomie par exemple.

On démarre de x_0 pour $n=0$, on calcule $x_1 = g(x_0)$ ensuite $n=1$, on calcule $x_2 = g(x_1), \dots, x_{n+1} = g(x_n)$. Sous certaines conditions, la suite $\{x_n\}_{n=0,\infty}$ converge vers la solution \bar{x} point fixe de g et solution de l'équation $f(x)=0$.

Exemple : Ecrire l'équation $f(x)=0$ sous la forme $x = g(x)$ si $f(x) = x^2 + 3e^x - 12$.

On peut écrire $x = g_1(x) = x^2 + 3e^x - 12 + x$

$$x = g_2(x) = \sqrt{12 - 3e^x}$$

$$x = g_3(x) = \ln\left(\frac{12-x^2}{3}\right)$$

Pour pouvoir choisir la forme de g adéquate pour le calcul, un critère de convergence de cette méthode doit être vérifié.

3.1 Critère de convergence et d'arrêt de calculs pour la méthode des approximations successives.

Soit g une fonction dérivable définie de $[a, b] \rightarrow [a, b]$ tel que (condition suffisante) :

$$|g'(x)| \leq k < 1 \quad \forall x \in [a, b]$$

Alors la suite $\{x_n\}_{n=0,\infty}$ définie par $x_{n+1} = g(x_n)$ ($n=0,1,2,\dots$) converge indépendamment de la valeur de x_0 vers l'unique point fixe \bar{x} de g .

Si plusieurs formes de g vérifient cette condition, on aura plusieurs valeurs de k . On choisit celle avec la valeur minimale de k . En pratique, on calcule $k = \max_{x \in [a, b]} |g'(x)|$ qui doit être inférieure à l'unité pour que la méthode converge.

On arrête les calculs pour cette méthode lorsque la différence absolue entre deux itérations successives est inférieure à une certaine précision ε donnée.

$$|x_{n+1} - x_n| < \varepsilon$$

Exemple : Trouver la première racine de l'équation $\ln(x) - x^2 + 2 = 0$ qui appartient à $[0.1, 0.5]$ avec une précision $\epsilon=0.001$. On écrit cette équation sous la forme $x = g(x)$ et on vérifie les conditions de convergence. On peut écrire :

$$x = e^{x^2-2} = g_1(x) \text{ et}$$

$$x = \sqrt{\ln(x) + 2} = g_2(x)$$

Vérifions la condition de convergence pour cette méthode $k = \max_{x \in [a,b]} |g'(x)|$

$k_1 = \max_{x \in [0.1, 0.5]} |g_1'(x)| = \max_{x \in [0.1, 0.5]} |2xe^{x^2-2}|$ on a $g_1'(x)$ strictement croissante donc $k_1 = \max_{x=0.5} |2 * 0.5e^{0.5^2-2}| = 0.174 < 1$ cette forme converge.

On écrit : $x_{n+1} = g_1(x_n) = e^{x_n^2-2}$ ($n=0,1,2,\dots$)

Commençons $x_0=0.3$ le milieu de l'intervalle initial donné :

$$n = 0, \quad x_1 = g_1(x_0) = e^{x_0^2-2} = 0.148$$

On calcule $|x_1 - x_0| = 0.152 > \epsilon$;

On continue $n = 1, x_2 = g_1(x_1) = e^{x_1^2-2} = 0.138.$

On calcule $|x_2 - x_1| = 0.01 > \epsilon$

On continue $n = 2, x_3 = g_1(x_2) = e^{x_2^2-2} = 0.138.$

On calcule $|x_3 - x_2| = 0.00 < \epsilon$, **La solution est $x_2 = 0.138$**

4 Méthode de Newton-Raphson

C'est la méthode la plus efficace et la plus utilisée, elle repose sur le développement de Taylor. Si $f(x)$ est continue et continument dérivable dans le voisinage de \bar{x} solution de $f(x)=0$, alors le développement en série de Taylor autour d'un estimé x_n proche de \bar{x} s'écrit :

$$f(\bar{x}) = f(x_n) + \frac{(\bar{x}-x_n)}{1!} f'(x_n) + \frac{(\bar{x}-x_n)^2}{2!} f''(x_n) + \dots$$

Si x_n est un estimé proche de \bar{x} , alors le carré de l'erreur $\varepsilon_n = \bar{x} - x_n$ et les termes de degrés supérieurs sont négligeable. Sachant que $f(\bar{x}) = 0$ on obtient la relation approximative :

$$f(x_n) + (\bar{x} - x_n)f'(x_n) \approx 0$$

Donc

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

On peut écrire la $(n + 1)^{eme}$ itération approximant \bar{x} est :

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

Cette suite, si elle converge, doit converger vers la solution \bar{x} de $f(x)=0$. On remarque que $f'(x)$ doit être non nulle.

4.1 Critère de convergence de la méthode de Newton-Raphson

Soit une fonction f définie sur $[a,b]$ telle que :

- i. $f(a)f(b) < 0$
- ii. $f'(x)$ et $f''(x)$ sont non nulles et gardent un signe constant sur l'intervalle donné.

4.2 Critère d'arrêt de calcul pour la méthode de Newton-Raphson

Si la condition de convergence est vérifiée, le procédé itératif doit converger. Cela veut dire que chaque nouvelle itération est meilleure que la précédente, de ce fait on peut dire que si on a une précision ε , on arrête le calcul lorsque la différence absolue entre deux approximations successives est inférieure à la précision donnée. C'est-à-dire :

$$|x_{n+1} - x_n| \leq \varepsilon$$

Si cette condition est vérifiée on prend x_{n+1} comme solution de $f(x)=0$.

Exemple : Trouver la première racine de l'équation $f(x) = \ln(x) - x^2 + 2 = 0$ qui appartient à $[0.1, 0.5]$ avec une précision $\varepsilon=0.0001$. On calcule la dérivée première et seconde de f et on vérifie les conditions de convergence.

On a : $f'(x) = \frac{1}{x} - 2x$ qui est strictement décroissante et positive sur l'intervalle donné. $f'(x) > 0$ et $f''(x) = -\frac{1}{x^2} - 2$, $f''(x) < 0$ sur l'intervalle donné. La condition de convergence est vérifiée. On écrit donc : $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{\ln(x_n) - x_n^2 + 2}{\frac{1}{x_n} - 2x_n}$, ($n=0,1,2,\dots$).

Commençons $x_0=0.3$ le milieu de l'intervalle initial donné :

$$n = 0, \quad x_1 = x_0 - \frac{\ln(x_0) - x_0^2 + 2}{\frac{1}{x_0} - 2x_0} = 0.0417$$

On calcule $|x_1 - x_0| > \varepsilon$;

On continue $n = 1, x_2 = 0.0910$.

On calcule $|x_2 - x_1| > \varepsilon$

On continue $n = 2, x_3 = 0.1285$.

On calcule $|x_3 - x_2| > \varepsilon$;

On continue $n = 3, x_4 = 0.1376$.

On calcule $|x_4 - x_3| > \varepsilon$

On continue $n = 4, x_5 = 0.1379$.

On calcule $|x_5 - x_4| > \varepsilon$

On continue $n = 5, x_6 = 0.1379$. **La solution est $x_4 = 0.1379$**

Remarques :

- La méthode de bisection est inconditionnellement convergente, son inconvénient est sa lenteur pour obtenir la solution avec une grande précision. Elle peut servir pour démarrer d'autres méthodes plus performantes.
- La méthode des approximations successives est plus rapide que celle de bisection à condition qu'elle converge.
- La méthode de Newton-Raphson est la plus rapide, elle permet d'obtenir des solutions très précises en un nombre réduit d'itérations.

Chapitre II : Interpolation polynômiale

Soit par exemple une expérience où on enregistre la distance parcourue par un objet en fonction du temps, les résultats sont donnés dans le tableau suivant :

t(sec)	0	1	2	3	4
X(m)	0	5	15	0	3

On veut par exemple calculer la position de l'objet au temps $t=2.5 \text{ sec}$ ou la vitesse de l'objet à un temps donné. Pour cela, il faut avoir une forme analytique de x en fonction de t , $X(t)$. Cette forme doit au moins coïncider avec les points donnés dans le tableau. Ensuite on peut calculer $X(2.5)$, $\int_0^4 X(t)dt$ où bien $v(t) = \frac{dX(t)}{dt}$.

Dans ce chapitre, on va considérer l'approximation de $X(t)$ par une forme polynômiale c'est-à-dire :

$$X(t) = a_0 + a_1 t + a_2 t^2 + \dots + a_n t^n$$

Avec a_i ($i = 0, n$) sont des coefficients à déterminer.

Les polynômes que nous allons étudier diffèrent seulement par la façon de déterminer les coefficients a_i ($i = 0, n$), car pour un tableau de valeurs données le polynôme d'interpolation est unique.

2.1 Polynôme d'interpolation de Lagrange.

Soient $(n+1)$ points distincts x_0, x_1, \dots, x_n et f une fonction dont les valeurs sont $f(x_0), f(x_1), \dots, f(x_n)$. Alors, il existe un seul polynôme de degré inférieur ou égal à n et qui coïncide avec les points d'interpolation, i e :

$$f(x_k) = P_n(x_k), \quad k = 0, 1, 2, \dots, n$$

Ce polynôme est donné par :

$$P_n(x) = \sum_{i=0}^n f(x_i) L_i(x) = f(x_0) L_0(x) + f(x_1) L_1(x) + \dots + f(x_n) L_n(x)$$

$$\text{Avec} \quad L_k(x) = \sum_{i=0, i \neq k}^n \frac{(x-x_i)}{(x_k-x_i)} = \frac{(x-x_0)}{(x_k-x_0)} \frac{(x-x_1)}{(x_k-x_1)} \dots \frac{(x-x_{k-1})}{(x_k-x_{k-1})} \frac{(x-x_{k+1})}{(x_k-x_{k+1})} \dots \frac{(x-x_n)}{(x_k-x_n)} \quad (k=0, \dots, n)$$

$L_k(x)$ sont dits coefficients polynômes de Lagrange, ils sont orthogonaux c'est-à-dire $L_k(x_j) = 0$ et $L_k(x_k) = 1$.

Exemple : Reprenons la table donnée au début du chapitre et essayons de calculer le polynôme de Lagrange pour cette table. Notons que pour $n+1$ points le degré du polynôme est inférieure ou égal à n . Pour notre cas on a **5** points, cela nous donne un polynôme de de gré inférieur ou égal à **4**.

$$X(t) \approx P_4(t) = \sum_{i=0}^4 f(t_i)L_i(t) = f(t_0)L_0(t) + f(t_1)L_1(t) + f(t_2)L_2(t) + f(t_3)L_3(t) + f(t_4)L_4(t)$$

Avec les coefficients $f(t_i)$ sont les valeurs de $X(t_i)$ aux points donnés t_i , on remplace et on écrit donc :

$$X(t) \approx P_4(t) = 0 * L_0(t) + 5 * L_1(t) + 15 * L_2(t) + 0 * L_3(t) + 3 * L_4(t)$$

Ensuite, on calcule les coefficients polynômes de Lagrange :

$L_0(t) = \sum_{i=0, i \neq 0}^4 \frac{(t-t_i)}{(t_0-t_i)} = \frac{(t-t_1)(t-t_2)(t-t_3)(t-t_4)}{(t_0-t_1)(t_0-t_2)(t_0-t_3)(t_0-t_4)}$ Noter bien qu'il est inutile de calculer les coefficients polynômes $L_0(t)$ et $L_3(t)$ car ils seront multipliés par zéro dans le remplacement.

$$L_1(t) = \sum_{i=0, i \neq 1}^4 \frac{(t-t_i)}{(t_1-t_i)} = \frac{(t-t_0)(t-t_2)(t-t_3)(t-t_4)}{(t_1-t_0)(t_1-t_2)(t_1-t_3)(t_1-t_4)} = \frac{(t-0)(t-2)(t-3)(t-4)}{(1-0)(1-2)(1-3)(1-4)} = -\frac{1}{6}(t^4 - 9t^3 + 26t^2 - 24t)$$

$$L_2(t) = \sum_{i=0, i \neq 2}^4 \frac{(t-t_i)}{(t_2-t_i)} = \frac{(t-t_0)(t-t_1)(t-t_3)(t-t_4)}{(t_2-t_0)(t_2-t_1)(t_2-t_3)(t_2-t_4)} = \frac{(t-0)(t-1)(t-3)(t-4)}{(2-0)(2-1)(2-3)(2-4)} = \frac{1}{4}(t^4 - 8t^3 + 19t^2 - 12t)$$

$$L_4(t) = \sum_{i=0, i \neq 4}^4 \frac{(t-t_i)}{(t_4-t_i)} = \frac{(t-t_0)(t-t_1)(t-t_2)(t-t_3)}{(t_4-t_0)(t_4-t_1)(t_4-t_2)(t_4-t_3)} = \frac{(t-0)(t-1)(t-2)(t-3)}{(4-0)(4-1)(4-2)(4-3)} = \frac{1}{24}(t^4 - 6t^3 + 11t^2 - 6t)$$

Finalement on remplace les coefficients polynômes et on obtient :

$$X(t) \approx P_4(t) = -25.75t + 50.95833t^2 - 23.25t^3 + 3.04167t^4$$

Avec : $f[x_0, x_1] = \frac{f[x_1] - f[x_0]}{x_1 - x_0}$, $f[x_1, x_2] = \frac{f[x_2] - f[x_1]}{x_2 - x_1}$,

$$f[x_0, x_1, x_2] = \frac{f[x_1, x_2] - f[x_0, x_1]}{x_2 - x_0}, \quad f[x_1, x_2, x_3] = \frac{f[x_2, x_3] - f[x_1, x_2]}{x_3 - x_1}, \dots \dots$$

$$f[x_0, x_1, x_2, x_3] = \frac{f[x_1, x_2, x_3] - f[x_0, x_1, x_2]}{x_3 - x_0}, \quad f[x_1, x_2, x_3, x_4] = \frac{f[x_2, x_3, x_4] - f[x_1, x_2, x_3]}{x_4 - x_1}$$

$$f[x_0, x_1, x_2, x_3, x_4] = \frac{f[x_1, x_2, x_3, x_4] - f[x_0, x_1, x_2, x_3]}{x_4 - x_0}$$

2.4 Erreur d'interpolation

C'est l'erreur commise lorsqu'on remplace la fonction f par le polynôme d'interpolation équivalent. Elle est notée par $\varepsilon(x)$ car elle varie d'un point à un autre dans l'intervalle d'interpolation. Cette erreur doit être nulle aux points d'interpolation, $\varepsilon(x_i) = 0$, ($i=0, \dots, n$).

Si la fonction f est continue et $(n+1)$ fois dérivable sur l'intervalle d'interpolation $[a, b]$. Alors pour tout $x \in [a, b]$ il existe $z \in [a, b]$ tel que :

$$\varepsilon(x) = |f(x) - P_n(x)| = \prod_{i=0}^n \frac{(x - x_i)}{(n+1)!} f^{(n+1)}(z)$$

Si $|f^{(n+1)}(z)| \leq M \quad \forall z \in [a, b]$ on peut écrire :

$$\varepsilon(x) = |f(x) - P_n(x)| \leq \prod_{i=0}^n \frac{(x - x_i)}{(n+1)!} M$$

Dans ce cas M est majorant de la fonction $f^{(n+1)}(x)$ sur l'intervalle $[a, b]$.

Exemple : Reprenons la table donnée au début du chapitre et essayons de calculer le polynôme de Newton pour cette table. Notons que pour $n+1$ points le degré du polynôme est inférieure ou égal à n . Pour notre cas on a **5** points, cela nous donne un polynôme de de gré inférieure ou égal à **4**. Ecrivons les polynômes de Newton :

$$\left\{ \begin{array}{l} P_1(t) = a_0 + a_1(t - t_0) \\ P_2(t) = P_1(t) + a_2(t - t_0)(t - t_1) \\ P_3(t) = P_2(t) + a_3(t - t_0)(t - t_1)(t - t_2) \\ P_4(t) = P_3(t) + a_4(t - t_0)(t - t_1)(t - t_2)(t - t_3) \\ \text{Les } a_i \text{ sont les différences divisées d'ordre } i \end{array} \right.$$

Calculons la table des différences divisées.

t_k	$f(t_k) = f[t_k]$	DD^1	DD^2	DD^3	DD^4
0	0 = a₀				
1	5	5 = a₁			
2	15	10	2.5 = a₂		
3	0	-15	9	-5 = a₃	
4	3	3		7.1667	3.04167 = a₄

Remplaçant les a_i et les t_i par leurs valeurs dans les polynômes de Newton, on trouve :

$$P_1(t) = a_0 + a_1(t - t_0) = 0 + 5(t - 0) = 5t$$

$$P_2(t) = P_1(t) + a_2(t - t_0)(t - t_1) = 5t + 2.5(t - 0)(t - 1) = 2.5t^2 + 2.5t$$

$$P_3(t) = P_2(t) + a_3(t - t_0)(t - t_1)(t - t_2) = 2.5t^2 + 2.5t - 5(t - 0)(t - 1)(t - 2) = -5t^3 + 17.5t^2 - 7.5t$$

$$P_4(t) = P_3(t) + a_4(t - t_0)(t - t_1)(t - t_2)(t - t_3) = -5t^3 + 17.5t^2 - 6t + 3.0417(t - 0)(t - 1)(t - 2)(t - 3)$$

$P_4(t) = 3.04167t^4 - 23.25t^3 + 50.95833t^2 - 25.75t$ C'est le même polynôme que celui de Lagrange.

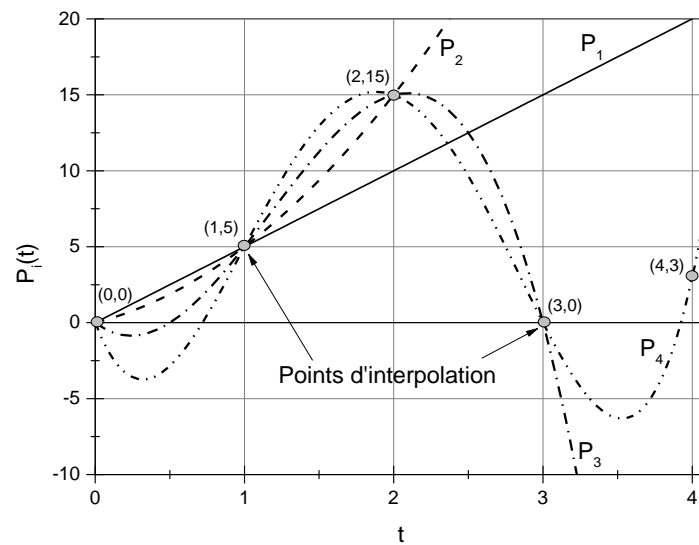


Fig. 2.1. Tracés des polynômes d'interpolation de Newton

Chapitre III : Intégration Numérique

Dans ce chapitre, on va étudier quelques méthodes approximatives pour le calcul des intégrales limitées. Aussi ces méthodes permettent le calcul des intégrales qui n'ont pas de solutions directes ou analytiques. On peut aussi calculer l'intégrale d'une fonction donnée sous forme tabulaire ou discrète.

3.1 Formule du trapèze

Cette formule est très simple, elle permet de remplacer la courbe $f(x)$ de la fonction à intégrer par une ligne droite qui relie les points $(a, f(a))$ et $(b, f(b))$ ce qui donne un trapèze (Fig. 3.1).

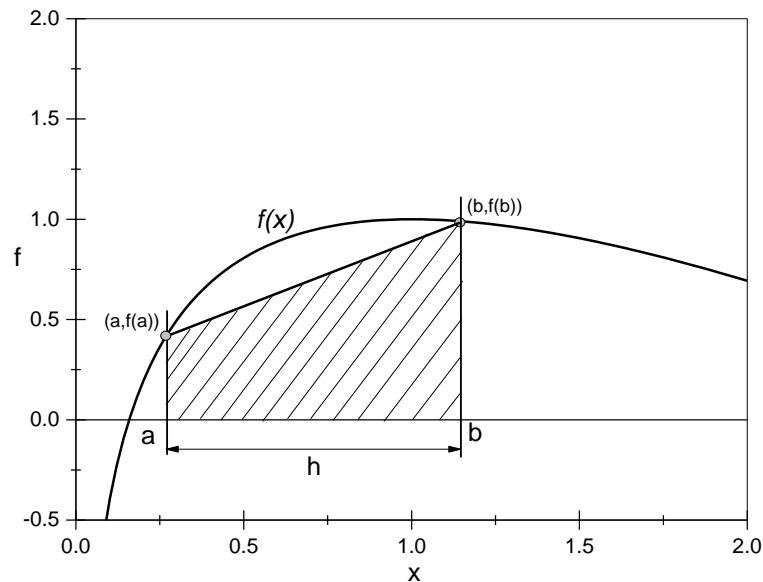


Fig. 3.1 : Méthode du trapèze

L'intégrale est donc remplacée par la surface du trapèze :

$$s = \int_a^b f(x) \approx \frac{h}{2} (f(a) + f(b))$$

Avec $h=b-a$ est dit pas d'intégration. On peut remarquer qu'il y'a une différence importante entre la courbe de la fonction et la ligne droite, cela veut qu'on commît une erreur de calcul. Pour minimiser cette erreur, on utilise une autre forme plus adaptée de cette formule.

3.1.1 Formule du trapèze généralisée.

On divise l'intervalle $[a, b]$ en plusieurs sous intervalles égaux et on applique la formule du trapèze à chaque sous intervalle (Fig. 3.2). On a donc les sous intervalles $[a = x_0, x_1] \cup [x_1, x_2] \cup \dots \cup [x_{n-1}, x_n = b]$, l'application de la formule du trapèze donne :

$$\int_a^b f(x) = \frac{h}{2}(f(x_0) + f(x_1)) + \frac{h}{2}(f(x_1) + f(x_2)) + \frac{h}{2}(f(x_2) + f(x_3)) + \dots + \frac{h}{2}(f(x_{n-1}) + f(x_n))$$

$$\int_a^b f(x) = \frac{h}{2} \left(f(x_0) + 2 \sum_{i=1}^{n-1} f(x_i) + f(x_n) \right) = \frac{h}{2} \left(f_0 + 2 \sum_{i=1}^{n-1} f_i + f_n \right)$$

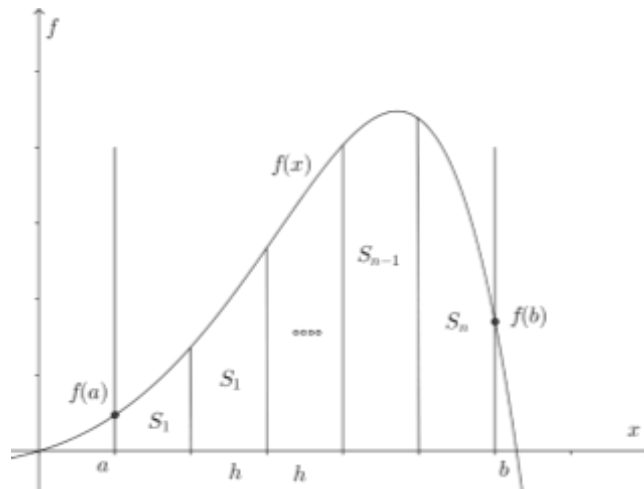


Fig. 3.2. Méthode du trapèze généralisée

3.1.2 Erreur d'intégration

C'est la différence entre l'intégrale exacte de la fonction et celle calculée par la méthode du trapèze, elle est notée par $R(f)$.

$$R(f) = \int_{x_0}^{x_n} f(x) dx - \frac{h}{2} (f_0 + 2 \sum_{i=1}^{n-1} f_i + f_n) = -\frac{b-a}{12} h^2 f''(z) \text{ avec } z \in [a, b]$$

Exemple : Soit à calculer l'intégrale $\int_0^1 e^{-x^2} dx$ avec une précision de 0.001 par la méthode du trapèze. On doit premièrement trouver le nombre de division à faire pour obtenir cette précision. L'erreur d'intégration s'écrit $R(f) = -\frac{b-a}{12} h^2 f''(z)$ sa valeur absolue doit être inférieure ou égale à la précision donnée (0.001), c.à.d. :

$$|R(f)| = \left| -\frac{b-a}{12} h^2 f''(z) \right| \leq 0.001$$

La fonction $f(x) = e^{-x^2}$ donc sa dérivée seconde est $f''(x) = 2(2x^2 - 1)e^{-x^2}$, cette fonction est strictement croissante dans l'intervalle donné (Fig. 3.3).

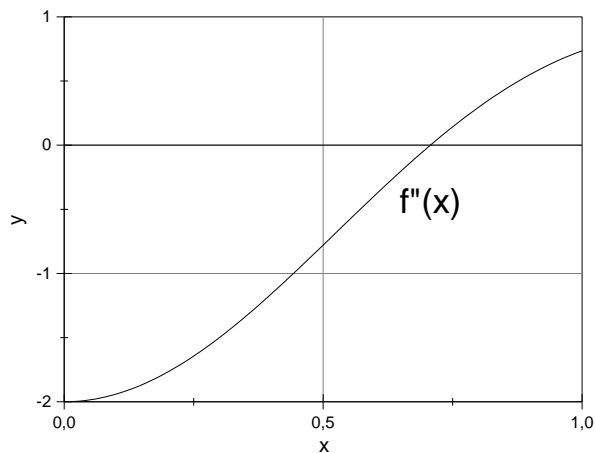


Fig. 3.3. Courbe de la dérivée seconde de e^{-x^2}

On calcule

$$M = \max |f''(z)| = 2 \text{ à } x = 0$$

$$\text{donc } |R(f)| = \left| -\frac{b-a}{12} h^2 f''(z) \right| \leq 0.001$$

$$\text{d'où } h \leq \sqrt{\frac{12 \cdot 0.001}{(1-0) \cdot 2}} = 0.0774 \text{ donc } n = \frac{1}{0.0774} = 12.91 \text{ on prend 13 divisions.}$$

Le pas d'intégration $h = \frac{1}{13}$.

$$\int_0^1 e^{-x^2} dx = \frac{1}{2 \cdot 13} \left(e^{-0^2} + 2 \sum_{i=1}^{12} e^{-\left(\frac{i}{13}\right)^2} + e^{-1^2} \right) = 0.74646$$

3.2 Formules de Simpson

Dans cette formule on ne remplace pas la fonction par une droite mais par une parabole de degré n inférieure ou égale à deux. Cette dernière doit passer par trois points (x_0, y_0) , (x_1, y_1) et (x_2, y_2) ce qui fait que cette méthode n'est applicable que pour un nombre pair de tranches (une tranche c'est l'intervalle entre deux points) (Fig. 3.4.). La formule de Simpson s'écrit :

$$\int_a^b f(x) \cong \frac{h}{3} (f(x_0) + 4f(x_1) + f(x_2))$$

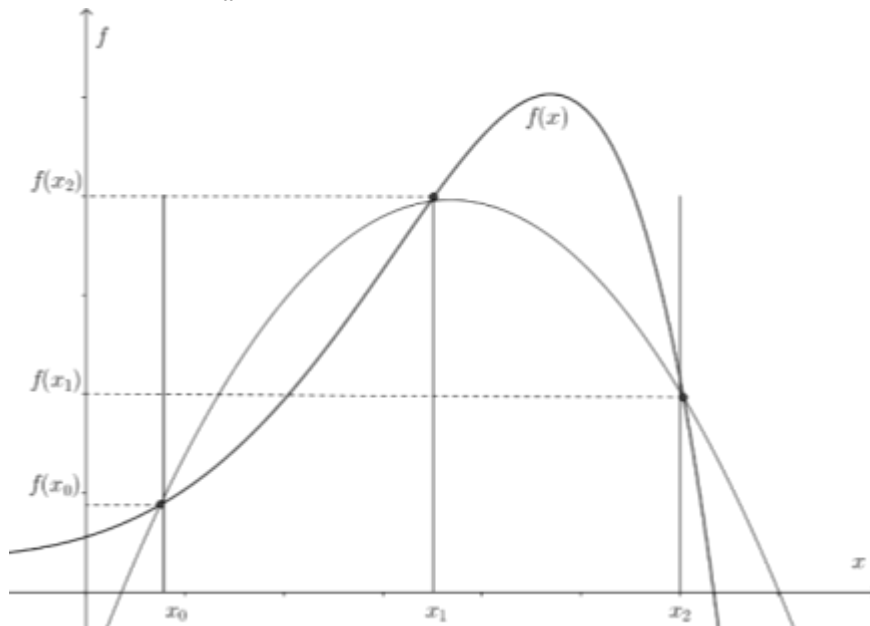


Fig. 3.4. Méthode de Simpson

Si on généralise la formule de Simpson pour $2n$ sous intervalles avec un pas d'intégration $h = \frac{b-a}{2n}$, $a = x_0 < x_1 < \dots < x_{2n} = b$ et $x_k = a + hk$ pour $k=0,1,2,\dots,2n$.

La formule de Simpson généralisée s'écrit :

$$\int_a^b f(x) \cong \frac{h}{3} \left(f(x_0) + 2 \sum_{i \text{ pair}} f(x_i) + 4 \sum_{i \text{ impair}} f(x_i) + f(x_{2n}) \right)$$

L'erreur d'interpolation de la formule de Simpson s'écrit :

$$R(f) = -\frac{b-a}{180} h^4 f^{(4)}(z) \quad \text{avec } z \in [a, b]$$

Exemple : Soit à calculer l'intégrale $\int_0^1 e^{-x^2} dx$ avec une précision de 0.001 par la méthode de Simpson. On doit premièrement trouver le nombre de division à faire pour obtenir cette précision.

L'erreur d'intégration s'écrit :

$$R(f) = -\frac{(b-a)}{180} h^4 f^{(4)}(z)$$

sa valeur absolue doit être inférieure ou égale à la précision donnée (0.001), c.à.d. :

$$|R(f)| = \left| -\frac{(b-a)}{180} h^4 f^{(4)}(z) \right| \leq 0.001$$

La fonction $f(x) = e^{-x^2}$ sa dérivée quatrième est $f^{(4)}(x) = (16x^4 - 48x^2 + 12)e^{-x^2}$, cette fonction est non monotone dans l'intervalle donné. On calcule son maximum par le traceur Origin (Fig. 3.3.).

$$M = \max |f^{(4)}(z)| = 12 \text{ à } x = 0.$$

D'ou
$$h \leq \sqrt[4]{\frac{180 \cdot 0.001}{(1-0) \cdot 12}} = 0.35$$

donc $2n = \frac{1}{0.35} = 2.85$ on prend 4 divisions, le pas d'intégration $h = \frac{1}{4} = 0.25$.

On trouve : $\int_0^1 e^{-x^2} dx = \frac{0.25}{3} (e^{-0^2} + 4(e^{-0.25^2} + e^{-0.75^2}) + 2e^{-0.5^2} + e^{-1^2}) = 0.7469$

3.3 Méthode de quadrature

Cette méthode permet l'élaboration de formules d'intégration numériques en se basant sur le polynôme de Lagrange. On peut par exemple retrouver les formules du trapèze et de Simpson par cette méthode ou bien de construire d'autres formules d'intégration plus performantes. Dans cette méthode on remplace la fonction par un polynôme de Lagrange puis on intègre le polynôme trouvé, on écrit donc :

$$f(x) \cong P_n(x) = \sum_{i=0}^n f_i L_i(x)$$

On intègre :

$$\int_a^b f(x) dx \cong \int_a^b P_n(x) dx = \sum_{i=0}^n f_i \int_a^b L_i(x) dx$$

Si on pose :

$$A_i = \int_a^b L_i(x) dx \text{ pour } i=0,1,\dots,n$$

On obtient la formule de quadrature :

$$\int_a^b f(x) dx \cong \sum_{i=0}^n f_i A_i$$

On doit maintenant choisir la forme de la fonction $f(x)$, dans notre cas on prend :

$$f(x) = x^k \text{ avec } k=0,1,2,\dots,n$$

On remplace dans l'intégrale :

$$\int_a^b x^k dx \cong \sum_{i=0}^n f_i A_i = \frac{b^{k+1}-a^{k+1}}{k+1} \quad k=0,1,2,\dots,n$$

En variant i de 0 à n pour chaque valeur de k , on aura:

pour $k=0$:

$$x_0^0 A_0 + x_1^0 A_1 + \dots + x_n^0 A_n = \frac{b^1-a^1}{1}$$

pour $k=1$:

$$x_0^1 A_0 + x_1^1 A_1 + \dots + x_n^1 A_n = \frac{b^2-a^2}{2}$$

pour $k=2$:

$$x_0^2 A_0 + x_1^2 A_1 + \dots + x_n^2 A_n = \frac{b^3-a^3}{3}$$

.....

pour $k=n$:

$$x_0^n A_0 + x_1^n A_1 + \dots + x_n^n A_n = \frac{b^{n+1}-a^{n+1}}{n+1}$$

D'où on obtient pour toutes les valeurs de i et k le système suivant :

$$\begin{bmatrix} 1 & 1 & 1 & \dots & \dots & 1 \\ x_0 & x_1 & x_2 & \dots & \dots & x_n \\ x_0^2 & x_1^2 & x_2^2 & \dots & \dots & x_n^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_0^n & x_1^n & x_2^n & \dots & \dots & x_n^n \end{bmatrix} \begin{bmatrix} A_0 \\ A_1 \\ A_2 \\ \dots \\ A_n \end{bmatrix} = \begin{bmatrix} I_0 \\ I_1 \\ I_2 \\ \dots \\ I_n \end{bmatrix} \text{ avec } I_k = \frac{b^{k+1}-a^{k+1}}{k+1}$$

Le déterminant de la matrice du système trouvé est dit de "Von-Dermonde" il est non nul d'où la solution de ce système existe et unique (A_0, A_1, \dots, A_n) .

Exemple : Soit à calculer l'intégrale $\int_0^1 e^{-x^2} dx$ avec une formule qui a la forme suivante : $\int_0^1 e^{-x^2} dx = A_0 f(0) + A_1 f(0.25) + A_2 f(0.5) + A_3 f(0.75) + A_4 f(1.00)$. Dans ce cas on cherche premièrement les constantes A_i puis on calcule l'intégrale. Ces constantes sont données par le système d'équations suivant :

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 0.25 & 0.5 & 0.75 & 1 \\ 0 & 0.0625 & 0.25 & 0.5626 & 1 \\ 0 & 0.015625 & 0.125 & 0.421875 & 1 \\ 0 & 0.00390625 & 0.0625 & 0.31640625 & 1 \end{bmatrix} \begin{bmatrix} A_0 \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.5 \\ 0.333 \\ 0.250 \\ 0.200 \end{bmatrix}$$

La solution du système est (0.0691, 0.3812, 0.1052, 0.3694, 0.0751) donc :

$$\int_0^1 e^{-x^2} dx = 0.0691e^0 + 0.3812e^{-0.25^2} + 0.1052e^{-0.5^2} + 0.3694e^{-0.75^2} + 0.0751e^{-1^2} = 0.7472$$

Chapitre VI : Résolution numérique des équations différentielles ordinaires : Problème de Cauchy

Les équations différentielles sont utilisées dans la modélisation mathématique de la quasi- totalité des phénomènes physiques. Une équation différentielle est une relation entre une variable et sa ou ses dérivées de divers ordres. Dans ce chapitre, on va considérer les équations différentielles ordinaires du premier ordre avec une condition initiale imposée (problème de Cauchy).

Le problème de Cauchy est défini par la solution de l'équation différentielle suivante :

$$\begin{cases} \frac{dy(t)}{dt} = y' = f(t, y(t)) \\ y(t_0) = y_0 \end{cases} \quad y(t_0) = y_0 \text{ est une condition initiale imposée.}$$

On remarque que la solution doit obligatoirement passer par le point (t_0, y_0) .

La solution du problème de Cauchy existe et unique si la fonction $f(t, y(t))$ satisfait la condition de Lipschitz en y sur un rectangle R défini par $a \leq t \leq b$ et $c \leq y \leq d$. Cette condition nécessite que $|f(t, y_1) - f(t, y_2)| \leq L|y_1 - y_2|$.

En pratique pour vérifier cette condition, on calcule $\text{Max} \left| \frac{\partial f(t, y)}{\partial y} \right| \leq L$ sur R

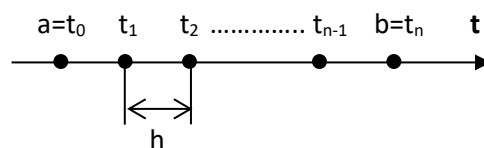
4.1 Méthode d'Euler

Soit un intervalle $[a, b]$ sur lequel on cherche la solution d'un problème de Cauchy

$$\begin{cases} \frac{dy(t)}{dt} = y' = f(t, y(t)) \\ y(t_0 = a) = y_0 \end{cases}$$

La fonction $f(t, y(t))$ vérifie la condition de Lipschitz en y sur le rectangle R .

La première étape consiste à diviser l'intervalle donné en n points équidistants ce qui donne un pas d'intégration $h = \frac{b-a}{n}$, le point d'abscisse t_i est donné par $t_i = a + ih$ (pour $i=1, 2, \dots, n$).



On va donc résoudre le problème sur l'intervalle $[a = t_0, b = t_n]$ avec $y(t_0 = a) = y_0$. Si les fonctions $y(t), y'(t)$ et $y''(t)$ sont continues, on peut écrire le développement en série de Taylor pour $y(t)$ au voisinage de t_0 .

$$y(t) = y(t_0) + y'(t_0) \frac{(t - t_0)}{1!} + y''(t_0) \frac{(t - t_0)^2}{2!} + \dots$$

Sachant que $y'(t_0) = f(t_0, y(t_0))$ et $h = t_1 - t_0$,

Ecrivons
$$y(t_1) = y(t_0) + f(t_0, y(t_0))h + y''(t_0) \frac{h^2}{2!} + \dots$$

Si h est suffisamment petit alors on peut négliger h^2 et donc les termes d'ordre deux et plus, on obtient donc :

$$y(t_1) = y(t_0) + hf(t_0, y(t_0)) \rightarrow y_1 = y_0 + hf(t_0, y_0)$$

C'est l'approximation d'Euler d'ordre 1, en répétant le procédé on génère une séquence de points $y_1, y_2, y_3, \dots, y_{n-1}, y_n$ approximant $y = y(t)$, en général :

$$\begin{cases} t_i = a + ih, & i = 0, 1, 2, \dots, n-1 \\ y_{i+1} = y_i + hf(t_i, y_i), & y_0 = y(t_0) \end{cases}$$

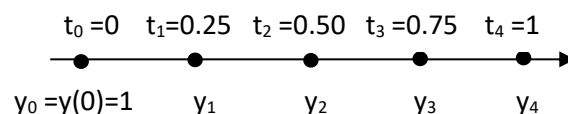
Exemple : Résoudre le problème de Cauchy suivant par la méthode d'Euler en prenant un pas d'intégration $h=0.25$.

$$\begin{cases} y' = 2 - ty^2 & t \in [0,1], y \in [0,1] \\ y(0) = 1 \end{cases}$$

Vérifions la condition de Lipchitz sur le rectangle R défini par $t \in [0,1], y \in [0,1]$.

$$\text{Max} \left| \frac{\partial f(t,y)}{\partial y} \right| = \text{Max} \left| \frac{\partial (2-ty^2)}{\partial y} \right| = |-2ty| = 2 < L \text{ Condition vérifiée}$$

On divise l'intervalle de t , $[0,1]$ avec une pas $h=0.25$, soit 0, 0.25, 0.50, 0.75 et 1.



On écrit :
$$\begin{cases} t_i = ih, & i = 0, 1, 2 \text{ et } 3 \\ y_{i+1} = y_i + hf(t_i, y_i) = y_i + 0.25(2 - t_i y_i^2), & y_0 = 1 \end{cases}$$

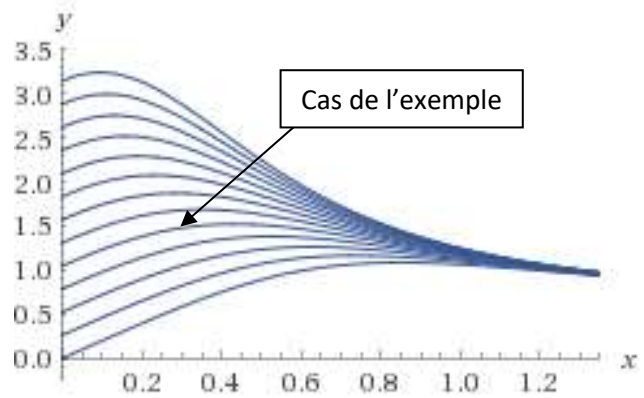
$$\begin{aligned}
 i = 0, \quad y_1 &= y_0 + hf(t_0, y_0) = 1 + 0.25(2 - 0 * 1^2) &= 1.5 \\
 i = 1, \quad y_2 &= y_1 + hf(t_1, y_1) = 1.5 + 0.25(2 - 0.25 * 1.5^2) &= 1.8594 \\
 i = 2, \quad y_3 &= y_2 + hf(t_2, y_2) = 1.859 + 0.25(2 - 0.5 * 1.859^2) &= 1.9272 \\
 i = 3, \quad y_4 &= y_3 + hf(t_3, y_3) = 1.927 + 0.25(2 - 0.75 * 1.927^2) &= 1.7308
 \end{aligned}$$

La solution exacte est :

$$y(t) = \frac{2^{\frac{3}{2}} c_1 t Ai(2^{\frac{2}{3}} t) + 2^{\frac{3}{2}} t Bi(2^{\frac{2}{3}} t)}{2t(c_1 Ai'(2^{\frac{2}{3}} t) + Bi'(2^{\frac{2}{3}} t))}$$

Avec **Ai** et **Bi** sont les fonctions d'Airy et d'Airy de seconde espèce.

Fig. 4.1. Solutions **y(t)** pour différentes conditions initiales **y(0)**.



4.2 Méthode d'Euler améliorée(Huns)

D'après le développement en série de Taylor, on a :

$$y(t) = y(t_0) + y'(t_0) \frac{(t-t_0)}{1!} + y''(t_0) \frac{(t-t_0)^2}{2!} + y'''(t_0) \frac{(t-t_0)^3}{3!} + \dots$$

Si on prend les trois premiers termes et on néglige les autres d'ordre supérieur, on obtient :

$$y(t) \cong y(t_0) + y'(t_0) \frac{(t-t_0)}{1!} + y''(t_0) \frac{(t-t_0)^2}{2!} + \dots$$

En remplaçant t par t_1 cela donne :

$$y(t_1) = y(t_0) + y'(t_0) \frac{(t_1-t_0)}{1!} + y''(t_0) \frac{(t_1-t_0)^2}{2!} = y(t_0) + y'(t_0)h + y''(t_0) \frac{h^2}{2!}$$

Or $y''(t_0) = \frac{y'(t_1)-y'(t_0)}{h}$ on aura $y(t_1) = y(t_0) + \frac{h}{2}[f(t_1, y_1) + f(t_0, y_0)]$

En général la formule d'Euler améliorée s'écrit :

$$y_{i+1} = y_i + \frac{h}{2}[f(t_{i+1}, y_{i+1}^E) + f(t_i, y_i)]$$

On remarque que cette formule donne y_{i+1} en fonction de y_{i+1}^E qui doit être calculée par la méthode d'Euler.

Exemple : Résoudre le problème de Cauchy précédant par la méthode d'Euler améliorée en prenant un pas d'intégration $h=0.25$.

On a :

$$\begin{cases} t_i = ih, & i = 0, 1, 2 \text{ et } 3 \\ y_{i+1} = y_i + \frac{h}{2}[f(t_{i+1}, y_{i+1}^E) + f(t_i, y_i)] = y_i + \frac{h}{2}[(2-t_i y_i^2) + (2-t_{i+1} y_{i+1}^E)], & y_0 = 1 \end{cases}$$

Avec : $y_{i+1}^E = y_i + h(2-t_i y_i^2)$ la solution obtenue par la méthode d'Euler. On

remplace y_{i+1}^E , on aura : $y_{i+1} = y_i + \frac{h}{2}[(2-t_i y_i^2) + (2-t_{i+1} (y_i + h(2-t_i y_i^2))^2)]$,

$y_1 = 1.4297, y_2 = 1.6629, y_3 = 1.6805, y_4 = 1.5750$

4.3 Méthode de Runge-Kutta d'ordre 4

C'est la méthode la plus précise et la plus utilisée, elle est d'ordre quatre. L'intervalle $[a, b]$ est divisé en n sous intervalles de largeur h , cette formule s'écrit :

$$y_{i+1} = y_i + \frac{h}{6}(K_1 + 2(K_2 + K_3) + K_4)$$

$$K_1 = f(t_i, y_i)$$

$$K_2 = f\left(t_i + \frac{h}{2}, y_i + \frac{h}{2}K_1\right)$$

$$K_3 = f\left(t_i + \frac{h}{2}, y_i + \frac{h}{2}K_2\right)$$

$$K_4 = f(t_i + h, y_i + hK_3)$$

Exemple : Résoudre le problème de Cauchy précédant par la méthode de Runge-Kutta d'ordre 4 en prenant un pas d'intégration $h=0.25$.

$$y_{i+1} = y_i + \frac{h}{6}(K_1 + 2(K_2 + K_3) + K_4)$$

$$K_1 = f(t_i, y_i) = 2 - t_i y_i^2$$

$$K_2 = f\left(t_i + \frac{h}{2}, y_i + \frac{h}{2}K_1\right) = 2 - \left(t_i + \frac{h}{2}\right)\left(y_i + \frac{h}{2}K_1\right)^2$$

$$K_3 = f\left(t_i + \frac{h}{2}, y_i + \frac{h}{2}K_2\right) = 2 - \left(t_i + \frac{h}{2}\right)\left(y_i + \frac{h}{2}K_2\right)^2$$

$$K_4 = f(t_i + h, y_i + hK_3) = 2 - (t_i + h)(y_i + hK_3)^2$$

$$i=0, K_1 = 2, K_2 = 1.8047, K_3 = 1.8122, K_4 = 1.4722, y_1 = 1.4461$$

$$y_2=1.7028, y_3=1.7317, y_4=1.6148$$

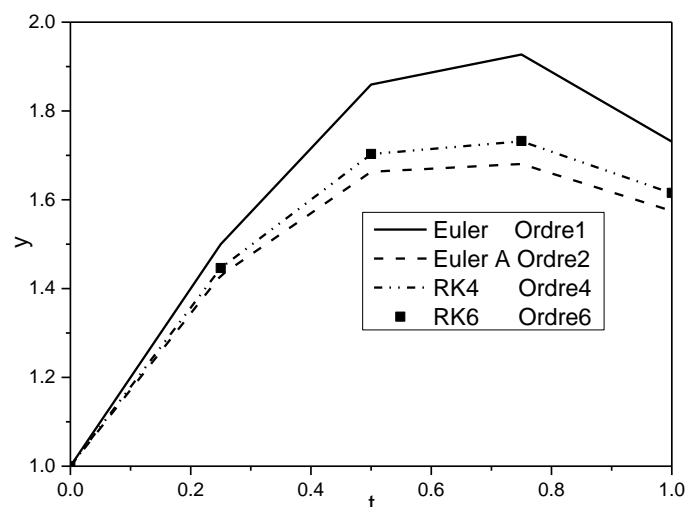


Fig. 4.2. Comparaison entre différentes méthodes

Pour le calcul de x_i , les composantes $x_n, x_{n-1}, x_{n-2}, \dots, x_{i+1}$ sont connues, on les remplace dans l'équation i , ce qui donne :

$$x_i = (y_i - \sum_{j=i+1}^n u_{ij}x_j)/u_{ii} \quad i=n-1, n-2, \dots, 1.$$

Le déterminant d'une matrice triangulaire (supérieure ou inférieure) est le produit des éléments de la diagonale (du pivot) de cette matrice.

$$\det(U) = \prod_{i=1}^n u_{ii} = u_{11}u_{22} \dots u_{nn}$$

Exemple : Soit le système d'équations suivant :

$$\begin{cases} x - y + z = 1 \\ 3z = 3 \\ 2y - z = 1 \end{cases}$$

1. Ecrire le système sous la forme matricielle.
2. Trouver la solution du système.
3. Calculer le déterminant de la matrice du système.

Solution

1. Ecriture du système sous la forme matricielle en permutant les lignes 2 et 3 pour avoir un système triangulaire supérieur. N'oublions pas que le déterminant sera multiplié par -1 à cause de la permutation.

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$$

2. La substitution en arrière donne :

$$eq. 3 : z = \frac{3}{3} = 1, \quad eq. 2: y = \frac{1+1}{2} = 1 \quad et \quad x = 1 + 1 - 1 = 1$$

3. $\det(A) = (-1) * 1 * 2 * 3 = -6.$

5.1 Méthode d'élimination de GAUSS

Dans ce chapitre, on va commencer par la méthode d'élimination de GAUSS qui est similaire à celle de l'élimination avec une modification très importante qui facilite énormément la résolution même de systèmes à grandes tailles (des milliers où même des millions). Cette modification transforme un système à matrice A pleine, en un autre à système à matrice U triangulaire supérieure, de telle façon que les deux systèmes $AX=B$ et $UX=Y$ soient équivalents c. à. d. ont la même solution.

L'algèbre linéaire montre que certaines transformations apportées aux systèmes d'équations ne changent pas leurs solutions, dans notre cas les opérations qu'on va appliquer sont les suivantes :

- Multiplication de l'équation E_i par une constante α non nulle, la nouvelle équation obtenue $E_{in} = \alpha E_i$ remplacera l'ancienne E_i .
- Multiplication de l'équation E_j par α non nulle et son ajout à E_i , $E_{in} = E_i + \alpha E_j$, l'équation obtenue E_{in} remplacera E_i .
- Permutation des équations E_i et E_j .

L'application d'une série de ces opérations transformera le système $AX=B$ en $UX=Y$ puis une substitution en arrière donnera la solution du système.

5.1.1 Description de la méthode d'élimination de GAUSS

On va montrer comment appliquer les transformations au système $\mathbf{AX}=\mathbf{B}$, pour cela le second membre \mathbf{B} sera considéré comme la colonne $(\mathbf{n}+1)$ et sera aussi affecté par les opérations. On divise le travail en $(\mathbf{n}-1)$ étapes chacune d'elle annule les éléments au-dessous du pivot de la colonne $(a_{ij}$ pour $i > j$). Au début chaque étape, on vérifie que le pivot est non nul. Pour l'étape i le pivot est $a_{ii}^{(i-1)}$ à l'étape $(i-1)$.

Le système à l'état initial où à l'étape (0) est donné par :

$$\begin{bmatrix} a_{11} & a_{12} & a_{1n} \\ a_{21} & a_{22} & a_{2n} \\ \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{nn} \end{bmatrix}^{(0)} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}^{(0)} = \begin{bmatrix} a_{1n+1} \\ a_{2n+1} \\ \vdots \\ a_{nn+1} \end{bmatrix}^{(0)}$$

Première étape : On vérifie tout d'abord que le pivot de la première étape qui est $a_{ii}^{(i-1)} \neq 0$ pour $i = 1$ i.e $a_{11}^{(0)} \neq 0$.

Pour annuler l'élément $a_{21}^{(0)}$ de la deuxième ligne, on multiplie la première équation par $a_{21}^{(0)}$ et on la divise par $a_{11}^{(0)}$ puis on fait la différence de cette nouvelle équation avec la deuxième. L'équation obtenue remplacera la deuxième. $E_2^{(1)} = E_2^{(0)} - E_1^{(0)} \frac{a_{21}^{(0)}}{a_{11}^{(0)}}$

Cette opération donne $a_{21}^{(1)} = 0$, $a_{22}^{(1)} = a_{22}^{(0)} - a_{12}^{(0)} \frac{a_{21}^{(0)}}{a_{11}^{(0)}}$, ..., $a_{2n+1}^{(1)} = a_{2n+1}^{(0)} - a_{1n+1}^{(0)} \frac{a_{21}^{(0)}}{a_{11}^{(0)}}$

En général, on écrit : $a_{2j}^{(1)} = a_{2j}^{(0)} - a_{1j}^{(0)} \frac{a_{21}^{(0)}}{a_{11}^{(0)}}$ Pour $j = 2, n + 1$

On continue cette procédure avec les lignes 3,4,..., pour la ligne i on a :

$$E_i^{(1)} = E_i^{(0)} - E_1^{(0)} \frac{a_{i1}^{(0)}}{a_{11}^{(0)}}$$

Cette opération donne $a_{i1}^{(1)} = 0$, $a_{i2}^{(1)} = a_{i2}^{(0)} - a_{i2}^{(0)} \frac{a_{i1}^{(0)}}{a_{11}^{(0)}}$, \dots , $a_{in+1}^{(1)} = a_{in+1}^{(0)} - a_{1n+1}^{(0)} \frac{a_{i1}^{(0)}}{a_{11}^{(0)}}$

En général, on écrit : $a_{ij}^{(1)} = a_{ij}^{(0)} - a_{1j}^{(0)} \frac{a_{i1}^{(0)}}{a_{11}^{(0)}}$ Pour $i = 2, n$ et $j = 2, n + 1$

A la fin de la première étape, on obtient des éléments nuls au-dessous du pivot de la première étape. Le système s'écrit :

$$\begin{bmatrix} a_{11}^{(0)} & a_{12}^{(0)} & a_{1n}^{(0)} \\ 0 & a_{22}^{(1)} & a_{2n}^{(1)} \\ \vdots & \vdots & \vdots \\ 0 & a_{n2}^{(1)} & a_{nn}^{(1)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{1n+1}^{(0)} \\ a_{2n+1}^{(1)} \\ \vdots \\ a_{nn+1}^{(1)} \end{bmatrix}$$

Deuxième étape : $a_{ii}^{(i-1)} \neq 0$ pour $i = 2$ i.e $a_{22}^{(1)} \neq 0$.

De la même façon, on obtient pour le cas général

$$a_{ij}^{(2)} = a_{ij}^{(1)} - a_{2j}^{(1)} \frac{a_{i2}^{(1)}}{a_{22}^{(1)}} \text{ Pour } i = 3, n \text{ et } j = 3, n + 1$$

Etape k : $a_{ii}^{(i-1)} \neq 0$ pour $i = k$ i.e $a_{kk}^{(k-1)} \neq 0$.

Pour une étape k quelconque, on a :

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - a_{kj}^{(k-1)} \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}} \text{ Pour } k = \overline{1, n-1}, \quad i = \overline{k+1, n} \text{ et } j = \overline{k+1, n+1}$$

A la fin de la procédure, on obtient un système à matrice triangulaire supérieure qui s'écrit :

$$\begin{bmatrix} a_{11}^{(0)} & a_{12}^{(0)} & a_{1n}^{(0)} \\ 0 & a_{22}^{(1)} & a_{2n}^{(1)} \\ \vdots & \vdots & \vdots \\ 0 & 0 & a_{nn}^{(n-1)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{1n+1}^{(0)} \\ a_{2n+1}^{(1)} \\ \vdots \\ a_{nn+1}^{(n-1)} \end{bmatrix}$$

La résolution de ce système se fait par substitution en arrière.

5.1.2 Le nombre d'opérations nécessaires pour l'application de l'algorithme de

GAUSS est :

Nombre de multiplications et d'additions :

$$nm = na = \frac{n(n-1)(2n+5)}{6}$$

Nombre de divisions :

$$nd = \frac{n(n+1)}{2}$$

5.1.3 Applications de la méthode de GAUSS.**a) Calcul du déterminant d'une matrice triangulaire :**

Le déterminant d'une matrice triangulaire est donné par :

$$\det(U) = (-1)^P \prod_{i=1}^n u_{ii} = (-1)^P u_{11} u_{22} \dots u_{nn}$$

Avec P le nombre de permutation de lignes ou de colonnes effectuées lors de l'application de l'algorithme de GAUSS.

Exemple : Soit le système d'équations suivant :

$$\begin{cases} 2x + y + 2z = 10 \\ 3x + 5y + z = 16 \\ -x + 4y + 7z = 28 \end{cases}$$

1. Calculer le nombre d'opérations élémentaires pour la méthode de Gauss.
2. Calculer le déterminant de la matrice du système.
3. Résoudre le système par l'élimination de Gauss.
4. Recalculer le déterminant de la matrice du système.

b) Résolution simultanée de plusieurs systèmes à même matrice A :

Dans la pratique, on rencontre souvent le cas de plusieurs systèmes d'équations qui ne diffèrent que par le second membre B. On peut appliquer l'algorithme de GAUSS sur la matrice A augmentée par tous les seconds membres. De cette façon, on fait les calculs une seule fois sur la matrice A, la substitution se fait avec chaque second membre obtenu à part.

$$\begin{bmatrix} a_{11} & a_{12} & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} b_1 & c_1 & z_1 \\ b_2 & c_2 & \dots & z_2 \\ \vdots & \vdots & & \vdots \\ b_n & c_n & \dots & z_n \end{bmatrix}$$

c) Calcul de l'inverse d'une matrice :

Si **A** est une matrice d'ordre **n**, la matrice **A⁻¹** tel que **A.A⁻¹ = I** est dite matrice inverse de **A**.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{1n} \\ a_{21} & a_{22} & \dots a_{2n} \\ \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots a_{nn} \end{bmatrix} \text{ et } A^{-1} = \begin{bmatrix} x_{11} & x_{12} & x_{1n} \\ x_{21} & x_{22} & \dots x_{2n} \\ \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \dots x_{nn} \end{bmatrix}$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{1n} \\ a_{21} & a_{22} & \dots a_{2n} \\ \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots a_{nn} \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} & x_{1n} \\ x_{21} & x_{22} & \dots x_{2n} \\ \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \dots x_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & \dots 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & \dots 1 \end{bmatrix}$$

Exemple : Soit la matrice suivante :

$$\begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 1 \end{bmatrix}$$

1. Utiliser la méthode de Gauss pour calculer la matrice inverse.
2. Calculer le nombre d'opérations.

5.2. Utilisation de la pivotation.

Dans l'élaboration de l'algorithme de GAUSS, on a supposé que le pivot ne soit pas nul, ce n'est pas le cas toujours. Parfois le pivot est très petit comparativement aux autres termes ou même nul, dans ce cas on peut utiliser la technique de la pivotation soit partielle ou totale.

Pivotation partielle :

Dans ce cas on choisit comme pivot l'élément $a_{ik}^{(k-1)}$ tel que :

$$a_{ik}^{(k-1)} = \max_{i \in [k, n]} |a_{ik}^{(k-1)}|$$

$$\begin{bmatrix} a_{11} & \dots & a_{1k} & a_{1n} \\ 0 & \dots & \mathbf{a_{kk}} & a_{kn} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \mathbf{a_{nk}} & a_{nn} \end{bmatrix}$$

Dans la pivotation partielle, on utilise la permutation des lignes ceci n'a aucun effet sur la solution du système.

Pivotation totale :

Dans la pivotation totale le choix du pivot se fait à partir d'une sous matrice incluant la permutation des lignes et des colonnes tel que :

$$a_{lm}^{(k-1)} = \max_{i, j \in [k, n]} |a_{ij}^{(k-1)}|$$

On a donc $AX=B$ (1) et $A=LU$ donc $LUX=B$, on pose $UX=Y$ (Y vecteur inconnu), cela donne :

$$\begin{cases} LY = B \\ UX = Y \end{cases} \quad (2)$$

Le système (1) est décomposé en deux systèmes triangulaires faciles à résoudre (2). Le système à matrice triangulaire supérieure est résolu par substitution en arrière, celui à matrice triangulaire inférieure par substitution en avant.

5.3.1 Détermination des matrices L et U

$$L = \begin{bmatrix} l_{11} & 0 & \dots & 0 \\ l_{21} & l_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n-1,1} & l_{n-1,2} & \dots & l_{n-1,n-1} \\ l_{n1} & l_{n2} & \dots & l_{nn} \end{bmatrix} \quad \text{et} \quad U = \begin{bmatrix} 1 & u_{12} & \dots & u_{1n-1} & u_{1n} \\ 0 & 1 & \dots & u_{2n-2} & u_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 1 & u_{n-1n} \\ 0 & \dots & 0 & 1 \end{bmatrix}$$

Les éléments de chaque matrice sont donnés par :

$$\begin{cases} l_{ki} = a_{ki} - \sum_{j=1}^{i-1} l_{kj}u_{ji} \\ u_{ik} = [a_{ik} - \sum_{j=1}^{i-1} l_{ij}u_{jk}] / l_{ii} \end{cases} \quad i = 2, 3, \dots, n \quad \text{et} \quad k = i, i+1, \dots, n$$

En pratique pour calculer les éléments des deux matrices, on divise la tâche en plusieurs étapes. Par exemple dans l'étape i on détermine :

- La colonne i de L en multipliant L par la colonne i de U .
- La ligne i de U en multipliant la ligne i de L par U .

Exemple : Utiliser la méthode de factorisation LU pour résoudre le système suivant :

$$\begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 9 \\ 8 \\ 7 \end{bmatrix}$$

5.4 Méthode de Choleski :

Cette méthode est applicable aux systèmes à matrices symétriques définies positives ($\text{Det}(A) > 0$ et a_{ij} réels). Cherchons une matrice M telle que $A = MM^t$ ou M est triangulaire inférieure et M^t la matrice transposée de M .

$$M = \begin{bmatrix} m_{11} & m_{12} & \dots & m_{1n-1} & m_{1n} \\ 0 & m_{22} & & m_{2n-2} & m_{2n} \\ \vdots & & \ddots & & \vdots \\ 0 & & \dots & m_{n-1n-1} & m_{n-1n} \\ & & & 0 & m_{nn} \end{bmatrix}$$

Les éléments de la matrice M sont donnés par :

$$\begin{cases} m_{ii} = \sqrt{a_{ii} - \sum_{j=1}^{i-1} m_{ij}^2} \\ m_{ji} = \left[a_{ij} - \sum_{k=1}^{i-1} m_{ik} m_{jk} \right] / m_{ii} \end{cases} \quad i = \overline{1, n} \text{ et } j = \overline{i+1, n}$$

Exemple : Utiliser la méthode de factorisation de Choleski pour résoudre le système suivant :

$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 4 \end{bmatrix}$$

6.4 Critère d'arrêt de calcul pour la méthode de Jacobi et Gauss-Seidel

On arrête les calculs pour cette méthode lorsque la différence absolue entre deux itérations successives soit inférieure à une certaine précision ε donnée.

$$\left| X^{(n+1)} - X^{(n)} \right| < \varepsilon \quad (6)$$

Ici, il faut vérifier la différence pour toutes les composantes une par une.

$$\left| x_1^{(n+1)} - x_1^{(n)} \right| < \varepsilon, \quad \left| x_2^{(n+1)} - x_2^{(n)} \right| < \varepsilon, \dots, \left| x_n^{(n+1)} - x_n^{(n)} \right| < \varepsilon$$