

Analyse numérique

A. Blouza

email : adel.blouza@univ-rouen.fr

bureau : M.1.29

ESITech – Université de Rouen

Plan du cours

1. Interpolation polynomiale
 - 2.1 Méthode de Lagrange
 - 2.2 Méthode de Newton
2. Intégration
 - 3.1 Rappels d'intégration
 - 3.2 Méthodes d'approximation d'une intégrale
 - 3.3 Intégrales doubles
3. Résolution numérique d'équations non linéaires
 - 4.1 Méthode de la dichotomie
 - 4.2 Méthode de point fixe
 - 4.3 Méthode de Newton
4. Résolution de systèmes linéaires
 - 5.1 Conditionnement
 - 5.2 Décomposition Cholesky, LU

Références bibliographiques

- Analyse numérique :
 - ▶ A. Fortin, *Analyse numérique pour ingénieurs*, Broché.
 - ▶ F. Filbet, *Analyse numérique - algorithmes et étude mathématique. Cours et exercices corrigés*, Dunod.
- Pour aller plus loin,
 - ▶ W. Rudin, *Principes d'analyse mathématique : cours et exercices*, Dunod.
 - ▶ M. Lefebvre, *Equations différentielles*, Collection Paramètres .
 - ▶ Allaire G. et Kaber S.M., *Algèbre linéaire numérique : Cours et exercices*, Ellipses.
 - ▶ Demailly J-P., *Analyse numérique et équations différentielles*, Presses Universitaires de Grenoble.
 - ▶ Quarteroni A., Sacco R. , Saleri F., *Méthodes numériques pour le calcul scientifique : programmes en MATLAB*, Springer.

Objectif : Approcher une fonction dont on ne connaît les valeurs qu'en certains points.

- Lorsqu'une fonction connue analytiquement est difficile à évaluer, différencier ou intégrer par ordinateur.
- Lorsque l'on dispose d'un nombre fini de valeurs obtenues expérimentalement (étalonnage en métrologie, relevé de la température d'une réaction chimique au cours du temps, ...)

Pourquoi une approximation polynomiale ?

- Toute fonction continue peut-être approchée par un polynôme,

Théorème 1 (d'approximation de Weirstrass)

Supposons que f est définie et continue sur $[a, b]$. Pour tout $\varepsilon > 0$, il existe un polynôme $P(x)$ tel que :

$$|f(x) - P(x)| < \varepsilon, \quad \forall x \in [a, b].$$

- Calculs de dérivées et d'intégrales de polynômes sont plus aisés.

Soient a un réel et f une fonction définie au voisinage de a , sauf éventuellement en a , et à valeurs dans \mathbb{R} . On note \mathcal{D}_f le domaine de définition de f .

Définition 1

On dit que f **tend vers le réel ℓ quand x tend vers a** , (ou que f a pour limite ℓ en a) si

$$\forall \epsilon > 0, \exists \eta > 0, \quad 0 < |x - a| < \eta \Rightarrow |f(x) - \ell| \leq \epsilon.$$

On note $\lim_{x \rightarrow a} f(x) = \ell$ ou $f(x) \rightarrow \ell$ quand $x \rightarrow a$.

Tout intervalle centré en ℓ contient toutes les valeurs $f(x)$, pour x suffisamment proche de a .

Proposition 1

La fonction f tend vers ℓ quand x tend vers a , si et seulement si, pour toute suite (x_n) , à valeurs dans $\mathcal{D}_f \setminus \{a\}$ et convergeant vers a , la suite $(f(x_n))$ converge vers ℓ .

Définition 2

La fonction f est **continue en a** si f admet une limite au point a et que

$$\lim_{x \rightarrow a} f(x) = f(a).$$

Autrement dit,

$$\forall \epsilon > 0 \quad \exists \eta > 0, \quad |x - a| < \eta \Rightarrow |f(x) - f(a)| \leq \epsilon.$$

Si f est continue en tout point d'un intervalle \mathcal{I} , alors on dit que f est **continue sur \mathcal{I}** .

Proposition 2

La fonction f est continue en a , si et seulement si, pour toute suite (x_n) , à valeurs dans $\mathcal{D}_f \setminus \{a\}$ et convergant vers a , la suite $(f(x_n))$ converge vers $f(a)$.

Définition 3

Soit f une fonction d'un intervalle \mathcal{I} dans \mathbb{R} . On dit que f est **dérivable au point $x \in \mathcal{I}$** si le taux d'accroissement

$$\tau_x(h) = \frac{f(x+h) - f(x)}{h},$$

admet une limite finie lorsque h tend vers 0. On appelle **nombre dérivé de f au point x** cette limite, on la note

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}.$$

On dit que f est dérivable sur \mathcal{I} si elle admet une dérivée en tout point de \mathcal{I} .

Proposition 3

Si f est dérivable en x alors f est continue en x .

REMARQUE. La réciproque est fautive en général.

Définition 4

Soit $n \in \mathbb{N}$.

- On définit, si elle existe, la dérivée n -ème de f ,

$$f^{(p)} = (f^{(p-1)})', \quad \text{pour } 1 \leq p \leq n,$$

en posant par convention $f^0 = f$.

- Si la dérivée d'ordre n de f , $f^{(n)}$ existe, on dit que f est **n fois dérivable sur \mathcal{I}** ,
- si $f^{(n)}$ est continue sur I on dit que f est de **classe C^n sur I** .
- On dit que f est de **classe C^∞ sur \mathcal{I}** si pour tout entier n , f est de classe C^n sur \mathcal{I}

Théorème 2 (de Rolle.)

Soit f une fonction continue sur $[a, b]$, dérivable sur un intervalle $]a, b[$ telle que $f(a) = f(b)$. Alors il existe au moins un $c \in]a, b[$ tel que $f'(c) = 0$.

Proposition 4 (Formule des accroissements finis)

Soit f une fonction continue sur $[a, b]$, dérivable sur $]a, b[$. Il existe $c \in]a, b[$ tel que

$$f(b) = f(a) + (b - a)f'(c).$$

On écrit parfois $c = a + \theta(b - a)$ avec $\theta \in]0, 1[$.

Corollaire 1 (Inégalité des accroissements finis)

Soit f une fonction dérivable sur un intervalle $[a, b]$, on suppose qu'il existe un réel M tel que

$$\forall x \in [a, b], \quad |f'(x)| \leq M.$$

Alors,

$$\forall x_1, x_2 \in [a, b], \quad |f(x_2) - f(x_1)| \leq M|x_2 - x_1|.$$

Formule et inégalité de Taylor-Lagrange sont des généralisations du théorème et de l'inégalité des accroissements finis.

Définition 5

Soit f une fonction définie sur un intervalle ouvert \mathcal{I} contenant un point a , dérivable $n - 1$ fois sur \mathcal{I} , et dont la dérivée n -ième en a existe. On appelle **polynôme de Taylor de degré n en a de f** , le polynôme :

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

On appelle **reste de Taylor d'ordre n en a de f** , la fonction R_n qui à $x \in \mathcal{I}$ associe :

$$R_n(x) = f(x) - P_n(x).$$

Théorème 3 (Taylor-Lagrange)

Soit f une fonction de classe C^n sur $[a, b]$, dont la dérivée $(n + 1)$ ème existe sur $]a, b[$. Il existe c_x entre a et x tel que

$$R_n(x) = |x - a|^{n+1} \frac{f^{(n+1)}(c_x)}{(n + 1)!}.$$

Corollaire 2 (Inégalité de Taylor-Lagrange)

Soit f une fonction de classe C^n sur $[a, b]$, dont la dérivée $(n + 1)$ ème existe sur $]a, b[$. On suppose qu'il existe un réel M tel que pour tout x dans $]a, b[$, $|f^{(n+1)}(x)| \leq M$. Alors,

$$\left| f(x) - \sum_{k=0}^n \frac{(x - \alpha)^k}{k!} f^{(k)}(\alpha) \right| \leq M \frac{(x - \alpha)^{n+1}}{(n + 1)!}.$$

2.2 APPROXIMATION D'UNE FONCTION PAR SON POLYNÔME DE TAYLOR AU VOISINAGE D'UN POINT

Le polynôme de Taylor de degré n en a de f est une approximation de f au voisinage de a . Si l'on sait estimer l'erreur R_n , on obtient la précision de l'approximation.

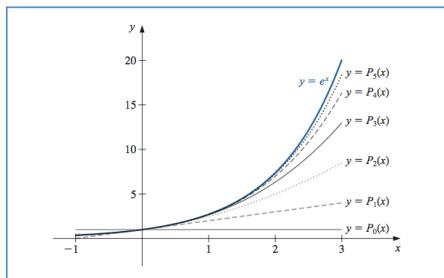
EXEMPLE. Polynômes de Taylor de degré $n = 1, \dots, 5$, de $f(x) = e^x$ au point 0.

$$P_1(x) = 1, \quad P_2(x) = 1 + x,$$

$$P_3(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6},$$

$$P_4(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24},$$

$$P_5(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120}.$$



	$P_2(x)$	$P_4(x)$	$P_6(x)$	$P_8(x)$	$P_{10}(x)$	e^x
$x = 0.2$	1.220000	1.221400	1.221403	1.221403	1.221403	1.221403
$x = 3$	8.500000	16.375000	19.412500	20.009152	20.079665	20.085537

EXEMPLE. Polynômes de Taylor de degré n de $f(x) = \frac{1}{x}$ au point 1.

$$P_n(x) = \sum_{k=0}^n (-1)^k (x-1)^k.$$

L'approximation de $f(3) = \frac{1}{3}$ à l'aide de $P_n(x)$ conduit à :

n	0	1	2	3	4	5	6	7
$P_n(3)$	1	-1	3	-5	11	-21	43	-85

- Le polynôme de Taylor donne une approximation précise d'une fonction en un point spécifique.
- Comment obtenir une approximation sur l'intervalle tout entier ?

Formulation mathématique : étant donnés $(n+1)$ couples (x_i, y_i) le problème consiste à trouver une fonction $\Phi(x)$ telle que

$$\Phi(x_i) = y_i, \quad i = 1, \dots, m,$$

où les y_i sont donnés. On dit alors que Φ **interpole** $\{y_i\}_i$ aux **nœuds** $\{x_i\}_i$.

- Lorsque Φ est un polynôme on parle d'**interpolation polynomiale**.
- Lorsque Φ est un polynôme trigonométrique on parle d'**interpolation trigonométrique**.
- Lorsque Φ est un polynôme par morceaux on parle d'**interpolation polynomiale par morceaux** (ou d'**interpolation par fonctions splines**).

Soient $(n + 1)$ couples (x_i, y_i) . On cherche un polynôme Π_m de degré inférieur ou égal à m tel que

$$\Pi_m(x_i) = a_m x_i^m + \dots + a_1 x_i + a_0 = y_i, \quad i = 0, \dots, n.$$

Définition 6

- Le polynôme Π_m est appelé **polynôme d'interpolation** (ou polynôme interpolant).
- Les points x_i sont appelés **nœuds d'interpolation**.

Notation : Lorsque $y_i = f(x_i)$, f étant une fonction donnée, le polynôme d'interpolation $\Pi_n(x)$ est noté $\Pi_n f(x)$.

Dans tout ce qui suit on considèrera le cas $n = m$.

Théorème 4 (Existence et unicité)

Etant donné $(n + 1)$ points distincts x_0, \dots, x_n et $(n + 1)$ valeurs correspondantes y_0, \dots, y_n , il existe un unique polynôme Π_n de degré inférieur ou égal à n tel que $\Pi_n(x_i) = y_i$ pour $i = 0, \dots, n$.

COMMENT TROUVER UN TEL POLYNÔME ?

- ① Méthode générale : Remplacer les coordonnées des points dans l'expression du polynôme et résoudre le système linéaire.
 - ▶ Procédure coûteuse.
 - ▶ Systèmes mal conditionnés.
- ② Méthodes ad hoc :
 - ① Méthode de Lagrange.
 - ② Méthode de Newton.
 - ★ Même polynôme que par la méthode de Lagrange.
 - ★ Coût de calcul moins élevé que par la méthode de Lagrange.
 - ③ Méthode de Hermite
 - ④ ...

Définition 7

On appelle **polynômes caractéristiques** de Lagrange les polynômes ℓ_i définis pour $i = 0, \dots, n$ par

$$\ell_i(x) = \prod_{j=1, j \neq i}^n \frac{x - x_j}{x_i - x_j}.$$

Notation. Soient u_1, u_2, \dots, u_N , N réels. On note le produit $u_1 \times u_2 \times \dots \times u_N$ par : $\prod_{i=1}^N u_i$. Ainsi, les polynômes caractéristiques de Lagrange s'écrivent :

$$\begin{aligned} \ell_0(x) &= \frac{(x - x_1) \cdots (x - x_i) \cdots (x - x_n)}{(x_0 - x_1) \cdots (x_0 - x_i) \cdots (x_0 - x_n)}, \\ \ell_1(x) &= \frac{(x - x_0)(x - x_2) \cdots (x - x_i) \cdots (x - x_n)}{(x_1 - x_0)(x_1 - x_2) \cdots (x_1 - x_i) \cdots (x_1 - x_n)}, \\ \ell_i(x) &= \frac{(x - x_0) \cdots (x - x_{i-1})(x - x_{i+1}) \cdots (x - x_n)}{(x_i - x_0) \cdots (x_i - x_{i-1})(x_i - x_{i+1}) \cdots (x_i - x_n)}, \quad i \neq 0, 1, n \\ \ell_n(x) &= \frac{(x - x_0) \cdots (x - x_i) \cdots (x - x_{n-1})}{(x_n - x_0) \cdots (x_n - x_i) \cdots (x_n - x_{n-1})}. \end{aligned}$$

Les polynômes caractéristiques de Lagrange :

- sont de degré n ,
- sont tels que $\ell_i(x_i) = 1$, $i = 0, \dots, n$ et $\ell_i(x_j) = 0$ pour tout $j \neq i$.

Proposition 5

Les polynômes caractéristiques $\{\ell_i\}_{i=0,\dots,n}$ forment une base de l'ensemble des polynômes de degré inférieur ou égal à n .

Ainsi,

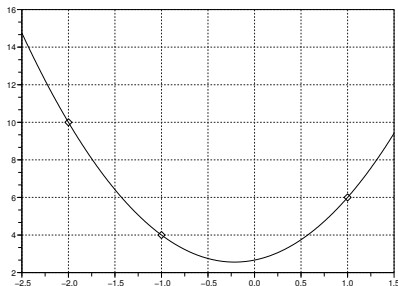
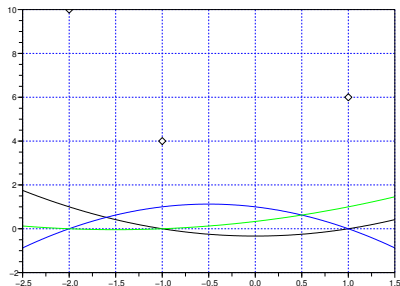
Théorème 5

Le polynôme interpolant $\{y_i\}_{i=0,\dots,n}$ aux nœuds $\{x_i\}_{i=0,\dots,n}$ dans la base $\{\ell_i\}_{i=0,\dots,n}$ s'écrit

$$\Pi_n(x) = \sum_{i=0}^n y_i \ell_i(x).$$

Ce polynôme est appelé **polynôme d'interpolation de Lagrange**.

EXEMPLE. Déterminer le polynôme de Lagrange interpolant les points $y_0 = 10, y_1 = 4$ et $y_2 = 6$ aux nœuds $x_0 = -2, x_1 = -1$ et $x_2 = 1$.



À gauche, *Polynômes caractéristiques de Lagrange, l_0, l_1, l_2* , à droite *Polynôme d'interpolation de Lagrange Π_2* .

EXEMPLE. Soient $f(x) = \cos(x)$ et $q_0 = (0, 1), q_1 = (\pi/16, \cos(\pi/16))$ et $q_2 = (\pi/8, \cos(\pi/8))$.

1. Calculer le polynôme d'interpolation de f passant par ces 3 points et en déduire une approximation de $\cos(\pi/32)$.
2. Calculer le développement de Taylor de f de degré 2 de la fonction $f(x) = \cos(x)$ au voisinage de 0 et en déduire une approximation de $\cos(\pi/32)$.

Définition 8

On appelle **polynôme nodal** de degré $(n + 1)$ le polynôme défini par $\omega_0 = 1$

$$\omega_{n+1}(x) = \prod_{i=0}^n (x - x_i) = (x - x_0)(x - x_1) \dots (x - x_n), \quad n \geq 0.$$

Théorème 6

Soient x_0, \dots, x_n , $(n + 1)$ nœuds distincts et x un point appartenant au domaine de définition de f . Si $f \in \mathcal{C}^{n+1}(I_x)$, où I_x est le plus petit intervalle contenant les nœuds x_0, \dots, x_n et le point x . Alors, l'erreur d'interpolation au point x est donnée par :

$$E_n(x) = f(x) - \Pi_n f(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega_{n+1}(x), \quad \text{avec } \xi \in I_x.$$

EXEMPLE. Soit la fonction $f(x) = 2xe^{-(4x+2)}$ définie sur l'intervalle $[0.2, 1]$.

1. Calculer le polynôme d'interpolation de Lagrange interpolant f aux nœuds $x = 0, 2$ et $x = 1$.
2. Pour quelle valeur de $x \in [0.2, 1]$ l'erreur d'interpolation $|E_2(x)|$ est-elle maximale ?

Objectif : Etant données $(n + 1)$ paires (x_i, y_i) , $i = 0, \dots, n$, écrire Π_n tel que

$$\Pi_n(x) = \Pi_{n-1}(x) + q_n(x),$$

où q_n est un polynôme de degré n ne dépendant que des nœuds x_i et d'un seul coefficient inconnu. Alors :

$$q_n(x) = a_n(x - x_0)(x - x_1) \cdots (x - x_{n-1}) = a_n \omega_n(x),$$

avec a_0, \dots, a_n des réels.

Puisque $q_n(x_i) = \Pi_n(x_i) - \Pi_{n-1}(x_i) = 0$ pour $i = 0, \dots, n - 1$, on a nécessairement

$$q_n(x) = a_n(x - x_0) \cdots (x - x_{n-1}) = a_n \omega_n(x).$$

Pour déterminer le coefficient a_n , supposons que $y_i = f(x_i)$, $i = 0, \dots, n$, où f est une fonction donnée, pas nécessairement sous forme explicite. Puisque $\Pi_n f(x_n) = f(x_n)$, on déduit que

$$a_n = \frac{f(x_n) - \Pi_{n-1} f(x_n)}{\omega_n(x_n)}.$$

Définition 9

Le coefficient a_n est appelé **n -ème différence divisée de Newton** et est souvent noté

$$a_n = f[x_0, x_1, \dots, x_n].$$

Proposition 6

Les polynômes nodaux $\{\omega_i\}_{i=0,\dots,n}$ forment une base de l'ensemble des polynômes de degré inférieur ou égal à n .

Théorème 7

En posant $y_0 = f(x_0) = f[x_0]$, et $\omega_0 = 1$, on a

$$\Pi_n f(x) = \sum_{k=0}^n \omega_k(x) f[x_0, \dots, x_k].$$

Cette expression est appelée **formule des différences divisées de Newton** du polynôme d'interpolation.

REMARQUE. Par unicité du polynôme d'interpolation cette expression définit le même polynôme que la formule de Lagrange.

Propriétés

- 1 La valeur prise par la différence divisée est invariante par permutation des nœuds.
- 2 On a la formule de récurrence suivante :

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

Des propriétés ci-dessus on peut déduire le *tableau des différences divisées* :

x	$f(x)$	1ère diff. divisées	2ème diff. divisées	nème diff. divisées	
x_0	$f[x_0]$				
x_1	$f[x_1]$	$f[x_0, x_1]$			
x_2	$f[x_2]$	$f[x_1, x_2]$	$f[x_0, x_1, x_2]$		
\vdots	\vdots		\vdots	\ddots	
x_n	$f[x_n]$	$f[x_{n-1}, x_n]$	$f[x_{n-2}, x_{n-1}, x_n]$	\dots	$f[x_0, \dots, x_n]$

REMARQUE.

- Pour $n + 1$ points il est nécessaire de calculer une matrice triangulaire inférieure de taille n laquelle a $n(n + 1)/2$ éléments différents de zéro.
- $n(n + 1)$ additions et $n(n + 1)/2$ divisions sont nécessaires pour construire la matrice triangulaire inférieure des différences divisées.
- Pour construire Π_{n+1} à partir de Π_n , $(n + 1)$ divisions et $2(n + 1)$ additions sont nécessaires. Ceci n'est pas le cas pour la méthode de Lagrange où il est nécessaire de répéter toute la procédure.

EXEMPLE.

- Étant donné trois points $(0, 1)$, $(2, 5)$ et $(4, 17)$, déterminer le polynôme d'interpolation passant par ces points.
- Soit f une fonction passant par les points $q_1 = (0, 3)$, $q_2 = (2, -1)$ et $q_3 = (5, 8)$.
 1. Donner la forme de Newton du polynôme d'interpolation de f passant par les points q_1 , q_2 et q_3 et donner une approximation de $f(3)$.
 2. Sachant que $f(6) = 7$, donner une approximation de l'erreur commise.
 3. On sait aussi que $f'(0) = 6$. Calculer le polynôme d'interpolation de degré minimal passant par les points q_1 , q_2 et q_3 , dont la dérivée en $x = 0$ est égale à 6.
- Une voiture roulant à 60km/h accélère au temps $t = 0$ s et sa vitesse v (en km/h) est mesurée régulièrement :

$t [s]$	0.0	0.7	1.4	2.1	2.8
$v [km/h]$	60	72.4	81.5	87.2	95.9

1. À l'aide d'un polynôme d'interpolation de degré inférieur ou égal à 2, donner une approximation de la vitesse (en km/h) à $t = 1.2$ s.
2. Donner l'expression analytique de l'erreur commise.
3. Obtenir une approximation de cette erreur.

CHAPITRE 2 : INTÉGRATION

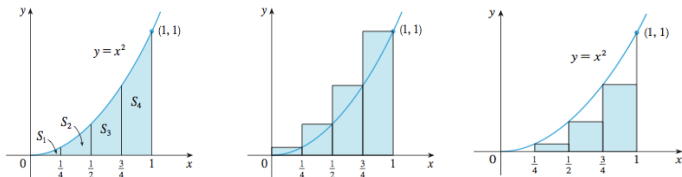
Objectif : On cherche à calculer l'aire de la partie du plan de \mathbb{R}^2 définie par

$$\mathcal{E}_{a,b}(f) = \{(x, y) \in \mathbb{R}^2, a \leq x \leq b, 0 \leq y \leq f(x)\},$$

où f est une fonction de $[a, b]$ dans \mathbb{R}^+ .

Idée : Décomposer $\mathcal{E}_{a,b}(f)$ en bandes, puis approcher chaque bande par un rectangle de même base.

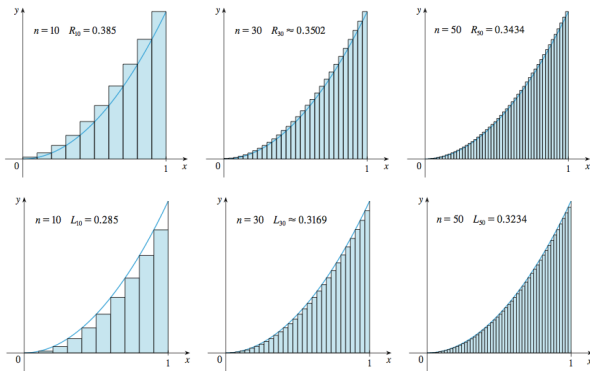
EXEMPLE. Calculer l'aire de $\mathcal{E}_{0,1}(f)$ avec $f(x) = x^2$.



Décomposition de $\mathcal{E}_{0,1}(f)$ en bandes, figure de gauche ; approximations de chaque bande par un rectangle de même base, figure du milieu et de droite.

Soient $[x_{i-1}, x_i]$, $i = 0, \dots, n$, n sous-intervalles de $[0; 1]$ avec $\cap_{i=1}^n]x_{i-1}, x_i[= \emptyset$ et $x_0 = 0, x_1 = 1$. Soient R_n la somme des n rectangles approchant $\mathcal{E}_{0,1}(f)$ "par la droite" et L_n la somme des n rectangles approchant $\mathcal{E}_{0,1}(f)$ "par la gauche" :

$$R_n = \sum_{i=1}^n (x_i - x_{i-1})f(x_{i-1}), \quad L_n = \sum_{i=1}^n (x_i - x_{i-1})f(x_i).$$



On a $\text{Aire}(\mathcal{E}_{0,1}(f)) = \lim_{n \rightarrow +\infty} R_n = \frac{1}{3}$ et $\text{Aire}(\mathcal{E}_{0,1}(f)) = \lim_{n \rightarrow +\infty} L_n = \frac{1}{3}$.

REMARQUE. On peut choisir n'importe quel point $x_i^* \in [x_{i-1}, x_i]$ et montrer que

$$\text{Aire}(\mathcal{E}_{0,1}(f)) = \lim_{n \rightarrow +\infty} \sum_{i=1}^n (x_i - x_{i-1})f(x_i^*).$$

Définition 10

On appelle **subdivision** σ de l'intervalle $[a, b]$ un ensemble de $(n + 1)$ réels x_0, x_1, \dots, x_n tels que $a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$.

Définition 11

Soit f une fonction continue sur $[a, b]$ et soient σ une subdivision quelconque de $[a, b]$ et $(\eta_i)_{1 \leq i \leq n}$ un ensemble de points tels que $\eta_i \in [x_{i-1}, x_i]$. Le réel

$$S_n = \sum_{i=1}^n f(\eta_i)(x_i - x_{i-1}),$$

est appelé **somme de Riemann** de f relative à σ et à $(\eta_i)_{1 \leq i \leq n}$.

Proposition 7

Si $\max_{i=1, \dots, n} (x_{i+1} - x_i)$ tend vers 0 quand n tend vers $+\infty$ et si S_n admet une limite finie pour toute subdivision σ et tout choix η_i , alors

$$\lim_n \sum_{i=1}^n f(\eta_i)(x_i - x_{i-1}) = \int_a^b f(t)dt.$$

On dit que f est **intégrable au sens de Riemann**.

Proposition 8

- Si f est une fonction monotone sur $[a, b]$, alors f est intégrable sur $[a, b]$.
- Si f est une fonction continue sur $[a, b]$ elle est intégrable sur $[a, b]$.

EXEMPLE. Calculer $\int_0^3 (x^3 - 6x)dx$.

Proposition 9 (Propriétés de l'intégrale de Riemann)

- 1 Si f une fonction intégrable sur l'intervalle $[a, b]$ alors, $|f|$ est intégrable sur $[a, b]$.
- 2 (Relation de Chasles.) Soit $c \in]a, b[$. f est intégrable sur $[a, b]$ si et seulement si f est intégrable sur $[a, c]$ et $[c, b]$. Alors, on a

$$\int_a^b f(t)dt = \int_a^c f(t)dt + \int_c^b f(t)dt.$$

- 3 (Linéarité.) Pour tous $\lambda, \mu \in \mathbb{R}$ et pour toutes fonctions f et g intégrables sur $[a, b]$ on a

$$\int_a^b (\lambda f(t) + \mu g(t))dt = \lambda \int_a^b f(t)dt + \mu \int_a^b g(t)dt.$$

- 4 (Convention.) Pour une fonction f intégrable sur $[a, b]$ on pose

$$\int_a^b f(t)dt = - \int_b^a f(t)dt.$$

Théorème 8 (Formules de la moyenne)

- ① Soit f continue sur $[a, b]$. On pose $m = \inf_{x \in [a, b]} f(x)$ et $M = \sup_{x \in [a, b]} f(x)$. Alors,

$$m(b-a) \leq \int_a^b f(t)dt \leq M(b-a).$$

- ② **Première formule de la moyenne.**

Soient f et g continues sur $[a, b]$ avec g positive. Alors il existe $c \in]a, b[$ tel que

$$\int_a^b f(t)g(t)dt = f(c) \int_a^b g(t)dt.$$

- ③ **Deuxième formule de la moyenne.**

Soient f et g continues sur $[a, b]$ avec f positive et décroissante. Alors il existe $c \in]a, b[$ tel que

$$\int_a^b f(t)g(t)dt = f(a) \int_a^c g(t)dt.$$

EXEMPLE.

- Montrer que la vitesse moyenne d'une voiture pendant l'intervalle de temps $[t_1; t_2]$ est la même que la moyenne de ses vitesses sur tout le trajet. ??????
- Soit f une fonction continue au voisinage de 0. Calculer $\frac{1}{x^2} \lim_{x \rightarrow 0} \int_0^x tf(t)dt$ et

$$\lim_{x \rightarrow 0^+} \int_x^{2x} \frac{f(t)}{t} dt$$

Proposition 10

Soient f et g deux fonctions intégrables sur $[a, b]$. Alors,

- ① si $f > 0$. $\int_a^b f(t)dt > 0$.

Théorème 9

Soit f une fonction continue sur $[a, b]$ positive ou nulle, si $\int_a^b f(t)dt = 0$ alors $f = 0$.

3.2 RELATION ENTRE PRIMITIVE ET INTÉGRALE

Définition 12

Soit $f : [a, b] \rightarrow \mathbb{R}$. On dit que G est **une primitive** de f si

$$\forall x \in [a, b], \quad G'(x) = f(x).$$

Théorème 10

Soit f une fonction intégrable sur $[a, b]$.

- La fonction F définie par $F(x) = \int_a^x f(t) dt$, est continue sur $[a, b]$, dérivable sur $]a, b[$ et $F'(x) = f(x)$.
- Soit G une primitive de f . Alors,

$$\int_a^b f(x) dx = G(b) - G(a) = [G(x)]_a^b.$$

Notation :

- $\int f(x) dx$ désigne une primitive de f .
- de Leibniz : $\frac{d}{dx} \int_a^x f(t) dt = f(x)$.

EXEMPLE. Une particule bouge lelong d'une ligne droite à la vitesse (en m/s)
 $v(t) = t^2 - t - 6$.

- Calculer le déplacement de la particule pendant la période $1 \leq t \leq 4$.
- Calculer la distance parcourue durant cette période.

CALCUL À L'AIDE DE PRIMITIVES

Théorème 11 (Intégration par parties)

Soient u et v des fonctions de classe \mathcal{C}^1 sur $[a, b]$, alors

$$\int_a^b u'(x)v(x) dx = \{u(b)v(b) - u(a)v(a)\} - \int_a^b u(x)v'(x) dx.$$

Théorème 12 (Changement de variable)

Soit $\phi : [\alpha, \beta] \rightarrow [a, b]$ une fonction de classe C^1 telle que $\phi([\alpha, \beta]) \subset [a, b]$ et soit f une fonction continue sur $[a, b]$. Alors,

$$\int_a^b f(\phi(t))\phi'(t) dt = \int_{\phi(a)}^{\phi(b)} f(x) dx,$$

et

$$\int f(\phi(t))\phi'(t) dt = \int f(x) dx.$$

QUELQUES PRIMITIVES USUELLES

$$\int x^m dx = \frac{x^{m+1}}{m+1} + C, \quad m \in \mathbb{R} \setminus \{-1\} \quad \int \sin(x) dx = -\cos(x) + C$$

$$\int \frac{1}{x} dx = \ln|x| + C \quad \int \cos(x) dx = \sin(x) + C$$

$$\int e^{ax} dx = \frac{1}{a} e^{ax} + C, \quad a \in \mathbb{R}^* \quad \int \frac{1}{1+x^2} dx = \arctan(x) + C$$

$$\int a^x dx = \frac{a^x}{\ln a} + C, \quad a \in \mathbb{R} \setminus \{1\},$$

avec C une constante.

Définition 13

On appelle **arc paramétré** de classe \mathcal{C}^k un couple (I, γ) où I est un intervalle de \mathbb{R} et γ une application de classe \mathcal{C}^k de I dans \mathbb{R}^2 .

Théorème 13 (Longueur d'un arc de courbe)

Soit $(I = [a, b], \gamma)$ un arc paramétré de classe \mathcal{C}^1 . Sa longueur est donné par

$$\int_a^b \|\gamma'(t)\| dt.$$

- Si l'arc est donné en coordonnées cartésiennes $(x(t), y(t))$ sa longueur est

$$\int_a^b \sqrt{x'(t)^2 + y'(t)^2} dt.$$

- Si l'arc est donné en coordonnées polaires sa longueur est

$$\int_{\theta_1}^{\theta_2} \sqrt{r(\theta)^2 + r'(\theta)^2} d\theta.$$

EXEMPLE. Calculer la longueur de la cardioïde définie par $r = 2(1 - \cos t)$.

Objectif : Etant donnée une fonction f continue sur un intervalle $[a, b] \subset \mathbb{R}$ calculer de façon approchée l'intégrale $I(f) = \int_a^b f(x)dx$ à l'aide de méthodes dites de **quadrature numérique** ou d'**intégration numérique**.

Principe : Approcher f par $f_n, n \geq 0$ et calculer l'intégrale $I(f_n)$ au lieu de $I(f)$. En posant $I(f_n) = I_n(f)$ on a

$$I_n(f) = \int_a^b f_n(x)dx.$$

L'approximation f_n doit donc être facilement intégrable. Considérons une fonction f continue.

Alors l'**erreur de quadrature** $E_n(f) = I(f) - I_n(f)$ vérifie l'inégalité

$$|E_n(f)| \leq \sup_{x \in [a, b]} |f(x) - f_n(x)|(b - a).$$

On choisit f_n comme étant le polynôme d'interpolation de Lagrange, $\Pi_n f$, de f en $(n + 1)$ points distincts $x_i, i = 0, \dots, n$. Alors, $I_n(f) = \sum_{i=0}^n f(x_i) \int_a^b l_i(x) dx$, où l_i est le polynôme caractéristique de Lagrange de degré n associé au nœud x_i .

Plus généralement, on appelle **formule de quadrature à poids** la quantité définie par

$$I_n(f) = \sum_{i=0}^n \omega_i f(x_i).$$

Définition 14

- Les réels $\omega_i, i = 0, \dots, n$ sont appelés **poids de la quadrature**,
- les points $x_i, i = 0, \dots, n$ sont appelés **nœuds de la quadrature**.
- Le plus grand entier k tel que pour tout polynôme P de degré k , $I_n(P) = I(P)$, est appelé **ordre de quadrature**.

Cette formule est obtenue en remplaçant f par une constante égale à la valeur de $f(a)$, autrement dit

$$I_0(f) = (b - a)f(a).$$

Proposition 11

Si $f \in \mathcal{C}^1([a, b])$ l'erreur de quadrature est

$$E_0(f) = \frac{(b - a)^2}{2} f'(\xi), \quad \xi \in]a, b[.$$

Corollaire 3

La formule du rectangle est d'ordre 0.

Cette formule est obtenue en remplaçant f par une constante égale à la valeur de f au point milieu de $[a, b]$, autrement dit

$$I_0(f) = (b - a)f\left(\frac{a + b}{2}\right).$$

Proposition 12

Si $f \in \mathcal{C}^2([a, b])$ l'erreur de quadrature est

$$E_0(f) = \frac{(b - a)^3}{24} f''(\xi), \quad \xi \in]a, b[.$$

Corollaire 4

La formule du point milieu est d'ordre 1.

Cette formule est obtenue en remplaçant f par son polynôme d'interpolation de Lagrange de degré 1 aux nœuds a et b , autrement dit

$$I_1(f) = \frac{b-a}{2}(f(a) + f(b)).$$

Proposition 13

Si $f \in \mathcal{C}^2([a, b])$ l'erreur de quadrature est

$$E_1(f) = -\frac{(b-a)^3}{12}f''(\xi), \quad \xi \in]a, b[.$$

Corollaire 5

La formule du trapèze est d'ordre 1.

Cette formule est obtenue en remplaçant f par son polynôme d'interpolation de Lagrange de degré 2 aux nœuds a , b et $(a + b)/2$, autrement dit

$$I_2(f) = \frac{b-a}{6} \left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right).$$

Proposition 14

Si $f \in \mathcal{C}^4([a, b])$ l'erreur de quadrature est

$$E_2(f) = -\frac{(b-a)^5}{2880} f^{(4)}(\xi), \quad \xi \in]a, b[.$$

Corollaire 6

La formule de Simpson est d'ordre 3.

Objectif : Améliorer la précision des méthodes de quadratures en subdivisant l'intervalle $[a, b]$ en m sous-intervalles $I_j = [y_j, y_{j+1}]$ avec $y_{j+1} = a + jh$ où $h = \frac{b-a}{m}$ et $j = 0, \dots, m$.

Principe : Sur chaque sous-intervalle I_j appliquer une quadrature interpolatoire de nœuds $\{x_k^{(j)}, k = 0, \dots, n\}$ et de poids $\{\omega_k^{(j)}, k = 0, \dots, n\}$.

Une quadrature interpolatoire composite est obtenue en remplaçant $I(f)$ par

$$I_{n,m}(f) = \sum_{j=0}^{m-1} \sum_{k=0}^n \omega_k^{(j)} f(x_k^{(j)}).$$

En introduisant les de nœuds de quadrature $x_k = a + (2k + 1)\frac{h}{2}$ $k = 0, \dots, m - 1$,

$$I_{0,m} = h \sum_{k=0}^{m-1} f(x_k).$$

Proposition 15

- Si $f \in \mathcal{C}^2([a, b])$ l'erreur de quadrature est donnée par

$$E_{0,m}(f) = \frac{(b-a)}{24} h^2 f''(\xi), \text{ avec } \xi \in]a, b[.$$

- La méthode est d'ordre 1.

En introduisant les de nœuds de quadrature $x_k = a + kh, k = 0, \dots, m,$

$$I_{1,m}(f) = \frac{h}{2} \sum_{k=0}^{m-1} (f(x_k) + f(x_{k+1})), \text{ où } x_0 = a \text{ et } x_m = b.$$

Proposition 16

- Si $f \in \mathcal{C}^2([a, b])$, l'erreur de quadrature est donnée par

$$E_{1,m}(f) = -\frac{(b-a)}{12} h^2 f''(\xi), \text{ avec } \xi \in]a, b[.$$

- La méthode est d'ordre 1.

En introduisant les de nœuds de quadrature $x_k = a + k\frac{h}{2}$, $k = 0, \dots, m$,

$$I_{2,m}(f) = \frac{h}{6} \left(f(x_0) + 2 \sum_{r=0}^{m-1} f(x_{2r}) + 4 \sum_{s=0}^{m-1} f(x_{2s+1}) + f(x_{2m}) \right),$$

où $x_0 = a$ et $x_{2m} = b$.

Proposition 17

- Si $f \in \mathcal{C}^4([a, b])$, l'erreur de quadrature est donnée par

$$E_{2,m}(f) = -\frac{(b-a)}{2880} h^4 f^{(4)}(\xi), \text{ avec } \xi \in]a, b[.$$

- La méthode est d'ordre 3.

EXEMPLE

Calcul approché de $\int_1^2 \frac{1}{x} dx$ par les formules composites du rectangle à gauche, I_n^G , à droite, I_n^D , du point milieu, I_n^M , du trapèze, I_n^T , et de Simpson, I_n^S .

On note $E_n^G, E_n^D, E_n^M, E_n^T$ et E_n^S les erreurs respectives.

On a $\int_1^2 \frac{1}{x} dx = \ln(2) \approx 0.6931472$.

n	I_n^G	I_n^D	I_n^M	I_n^T	I_n^S
5	0.745635	0.645635	0.668771	0.680803	0.693150
10	0.718771	0.695635	0.693771	0.693303	0.693147
20	0.705803	0.691908	0.692835	0.693069	0.693147

n	E_n^G	E_n^D	E_n^M	E_n^T	E_n^S
5	-0.052488	0.047512	-0.002488	0.001239	-0.0000031
10	-0.025624	0.024376	-0.000624	0.000312	-0.0000002
20	-0.012656	0.012344	-0.000156	0.000078	-1.21910 ⁻⁸

Remarques.

- Plus n est grand, plus l'approximation est précise.
- Les erreurs des formules du rectangle à gauche et à droite sont, de signe opposé, et décroissent d'un facteur 2 lorsque n est multiplié par 2.
- Les erreurs des formules du point milieu et du trapèze sont, de signe opposé, et décroissent d'un facteur 4 lorsque n est multiplié par 2.
- L'erreur de la formule de Simpson décroît d'un facteur 16 lorsque n est multiplié par 2.

Combien faut-il prendre de sous-intervalles de $[1, 2]$ pour que l'approximation de $\int_1^2 \frac{1}{x} dx$ ait une précision de l'ordre de 10^{-4} ?

Données : a, b : extrémité de l'intervalle d'intégration $[a, b]$,
 m : nombre de subdivision de l'intervalle $[a, b]$,
 f : fonction dont on cherche à approcher l'intégrale.

Résultat : I : le réel $I = I_{1,m}(f)$.

```
1: Fonction TRAPEZE(a,b,m,f)
2:    $h \leftarrow (b - a)/m$ 
3:    $I \leftarrow 0$ 
4:   Pour  $j \leftarrow 1$  à  $m + 1$  faire
5:      $x \leftarrow a + h * (j - 1)$ 
6:      $y \leftarrow f(x)$ 
7:      $I \leftarrow I + 0.5 * h * y$ 
8:   Fin Pour
9: Fin Fonction
```

Objectif : Pour une fonction $f :]a, b[\subset \mathbb{R} \longrightarrow \mathbb{R}$ on cherche $p \in]a, b[$ tel que

$$f(p) = 0.$$

Généralement la racine (le zéro) p ne peut être calculée explicitement. On peut alors utiliser une méthode itérative pour approcher p .

Principe d'une méthode itérative : Construire une suite (x_k) telle que

$$\lim_{k \rightarrow +\infty} x_k = p.$$

Définition 15 (Convergence)

On dit qu'une suite (x_k) construite par une méthode itérative converge vers p à l'ordre $q \geq 1$ si

$$\exists C > 0, \quad \frac{|x_{k+1} - p|}{|x_k - p|^q} \leq C, \quad \forall k \geq k_0,$$

avec $k_0 \in \mathbb{N}$. Le réel C est appelé **facteur de convergence**.

Remarques :

- Si $q = 1$ il est nécessaire que $C < 1$ pour que (x_n) converge vers p .
- Lorsque $q = 2$ on dit que la convergence est **quadratique**.

La convergence de ces méthodes dépend souvent du choix de la donnée initiale x_0 .

Définition 16

- Lorsque la méthode converge pour un x_0 "suffisamment proche" de p on dit que la méthode **converge localement**.
- Lorsque la méthode converge pour tout x_0 on dit que la méthode **converge globalement**.

Dans la pratique, on ne peut pas faire tendre k vers $+\infty$, il faut donc définir un ou plusieurs critères d'arrêt de la méthode itérative : On note $e_k = p - x_k$ l'erreur à l'itération k .

Exemples de critères d'arrêt :

- On se fixe une tolérance $\varepsilon > 0$ et on arrête les itérations s'arrêtent lorsque l'**incrément** devient "petit" : $|x_{m+1} - x_m| \leq \varepsilon$.
- On se fixe une tolérance $\varepsilon > 0$ et ainsi les itérations s'arrêtent lorsque $|f(x_m)| \leq \varepsilon$.
- Lorsque l'erreur $|e_m|$ peut-être majorée par un réel M (ne dépendant pas de la solution exacte), on se fixe une tolérance $\varepsilon > 0$ et ainsi les itérations s'arrêtent lorsque $|e_m| \leq M \leq \varepsilon$.
- On se donne un nombre maximal d'itérations.

5.1 MÉTHODE DE DICHOTOMIE OU DE LA BISSECTION

Soit $\mathcal{I}_0 = [a, b]$ avec a et b tels que $f(a)f(b) < 0$ et f continue.

Alors par le théorème des valeurs intermédiaires il existe $c \in]a, b[$ tel que $f(c) = 0$.

Principe : Construire une suite de sous-intervalles $\mathcal{I}_k = [a_k, b_k]$ tels que $\mathcal{I}_{k+1} \subset \mathcal{I}_k$ et $f(a_k)f(b_k) < 0$ encadrant la racine p .

Méthode : On pose $a^{(0)} = a, b^{(0)} = b$ et $x_0 = \frac{a^{(0)} + b^{(0)}}{2}$ alors pour $k \geq 0$

Si $f(x_k)f(a_k) < 0$, on pose $a_{k+1} = a_k$, et $b_{k+1} = x_k$,

Si $f(x_k)f(b_k) < 0$, on pose $a_{k+1} = x_k$, et $b_{k+1} = b_k$,

et

$$x_{k+1} = \frac{a_{k+1} + b_{k+1}}{2}.$$

Proposition 18 (Convergence)

- Pour tout $k \geq 0$, on a $|e_k| \leq \frac{b-a}{2^{k+1}}$.
- La méthode de dichotomie est globalement convergente.

$$\lim_{k \rightarrow +\infty} |e_k| = 0.$$

Propriété : La méthode de dichotomie permet de déterminer à l'avance le nombre d'itérations nécessaires pour atteindre la précision souhaitée, i.e. pour avoir $|x_m - p| \leq \varepsilon$ il faut prendre

$$m \geq \frac{\ln((b-a)/\varepsilon)}{\ln 2} - 1.$$

Exemple : L'équation $f(x) = x^3 + 4x^2 - 10 = 0$ a une racine dans $[1, 2]$ puisque $f(1) = -5$ et $f(2) = 14$. La racine est $p = 1.365230013$.

Résultats de la méthode de dichotomie où le critère d'arrêt est : $|e_m| \leq \varepsilon$.

n	a_n	b_n	x_n	$f(x_n)$
1	1	2	1.5	2.375
2	1	1.5	1.25	-1.796875
3	1.25	1.5	1.375	0.1621094
4	1.25	1.375	1.3125	-0.8483887
5	1.3125	1.375	1.34375	-0.3509827
9	1.36328125	1.3671875	1.365234375	0.000072
10	1.36328125	1.365234375	1.364257813	-0.01605

Remarque : La méthode de dichotomie n'est pas une méthode d'ordre 1.

5.2 MÉTHODE DU POINT FIXE

Définition 17

Pour une fonction Φ donnée on appelle **point fixe** d'une fonction Φ un réel p tel que $\Phi(p) = p$.

Pour toute fonction $f : [a, b] \rightarrow \mathbb{R}$ le problème $f(x) = 0$ peut se mettre sous la forme $x - \Phi(x) = 0$ où $\Phi : [a, b] \rightarrow \mathbb{R}$ est telle que $\Phi(p) = p$ lorsque $f(p) = 0$.

Réciproquement, si Φ a un point fixe en p , alors p est une racine de la fonction $f(x) = x - \Phi(x)$.

Exemple : La fonction $\Phi(x) = x^2 - 2$ a deux points fixes dans $[-2, 3]$: $x = -1$ et $x = 2$.

Principe de la méthode de point fixe.

Pour approcher le point fixe d'une fonction Φ :

- on choisit une approximation initiale x_0 , puis
- on génère une suite (x_n) en posant $x_n = \Phi(x_{n-1})$ pour tout $n \geq 1$.

Ainsi, si la suite (x_n) converge vers p et si Φ est continue alors

$$p = \lim_n x_n = \lim_n \Phi(x_{n-1}) = \Phi(p).$$

Cette technique est appelée **itération de point fixe**.

ALGORITHME DU POINT FIXE

Données : Φ : fonction dont on cherche un point fixe dans $[a, b]$
 a, b : bornes de l'intervalle
 x_0 : donnée initiale
 tol : tolérance de la méthode
 $nmax$: nombre maximal d'itérations.

Résultats : p : le réel tel que $\Phi(p) \approx p$
 $niter$: nombre d'itérations.

```
1: Fonction POINTFIXE( $f, a, b, x_0, tol, nmax$ )
2:    $niter \leftarrow 1$ 
3:   Tantque  $niter < nmax$  &  $err > tol$  faire
4:      $p \leftarrow \Phi(x_0)$ 
5:      $niter \leftarrow niter + 1$ 
6:      $err \leftarrow |p - x_0|$ 
7:      $x_0 \leftarrow p$ 
8:   Fin Tantque
9: Fin Fonction
```

Exemple : Approchons la racine $p = 1.365230013$ de l'équation $f(x) = x^3 + 4x^2 - 10 = 0$ dans $[1, 2]$ par la méthode du point fixe.

Plusieurs choix possibles pour Φ :

a. $\Phi_1(x) = x - x^3 - 4x^2 + 10$

b. $\Phi_2(x) = \left(\frac{10}{x} - 4x\right)^{\frac{1}{2}}$

c. $\Phi_3(x) = \frac{1}{2}(10 - x^3)^{\frac{1}{2}}$

d. $\Phi_4(x) = \left(\frac{10}{4+x}\right)^{\frac{1}{2}}$

e. $\Phi_5(x) = x - \frac{x^3 + 4x^2 - 10}{3x^2 + 8x}$

Valeurs obtenues de $x_{n+1}^i = \Phi(x_n^i)$ lorsque $x_0 = 1.5$ et le critère d'arrêt choisit est : $|x_{n+1} - x_n| \leq \varepsilon = 10^{-8}$.

n	$\Phi_1(x_n^1)$	$\Phi_2(x_n^2)$	$\Phi_3(x_n^3)$	$\Phi_4(x_n^4)$	$\Phi_5(x_n^5)$
0	-0.875	0.8165	1.286953768	1.348399725	1.373333333
2	-469.7	"(-8.65) ^{1/2} "	1.345458374	1.364957015	1.365230014
3	1.03×10^8		1.375170253	1.365264748	1.365230013
4			1.360094193	1.365225594	
14			1.365223680	1.365230013	
29			1.365230013		

Comment choisir Φ ?

Théorème 14 (point fixe)

1. Si $\Phi \in \mathcal{C}([a, b])$ et $\Phi(x) \in [a, b]$ pour tout $x \in [a, b]$, alors Φ a un point fixe dans $[a, b]$.
2. Si de plus, Φ' existe sur $]a, b[$ et s'il existe une constante $k < 1$ telle que

$$|\Phi'(x)| \leq k, \quad \text{pour tout } x \in]a, b[,$$

alors le point fixe dans $[a, b]$ est unique.

Alors, pour tout réel x_0 dans $[a, b]$ la suite récurrente définie pour $n \geq 1$ par $x_{n+1} = \Phi(x_n)$, converge vers l'unique point fixe, p , de Φ . De plus, on a

$$\lim_k \frac{x_{k+1} - p}{x_k - p} = \Phi'(p).$$

La quantité $|\Phi'(p)|$ est appelée **facteur de convergence asymptotique**.

Définition 18

Une fonction f définie sur un intervalle $[a, b]$ est dite **contractante** lorsqu'il existe un réel $0 < k < 1$ tel que pour tous x, y dans $[a, b]$, $|f(x) - f(y)| \leq k|x - y|$.

Si f est dérivable sur $[a, b]$ et si $f'(x) < 1$ pour tout x dans $[a, b]$, alors f est contractante.

Proposition 19

La méthode du point fixe converge à l'ordre 1 lorsque $\Phi'(p) \neq 0$.

Que dire lorsque $\Phi'(p) = 0$?

Supposons que $\Phi \in \mathcal{C}^2([a, b])$ et qu'il existe $M > 0$ tel que $|\Phi''(x)| \leq M$ pour tout x dans $[a, b]$. Le développement de Taylor-Lagrange de Φ en p à l'ordre 1 conduit à :

$$\begin{aligned}\Phi(x) &= \Phi(p) + (x - p)\Phi'(p) + \frac{(x - p)^2}{2}\Phi''(\xi), \quad \text{avec } \xi \text{ entre } x \text{ et } p \\ &= p + \frac{(x - p)^2}{2}\Phi''(\xi),\end{aligned}$$

d'où $|\Phi(x) - p| \leq \frac{M}{2}|x - p|^2$, et donc en prenant $x = x_n$ on obtient :

$$|x_{n+1} - p| \leq \frac{M}{2}|x_n - p|^2.$$

La suite (x_n) converge quadratiquement vers p .

EXERCICE. Montrer que $|x_{n+1} - p| \leq \frac{2}{M} \left(\frac{M}{2}(x_n - p) \right)^{2^{n+1}}$.

Remarques.

- Si $|\Phi'(p)| > 1$ et si x_n est proche de p avec $|\Phi'(x_n)| > 1$, alors par un développement de Taylor de Φ en p à l'ordre 1 on obtient que $|x_{n+1} - p| > |x_n - p|$. Il ne peut donc y avoir convergence de la méthode!
- Dans le cas où $|\Phi'(p)| = 1$, on ne peut en général tirer aucune conclusion : selon le problème considéré, il peut y avoir convergence ou divergence.

Soit $\Phi(x) = x - x^3$ qui admet $p = 0$ comme point fixe. Bien que $\Phi'(p) = 1$, si $x_0 \in [-1, 1]$ alors $x_n \in]-1, 1[$ pour $n \geq 1$ et la suite converge (très lentement) vers p (en prenant $x_0 = \frac{1}{2}$, l'erreur absolue après 2000 itérations, $|x_{2000} - p|$, vaut 0.0158).

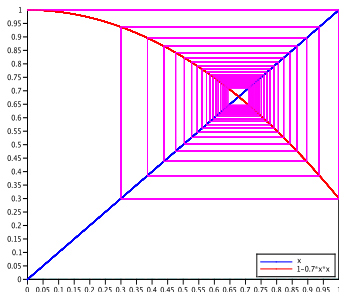
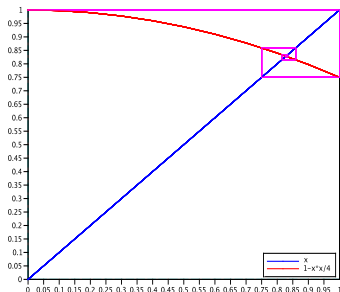
Considérons maintenant $\Phi(x) = x + x^3$ qui a aussi $p = 0$ comme point fixe. À nouveau, $|\Phi'(p)| = 1$ mais dans ce cas la suite x_n diverge pour tout choix $x_0 \neq 0$.

Définition 19

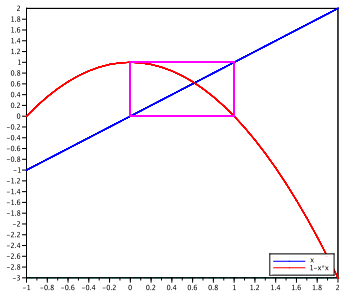
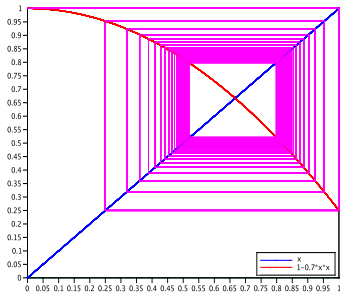
- Un point fixe p tel que $|\Phi'(p)| < 1$ est dit **attractif**.
- Un point fixe p tel que $|\Phi'(p)| > 1$ est dit **répulsif**.

Exemple : Approchons la solution de $g(x) = 1 - \beta x^2$ sur $[-1, 1]$.

- Lorsque $0 \leq \beta \leq 2$, l'intervalle $[-1, 1]$ est stable par g , i.e. $g([-1, 1]) \subset [-1, 1]$.
- g est contractante lorsque $0 < \beta < \frac{1}{2}$.
- Lorsque $0 < \beta < 2$, $p = \frac{\sqrt{4\beta+1}-1}{2\beta}$ est un point fixe de g . Lorsque $\beta = 2$, g a deux points fixes dans $[-1, 1]$: -1 et $\frac{1}{2}$.
- p est un point fixe attractif lorsque $0 < \beta \leq \frac{3}{4}$.
- Les points fixes -1 et $\frac{1}{2}$ sont répulsifs.
- Lorsque $0 < \beta < \frac{1}{2}$, g est strictement contractante, ainsi par le théorème de point fixe, la suite (x_n) définie par $x_{n+1} = g(x_n)$, x_0 donné, converge vers p .
- Lorsque $\frac{1}{2} \leq \beta < \frac{3}{4}$, g n'est pas strictement contractante. Le théorème du point fixe ne peut s'appliquer. Les sous-suites $x_{2n+2} = g \circ g(x_{2n})$ et $x_{2n+1} = g \circ g(x_{2n-1})$ sont monotones et bornées donc convergentes. Elles convergent l'une et l'autre vers un point fixe de $g \circ g$, qui est aussi un point fixe de g .



- Lorsque $\beta = \frac{3}{4}$, $|g'(p)| = 1$. Dans ce cas, si l'on se réfère à la figure ci-dessous, la méthode de point fixe semble converger vers p *très lentement*.
- Lorsque $\beta > \frac{3}{4}$, le point fixe p est répulsif et la suite (x_n) ne converge pas, sauf si $x_0 = p$!



Dans la pratique, il est souvent difficile de déterminer a priori l'intervalle $[a, b]$ sur lequel Φ serait contractante.

Théorème 15 (d'Ostrowski)

Soit p un point fixe d'une fonction Φ continue et différentiable dans un voisinage J de p . Si $|\Phi'(p)| < 1$ alors il existe $\delta > 0$ tel que la suite (x_k) converge vers p pour tout x_0 tel que $|x_0 - p| < \delta$.

Proposition 20

Soient J un voisinage de p et q un entier supérieur ou égal à 1. Si $\Phi \in \mathcal{C}^{q+1}(J)$ et si $\Phi^i(p) = 0$ pour $1 \leq i \leq q$ et $\Phi^{q+1}(p) \neq 0$, alors la méthode de point fixe est d'ordre $q + 1$ et

$$\lim_k \frac{x_{k+1} - p}{(x_k - p)^{q+1}} = \frac{\Phi^{(q+1)}(p)}{(q+1)!}.$$

5.4 MÉTHODE DE NEWTON (NEWTON-RAPHSON OU DE LA TANGENTE)

Soit f une fonction définie sur $[a, b]$ n'ayant qu'un seul zéro dans $[a, b]$.

On cherche à calculer de façon approchée ce zéro. L'idée est de remplacer la courbe représentative de f par sa tangente.

Approche géométrique.

- On se donne un point x_0 de l'intervalle $[a, b]$, et on considère la tangente à la courbe représentative de f en $(x_0, f(x_0))$.
- On note x_1 l'abscisse de l'intersection de la tangente avec l'axe des abscisses.
Puisque la tangente est proche de la courbe, on peut espérer que x_1 donne une meilleure approximation d'une solution de l'équation $f(x) = 0$ que x_0 .
On considère la tangente à la courbe représentative de f en $(x_1, f(x_1))$.
- On note x_2 l'abscisse de l'intersection de la tangente avec l'axe des abscisses.
Puisque la tangente est proche de la courbe, on peut espérer que x_2 donne une meilleure approximation d'une solution de l'équation $f(x) = 0$ que x_1 .
- On recommence alors le procédé à partir de x_2 , et on construit par récurrence une suite (x_n) définie par :

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

Approche par le calcul. Supposons $f \in \mathcal{C}^2([a, b])$. Soit \bar{x} une approximation de p telle que $f'(\bar{x}) \neq 0$ et $|\bar{x} - p| \ll 1$. Le développement de Taylor à l'ordre 1 de f en \bar{x}

$$f(x) = f(\bar{x}) + (x - \bar{x})f'(\bar{x}) + \frac{(x - \bar{x})^2}{2} f''(\zeta_x), \text{ avec } \zeta_x \text{ compris entre } x \text{ et } \bar{x}.$$

conduit à

$$0 = f(p) = f(\bar{x}) + (p - \bar{x})f'(\bar{x}) + \frac{(p - \bar{x})^2}{2} f''(\zeta_p).$$

Comme $|\bar{x} - p| \ll 1$ on peut écrire

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

ainsi

$$p \approx \bar{x} - \frac{f(\bar{x})}{f'(\bar{x})}.$$

De cette expression, on construit par récurrence une suite (x_n) permettant d'approcher p : On se donne une donnée initiale x_0 , et on définit la suite (x_n) par

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

Proposition 21

La méthode de Newton est une méthode de point fixe où la fonction point fixe est :

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

Théorème 16 (Convergence locale)

Soit p une racine de f . Supposons $f \in \mathcal{C}^2(]a, b[)$ et $f'(p) \neq 0$. Si la donnée initiale x_0 est assez proche de p , alors la suite (x_n) définie par

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

converge vers p à l'ordre 2.

REMARQUE. Cette méthode nécessite le calcul à chaque itération de $f(x_n)$ et $f'(x_n)$.

Exemple : Approchons la racine de $f(x) = x^2 - 4$ dans $[-1, 3]$ avec la méthode de Newton.

n	x_n	x_n
0	0.5	-1
1	4.25	-2.5
2	2.5955882	-2.05
3	2.0683324	-2.0006098
4	2.0011288	-2.0000001
5	2.0000003	
6	2.	

ALGORITHME DE LA MÉTHODE DE NEWTON

Données : f : fonction dont on cherche une racine
 df : dérivée de la fonction f
 x^0 : donnée initiale
 tol : tolérance de la méthode
 $nmax$: nombre maximal d'itérations.

Résultats : x : le réel tel que $f(x) \approx 0$
 err : l'erreur commise
 $niter$: nombre d'itérations.

```
1: Fonction NEWTON( $f,df,x^0,tol,nmax$ )
2:    $err \leftarrow tol+1$ 
3:    $niter \leftarrow 0$ 
4:    $x \leftarrow x^0$ 
5:    $fx \leftarrow f(x)$ 
6:    $dfx \leftarrow df(x)$ 
7:   Tantque  $niter < nmax$  &  $err > tol$  faire
8:      $niter \leftarrow niter+1$ 
9:     Si  $dfx == 0$  alors STOP
10:    Sinon
11:       $x_n \leftarrow x - fx/dfx$ 
12:       $err \leftarrow |x_n - x|$ 
13:       $x \leftarrow x_n$ 
14:       $fx \leftarrow f(x)$ 
15:       $dfx \leftarrow df(x)$ 
16:    Fin Si
17:  Fin Tantque
18: Fin Fonction
```

On cherche à résoudre des systèmes linéaires de la forme $\mathbb{A}\mathbf{x} = \mathbf{b}$ avec \mathbb{A} une matrice carrée d'ordre n .

6.1 RAPPELS

- **Produit matriciel.** Soient $\mathbb{A} = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq m}$ et $\mathbb{B} = (b_{ij})_{1 \leq i \leq m, 1 \leq j \leq p}$ deux matrices de taille $n \times m$ et $m \times p$ resp. Alors les coefficients de la matrice $\mathbb{C} = \mathbb{A}\mathbb{B}$ sont :

$$C_{ij} = \sum_{k=1}^m a_{ik} b_{kj}.$$
- Une matrice carrée est dite **triangulaire supérieure** lorsque $a_{ij} = 0$ dès que $i > j$ et est dite **triangulaire inférieure** lorsque $a_{ij} = 0$ dès que $i < j$.
- Une matrice carrée \mathbb{A} d'ordre n est dite **invertible ou régulière ou non singulière** s'il existe une matrice \mathbb{B} d'ordre n telle que $\mathbb{A}\mathbb{B} = \mathbb{B}\mathbb{A} = \mathbb{I}_n$, \mathbb{I}_n étant la matrice identité d'ordre n . On note alors \mathbb{B} , \mathbb{A}^{-1} . Une matrice qui n'est pas invertible est dite **singulière**.
- Le déterminant d'une matrice $\mathbb{A} = (a_{ij})$ d'ordre n est : $\det(\mathbb{A}) = \sum_{j=1}^n (-1)^{1+j} a_{1j} \det \mathbb{A}_{1j}$ avec \mathbb{A}_{1j} les sous-matrices d'ordre n .

Soient \mathbb{A} une matrice réelle d'ordre n et $\mathbf{b} \in \mathbb{R}^n$. On considère le système linéaire

$$\mathbb{A}\mathbf{x} = \mathbf{b}. \quad (S)$$

Proposition 22 (Existence et unicité.)

Le système (S) admet une unique solution si et seulement si l'une des propositions ci-dessous est vraie.

- 1 La matrice \mathbb{A} est inversible.
- 2 $\text{rg}(\mathbb{A}) = n$.
- 3 Le système homogène $\mathbb{A}\mathbf{x} = \mathbf{0}$ admet uniquement la solution nulle.

Ces trois propositions sont équivalentes.

Proposition 23

La solution du système (S) est donnée par la formule de Cramer :

$$x_j = \frac{D_j}{\det(\mathbb{A})}, \quad j = 1, \dots, n.$$

D_j est le déterminant de la matrice obtenue en remplaçant la j -ème colonne de \mathbb{A} par \mathbf{b} .

On peut montrer que le nombre de multiplications requises pour calculer le déterminant d'une matrice carrée $n \times n$ est supérieur à $n!$.

Si $n = 50$, résoudre un système linéaire $n \times n$ par la méthode de Cramer sur un ordinateur 1 gigaflop nécessite environ $4,8 \cdot 10^{49}$ années !!

Objectif : Réduire le temps de calcul pour résoudre un système linéaire en utilisant des méthodes de résolution exacte ou approchée.

- **méthodes directes** de résolution d'un système linéaire sont des méthodes numériques permettant d'obtenir une solution en un nombre fini d'itérations.
- **méthodes itératives** de résolution d'un système linéaire sont des méthodes numériques permettant d'obtenir une solution en un nombre infini (théoriquement) d'itérations.

6.2 MÉTHODE DE GAUSS

Objectif : Transformer le système $Ax = b$ en un système équivalent de la forme $Ux = \hat{b}$, où U est une matrice triangulaire supérieure et \hat{b} est le second membre convenablement modifié.

Définition 20

On appelle **opérations élémentaires** sur les lignes d'une matrice les trois opérations suivantes :

- i) **Échange** de deux lignes ($L_i \leftrightarrow L_j$).
- ii) **Multiplication** d'une ligne par une constante non nulle ($L_i \leftarrow \lambda L_i$).
- iii) **Substitution** : remplacer une ligne par elle-même plus un multiple d'une autre ligne ($L_i \leftarrow L_i + \lambda L_j$).

On rappelle que deux systèmes sont dits **équivalents** s'ils ont le même ensemble de solutions.

Théorème 17

On ne modifie pas l'ensemble des solutions d'un système linéaire en appliquant une ou plusieurs opérations élémentaires à sa matrice augmentée.

Méthode (du pivot de Gauss)

Soit $A = (a_{ij})$.

Itération k : en permutant éventuellement deux lignes du système (S) , on peut supposer que $a_{kk} \neq 0$.

On transforme toutes les lignes L_i avec $i > k$ de la façon suivante :

$$L_i \leftarrow L_i - \frac{a_{ik}}{a_{kk}} L_k.$$

Le terme a_{kk} est appelé **pivot**.

Données :

$A = (a_{ij})$: matrice carrée d'ordre n , telle que $a_{ii} \neq 0$,

$b = (b_i)$: vecteur b de \mathbb{R}^n ,

Résultat :

A : la matrice A échelonnée.

Données : $A = (a_{ij})$: matrice carrée triangulaire supérieure

d'ordre n telle que $a_{ii} \neq 0$,

$b = (b_i)$: vecteur de \mathbb{R}^n .

Résultat :

$x = (x_j)$: le vecteur solution.

```
1: Fonction GAUSS(A,b)
2:   Pour  $k \leftarrow 1$  à  $n - 1$  faire
3:     Pour  $i \leftarrow k + 1$  à  $n$  faire
4:        $m_{ik} \leftarrow a_{ik} / a_{kk}$ 
5:        $b_i \leftarrow b_i - m_{ik} * b_k$ 
6:       Pour  $j \leftarrow k + 1$  à  $n$  faire
7:          $a_{ij} \leftarrow a_{ij} - m_{ik} * a_{kj}$ 
8:       Fin Pour
9:     Pour  $j \leftarrow 1$  à  $k$  faire
10:       $a_{ij} \leftarrow 0$ 
11:    Fin Pour
12:  Fin Pour
13: Fin Fonction
```

```
1: Fonction REMONTEE(A,b)
2:    $x_n \leftarrow b_n / a_{nn}$ 
3:   Pour  $i \leftarrow n - 1$  à 1 faire
4:      $s \leftarrow 0$ 
5:     Pour  $j \leftarrow i + 1$  à  $n$  faire
6:        $s \leftarrow s + a_{ij} * x_j$ 
7:     Fin Pour
8:      $x_i \leftarrow (b_i - s) / a_{ii}$ 
9:   Fin Pour
10: Fin Fonction
```

Remarques

- L'algorithme de Gauss ci-dessus ne prend pas en compte la recherche d'un pivot non nul ! Dans la pratique il faut rajouter cette étape.
- Le nombre d'opérations nécessaires à la résolution d'un système de taille n

▶ pour l'algorithme de Gauss à l'itération k est :

- ★ $(n - k) + (n - k)(n - k + 1) = (n - k)(n - k + 2)$ multiplications/divisions,
- ★ $(n - k)(n - k + 1)$ additions/soustractions,

le nombre total d'opérations pour l'algorithme de Gauss est donc $\frac{1}{6}(2n^3 + 3n^2 - 5n) + \frac{1}{3}(n^3 - n)$.

▶ pour l'algorithme de substitution rétrograde :

- ★ $\frac{1}{2}(n^2 + n)$ multiplications/divisions,
- ★ $\frac{1}{2}(n^2 - n)$ additions/soustractions.

Le nombre total d'opérations est : $(\frac{n^3}{3} + n^2 - \frac{n}{3}) + (\frac{n^3}{3} + \frac{n^2}{2} - \frac{5n}{6})$, c'est-à-dire de l'ordre de $O(\frac{2}{3}n^3)$.

Définition 21

Une matrice est une **matrice élémentaire** si elle résulte d'une unique opération élémentaire effectuée sur les lignes de la matrice identité.

Il existe donc trois types de matrices élémentaires, correspondant aux trois types d'opérations élémentaires :

- \mathbb{E}_a : multiplier une (ou plusieurs) ligne(s) de la matrice identité par $a \neq 0$,
- $\tilde{\mathbb{E}}$: permuter deux lignes de la matrice identité,
- $\mathbb{E}_j(c)$: ajouter c fois la j -ème ligne de la matrice identité à une (ou plusieurs) de ses autres lignes.

Proposition 24

Soient \mathbb{A} une matrice quelconque et \mathbb{E} une matrice élémentaire résultant d'une certaine opération élémentaire effectuée sur une lignes de la matrice identité. Alors, la matrice $\mathbb{E}\mathbb{A}$ est le résultat de la même opération élémentaire appliquée aux lignes de \mathbb{A} .

Proposition 25

Toute matrice élémentaire est inversible, et on a :

- $[\mathbb{E}_a]^{-1} = \mathbb{E}_{\frac{1}{a}}$, $a \neq 0$,
- $[\tilde{\mathbb{E}}]^{-1} = \tilde{\mathbb{E}}$,
- $[\mathbb{E}_j(c)]^{-1} = \mathbb{E}_j(-c)$.

Soient $a_{ij}^{(k)}$ les coefficients de la matrice obtenue à la k -ème itération de l'algorithme de Gauss appliquée à \mathbb{A} d'ordre n .

Supposons que tous les pivots $a_{kk}^{(k)}$ sont non nuls.

La matrice élémentaire, $\mathbb{E}^{(k)}$, correspondant à la k -ème itération de l'algorithme de Gauss vaut :

$$\mathbb{E}^{(k)} = \begin{pmatrix} 1 & 0 & & & & & & \\ & \ddots & & & & & & \\ 0 & & \ddots & & & & & \\ \vdots & & & 1 & & & & 0 \\ \vdots & & & & & & & \\ \vdots & & -m_{k+1 k} & & 1 & & & 0 \\ \vdots & & & & & & & \\ \vdots & & & & & \ddots & & 0 \\ 0 & & -m_{nk} & & & & & 1 \end{pmatrix}, \text{ avec } m_{ik} = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}, \text{ pour } i = k + 1, \dots, n.$$

La matrice triangulaire supérieure, \mathbb{U} , obtenue après (au plus) n itérations de l'algorithme de Gauss s'écrit :

$$\mathbb{U} = \mathbb{E}^{(n-1)} \mathbb{E}^{(n-2)} \dots \mathbb{E}^{(1)} \mathbb{A}.$$

6.3 DÉCOMPOSITION (FACTORISATION) LU

Principe : Écrire la matrice \mathbb{A} comme produit d'une matrice \mathbb{L} triangulaire inférieure et d'une matrice \mathbb{U} triangulaire supérieure.

$$\mathbb{L} = \begin{pmatrix} l_{11} & 0 & 0 & \dots & 0 \\ l_{21} & l_{22} & 0 & \dots & 0 \\ l_{31} & l_{32} & l_{33} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \dots & l_{nn} \end{pmatrix} \quad \mathbb{U} = \begin{pmatrix} u_{11} & u_{12} & u_{13} & \dots & u_{1n} \\ 0 & u_{22} & u_{23} & \dots & u_{2n} \\ 0 & 0 & u_{33} & \dots & u_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & u_{nn} \end{pmatrix}$$

Ainsi la solution du système $\mathbb{A}\mathbf{x} = \mathbf{b}$ est obtenue en résolvant successivement les deux systèmes triangulaires $\mathbb{L}\mathbf{y} = \mathbf{b}$ et $\mathbb{U}\mathbf{x} = \mathbf{y}$.

- Les matrices \mathbb{L} et \mathbb{U} ainsi définies ne sont pas uniques. En effet, il y a n^2 coefficients dans la matrice \mathbb{A} et $n^2 + n$ coefficients inconnus dans les matrices \mathbb{L} et \mathbb{U} .
- On impose par exemple aux coefficients diagonaux de \mathbb{L} d'être égaux à 1.

Appliquons l'algorithme de Gauss à la matrice \mathbb{A} et supposons que tous les pivots $a_{kk}^{(k)}$ sont non nuls. L'écriture matricielle de l'algorithme de Gauss, $\mathbf{U} = \mathbb{E}^{(n-1)}\mathbb{E}^{(n-2)} \dots \mathbb{E}^{(1)}\mathbb{A}$, conduit à

$$\mathbb{A} = [\mathbb{E}^{(n-1)}\mathbb{E}^{(n-2)} \dots \mathbb{E}^{(1)}]^{-1}\mathbf{U}.$$

En posant $\mathbf{L} = [\mathbb{E}^{(n-1)}\mathbb{E}^{(n-2)} \dots \mathbb{E}^{(1)}]^{-1} = (\mathbb{E}^{(1)})^{-1} \dots (\mathbb{E}^{(n-1)})^{-1}$, on obtient bien la décomposition $\mathbb{A} = \mathbf{L}\mathbf{U}$. On vérifie que :

- la matrice \mathbf{U} est triangulaire supérieure,
- la matrice \mathbf{L} ainsi définie est une matrice triangulaire inférieure,
- la k -ième colonne de \mathbf{L} est la k -ième colonne de $(\mathbb{E}^{(k)})^{-1}$ vaut

$$(0 \dots 1 \ m_{k+1k} \ \dots \ m_{nk})^t.$$

Quelles sont les matrices ne nécessitant pas de permutation de lignes dans l'algorithme de Gauss ?

Définition 22

Soit $\mathbb{A} = (a_{ij})_{1 \leq i, j \leq n}$ une matrice carrée $n \times n$. La matrice \mathbb{A} est dite à **diagonale strictement dominante** si $|a_{ii}| > \sum_{j=1, j \neq i}^n |a_{ij}|, i = 1, \dots, n$.

Proposition 26

Une matrice à diagonale strictement dominante est inversible.
Dans ce cas, l'algorithme de Gauss peut-être mené sans permutation de lignes.

Définition 23

On appelle **sous-matrices diagonales d'ordre k** les matrices définies pour $k = 1, \dots, n$ par

$$\Delta^k = \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \ddots & \dots \\ a_{k1} & \dots & a_{kk} \end{pmatrix}.$$

Théorème 18 (Existence et Unicité de la décomposition)

Soit $\mathbb{A} = (a_{i,j})_{1 \leq i,j \leq n}$ une matrice d'ordre n . Si les sous-matrices diagonales d'ordre k de \mathbb{A} sont inversibles, alors la décomposition LU de \mathbb{A} existe et est unique.

Que faire si un pivot est nul ?

On échange deux lignes de la matrice \mathbb{A} , autrement dit on applique une matrice de permutation à \mathbb{A} .

Proposition 27

Pour toute matrice \mathbb{A} , il existe au moins une matrice de permutation \mathbb{P} telle que la matrice $\mathbb{P}\mathbb{A}$ admette une décomposition LU.

1. De la définition du produit matriciel et du fait que \mathbb{L} et \mathbb{U} sont des matrices triangulaires on a pour $1 \leq i, j \leq n$

$$a_{ij} = \sum_{s=1}^n l_{is} u_{sj} = \sum_{s=1}^{\min(i,j)} l_{is} u_{sj}.$$

On en déduit que

$$a_{kj} = \sum_{s=1}^{k-1} l_{ks} u_{sj} + l_{kk} u_{kj}, \quad j = k, \dots, n,$$

$$a_{ik} = \sum_{s=1}^{k-1} l_{is} u_{sk} + l_{ik} u_{kk}, \quad i = k+1, \dots, n.$$

2. On obtient d'abord la k -ième ligne de \mathbb{U}

$$u_{kj} = a_{kj} - \sum_{s=1}^{k-1} l_{ks} u_{sj}, \quad j = k, \dots, n,$$

3. puis la k -ième colonne de \mathbb{L}

$$l_{ik} = \frac{1}{u_{kk}} \left(a_{ik} - \sum_{s=1}^{k-1} l_{is} u_{sk} \right), \quad i = k+1, \dots, n.$$

6.4 DÉCOMPOSITION DE CHOLESKY

Principe : Pour une matrice \mathbb{A} symétrique définie positive faire une décomposition $\mathbb{L}\mathbb{U}$ où $\mathbb{U} = \mathbb{L}^T$.

Définition 24

La matrice \mathbb{A} est dite **symétrique définie positive** si elle est symétrique et si $x\mathbb{A}x^T > 0$ pour tout vecteur x de \mathbb{R}^n non nul.

Si \mathbb{A} est une matrice définie positive alors

- i) \mathbb{A} est inversible.
- ii) $a_{ii} > 0$ pour $i = 1, \dots, n$.
- iii) $\max_{i \leq k, j \leq n} |a_{kj}| \leq \max_{1 \leq i \leq n} |a_{ii}|$.

Proposition 28

Une matrice est définie positive si et seulement si ses sous-matrices diagonales ont un déterminant strictement positif.

Lorsque \mathbb{A} est définie il existe une matrice \mathbb{M} inversible telle que $\mathbb{A} = \mathbb{M}\mathbb{M}^T$. Ici \mathbb{M} n'est pas unique.

Théorème 19

Soit $\mathbb{A} \in \mathbb{R}^{n \times n}$ une **matrice symétrique définie positive**. Alors, il existe une unique matrice triangulaire inférieure \mathbb{H} dont les coefficients diagonaux sont strictement positifs telle que

$$\mathbb{A} = \mathbb{H}\mathbb{H}^T.$$

COMMENT CALCULE-T-ON LA MATRICE \mathbb{H} ?

De la définition du produit matriciel et du fait que \mathbb{L} et \mathbb{U} sont des matrices triangulaires on a

$$a_{ij} = \sum_{s=1}^n h_{is} h_{sj} = \sum_{s=1}^{\min(i,j)} h_{is} h_{sj}.$$

On en déduit que $h_{11} = \sqrt{a_{11}}$ et pour $i = 2, \dots, n$

$$h_{ij} = \left(a_{ij} - \sum_{k=1}^{j-1} h_{jk} h_{ik} \right) / h_{jj}, \quad j = 1, \dots, i-1,$$

$$h_{ii} = \sqrt{ \left(a_{ii} - \sum_{k=1}^{i-1} h_{ik}^2 \right) }.$$

POURQUOI FAIRE UNE DÉCOMPOSITION LU ?

- ① La décomposition LU nécessite un nombre d'opérations de l'ordre de $\frac{n^3}{3}$.
- ② La décomposition LU permet de calculer \mathbb{A}^{-1} , $\det(\mathbb{A})$, $\text{rg}(\mathbb{A})$, $\text{Ker}(\mathbb{A})$
- ③ Une approximation de \mathbb{L} et de \mathbb{U} fournit un préconditionneur de \mathbb{A} .

POURQUOI FAIRE UNE DÉCOMPOSITION DE CHOLESKY?

- ① La décomposition de Cholesky nécessite un nombre d'opérations de l'ordre de $\frac{n^3}{6}$, alors que la méthode de Gauss nécessite un nombre d'opérations de l'ordre de $\frac{n^3}{3}$.
- ② La décomposition de Cholesky permet de calculer \mathbb{A}^{-1} , $\det(\mathbb{A})$ de calculer le conditionnement de \mathbb{A} , de calculer un préconditionneur.