ENSIM
École d'ingénieurs
Le Mans Université

Le Mans Université

$3^{\text{éme}}$ **Année**

Spécialité Acoustique et Instrumentation

Mathématiques de l'ingénieur 2

Responsable : Y.SERRESTOU

ENSEIGNANTS : D.CACITTI-HOLLAND,Y.ESSTAFA, Y.SERRESTOU

ENSIM - Université du Maine - Rue Aristote - 72085 LE MANS Cedex 9 - France
Tél : 02 43 83 35 93 - Télécopie : 02 43 83 37 94

Table des matières

Introduction

Dans plusieurs domaines d'ingénierie (mécanique, physique, électrotechnique, électronique, traitement du signal, analyse des données, etc.), des problèmes concrets et complexes, nécessitent pour leur résolution le passage par le «calcul numérique ». Une définition, de cette discipline, est donnée par Nick Trefethen d'Oxford : «le calcul numérique est une discipline qui traite de la définition, l'analyse et l'implémentation d'algorithmes pour la résolution numérique des problèmes mathématiques continus qui proviennent de la modélisation des phénomènes réels».

Ce cours, s'inscrivant dans ce cadre, présente quelques méthodes et algorithmes numériques pour la résolution de problèmes, que nous avons traités d'un point de vie analytique en premier semestre. L'analyse de ces algorithmes permet d'aborder les notions de complexité, de convergence et de stabilité numérique. Ce cours est axé sur les trois problèmes centraux suivants :

— Résolution de systèmes d'équations linéaires

— Intégration numérique et résolution numérique d'équations différentielles

— Calcul de la transformée de Fourier

Dans une première partie de ce document, nous aborderons quelques méthodes de résolution de systèmes d'équations linéaires. Ils seront ainsi présentés l'algorithme du pivot de Gauss, l'algorithme de factorisation LU et l'algorithme de factorisation de Cholesky. Dans la seconde partie , des méthodes d'intégration et de résolution numérique du problème de Cauchy seront présentées. Dans la troisième partie, nous revenons sur la transformée de Fourier en présentant des algorithmes de calcul de sa version discrète et ses applications. Ce cours ne vaut pas être exhaustif pour recouvrir les différentes thèmes d'analyse numérique, mais il permet d'avoir un aperçu de cette discipline. Ce cours est complété par des exercices reportés à la fin du polycopié et trois TP.

⋆

Première partie

Résolution numérique de systèmes d'équations linéaires

1 Pivot de Gauss

1.1 Introduction et remarques

On s'intéresse à la résolution numérique de systèmes linéaires, i.e. de la forme avec

$$\left\{\begin{array}{ccccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \dots & + & a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \dots & + & a_{2,n}x_n & = & b_2 \\ \vdots & + & \vdots & + & \dots & + & \vdots & = & \vdots \\ a_{n,1}x_1 & + & a_{n,2}x_2 & + & \dots & + & a_{n,n}x_n & = & b_n \end{array}\right. , \tag{1}$$

où $a_{11}, ..., a_{nn}$et $b_1, ..., b_n$ sont les paramètres du système et $x_1, ..., x_n$ sont les inconnues que l'on cherche à déterminer.

Exemple 1.

$$\left\{\begin{array}{ccccccc} x_1 & + & 2x_2 & + & 3x_3 & = & 8 \\ 3x_1 & - & 2x_2 & + & 3x_3 & = & 6 \\ -x_1 & + & 3x_2 & + & 5x_3 & = & 1. \end{array}\right. \tag{2}$$

Remarque 1. Le système de 1 peut être écrit sous forme matricielle :

$$Ax = b$$

avec A est une matrice carrée inversible de taille n et x et b deux vecteurs colonnes de taille n :

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{et} \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

Exemple 2. Pour notre exemple 2 on a :

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & -2 & 3 \\ -1 & 3 & 5 \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad b = \begin{pmatrix} 8 \\ 6 \\ 1 \end{pmatrix}.$$

Remarque 2. En théorie nous avons directement

$$x = A^{-1}b,$$

avec A^{-1} la matrice inverse de la matrice A. Cependant le calcul direct de la matrice inverse est très coûteux en termes d'opérations arithmétiques. Il existe des méthodes plus optimales en termes d'opérations.

1.2 Résolution des systèmes triangulaires

1.2.1 Système triangulaire supérieur

La méthode du pivot de Gauss est basée sur la propriété suivante :

Proposition 1. Si le système est triangulaire supérieur i.e. de la forme

$$\left\{\begin{array}{ccccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \dots & + & a_{1,n}x_n & = & b_1 \\ & & a_{2,2}x_2 & + & \dots & + & a_{2,n}x_n & = & b_2 \\ & & & & \ddots & & \vdots & = & \vdots \\ & & & & a_{n-1,n-1}x_{n-1} & + & a_{n-1,n}x_n & = & b_{n-1} \\ & & & & & & a_{n,n}x_n & = & b_n \end{array}\right. , \tag{3}$$

alors, d'après la dernière équation on obtient directement $x_n = \dfrac{b_n}{a_{nn}}$ puis x_{n-1} grâce à l'avant-dernière ligne, ainsi de suite jusqu'à x_1. Cette méthode s'appelle la méthode de remontée :

$$\left\{\begin{array}{ccc} x_n & = & \dfrac{b_n}{a_{n,n}} \\ x_{n-1} & = & \dfrac{b_{n-1} - a_{n-1,n}x_n}{a_{n-1,n-1}} \\ \vdots & & \vdots \\ x_1 & = & \dfrac{b_1 - a_{1,2}x_2 - ... - a_{1,n}x_n}{a_{1,1}} \end{array}\right.$$

Remarque 3. Il n'y a pas de problème de division par 0 car le système est supposé résoluble i.e. la matrice triangulaire supérieure A associée est inversible :

$$\det(A) = a_{11} \times ... \times a_{nn} \neq 0.$$

Exemple 3. Si on considère le système

$$\left\{\begin{array}{ccccccccc} x_1 & + & 2x_2 & + & 3x_3 & = & 8 \\ & - & 8x_2 & - & 6x_3 & = & -18 \\ & & & & \dfrac{17}{4}x_3 & = & -\frac{9}{4}. \end{array}\right. , \tag{4}$$

alors la solution (x_1, x_2, x_3) est donnée par

$$\left\{\begin{array}{ccccccc} x_3 & = & -\dfrac{9}{17} & & & & \\ x_2 & = & \dfrac{-18 + 6x_3}{-8} & = & \dfrac{18 \times 17 + 6 \times 9}{8 \times 17} & = & \dfrac{360}{136} = \dfrac{45}{17} \\ x_1 & = & 8 - 2x_2 - 3x_3 & = & \dfrac{8 \times 17 - 2 \times 45 + 3 \times 9}{17} & = & \dfrac{73}{17} \end{array}\right. \tag{5}$$

1.2.2 Système triangulaire inférieur

Le même raisonnement peut être fait pour un système triangulaire inférieur i.e. de la forme :

$$\left\{\begin{array}{ccccccccc} a_{1,1}x_1 & & & & & & & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & & & & & = & b_2 \\ & & & & \ddots & & \vdots & = & \vdots \\ a_{n-1,1}x_1 & + & \dots & + & a_{n-1,n-1}x_{n-1} & + & & = & b_{n-1} \\ a_{n,1}x_1 & + & \dots & + & a_{n,n-1}x_2 & + & a_{n,n}x_n & = & b_n \end{array}\right. , \tag{6}$$

Dans ce cas on parle de la méthode de descente :

$$\left\{\begin{array}{lcl} x_1 & = & \dfrac{b_1}{a_{1,1}} \\ x_2 & = & \dfrac{(b_2 - a_{2,1}x_1)}{a_{2,2}} \\ \vdots & & \vdots \\ x_1 & = & \dfrac{(b_n - a_{n,1}x_1 - \ldots - a_{n,n-1}x_{n-1})}{a_{n,n}} \end{array}\right. , \tag{7}$$

1.2.3 Complexité de ces méthodes

Ces méthodes nécessitent :

- n divisions,
- $0 + 1 + 2 + .. + (n-1) = \frac{n(n-1)}{2}$ additions(soustractions),
- $0 + 1 + 2 + .. + (n-1) = \frac{n(n-1)}{2}$ multiplications.

1.3 Méthode du pivot de Gauss

1.3.1 Principe général

La méthode de Gauss pour résoudre le système linéaire :

$$\left\{\begin{array}{ccccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \ldots & + & a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \ldots & + & a_{2,n}x_n & = & b_2 \\ \vdots & + & \vdots & + & \ldots & + & \vdots & = & \vdots \\ a_{n,1}x_1 & + & a_{n,2}x_2 & + & \ldots & + & a_{n,n}x_n & = & b_n \end{array}\right. , \tag{8}$$

que l'on note sous forme matricielle

$$Ax = b\,. \tag{9}$$

consiste à :

1. chercher un système triangulaire équivalent de la forme

$$MAx = Mb \tag{10}$$

où M est une matrice inversible telle que la matrice MA soit triangulaire supérieure ;

2. résoudre le système linéaire 10 par la méthode de la remontée.

Remarque 4. En pratique, on ne calcule pas explicitement la matrice M, mais la matrice MA et le vecteur Mb. L'introduction de la matrice M est une commodité d'écriture.

1.3.2 Méthode de triangularisation

On considère le système initial :

$$\underbrace{\begin{pmatrix} a_{1,1}^{(0)} & a_{1,2}^{(0)} & \ldots & a_{1,n}^{(0)} \\ a_{2,1}^{(0)} & a_{2,2}^{(0)} & \ldots & a_{2,n}^{(0)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1}^{(0)} & a_{2,1}^{(0)}a_{n,2}^{(0)} & \ldots & a_{n,n}^{(0)} \end{pmatrix}}_{A=A^0} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \underbrace{\begin{pmatrix} b_1^0 \\ b_2^0 \\ \vdots \\ b_n^0 \end{pmatrix}}_{b=b^0} \tag{11}$$

1. On choisit un indice $i_1 \in [\![1, n]\!]$ tel que $a_{i_1,1} \neq 0$, qu'on appelle le pivot de l'élimination. Ceci existe car au moins un des éléments $a_{i,1}$,$i \in [\![1, n]\!]$, n'est pas nul sans quoi la matrice A ne serait pas inversible.
2. On échange la première ligne du système 11 avec la ligne i du pivot. Cela revient à multiplier la matrice A à gauche par une matrice $T(1, i)$ de transposition de la forme

$$T_1 = T(1,i) = \begin{pmatrix} 0 & \dots & \dots & \dots & 1 & & & \\ \vdots & 1 & & & \vdots & & & \\ \vdots & & 1 & & \vdots & & & \\ \vdots & & & 1 & \vdots & & & \\ 1 & \dots & \dots & \dots & 0 & & & \\ & & & & & 1 & & \\ & & & & & & 1 & \\ & & & & & & & 1 \end{pmatrix}.$$

3. Pour $i \in [\![2, n]\!]$ on remplace la ligne L_i par et en remplaçant

$$L_i \longleftarrow L_i - \ell_{i,1} \times L_1, \ o\grave{u}, \ell_{i,1} = \frac{a^{(0)}_{i,1}}{a^{(0)}_{1,1}}.$$

On obtient le système équivalent

$$\underbrace{\begin{pmatrix} a^{(1)}_{1,1} & a^{(1)}_{1,2} & \dots & a^{(1)}_{1,n} \\ 0 & a^{(1)}_{2,2} & \dots & a^{(1)}_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a^{(1)}_{n,2} & \dots & a^{(1)}_{n,n} \end{pmatrix}}_{A^1} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \underbrace{\begin{pmatrix} b^1_1 \\ b^1_2 \\ \vdots \\ b^1_n \end{pmatrix}}_{b^1}, \tag{12}$$

avec

$$a^{(1)}_{i,j} = a^{(0)}_{i,j} - \ell_{i,1} a^{(0)}_{1,j} \quad \text{et} \quad b^{(1)}_i = b^{(0)}_i - \ell_{i,1} b^{(0)}_1$$

En écriture matricielle, cela revient à multiplier la matrice $T_1 A$ à gauche par la matrice

$$E_1 = \begin{pmatrix} 1 & & & & \\ -\frac{a^{(0)}_{2,1}}{a^{(0)}_{1,1}} & 1 & & & \\ -\frac{a^{(0)}_{3,1}}{a^{(0)}_{1,1}} & & 1 & & \\ \vdots & & & \ddots & \\ -\frac{a^{(0)}_{n,1}}{a^{(0)}_{1,1}} & & & & 1 \end{pmatrix}$$

4. On recommence les étapes précédentes avec les autres lignes pour obtenir, après $n - 1$

étapes, un système triangulaire supérieur :

$$\left\{\begin{array}{ccccccccc} a_{1,1}^{(n)}x_1 & + & a_{1,2}^{(n)}x_2 & + & \dots & + & a_{1,n}^{(n)}x_n & = & b_1^{(n)} \\ & & a_{2,2}^{(n)}x_2 & + & \dots & + & a_{2,n}^{(n)}x_n & = & b_2^{(n)} \\ & & & & \ddots & + & \vdots & = & \vdots \\ & & & & & & a_{n,n}^{(n)}x_n & = & b_n^{(n)} \end{array}\right. , \tag{13}$$

Matriciellement nous sommes arrivés à :

$$\underbrace{E_{n-1}T_{n-1}\dots E_2T_2E_1T_1}_{M}Ax = \underbrace{E_{n-1}T_{n-1}\dots E_2T_2E_1T_1}_{M}b$$
$$MAx = Mb$$

avec MA est une matrice triangulaire supérieure.

Remarque 5.

1. la matrice qui permet la permutation de deux lignes vérifie

$$\det(T(i,i)) = 1 \quad \text{et} \quad \det(T(i,j)) = -1,\ i \neq j\,.$$

2. selon que l'on a effectué un nombre pair d'échanges de lignes $(+)$ ou un nombre impair $(-)$, on obtient, au passage, un procédé rapide du calcul du déterminant.

$$\det(A) = \pm a_{1,1}^{(n)}a_{2,2}^{(n)}\dots a_{n-1,n-1}^{(n-1)}a_{n,n}^{(n)}$$

Exemple 4. On considère l'exemple précédent 2

$$(S)\begin{cases} x_1 & + & 2x_2 & + & 3x_3 & = & 8 \\ 3x_1 & - & 2x_2 & + & 3x_3 & = & 6 \\ -x_1 & + & 3x_2 & + & 5x_3 & = & 1. \end{cases}$$

1. **Itération 1** : le coefficient $a_{11} = 1$ est non nul, on peut le prendre comme pivot. Autrement dit $T_1 = I_3$ est la matrice identité. Ensuite on effectue les opérations

$$\begin{array}{rcl} L_2 & \longleftarrow & L_2 - 3L_1 \\ L_3 & \longleftarrow & L_3 + L_1. \end{array}$$

Pour obtenir le système

$$(S_1)\begin{cases} x_1 & + & 2x_2 & + & 3x_3 & = & 8 \\ & - & 8x_2 & - & 6x_3 & = & -18 \\ & & 5x_2 & + & 8x_3 & = & 9. \end{cases}$$

Matriciellement on a

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix},\ A^{(1)} = E_1 A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -8 & -6 \\ 0 & 5 & 8 \end{pmatrix}, b^{(1)} = E_1 b = \begin{pmatrix} 8 \\ -18 \\ 9 \end{pmatrix}$$

et

$$Ax = b \Longleftrightarrow A^{(1)}x = b^1$$

2. **Itération** 2 : le coefficient $a_{22}^{(1)} = -8$ est non nul, on peut le prendre comme pivot. Autrement dit $T_2 = I_3$ encore. Ensuite on effectue l'opération

$$L_3 \longleftarrow L_3 + \frac{5}{8}L_2.$$

Pour obtenir le système

$$(S_2)\begin{cases} x_1 & + & 2x_2 & + & 3x_3 & = & 8 \\ & - & 8x_2 & - & 6x_3 & = & -18 \\ & & & & \frac{17}{4}x_3 & = & -\frac{9}{4} \end{cases}$$

Matriciellement on a

$$E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{5}{8} & 1 \end{pmatrix},\ A^{(2)} = E_1 A^{(1)} = E_2 E_1 A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -8 & -6 \\ 0 & 0 & \frac{17}{4} \end{pmatrix}$$

$$b^{(2)} = E_1 b^{(1)} = E_2 E_1 b = \begin{pmatrix} 8 \\ -18 \\ \frac{-9}{4} \end{pmatrix}$$

3. **Résolution** : on obtient le système triangulaire supérieure de l'exemple 4 que l'on a résolu précédemment 5.

1.3.3 Choix du pivot

Parmi les coefficients non nuls de la première colonne que l'on considère, il peut y en avoir plusieurs non nuls. Il faudra faire un choix parmi eux. Cependant à causse des arrondis effectués par l'ordinateur pour écrire un nombre, il peut y avoir des erreurs de calcul avec plus ou moins d'impact. Pour comprendre ce phénomène, nous donnons tout d'abord un petit aperçu sur la représentation des nombres en virgule flottante, puis un exemple pour illustrer la propagtion des erreurs d'arrondi et l'impact du choix de pivot sur la précision de la résolution.

Aperçu sur la représentation des nombres en machine

Définition 1. La représentation d'un nombre flottant x en 32 bits est donnée par :

$$x = \pm a \times 2^e = \pm(1+m)2^e,$$

avec $m \in [0,1[$ est la mantisse, $a = 1+m \in [1,2[$ et $e \in \mathbb{Z}$ l'exposant tous deux écrits en système binaire. Pour le coder, l'ordinateur utilise :

— 1 bit pour le signe : 0 pour +, 1 pour −,

— 8 bits pour l'exposant : ce qui correspond à 256 valeurs possibles. Le 0 est réservé au nombre 0 et 255 pour l'infini, il reste donc 254 valeurs possibles pour l'exposant e. Pour que e puisse être négatif, nous avons donc

$$e \in [\![-126, 127]\!].$$

— 23 bits pour la mantisse.

Exemple 5. Si l'on considère le nombre

$$x = -118,625.$$

Alors le bit pour le signe est $s = 1$. De plus nous avons

$$118,625 = 118 + 0,625,$$

avec

$$118 = 64 + 54 = 2^6 + 32 + 22 = 2^6 + 2^5 + 16 + 6 = 2^6 + 2^5 + 2^4 + 2^2 + 2^1,$$

et

$$0,625 = 0,5 + 0,125 = 2^{-1} + 2^{-3}.$$

Ainsi

$$118,625 = 2^6 + 2^5 + 2^4 + 2^2 + 2^1 + 2^{-1} + 2^{-3} = 2^6(1 + 2^{-1} + 2^{-2} + 2^{-4} + 2^{-5} + 2^{-7} + 2^{-9}),$$

autrement dit

$$e = 6, \quad m = 2^{-1} + 2^{-2} + 2^{-4} + 2^{-5} + 2^{-7} + 2^{-9}$$

Enfin $x = -118,625$ est codé par

$$\underbrace{1}_{s}\ \underbrace{00000111}_{e}\ \underbrace{101101010...0}_{m}.$$

Remarque 6. Le plus petit nombre positif est donc $2^{-126}(1 + 2^{-23})$ et le plus grand est

$$2^{127}(1 + 2^{-1} + ... + 2^{-23}) = 2^{127}(2 - 2^{-23}).$$

Exemple illustrant l'impact du choix de pivot sur la précision de la résolution

Exemple 6 (Exemple de Forsythe)**.**
Considérons le système linéaire

$$\left\{\begin{array}{rcrcl} 10^{-4}x_1 & + & x_2 & = & 1 \\ x_1 & + & x_2 & = & 2, \end{array}\right.$$

et supposons que la mantisse n'a que trois chiffres significatifs. On peut obtenir facilement la solution par méthode de substitution

$$x_1 = 1.0001, \quad x_2 = 0.9999.$$

Ainsi, avec seulement les trois chiffres significatifs, pour l'ordinateur la réponse exacte est

$$x_1 = 1, \quad x_2 = 1.$$

Essayons d'appliquer la méthode du pivot de Gauss avec 10^{-4} comme pivot. Nous arrivons à

$$\left\{\begin{array}{rcrcl} 10^{-4}x_1 & + & x_2 & = & 1 \\ & & -9999x_2 & = & -9998, \end{array}\right.$$

Ainsi

$$x_2 = \frac{9998}{9999}$$

ce qui donne avec les trois chiffres significatifs $x_2 = 1$. On en déduit $x_1 = 0$ ce qui est très éloigné de la vraie solution.
Cependant si on échange les deux lignes du système linéaire pour considérer $a_{21} = 1$ comme le pivot

$$\left\{\begin{array}{rcrcl} x_1 & + & x_2 & = & 1 \\ 10^{-4}x_1 & + & x_2 & = & 2. \end{array}\right.$$

Nous obtenons

$$\left\{\begin{array}{rcrcl} x_1 & + & x_2 & = & 1 \\ & & 0.999x_2 & = & 0.999. \end{array}\right.$$

Ainsi $x_2 = 1$ puis $x_1 = 1$ ce qui est très proche de la vraie solution.

Nous aboutissons donc à la méthode suivante pour choisir un bon pivot.

Proposition 2 (Choix du pivot)**.**
Au début de la k-ème étape, on choisit un pivot $a^{(k)}_{i_0,k}$ tel que

$$|a^{(k)}_{i_0,k}| = \max_{k\leq i\leq n} |a^{(k)}_{i,k}|.$$

Ainsi les erreurs d'arrondis dues à la division par un nombre très petit sont minimisées.

Proposition 3 (Choix du pivot total)**.**
Au début de la k-ème étape, on choisit un pivot $a^{(k)}_{i_0,j_0}$ tel que

$$|a^{(k)}_{i_0,j_0}| = \max_{k\leq i,j\leq n} |a^{(k)}_{i,j}|.$$

Pour se faire il faut non seulement échanger deux lignes de la matrice A mais également deux colonnes en multipliant à droite la matrice A par la matrice $T(k, j_0)$. On passe donc de $A^{(k)}x = b$ à

$$T(k, i_0)A^{(k)}T(k, j_0)x' = T(k, i_0)b, \quad x' = T(k, j_0)^{-1}x = T(k, j_0)x$$

où x' est x avec les lignes k et j_0 échangées. Dans ce cas les erreurs d'arrondis sont encore davantage minimisées.

1.3.4 Calcul du nombre d'opérations élémentaires

Soit $k \in [\![1, n-1]\!]$. Alors à la k-ième étape nous effectuons les $n-k$ opérations

$$L_i^{(k)} \longleftarrow L_i^{(k)} - \frac{a_{i,k}^{(k)}}{a_{k,k}^{(k)}} L_k^{(k)}, \quad k+1 \leq i \leq n.$$

Ce qui correspond pour le côté gauche du système avec les a_{ij} et les x_i à :

- $n-k$ divisions pour obtenir les termes $\dfrac{a_{i,k}^{(k)}}{a_{k,k}^{(k)}}$,
- $(n-k)^2$ additions (soustractions) car, pour tout $i \in [\![k+1, n]\!]$, $n-k$ soustractions sont effectuées pour la ligne $L_i^{(k)}$,
- $(n-k)^2$ multiplications pour la même raison.

Ainsi on obtient pour $n-1$ étapes :

- $\displaystyle\sum_{k=1}^{n-1}(n-k) = \frac{n(n-1)}{2}$ divisions,
- $\displaystyle\sum_{k=1}^{n-1}(n-k)^2 = \frac{n(n-1)(2n-1)}{6}$ additions,
- $\displaystyle\sum_{k=1}^{n-1}(n-k)^2 = \frac{n(n-1)(2n-1)}{6}$ multiplications.

Quant au côté droit correspondant aux opérations sur les b_i, nous avons :

- $\displaystyle\sum_{k=1}^{n-1}(n-k) = \frac{n(n-1)}{2}$ additions (soustractions),
- $\displaystyle\sum_{k=1}^{n-1}(n-k) = \frac{n(n-1)}{2}$ multiplications.

Il reste encore à compter les opérations effectuées dans la méthode de remontée :

$$\begin{cases} x_n &= \dfrac{b_n^{(n-1)}}{a_n^{(n-1)}} \\ x_{n-1} &= \dfrac{b_{n-1}^{(n-1)} - a_{n-1,n}^{(n-1)} x_n}{a_{n-1,n-1}^{(n-1)}} \\ \vdots & \vdots \\ x_1 &= \dfrac{b_1^{(n-1)} - a_{1,2}^{(n-1)} x_2 - \ldots - a_{1,n}^{(n-1)} x_n}{a_{1,1}^{(n-1)}} \end{cases}$$

Nous avons donc

- n divisions,
- $0 + 1 + \ldots + n - 1 = \dfrac{n(n-1)}{2}$ additions (soustractions),

- $0 + 1 + ... + n - 1 = \dfrac{n(n-1)}{2}$ multiplications.

Par conséquent, en effectuant la somme de toutes ces opérations, pour n tendant vers $+\infty$, nous obtenons le résultat suivant :

Proposition 4. La méthode de Gauss nécessite de l'ordre de

- $\frac{n^2}{2}$ divisions,
- $\frac{n^3}{3}$ multiplications,
- $\frac{n^3}{3}$ additions.

Remarque 7. Nous n'avons ici pas tenu compte du coût de calcul de la stratégie du pivot utilisée.

Nous pouvons également comparer ce résultat avec la méthode utilisant les formules de Cramer :

$$x_i = \frac{\det(B_i)}{\det(A)}, \quad B_i = \begin{pmatrix} a_{1,1} & \dots & a_{1,i-1} & b_1 & a_{1,i+1} & \dots & a_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,i-1} & b_n & a_{n,i+1} & \dots & a_{n,n} \end{pmatrix}.$$

Nous devons donc calculer $n+1$ déterminants et effectuer n divisions. Or le calcul d'un déterminant par la formule

$$\det(A) = \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} A_{i,\sigma(i)}$$

nécessite $n!$ additions et $n!n$ multiplications. Ainsi cette méthode requiert de l'ordre de

- n divisions,
- $(n+1)!$ additions,
- $(n+2)!$ multiplications.

Nous obtenons donc beaucoup plus d'opérations à effectuer.

1.4 Méthode de Gauss-Jordan

Remarquons que nous pouvons obtenir une matrice diagonale MA au lieu de triangulaire supérieure, en modifiant légèrement les matrices E_k. En effet si l'on effectue les opérations sur les lignes supérieures à la k-ième lors de la k-ième étape, on transforme E_k en

$$E_k = \begin{pmatrix} 1 & & (0) & -\frac{a_{1,k}^{(k-1)}}{a_{k,k}^{(k-1)}} & & \\ & \ddots & & \vdots & & (0) \\ & & 1 & -\frac{a_{k-1,k}^{(k-1)}}{a_{k,k}^{(k-1)}} & & \\ & & & 1 & & \\ & & & -\frac{a_{k+1,k}^{(k-1)}}{a_{k,k}^{(k-1)}} & 1 & \\ & (0) & & \vdots & & \ddots \\ & & & -\frac{a_{n,k}^{(k-1)}}{a_{k,k}^{(k-1)}} & (0) & & 1 \end{pmatrix}.$$

Ainsi on obtient à la fin un système simple à résoudre

$$\left\{\begin{array}{lllll} a_{1,1}^{(n-1)}x_1 & & & = & b_1^{(n-1)} \\ & \ddots & & & \vdots \\ & & a_{n-1,n-1}^{(n-1)}x_{n-1} & = & b_{n-1}^{(n-1)} \\ & & a_{n,n}^{(n-1)}x_n & = & b_n^{(n-1)}, \end{array}\right.$$

i.e., en multpliant les lignes par $\dfrac{1}{a_{i,i}^{(n-1)}}$,

$$\left\{\begin{array}{lllll} x_1 & & & = & \dfrac{b_1^{(n-1)}}{a_{1,1}^{(n-1)}} \\ & \ddots & & & \vdots \\ & & x_{n-1} & = & \dfrac{b_{n-1}^{(n-1)}}{a_{n-1,n-1}^{(n-1)}} \\ & & x_n & = & \dfrac{b_n^{(n-1)}}{a_{n,n}^{(n-1)}} \end{array}\right.$$

correspondant à la multiplication matricielle par

$$D = \begin{pmatrix} \dfrac{1}{a_{1,1}^{(n-1)}} & & \\ & \ddots & \\ & & \dfrac{1}{a_{n,n}^{(n-1)}} \end{pmatrix}$$

Remarque 8. La méthode de Gauss-Jordan permet également d'obtenir l'inverse matricielle. En effet nous avons

$$DE_{n-1}T_{n-1}...E_1T_1A = I_n,$$

i.e.

$$A^{-1} = DE_{n-1}T_{n-1}...E_1T_1.$$

Exemple 7. On considère la matrice

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & -2 & 3 \\ -1 & 3 & 5 \end{pmatrix}.$$

Alors $T_1 = I_3$ et

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

On obtient alors

$$E_1T_1A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -8 & -6 \\ 0 & 5 & 8 \end{pmatrix}$$

Puis $T_2 = I_3$ et

$$E_2 = \begin{pmatrix} 1 & \frac{2}{8} & 0 \\ 0 & 1 & 0 \\ 0 & \frac{5}{8} & 1 \end{pmatrix}.$$

On obtient alors

$$E_2T_2E_1T_1A = \begin{pmatrix} 1 & 0 & \frac{3}{2} \\ 0 & -8 & -6 \\ 0 & 0 & -\frac{17}{4} \end{pmatrix}.$$

Enfin $T_3 = I_3$ et

$$E_3 = \begin{pmatrix} 1 & 0 & -\frac{6}{17} \\ 0 & 1 & -\frac{24}{17} \\ 0 & 0 & 1 \end{pmatrix}.$$

Par conséquent

$$E_3T_3E_2T_2E_1T_1A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -8 & 0 \\ 0 & 0 & \frac{17}{4} \end{pmatrix}.$$

Il ne reste plus qu'à multiplier à gauche par

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{8} & 0 \\ 0 & 0 & \frac{4}{17} \end{pmatrix}.$$

Finalement nous avons

$$DE_3T_3E_2T_2E_1T_1A = I_3, \quad \text{i.e.} \quad A^{-1} = DE_3T_3E_2T_2E_1T_1.$$

Remarque 9. En pratique on écrit A et la matrice I_n à côté puis on effectue les opérations sur A et I_n. En effet nous avons

$$DE_3T_3E_2T_2E_1T_1I_3 = DE_3T_3E_2T_2E_1T_1 = A^{-1}.$$

Exemple 8. Les opérations pour la matrice A de l'exemple précédent sont les suivantes

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 3 & -2 & 3 & 0 & 1 & 0 \\ -1 & 3 & 5 & 0 & 0 & 1 \end{array}\right)$$

1. **Itération 1 :** $\begin{array}{l} L_2 \longleftarrow L_2 - 3L_1 \\ L_3 \longleftarrow L_3 + L_1 \end{array}$

$$\downarrow$$

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -8 & -6 & -3 & 1 & 0 \\ 0 & 5 & 8 & 1 & 0 & 1 \end{array}\right)$$

2. **Itération 2 :**
$$\begin{array}{l} L_1 \longleftarrow L_1 - \dfrac{2}{-8}L_2 = L_1 + \dfrac{1}{4}L_2 \\ L_3 \longleftarrow L_3 - \dfrac{5}{-8}L_2 = L_3 + \dfrac{5}{8}L_2 \end{array}$$

$$\downarrow$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & \frac{3}{2} & \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & -8 & -6 & -3 & 1 & 0 \\ 0 & 0 & \frac{17}{4} & -\frac{7}{8} & \frac{5}{8} & 1 \end{array}\right)$$

3. **Itération 3 :**
$$\begin{array}{l} L_1 \longleftarrow L_1 - \dfrac{\frac{3}{2}}{\frac{17}{4}}L_3 = L_1 - \dfrac{6}{17}L_3 \\ L_2 \longleftarrow L_2 - \dfrac{-6}{\frac{17}{4}}L_3 = L_2 + \dfrac{24}{17}L_3 \end{array}$$

$$\downarrow$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{19}{34} & \frac{1}{34} & -\frac{6}{17} \\ 0 & -8 & 0 & -\frac{72}{17} & \frac{32}{17} & \frac{24}{17} \\ 0 & 0 & \frac{17}{4} & -\frac{7}{8} & \frac{5}{8} & 1 \end{array}\right)$$

4. **Itération 4 :**
$$\begin{array}{l} L_2 \longleftarrow -\dfrac{1}{8}L_2 \\ L_3 \longleftarrow \dfrac{4}{17}L_3 \end{array}$$

$$\downarrow$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{19}{34} & \frac{1}{34} & -\frac{6}{17} \\ 0 & 1 & 0 & \frac{9}{17} & -\frac{4}{17} & -\frac{3}{17} \\ 0 & 0 & 1 & -\frac{7}{34} & \frac{5}{34} & \frac{4}{17} \end{array}\right).$$

Par conséquent
$$A^{-1} = \begin{pmatrix} \frac{19}{34} & \frac{1}{34} & -\frac{6}{17} \\ \frac{9}{17} & -\frac{4}{17} & -\frac{3}{17} \\ -\frac{7}{34} & \frac{5}{34} & \frac{4}{17} \end{pmatrix}$$

2 Factorisation LU

2.1 Théorème et remarque

Théorème 1. Soit A une matrice carrée d'ordre n telle que les n sous-matrices diagonales soient inversibles
$$\forall k \in [\![1, n]\!], \quad (a_{i,j})_{1 \leq i,j \leq k} \in GL_k(\mathbb{R}).$$

Alors il existe une unique matrice triangulaire inférieure L avec uniquement des 1 sur la diagonale et une matrice triangulaire supérieure U telles que

$$A = LU.$$

Remarque 10. Pour calculer la matrice U on peut utiliser la méthode du pivot de Gauss. En effet comme les n sous-matrices diagonales sont inversibles, il n'y a pas d'interversion de lignes, on a alors

$$U = A_n = E_{n-1}...E_1 A$$

triangulaire supérieure. Puis

$$L = (E_{n-1}...E_1)^{-1}$$

pour avoir $LU = A$ et L triangulaire inférieure avec uniquement des 1 sur la diagonale car

$$E_k^{-1} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & (0) & \\ & & 1 & & & \\ & & & 1 & & \\ & & & \frac{a_{k+1,k}^{(k-1)}}{a_{kk}^{(k-1)}} & 1 & \\ & (0) & & \vdots & & \ddots \\ & & & \frac{a_{nk}^{(k-1)}}{a_{kk}^{(k-1)}} & (0) & & 1 \end{pmatrix}$$

et ainsi

$$L = E_1^{-1}...E_{n-1}^{-1} = \begin{pmatrix} 1 & & & \\ \frac{a_{2,1}^{(0)}}{a_{1,1}^{(0)}} & 1 & & \\ \vdots & \ddots & \ddots & \\ \frac{a_{n,1}^{(0)}}{a_{1,1}^{(0)}} & \cdots & \frac{a_{n,n-1}^{(n-2)}}{a_{n-1,n-1}^{(n-2)}} & 1 \end{pmatrix}.$$

Remarque 11. L'intérêt de la factorisation LU est quand on à résoudre plusieurs systèmes linéaires avec la même matrice A. On effectue la factorisation LU puis il suffit pour chaque système

$$Ax = LUx = b$$

de résoudre par remontée ou descente les deux systèmes triangulaires

$$Ly = b, \quad Ux = y.$$

2.2 Matrices tridiagonales

Définition 2. On dit qu'une matrice A est tridiagonale si elle est de la forme

$$A = \begin{pmatrix} b_1 & c_1 & & (0) \\ a_2 & \ddots & \ddots & \\ & \ddots & \ddots & c_{n-1} \\ (0) & & a_n & b_n \end{pmatrix}.$$

Proposition 5. On considère A une matrice tridiagonale et la suite δ définie par récurrence

$$\delta_0 = 1, \quad \delta_1 = b_1, \quad \forall k \in [\![2, n]\!], \delta_k = b_k\delta_{k-1} - a_k c_{k-1}\delta_{k-2}.$$

Alors, en notant Δ_k la sous-matrice diagonale de taille k de A, nous avons

$$\delta_k = \det \Delta_k$$

et la factorisation $A = LU$ est donnée par

$$L = \begin{pmatrix} 1 & & & & \\ a_2\frac{\delta_0}{\delta_1} & 1 & & & \\ & a_3\frac{\delta_1}{\delta_2} & \ddots & & \\ & & \ddots & 1 & \\ & & & a_n\frac{\delta_{n-2}}{\delta_{n-1}} & 1 \end{pmatrix}, \quad U = \begin{pmatrix} \frac{\delta_1}{\delta_0} & c_1 & & \\ & \frac{\delta_2}{\delta_1} & \ddots & \\ & & \ddots & c_{n-1} \\ & & & \frac{\delta_n}{\delta_{n-1}} \end{pmatrix}.$$

Démonstration. Les termes diagonaux sont donnés par

$$a_{1,1} = 1 \times \frac{\delta_1}{\delta_0} = b_1$$

et pour tout $k \in [\![2, n]\!]$,

$$a_{k,k} = a_k\frac{\delta_{k-2}}{\delta_{k-1}} \times c_{k-1} + 1 \times \frac{\delta_k}{\delta_{k-1}} = b_k.$$

Puis pour les termes sur-diagonaux

$$a_{k,k+1} = 1 \times c_k = c_k, \quad k \in [\![1, n-1]\!],$$

et pour les termes sous-diagonaux

$$a_{k,k-1} = a_k\frac{a_{k-2}}{a_{k-1}} \times \frac{\delta_{k-1}}{\delta_{k-2}} = a_k, \quad k \in [\![2, n]\!].$$

Par conséquent, par unicité de L et U, nous avons bien les formules souhaitées. □

3 Factorisation de Cholesky

3.1 Théorème et remarque

Théorème 2. Soit A une matrice symétrique définie positive de taille $n \times n$. Alors il existe une unique matrice réelle B triangulaire inférieure dont tous les éléments diagonaux sont strictement positifs telle que

$$A = BB^t.$$

Démonstration. — **Existence :**

Soit $A = (a_{ij})_{1 \leq i,j \leq n}$ une matrice symétrique définie positive. On démontre tout d'abord que toutes les sous matrices diagonales de A sont symétriques définies positives et par conséquent elles sont inversibles.

Soit

$$\Delta_k = \begin{pmatrix} a_{11} & . & . & . & a_{1k} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ a_{k1} & . & . & . & a_{kk} \end{pmatrix}$$

une sous matrice diagonale de A. On va démontrer que la matrice Δ_k est symétrique définie positive, i.e. $\forall w \in \mathbb{R}^k : w^T \Delta_k w > 0$.

Soit $w \in \mathbb{R}^k$ et $v = \begin{pmatrix} w \\ 0 \\ . \\ . \\ . \\ 0 \end{pmatrix} \in \mathbb{R}^n$. On a :

$$w^T \Delta_k w = v^T A v.$$

Or A une matrice symétrique définie positive donc

$$w^T \Delta_k w = v^T A v > 0$$

Ce qui preuve que la matrice Δ_k est symétrique définie positive et inversible par conséquent. La matrice A vérifie, donc, la condition de la factorisation LU. Alors il existe une unique matrice triangulaire inférieure L (avec les élément diagonaux égaux à 1) et une unique matrice triangulaire supérieure U(avec des élément strictement positif sur la diagonal à cause du déterminant > 0) telles que :

$$A = LU$$

On considère ensuite la matrice diagonale

$$\Delta = \begin{pmatrix} \sqrt{U_{1,1}} & 0 & . & . & . & 0 \\ 0 & \sqrt{U_{2,2}} & . & . & . & . \\ . & 0 & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & 0 \\ 0 & 0 & . & . & 0 & \sqrt{U_{n,n}} \end{pmatrix}.$$

Nous avons alors

$$A = (L\Lambda).(\Lambda^{-1}U) = BC\,.$$

Comme A est symétrique, on a $BC = C^T B^T$ ou encore

$$C(B^T)^{-1} = B^{-1}C^T$$

et on calcule explicitement ces matrices (à gauche triangulaire supérieur, à droite triangulaire inférieure, chacune avec des 1 sur la diagonale). Légalité n'a lieu que si c'est l'identité ou encore $C = B^T$.

— **Unicité** : On suppose qu'il existe deux matrices réelles triangulaires inférieures B et C telles que :

$$A = BB^T = CC^T, \tag{14}$$

et on démontre que $B = C$

L'égalité 14 permet d'écrire : $C^{-1}B = C^T \left(B^T\right)^{-1} = C^T \left(B^{-1}\right)^T = \left(B^{-1}C\right)^T$.

Le terme $C^{-1}B$ est une matrice triangulaire inférieure et $\left(B^{-1}C\right)^T$ est une matrice triangulaire supérieure, l'égalité n'est possible que ces deux termes sont des matrices diagonales, c'est-à-dire :

$$C^{-1}B = \left(B^{-1}C\right)^T = D = \begin{pmatrix} d_{11} & 0 & . & . & . & 0 \\ 0 & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & 0 \\ 0 & . & . & . & . & d_{nn} \end{pmatrix},$$

avec $\forall i \in 1, .., n \ d_{ii} > 0$

Donc

$$\underbrace{\left(B^{-1}C\right)^T}_{D} = B^{-1}C = \underbrace{\left(C^{-1}B\right)^{-1}}_{D^{-1}}$$

et par conséquent

$$D = D^{-1} \Longleftrightarrow \forall i \in 1, .., n \ d_{ii} \in \{-1, 1\}.$$

Et vu que $\forall i \in 1, .., n \ d_{ii} > 0$, donc $\forall i \in 1, .., n \ d_{ii} = 1$, c'est-à-dire que

$$D = I_n$$

Enfin

$$D = I_n \Longleftrightarrow B^{-1}C = I_n \Longleftrightarrow B = C.$$

Ce qui prouve l'unicité de la décomposition

□

Proposition 6 (Méthode de Cholesky)**.** On considère A une matrice carrée de taille $n \times n$ symétrique définie positive et un vecteur colonne b de taille n. On note

$$B = \begin{pmatrix} b_{1,1} & & (0) \\ \vdots & \ddots & \\ b_{n,1} & \dots & b_{n,n} \end{pmatrix}$$

la matrice de la décomposition de Cholesky $A = BB^t$. Alors pour tout $i, j \in [\![1, n]\!]$ tels que $i \geq j$,

$$a_{i,j} = \sum_{k=1}^{j} b_{i,k} b_{j,k}.$$

Ainsi pour la première colonne $j = 1$ nous avons :

- pour $i = 1$: $a_{1,1} = b_{1,1}^2$ d'où $b_{1,1} = \sqrt{a_{1,1}}$,
- pour $i = 2$: $a_{2,1} = b_{2,1}b_{1,1}$ d'où $b_{2,1} = \dfrac{a_{2,1}}{b_{1,1}}$,
- $\vdots$
- pour $i = n$: $a_{n,1} = b_{1,1}b_{n,1}$ d'où $b_{n,1} = \dfrac{a_{n,1}}{b_{1,1}}$.

Nous avons donc la première colonne de la matrice B. Puis on procède par récurrence. On suppose connue les $j-1$ premières colonnes de la matrice B. Alors pour la j-ième colonne :

- pour $i=j$: $a_{j,j}=b_{j,1}^2+...+b_{j,j}^2$ d'où $b_{j,j}=\sqrt{a_{j,j}-b_{j,1}^2-...-b_{j,j-1}^2}$,
- pour $i=j+1$: $a_{j+1,j}=b_{j+1,1}b_{j,1}+...+b_{j+1,j}b_{j,j}$ d'où $b_{j+1,j}=\dfrac{a_{j+1,j}-b_{j+1,1}b_{j,1}-...-b_{j+1,j-1}b_{j,j-1}}{b_{j,j}}$,
- $\vdots$
- pour $i=n$: $a_{n,j}=b_{n,1}b_{j,1}+...+b_{n,j}b_{j,j}$ d'où $b_{n,j}=\dfrac{a_{n,j}-b_{n,1}b_{j,1}-...-b_{n,j-1},b_{j,j-1}}{b_{j,j}}$.

Ce qui détermine la matrice B. Il ne reste plus qu'à résoudre $By=b$ par la méthode de descente et $B^tx=y$ par la méthode de remontée.

Remarque 12. Nous avons

$$\det(A)=(b_{1,1}...b_{n,n})^2>0.$$

En particulier les $b_{j,j}$ sont non nuls.

Exemple 9. On considère la matrice

$$A=\begin{pmatrix}2&-1\\-1&2\end{pmatrix}.$$

Alors la matrice A est symétrique définie positive car symétrique et

$$\chi_A=(X-2)^2-1=(X-3)(X-1)\quad\text{i.e.}\quad \mathrm{Sp}(A)=\{1,3\}\subset\mathbb{R}_+^*.$$

Par décomposition de Cholesky il existe une unique matrice B triangulaire inférieure telle que $A=BB^t$. Puis, d'après la méthode de Cholesky, nous avons

$$b_{1,1}=\sqrt{a_{1,1}}=\sqrt{2},\quad b_{2,1}=\frac{a_{2,1}}{b_{1,1}}=-\frac{1}{\sqrt{2}},\quad b_{2,2}=\sqrt{a_{2,2}-b_{2,1}^2}=\frac{\sqrt{3}}{\sqrt{2}}.$$

Donc la décomposition de Cholesky de A est donnée par

$$A=\begin{pmatrix}\sqrt{2}&0\\-\frac{1}{\sqrt{2}}&\frac{\sqrt{3}}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}\sqrt{2}&-\frac{1}{\sqrt{2}}\\0&\frac{\sqrt{3}}{\sqrt{2}}\end{pmatrix}.$$

On considère un vecteur colonne $b=\begin{pmatrix}1\\0\end{pmatrix}$. Alors on résout $By=b$ par la méthode de descente

$$y_1=\frac{1}{\sqrt{2}},\quad y_2=\frac{b_2+\frac{1}{\sqrt{2}}y_1}{\frac{\sqrt{3}}{\sqrt{2}}}=\frac{1}{\sqrt{2}\sqrt{3}}.$$

Enfin on résout $B^tx=y$ par la méthode de remontée

$$x_2=\frac{y_2}{\frac{\sqrt{3}}{\sqrt{2}}}=\frac{1}{3},\quad x_1=\frac{y_1+\frac{1}{\sqrt{2}}x_2}{\sqrt{2}}=\frac{1}{2}+\frac{1}{6}=\frac{2}{3}.$$

Nous avons donc obtenu la solution $x=\frac{1}{3}\begin{pmatrix}2\\1\end{pmatrix}$ de $Ax=b$.

3.2 Calcul du nombre d'opérations élémentaires

Comptons le nombre d'opérations élémentaires rencontrées dans la méthode de Cholesky :

- n racines carrées,
- $n-1+...+1=\dfrac{n(n-1)}{2}$ divisions,
- $\dfrac{n^3-n}{6}$ multiplications et $\dfrac{n^3-n}{6}$ additions (soustractions) :

$$\sum_{j=1}^{n}\sum_{i=j}^{n}(j-1)=\sum_{j=1}^{n}(j-1)(n-j+1)=\sum_{k=0}^{n-1}k(n-k)=n\sum_{k=0}^{n-1}k-\sum_{k=0}^{n-1}k^2=\frac{n^2(n-1)}{2}-\frac{n(n-1)(2n-1)}{6}$$

$$=\frac{n(n-1)(3n-(2n-1)}{6}=\frac{n^3-n}{6}$$

- Pour les méthodes de remontée et de descente : $2n$ divisions, $n(n-1)$ additions et multiplications.

Par conséquent nous avons un ordre de n racines carrées, $\dfrac{n^2}{2}$ divisions, $\dfrac{n^3}{6}$ additions et multiplications ce qui est légèrement inférieur à l'ordre des opérations dans la méthode de Gauss si on ne compte pas les racines carrées. On favorisera donc la méthode de Cholesky pour les matrices symétriques définies positives.

Deuxième partie

Résolution numérique d'équations différentielles

1 Intégration numérique

Étant donné une fonction f continue sur un segment $[a, b]$, on cherche à calculer numériquement (car on ne connaît pas de primitive usuelle) l'intégrale de Riemann

$$\int_a^b f(x)dx\,. \tag{15}$$

La plupart des algorithmes procèdent en subdivisant le segment $[a, b]$ en plusieurs sous-intervalles

$$a = x_0 < x_1 < \ldots < x_N = b\,,$$

et en remarquant que

$$\begin{aligned}\int_a^b f(x)dx &= \sum_{i=0}^{N-1} \int_{x_i}^{x_{i+1}} f(x)dx \\ &= \sum_{i=0}^{N-1} h_i \int_0^1 f(x_i + th_i)dt\end{aligned}$$

où $h_i = x_{i+1} - x_i$. Chacune des intégrales de l'expression précédente doit être évaluée. On se concentre, dans la suite, en posant $g(t) = f(x_0 + th_0)$, sur la première intégrale

$$\int_0^1 g(t)dt \tag{16}$$

dont on cherche une valeur approchée.

1.1 Formules de quadratures

L'intégrale (16) peut-être approchée par la formule de quadrature

$$\sum_{j=1}^{s} b_j g(c_j) \tag{17}$$

où les $c_1, \ldots, c_s$ sont les nœuds de la formule de quadrature et $b_1, \ldots, b_s$, en sont les poids. Une formule de quadrature est d'ordre p si elle est exacte pour toutes les fonctions polynomiales de degré $p - 1$, i.e. pour tout polynôme P de degré inférieur ou égal à $p - 1$, on a

$$\int_0^1 P(t)dt = \sum_{j=0}^{s} b_j P(c_j).$$

Proposition 7. Une formule de quadrature est d'ordre p si et seulement si

$$\sum_{j=1}^{s} b_j c_j^{q-1} = \frac{1}{q} \quad \text{pour} \quad q = 1, \ldots, p \tag{18}$$

1.1.1 Formules des rectangles et du point-milieu, sommes de Riemann

Les formules de quadratures les plus simples consiste à approcher l'intégrale (16) par la valeur de la fonction g en un point ξ de l'intervalle $[0,1]$. On appelle formule des rectangles lorsque l'intégrale (16) est approchée par

$$g(0) \quad \text{ou} \quad g(1)$$

et formule du point-milieu lorsque cette intégrale est approchée par

$$g\left(\frac{1}{2}\right).$$

Pour une grille de discrétisation, on obtient pour l'approximation de l'intégrale (16) l'expression suivante

$$\sum_{i=0}^{N-1} (x_{i+1} - x_i)\, f(\xi_i) \quad \xi_i \in [x_i, x_{i+1}].$$

qui correspond à la somme de Riemann. On montre que ces sommes convergent vers

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

lorsque la finesse $\max_{i=0\ldots N-1} (x_{i+1} - x_i)$ tend vers 0.

Démonstration. On a, comme la fonction est continue, on a l'encadrement suivant

$$\sum_{i=0}^{N-1} (x_{i+1} - x_i)\, m_i \leq \sum_{i=0}^{N-1} (x_{i+1} - x_i)\, f(\xi_i) \leq \sum_{i=0}^{N-1} (x_{i+1} - x_i)\, M_i$$

où $m_i = \min_{x\in[x_i,x_{i+1}]} f(x)$ et $M_i = \max_{x\in[x_i,x_{i+1}]} f(x)$. On montre que, lorsque $\max_{i=0\ldots N-1} (x_{i+1} - x_i)$ tend vers 0,

$$\sum_{i=0}^{N-1} (x_{i+1} - x_i)\, (M_i - m_i)$$

tend vers 0 et que la limite de la somme de Riemann est la valeur de l'intégrale. □

— le théorème de Heine nous donne : toute fonction continue sur un segment est absolument continue, *i.e.*

$$\forall \varepsilon > 0, \quad \exists \eta > 0, \quad \forall x, y \quad \text{tq} \quad |x - y| < \eta, \quad |f(x) - f(y)| \leq \varepsilon.$$

— ce sont aussi deux fonctions en escaliers définissant l'intégrale de Riemann.

Numériquement, il est plus judicieux d'utiliser la formule du point milieu

$$\int_0^1 g(t)dt \approx g\left(\frac{1}{2}\right).$$

car elle est exacte pour les fonctions polynomiales de degré inférieur ou égal à 1 (*cf.* Proposition 8).

1.1.2 Formules de Newton-Cotes

En fixant les nœuds $c_1, \ldots, c_s$ distincts, $s \geq 2$, la condition (18) avec $p = s$ représente un système linéaire

$$\begin{pmatrix} 1 & 1 & \ldots & 1 \\ c_1 & c_2 & \ldots & c_s \\ \vdots & \vdots & & \vdots \\ c_1^{s-1} & c_2^{s-1} & & c_s^{s-1} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_s \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{1}{2} \\ \vdots \\ \frac{1}{s} \end{pmatrix}$$

qu'il faut résoudre. La matrice, qui est une matrice de Vandermonde, est inversible et la résolution du système nous donne une formule de quadrature d'ordre $p = s$.

Démonstration. On calcule le déterminant de la matrice

$$\det V = \begin{vmatrix} 1 & 1 & \ldots & 1 \\ c_1 & c_2 & \ldots & c_s \\ \vdots & \vdots & & \vdots \\ c_1^{s-1} & c_2^{s-1} & \ldots & c_s^{s-1} \end{vmatrix}$$

en effectuant l'opération

$$L_i \longleftarrow L_i - c_1 L_{i-1}$$

en partant de la n^{e} ligne et en remontant jusqu'à la seconde. Il vient,

$$\det(V) = \begin{vmatrix} 1 & 1 & \ldots & 1 \\ 0 & (c_2 - c_1) & \ldots & (c_s - c_1) \\ \vdots & \vdots & & \vdots \\ 0 & c_2^{s-2}(c_2 - c_1) & \ldots & c_s^{s-2}(c_s - c_1) \end{vmatrix}$$

en développant par rapport à la première colonne et en factorisant

$$\det V = (c_2 - c_1)(c_3 - c_1)\ldots(c_s - c_1) \begin{vmatrix} 1 & \ldots & 1 \\ c_2 & \ldots & c_s \\ \vdots & & \vdots \\ c_2^{s-2} & \ldots & c_s^{s-2} \end{vmatrix}.$$

Par récurrence, on obtient le résultat

$$\det(V) = \prod_{1 \leq i < j \leq n} (c_j - c_i)$$

□

Ainsi, en fixant les $c_j = \frac{(j-1)}{s-1}$, $j = 1, 2, \ldots, s$, on obtient les formules de quadrature de Newton-Cotes qui sont répertoriées dans le tableau suivant :

On remarque que pour la formule de Simpson, si il est évident que la condition

$$\sum_{j=1}^{s} b_j c_j^{q-1} = \frac{1}{q}$$

s	poids							nom	ordre
2	$\frac{1}{2}$	$\frac{1}{2}$						trapèze	2
3	$\frac{1}{6}$	$\frac{4}{6}$	$\frac{1}{6}$					Simpson	$3 \to 4$
4	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$				Newton	4
5	$\frac{7}{90}$	$\frac{32}{90}$	$\frac{12}{90}$	$\frac{32}{90}$	$\frac{7}{90}$			Boole	$5 \to 6$
6	$\frac{19}{288}$	$\frac{75}{288}$	$\frac{50}{288}$	$\frac{50}{288}$	$\frac{75}{288}$	$\frac{19}{288}$		–	6
7	$\frac{41}{840}$	$\frac{216}{840}$	$\frac{27}{840}$	$\frac{272}{840}$	$\frac{27}{840}$	$\frac{216}{840}$	$\frac{41}{840}$	Weddle	$7 \to 8$

TABLE 1 – Formules de quadrature de Newton-Cotes pour $2 \leq s \leq 7$.

est satisfaite pour $q = 1, 2, 3$, la même condition est satisfaite pour $q = 4$:

$$\frac{1}{6} \cdot 0^3 + \frac{4}{6} \cdot \left(\frac{1}{2}\right)^3 + \frac{1}{6}.1^3 = \frac{1}{4}$$

mais plus pour $q = 5$:

$$\frac{1}{6} \cdot 0^4 + \frac{4}{6} \cdot \left(\frac{1}{2}\right)^4 + \frac{1}{6}.1^4 = \frac{5}{24} \neq \frac{1}{5}\,.$$

Construite pour être d'ordre 3, la formule se révèle être d'ordre 4. Plus généralement,

Proposition 8. Si une formule de quadrature symétrique,*i.e.* pour tout $j = 1 \ldots s$,

$$c_j = 1 - c_{s+1-j} \quad \text{et} \quad b_j = b_{s+1-j}\,,$$

est exacte pour les fonctions polynomiales de degré inférieur ou égal à $2m-2$, $m \geq 1$, elle est automatiquement exacte pour les fonctions polynomiales de degré $2m-1$. Autrement dit, une formule de quadrature symétrique à toujours un ordre pair.

Démonstration. Un polynôme de degré $2m-1$ peut s'écrire sous la forme

$$g(t) = C\left(t - \frac{1}{2}\right)^{2m-1} + g_1(t)$$

avec $g_1(t)$ un polynôme de degré $\leq 2m-2$. Il suffit alors de montrer qu'une formule symétrique est exacte pour $\left(t - \frac{1}{2}\right)^{2m-1}$.

A cause de la symétrie de cette fonction, la valeur exacte vaut

$$\int_0^1 \left(t - \frac{1}{2}\right)^{2m-1} dt = 0\,.$$

L'approximation numérique satisfait

$$\begin{aligned}
\sum_{j=1}^{s} b_j (c_j - \frac{1}{2})^{2m-1} &= \sum_{j=1}^{s} b_{s+1-j} (1 - c_{s+1-j} - \frac{1}{2})^{2m-1} \\
&= -\sum_{j=1}^{s} b_{s+1-j} (c_{s+1-j} - \frac{1}{2})^{2m-1} \\
&= -\sum_{j=1}^{s} b_j (c_j - \frac{1}{2})^{2m-1} \quad \text{(changement d'indice)}
\end{aligned}$$

ou encore

$$\sum_{j=1}^{s} b_j (c_j - \frac{1}{2})^{2m-1} = 0.$$

□

Ainsi, la formule du point milieu à même ordre que la formule du trapèze, précisément 2 ; la formule de Simpson à même ordre que la formule de Newton, précisément 4, *etc.*.

1.2 Étude de l'erreur

On cherche dans cette section, un encadrement de l'erreur d'approximation de l'intégrale initiale par les méthodes par les formules de quadratures

$$e(f,N) = \int_a^b f(x)dx - \sum_{i=0}^{N-1} h_j \sum_{j=1}^{s} b_j f(x_i + c_j h_i)\,.$$

On se concentre sur l'erreur commise sur la première intégrale à évaluer (16),

$$e_0(f) = h_0 \int_0^1 f(x_0 + t h_0)dt - h_0 \sum_{j=1}^{s} b_j f(x_0 + c_j h_0)\,.$$

Proposition 9. Soit f une fonction k fois continûment dérivable et p l'ordre de la formule de quadrature tel que$p \geq k$. Alors

$$e_0(f) = h_0^{k+1} \int_0^1 N_k(\tau) f^{(k)}(x_0 + \tau h_0)d\tau \tag{19}$$

où $N_k(\tau)$ est le noyau de Peano, défini par

$$N_k(\tau) = \frac{(1-\tau)^k}{k!} - \sum_{j=1}^{s} b_j \frac{(c_j - \tau)_+^{k-1}}{(k-1)!}$$

et $h_0 = x_1 - x_0$.

Démonstration. On pose $g(t) = f(x_0 + t h_0)$. Nous pouvons écrire la formule de Taylor avec reste intégral,

$$g(t) = \sum_{j=0}^{k-1} \frac{t^j}{j!} g^{(j)}(0) + \underbrace{\int_0^t \frac{(t-\tau)^{k-1}}{(k-1)!} g^{(k)}(\tau)d\tau}_{h(t)}\,.$$

$$f(x_0 + t h_0) = \sum_{j=0}^{k-1} \frac{(t h_0)^j}{j!} f^{(j)}(x_0) + \underbrace{h_0^k \int_0^{t h_0} \frac{(t h_0 - \tau)^{k-1}}{(k-1)!} f^{(k)}(x_0 + \tau h_0)d\tau}_{h(t)}\,.$$

La formule de quadrature est d'ordre supérieur à k, le terme polynomial de degré inférieur ou égal à $k-1$ ne contribue pas à l'erreur car la formule de quadrature est exacte, il vient donc :

$$\begin{aligned}
e_0(f) &= h_0 \left(\int_0^1 h(t)dt - \sum_{j=1}^{s} b_j h(c_j) \right) \\
&= h_0^{k+1} \left(\int_0^1 \int_0^t \frac{(t-\tau)^{k-1}}{(k-1)!} f^{(k)}(x_0+\tau h_0)d\tau dt - \sum_{j=1}^{s} b_j \int_0^{c_j} \frac{(c_j-\tau)^{k-1}}{(k-1)!} f^{(k)}(x_0+\tau h_0)d\tau \right) \\
&= h_0^{k+1} \left(\int_0^1 \int_\tau^1 \frac{(t-\tau)^{k-1}}{(k-1)!} dt f^{(k)}(x_0+\tau h_0)d\tau - \sum_{j=1}^{s} b_j \int_0^1 \frac{(c_j-\tau)_+^{k-1}}{(k-1)!} f^{(k)}(x_0+\tau h_0)d\tau \right) \\
&= h_0^{k+1} \int_0^1 \left(\frac{(1-\tau)^k}{k!} - \sum_{j=1}^{s} b_j \frac{(c_j-\tau)_+^{k-1}}{(k-1)!} \right) f^{(k)}(x_0+\tau h_0)d\tau
\end{aligned}$$

□

Ainsi, en sommant les erreurs on obtient,

Proposition 10. Soit f une fonction k fois continûment dérivable et p l'ordre de la formule de quadrature tel que$p \geq k$. Alors

$$|e(f,N)| \leq h^k (b-a) \max_{x\in[a,b]} |f^{(k)}(x)| \int_0^1 |N_k(\tau)| d\tau$$

où $h = \max_{i=0,\ldots,N-1} h_i$.

Démonstration. La formule (19) donne

$$\begin{aligned}
|e_0(f)| &\leq h_0^{k+1} \int_0^1 |N_k(\tau)| \left| f^{(k)}(x_0+\tau h_0 \right| d\tau \\
&\leq h_0^{k+1} \max_{x\in[x_0,x_0+h_0]} \left| f^{(k)}(x) \right| \int_0^1 |N_k(\tau)| \, d\tau
\end{aligned}$$

et comme l'erreur est la somme des erreurs sur chacun des sous-intervalles de la division, on obtient

$$\begin{aligned}
|e(f,N)| &\leq \sum_{i=0}^{N-1} |e_i(f)| \\
&\leq \sum_{i=0}^{N-1} h_i^{k+1} \max_{x\in[x_i,x_i+h_i]} \left| f^{(k)}(x) \right| \int_0^1 |N_k(\tau)| \, d\tau \\
&\leq \sum_{i=0}^{N-1} h^k h_i \max_{x\in[a,b]} \left| f^{(k)}(x) \right| \int_0^1 |N_k(\tau)| \, d\tau \\
&\leq h^k \max_{x\in[a,b]} \left| f^{(k)}(x) \right| \int_0^1 |N_k(\tau)| \, d\tau \sum_{i=0}^{N-1} h_i
\end{aligned}$$

qui donne le résultat. □

Par exemple, pour la formule du point milieu on obtient la majoration de l'erreur suivante

$$|e(f,N)| \leq \frac{1}{24} h^2 (b-a) \max_{x\in[a,b]} |f''(x)|\,;$$

pour la formule du trapèze

$$|e(f,N)| \leq \frac{1}{12} h^2 (b-a) \max_{x\in[a,b]} |f''(x)|\,;$$

et pour la formule de Simpson

$$|e(f,N)| \leq \frac{1}{2880} h^4 (b-a) \max_{x\in[a,b]} |f^{(4)}(x)|\,.$$

Toutes les constantes ont été obtenus en utilisant le fait que pour une formule de quadrature d'ordre p

$$\begin{aligned}
\int_0^1 N_p(\tau)d\tau &= \int_0^1 \left(\frac{(1-\tau)^p}{p!} - \sum_{j=1}^s b_j \frac{(c_j-\tau)_+^{p-1}}{(p-1)!} \right) d\tau \\
&= \frac{1}{(p+1)!} - \sum_{j=1}^s b_j \int_0^1 \frac{(c_j-\tau)_+^{p-1}}{(p-1)!} d\tau \\
&= \frac{1}{(p+1)!} - \sum_{j=1}^s b_j \int_0^{c_j} \frac{(c_j-\tau)^{p-1}}{(p-1)!} d\tau \\
&= \frac{1}{p!} \left(\frac{1}{p+1} - \sum_{j=1}^s b_j c_j^p \right)
\end{aligned}$$

et la constance du signe de $N_p(\tau)$.

Par exemple pour la méthode de Simpson

$$\int_0^1 |N_4(\tau)|d\tau = -\frac{1}{4!}\left(\frac{1}{5} - \frac{5}{24}\right) = \frac{1}{2880}\,.$$

2 Méthodes d'Euler et méthode de Runge-Kutta

Dans cette section nous présentons des méthodes numériques de résolution d'équations différentielles "à un pas" où le calcul ne dépend que des résultats de l'étape précédente.

On cherche à résoudre numériquement l'équation différentielle du premier ordre

$$\left\{ \begin{array}{lcl} y'(x) & = & f(x, y(x)) \\ y(0) & = & y_0 \end{array} \right. \tag{20}$$

où $f : \mathbb{R} \to \mathbb{R}$ est une fonction connue et de classe$\mathcal{C}^1$ sur un voisinage de $(0, y_0)$, y_0 fixé. On admet qu'il existe un réel $\alpha > 0$ tel que le problème (20) possède une unique solution $y(x)$ de classe $\mathcal{C}^1$ sur $[0, \alpha[$.

Pour ce faire, on cherche une approximation de la solution en subdivisant l'intervalle $[0, \overline{x}]$, $\overline{x} < \alpha$, en sous-intervalles réguliers d'extrémités $0 = x_0 < x_1 < \ldots < x_i = ih < \ldots < x_N = \overline{x}$ avec $h = \frac{\overline{x}}{N}$: **le pas de discrétisation**.

2.1 Méthode d'Euler

En intégrant l'équation (20) sur $[0, h]$, on a

$$y(h) = y_0 + h \int_0^1 g(v)dv \tag{21}$$

avec $g(v) = f(vh, y(vh))$.

La méthode d'Euler consiste à approcher l'intégrale de l'équation (21) par la formule de quadrature des rectangles (à gauche)

$$\int_0^1 g(v)dv \simeq g(0).$$

On obtient le schéma numérique de la première étape

$$y_1 = y_0 + hf(0, y_0). \tag{22}$$

Ici y_1 est l'approximation de $y(x_1) = y(h)$. On montre que $y_1 - y(h) = \mathcal{O}(h^2)$.

On note, dans la suite, y_n l'approximation de $y(x_n) = y(nh)$. On applique aux étapes $1 \leq n \leq N - 1$ suivantes, le même schéma à un pas

$$y_{n+1} = y_n + hf(x_n, y_n). \tag{23}$$

On montre alors que pour h suffisamment petit il existe $\Lambda > 0$ et $C > 0$ tel que

$$|y(x_N) - y_N| \leq h^2 \frac{C}{\Lambda} \left(e^{\Lambda(x_N - x_0)} - 1 \right). \tag{24}$$

et que le schéma converge lorsque h tend vers 0.

2.2 Méthode de Runge

La méthode de Runge consiste à approcher l'intégrale par la formule de quadrature du point milieu

$$y(x_{n+1}) \simeq y_n + hf\left(x_n + \frac{h}{2}, y(x_n + \frac{h}{2})\right)$$

et de remplacer la valeur $y(x_n + \frac{h}{2})$ inconnue par sa valeur approchée par la méthode de d'Euler

$$y_{n+1} = y_n + hf\left(x_n + \frac{h}{2}, y_n + \frac{h}{2}f(x_n, y_n)\right). \tag{25}$$

2.3 Méthode de Heun

La méthode de Heun consiste à approcher l'intégrale par une formule de quadrature d'ordre 3

$$y(x_{n+1}) \simeq y_n + \frac{h}{4}\left(f(x_n, y_n) + 3f\left(x_n + \frac{2h}{3}, y\left(x_n + \frac{2h}{3}\right)\right)\right)$$

et de remplacer la valeur $y(x_n + \frac{2h}{3})$ inconnue par sa valeur approchée par la méthode de de Runge

$$y_{n+1} = y_n + \frac{h}{4}\left(f(x_n, y_n) + 3f\left(x_n + \frac{2h}{3}, y_n + \frac{2h}{3}f\left(x_n + \frac{h}{3}, y_n + \frac{h}{3}f(x_n, y_n)\right)\right)\right). \tag{26}$$

2.4 Euler implicite

La méthode d'Euler implicite consiste à approcher l'intégrale par les rectangle à droite

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

associé à la recherche de la valeur y_{n+1} en cherchant la racine de la fonction

$$h(y) = y_n + hf(x_{n+1}, y) - y.$$

Troisième partie

Calcul de la transformée de Fourier et applications

Dans le traitement de signaux, où on est confronté à une immense quantité (plusieurs milliers ou millions) de valeurs numériques, la transformée de Fourier est un outil inévitable. De plus, les données ont souvent une certaine périodicité ce qui rend la transformée plus efficace. Elle est, par exemple, très utilisée dans le traitement des sons ainsi que pour la compression d'images.

1 Transformée de Fourier (continue).

Une série trigonométrique (ou série de Fourier) est une série de la forme

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx) \tag{27}$$

En cas de convergence, elle représente une fonction 2π-périodique, c'est-à-dire pour tout $x \in \mathbb{R}$, $f(x + 2\pi) = f(x)$. Les formules deviennent plus simples en passant aux complexes. Grâce aux identités $e^{ix} = \cos x + i \sin x$, $\cos(x) = (e^{ix} + e^{-ix})/2$ et $\sin(x) = (e^{ix} - e^{-ix})/(2i)$, la série précédente devient simplement

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx}. \tag{28}$$

La clé fondamentale permettant le calcul des séries trigonométriques a été découverte par Euler et réside dans la *relation d'orthogonalité*

$$\int_0^{2\pi} e^{-ilx} e^{ikx} dx = \int_0^{2\pi} e^{i(k-l)x} dx = \begin{cases} 0 & \text{si } k \neq l \\ 2\pi & \text{si } k = l. \end{cases}$$

Cette propriété nous permet de calculer les coefficients. Il suffit de multiplier (28) par e^{-ikx} et intégrer terme par terme de 0 à 2π. Tous les termes, sauf un, disparaissent et on obtient

$$c_k = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} dx. \tag{29}$$

Pour marquer la dépendance de f, nous écrivons souvent pour les coefficients de Fourier $\hat{f}(k) = c_k$.

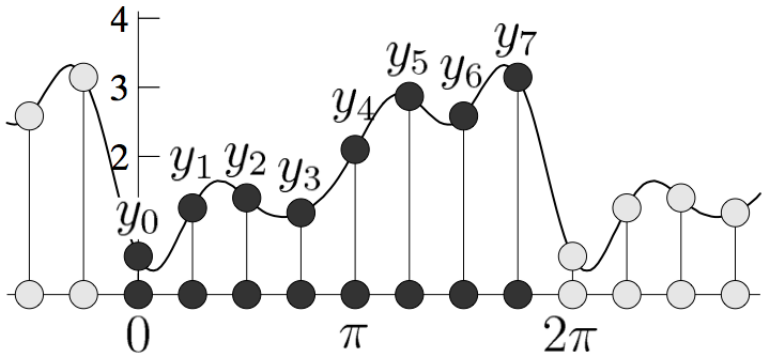
Proposition 1. Les coefficients $\hat{f}(k)$ convergent vers zéro pour k tendant vers $+\infty$, et on a $|\hat{f}(k)| = \mathcal{O}(|k|^{-p})$ si le prolongement 2π-périodique de f est p-fois continûment différentiable.

2 Transformée de Fourier discrète (DFT).

Supposons maintenant que la fonction f 2π-périodique soit seulement connue pour les x de la division équidistante

$$x_j = \frac{2\pi j}{N}, \qquad j = 0, 1, \ldots, N-1$$

et posons $y_j = f(x_j)$. Si nécessaire, on peut prolonger (y_j) à une suite N-périodique en posant $y_{j+N} = y_j$ pour tout entier j. Sur la figure suivante, on a $N = 8$.



Par analogie avec (28), on cherche à exprimer cette suite par

$$y_j = \sum_{k=0}^{N-1} z_k e^{ikx_j} = \sum_{k=0}^{N-1} z_k \omega^{kj}, \quad \text{avec} \quad \omega = e^{2\pi i/N}. \tag{30}$$

Cette fois la *relation d'orthogonalité discrète* (noter que $\omega^N = 1$)

$$\sum_{j=0}^{N-1} \omega^{-lj}\omega^{kj} = \sum_{j=0}^{N-1} \omega^{(k-l)j} = \begin{cases} \frac{1-\omega^{(k-l)N}}{1-\omega^{k-l}} = 0 & \text{si } k \neq l \text{ (modulo) } N \\ N & \text{si } k = l \text{ (modulo) } N \end{cases}$$

nous aide à trouver les z_k à partir de (30). Par parfaite analogie avec la preuve ci-dessus pour (29), on multiplie l'équation (30) par ω^{-kj} et on additionne de $j = 0$ à N. Tous les termes, sauf un, disparaissent et on obtient la transformée de Fourier discrète (DFT)

$$z_k = \frac{1}{N} \sum_{j=0}^{N-1} y_j \omega^{-kj}. \tag{31}$$

Si $y_j = f(x_j)$, nous écrivons aussi $\widehat{f_N}(k) = z_k$ pour les coefficients de la transformée de Fourier discrète. Comme $y_N = y_0$, la valeur z_k de (31) peut être interprétée comme le résultat de la règle du rectangle appliquée à l'intégrale dans (29). En effet pour k fixé :

$$\begin{aligned} c_k &= \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-ikx}dx \approx \frac{1}{2\pi}\sum_{j=0}^{N-1}(x_{j+1}-x_j)f(x_j)e^{-ikx_j} \\ &= \sum_{j=0}^{N-1}\frac{1}{N}y_j\omega^{-kj} = z_k. \end{aligned}$$

Toutefois, $z_k = \widehat{f_N}(k)$ n'est pas une approximation de $c_k = \hat{f}(k)$ pour tout k, car la suite (z_k) est N-périodique (ceci est une conséquence de $\omega^N = 1$) alors que $\hat{f}(k)$ converge vers zéro si k tend vers ∞.

Quelques remarques générales avant de continuer. Notons P_N l'espace des suites N-périodiques

$$P_N = \{(y_k)_{k\in\mathbb{Z}} | y_k \in \mathbb{C},\ y_{k+N} = y_k\}.$$

La transformée de Fourier discrète de $y = (y_k) \in P_N$ est la suite $z = (z_k)_{k\in\mathbb{Z}}$ où z_k est donné par (31).

Proposition 2. On note $z = \mathcal{F}_N(y)$.

1. Pour $y \in P_N$, $z = \mathcal{F}_N(y) \in P_N$.
2. L'application $\mathcal{F}_N : P_N \to P_N$ est linéaire et bijective.
3. l'application inverse de $\mathcal{F}_N$ est donnée par

$$\mathcal{F}_N^{-1} = N.\overline{\mathcal{F}_N}.$$

Étude de l'erreur.

Pour la division équidistante

$$x_j = \frac{2\pi j}{N}, \qquad j = 0, 1, \ldots, N-1$$

de l'intervalle $[0, 2\pi]$ et pour $y_0, y_1, \ldots, y_{N-1}$ donnés, on cherche un polynôme trigonométrique (une combinaison linéaire finie de fonctions e^{ikx}) passant par (x_j, y_j) pour $j = 0, \ldots, N-1$.

Théorème **1.** Pour $y_0, y_1, \ldots, y_{N-1}$ donnés, soit (z_k) sa transformée de Fourier discrète (31). Si N est pair, le polynôme trigonométrique

$$p_N(x) = \frac{1}{2}\left(z_{-N/2}e^{-iNx/2} + z_{N/2}e^{iNx/2}\right) + \sum_{|k|<-N/2} z_k e^{ikx}$$

satisfait $p_N(x_j) = y_j$ pour $j = 0, 1, \ldots, N-1$.

On remarque que si les y_k sont réels, alors p_N est une fonction réelle, c'est-à-dire une combinaison réelle de $\sin kx$ et $\cos kx$.

Supposons maintenant que $y_j = f(x_j)$ où

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx}, \qquad c_k = \hat{f}(k)$$

est une série de Fourier qui converge absolument pour tout x.

Théorème **2.** Alors pour tout $x \in \mathbb{R}$,

$$|f(x) - p_N(x)| \leq \left|\hat{f}\left(\frac{N}{2}\right)\right| + \left|\hat{f}\left(-\frac{N}{2}\right)\right| + 2\sum_{|k|>N/2}\left|\hat{f}\,(k)\right|.$$

De plus si f est p-fois dérivable (avec toutes ses dérivées continues), on a

$$|f(x) - p_N(x)| = \mathcal{O}\left(N^{-p+1}\right).$$

Ce théorème permet une interprétation intéressante. Considérons une fonction 2π-périodique de fréquence maximale M (c'est-à-dire, pour $|k| \geq M$, $\hat{f}(k) = 0$). Alors le polynôme trigonométrique donne le résultat exact ($p_N(x) = f(x)$ pour tout x) si $N > 2M$. Ce résultat (le théorème d'échantillonnage) nous donne une formule pour le nombre d'échantillons nécessaires pour une représentation exacte d'une telle fonction.

2.1 Compression des données

Étant donnée une suite N-périodique (y_j) et sa transformée de Fourier discrète (z_k), l'idée est de supprimer dans la représentation (30) pour y_j tous les termes dont la valeur absolue de z_k est en-dessous d'une certain seuil (par exemple, 3% du coefficient maximal).

Le premier dessin de la figure suivante montre la digitalisation d'une son. On a enregistré 22000 impulsions par seconde, dont 1024 sont dessinées (ceci correspond à 1024/22=46,5millisecondes). On observe bien une certaine périodicité des données.

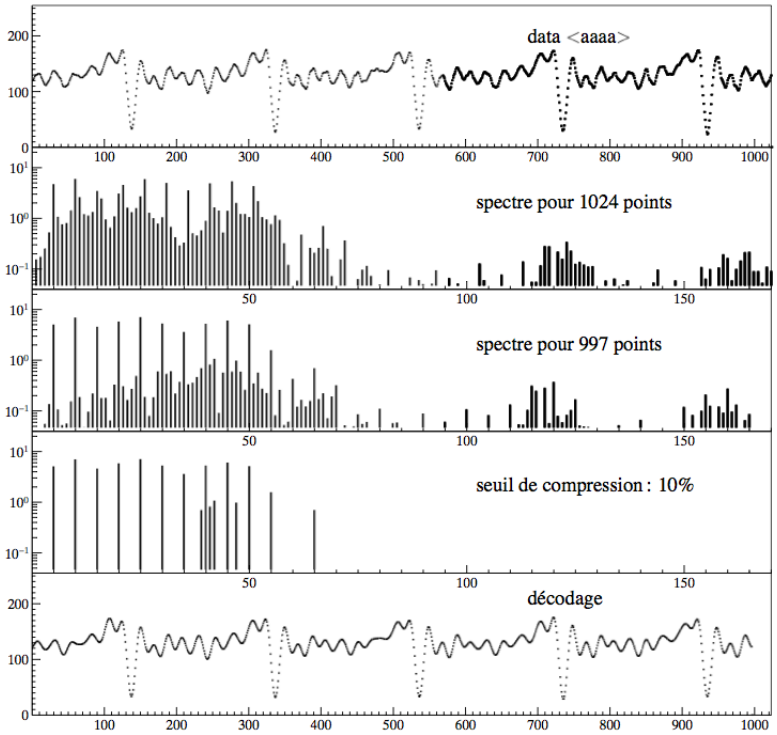
Pour les $N = 1024$ nombres (y_j) on a calculé la transformée de Fourier discrète z_k. La suite de leurs valeurs absolues est dessinée dans la deuxième image de la figure pour $k = 1, \ldots, 170$ (comme les y_j sont réels, on a $z_{N-k} = z_{-k} = \overline{z_k}$ et il n'est pas nécessaire de dessiner les valeurs pour $k \geq N/2$; pour $170 < k \leq N/2$ les z_k sont inférieurs à 0,072).

La théorie de ce paragraphe est basée sur le fait que $f(x)$ est une fonction périodique. Cependant, la période du signal n'est visiblement pas égale à $N = 1024$, mais elle est plutôt proche de $N = 997$. Si l'on calcule 1024 valeurs de $f(x)$ sur une période exacte (par interpolation linéaire) ainsi que leur transformée de Fourier discrète, on obtient la troisième image de la figure. Cette fois, on peut beaucoup mieux observer la fréquence principale ($5 \times 22000/997 \approx 110$ Hz) ainsi que les harmoniques (multiples de la fréquence principale).

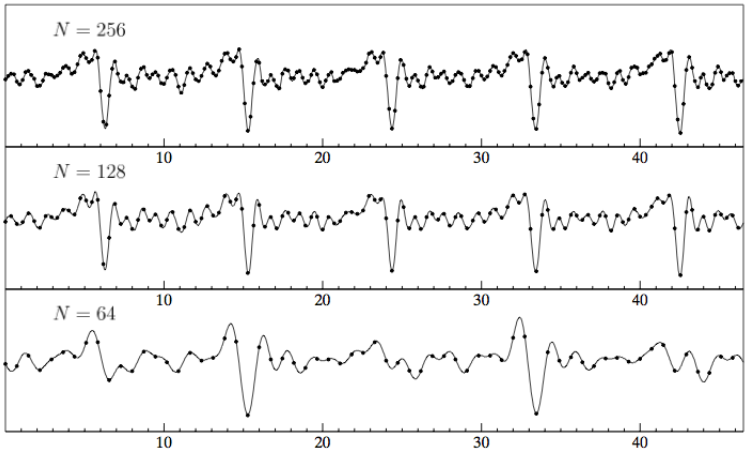
Maintenant nous supprimons tous les coefficients z_k dont la valeur absolue est inférieure à 10% de la valeur maximale. Les 16 coefficients restant sont dessinés dans le quatrième dessin de la figure. Ainsi, la vraie information contenue dans le signal ne contient que 16 nombres complexes au lieu des 997 valeurs réelles y_j.

Pour décoder le signal, nous utilisons la formule (30) avec les z_k restant après la compression. Le résultat (décodage) est dessiné en bas de la figure. On peut constater que le signal original est très bien reproduit.

La conclusion de cette expérience est la suivante : au lieu de stocker les 997 valeurs du signal original, il suffit de stocker quelques coefficients de la transformée de Fourier discrète sans perdre de l'informatisation visible.



Dans le deuxième dessin de cette figure on voit que les fréquences dominantes du son sont situées dans l'intervalle $|k| \leq 60$. Comme la longueur de l'intervalle dans la figure est de $1024/22000$ secondes, la valeur maximale M correspond à $60 \times 22000/1024 \approx 1390$ Hz. Alors, $N = 128$ échantillons sont suffisants pour représenter correctement le signal. Dans la figure suivante sont dessinés les polynômes trigonométriques $p_N(x)$ (pour $N = 64, 128$ et 256) passant par y_j pour $j = 0$ modulo (1024/N). où y_j sont les données de la figure précédente. On voit que la représentation est bonne à partir de $N = 128$. Il suffit alors d'utiliser chaque huitième échantillon (une autre possibilité de compression des données).



2.2 Transformée de Fourier rapide

Pour (y_j), $j = 0, \ldots, N-1$ donnés, un calcul direct de la transformée de Fourier discrète (z_k), $k = 0, \ldots, N-1$ (cf. (31)) nécessite N^2 multiplications et additions. Dans ce paragraphe, nous présentons un algorithme qui fait le même travail en $N \log_2(N)$ opérations. Cet algorithme est dû à Cooley & Tukey (1965). Nous supprimons le facteur $1/N$ dans (31) et nous utilisons la notation

$$z = \mathcal{F}_N(y), \qquad z_k = \sum_{j=0}^{N-1} y_j \omega_N^{-kj}, \quad \omega_N = e^{2\pi i/N}.$$

Faites attention à la redéfinition de $\mathcal{F}_N$ (par rapport à la proposition 2).

Lemme **1.** Soient $u = (u_0, u_1, \ldots, u_{N-1})$, $v = (v_0, v_1, \ldots, v_{N-1})$ et définissons

$$y = (u_0, v_0, u_1, v_1, \ldots, u_{N-1}, v_{N-1}).$$

Alors pour $k = 0, 1, \ldots, N-1$, on a avec $\omega_{2N} = e^{i\pi/N}$:

$$\begin{aligned} (\mathcal{F}_{2N}(y))_k &= (\mathcal{F}_N(u))_k + \omega_{2N}^{-k}(F_N(v))_k \\ (\mathcal{F}_{2N}(y))_{k+N} &= (\mathcal{F}_N(u))_k - \omega_{2N}^{-k}(\mathcal{F}_N(v))_k. \end{aligned} \tag{32}$$

La formule (32) nous permet de calculer (avec N multiplications et $2N$ additions) le vecteur $\mathcal{F}_{2N}(y)$ à partir de $\mathcal{F}_N(u)$ et $\mathcal{F}_N(v)$. La même procédure peut être appliquée récursivement aux suites u et v si elles ont une longueur paire. Si l'on suppose que $N = 2^m$, on obtient l'algorithme présenté dans le schéma suivant (pour $N = 8 = 2^3$)

$$\mathcal{F}_N \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \end{pmatrix} \begin{array}{l} \swarrow \\ \\ \\ \\ \\ \\ \\ \searrow \end{array} \begin{array}{l} \mathcal{F}_{N/2} \begin{pmatrix} y_0 \\ y_2 \\ y_4 \\ y_6 \end{pmatrix} \\ \\ \\ \\ \mathcal{F}_{N/2} \begin{pmatrix} y_1 \\ y_3 \\ y_5 \\ y_7 \end{pmatrix} \end{array} \begin{array}{l} \swarrow \\ \searrow \\ \\ \\ \swarrow \\ \searrow \end{array} \begin{array}{l} \mathcal{F}_{N/4} \begin{pmatrix} y_0 \\ y_4 \end{pmatrix} \\ \\ \mathcal{F}_{N/4} \begin{pmatrix} y_2 \\ y_6 \end{pmatrix} \\ \\ \mathcal{F}_{N/4} \begin{pmatrix} y_1 \\ y_5 \end{pmatrix} \\ \\ \mathcal{F}_{N/4} \begin{pmatrix} y_3 \\ y_7 \end{pmatrix} \end{array} \begin{array}{l} \swarrow \\ \searrow \\ \\ \\ \swarrow \\ \searrow \\ \\ \\ \swarrow \\ \searrow \\ \\ \\ \swarrow \\ \searrow \end{array} \begin{array}{l} \mathcal{F}_{N/8} y_0 = y_0 \\ \mathcal{F}_{N/8}\, y_4 = y_4 \\ \\ \\ \mathcal{F}_{N/8}\, y_2 = y_2 \\ \mathcal{F}_{N/8}\, y_6 = y_6 \\ \\ \\ \mathcal{F}_{N/8}\, y_1 = y_1 \\ \mathcal{F}_{N/8}\, y_5 = y_5 \\ \\ \\ \mathcal{F}_{N/8}\, y_3 = y_3 \\ \mathcal{F}_{N/8}\, y_7 = y_7 \end{array}$$

La programmation de cet algorithme se fait de droite à gauche. D'abord, on met les y_j dans d'ordre exigé par d'algorithme. Après, on effectue les opérations de (32) comme indiqué dans le schéma. Pour passer d'une colonne à une autre on a besoin de $N/2$ multiplications complexes et de N additions (ou soustractions). Comme $m = \log_2(N)$ passages sont nécessaires, on a le résultat suivant.

Théorème **3.** Pour $N = 2^m$, le calcul de $F_N(y)$ peut être effectué en $\frac{N}{2} \log_2(N)$ multiplications complexes et $N \log_2(N)$ additions complexes.

Pour mieux illustrer l'importance de cet algorithme, nous comparons dans le tableau suivant le nombre d'opérations nécessaires pour le calcul de $F_N(y)$ avec ou sans FFT.

N	N^2	$N \log_2 N$	quotient
$2^5 = 32$	$\approx 10^3$	160	≈ 6.4
$2^{10} \approx 10^3$	$\approx 10^6$	$\approx 10^4$	≈ 100
$2^{20} \approx 10^6$	$\approx 10^{12}$	$\approx 2 \cdot 10^7$	$\approx 5 \cdot 10^4$

L'inverser de la transformée de Fourier discrète. Pour le décodage il faut calculer les y_j à partir des z_k à l'aide de la formule (29) et de la proposition 2. Pour en obtenir un algorithme rapide, il suffit de remplacer ω_{2N}^k dans (32) par ω_{2N}^k.

Quatrième partie

Travaux dirigés

1 Pivot de Gauss

1.1 Méthodes de remontée et de descentes

Exercice 1. Trouver les solutions des systèmes suivants :

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \\ 1 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 6 \end{pmatrix}.$$

1.2 La méthode du pivot de Gauss

Exercice 2. On considère le système suivant :

$$\begin{pmatrix} 5 & 2 & 1 \\ 5 & -6 & 2 \\ -4 & 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 12 \\ 1 \\ 3 \end{pmatrix}.$$

1. Trouver la solution de ce système, en utilisant la méthode du pivot de Gauss.
2. Calculer le nombre exact d'opérations que vous avez effectuées.

Exercice 3. Trouvez la solution, en utilisant la méthode du pivot de Gauss, du système suivant :

$$\begin{pmatrix} 1 & 1 & 2 & 1 \\ 1 & 1 & 4 & 2 \\ 3 & 2 & 2 & 2 \\ 4 & 2 & 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 5 \\ 7 \\ 10 \\ 13 \end{pmatrix}.$$

1.3 Choix du pivot

Exercice 4. Reprendre le système précédent en essayant d'optimiser le choix du pivot.

$$\begin{pmatrix} 1 & 1 & 2 & 1 \\ 1 & 1 & 4 & 2 \\ 3 & 2 & 2 & 2 \\ 4 & 2 & 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 5 \\ 7 \\ 10 \\ 13 \end{pmatrix}.$$

1.4 Méthode de Gauss-Jordan

Exercice 5. Faire l'inversion des matrices suivantes par la méthode de Gauss-Jordan.

1. $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$
2. $A = \begin{pmatrix} 1 & 5 & 6 \\ 2 & 8 & 9 \\ 3 & 15 & 19 \end{pmatrix}$

2 Factorisation LU

Exercice 6. On considère la matrice A suivante

$$A = \begin{pmatrix} 9 & 6 & 3 \\ 6 & 3 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

1. Effectuez la décomposition LU de la matrice A, en utilisant le pivot de Gauss.
2. Quel est le déterminant de A ?
3. Résoudre, en utilisant la décomposition LU, le système $Ax = b$ où :

$$b = \begin{pmatrix} 9 \\ 5 \\ 1 \end{pmatrix}.$$

4. Calculer le coût de cette résolution en terme d'opérations.

Exercice 7. Effectuer la factorisation LU des matrices suivantes :

$$\begin{pmatrix} 5 & 2 & 1 \\ 5 & -6 & 2 \\ -4 & 2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 2 & 4 & 4 \\ 1 & 3 & 1 \\ 1 & 5 & 6 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} -2 & 1 & 0 & 0 & 0 \\ -4 & 5 & 2 & 0 & 0 \\ 0 & -3 & -1 & -1 & 0 \\ 0 & 0 & -2 & 4 & 1 \\ 0 & 0 & 0 & 2 & -2 \end{pmatrix}.$$

Exercice 8. Soit $A = (a_{ii})_{1 \leq i,j \leq n}$ définie par :

$$a_{ij} = \begin{cases} 1 & si\ i = j\ ou\ j = n \\ -1 & si \quad i > j \\ 0 & Sinon \end{cases}$$

1. Démontrer que A admet une décompostion LU.
2. Effectuer la décomposition LU de A pour $n = 2, 3, 4$.
3. Démontrer que pour tout $n \geq 2$, on a pour les matrices $L = (l_{ij})_{1 \leq i,j \leq n}$ et $U = (u_{ij})_{1 \leq i,j \leq n}$ de la décompoistion $A = LU$ vérifient :

$$l_{ij} = \begin{cases} 1 & si\ i = j \\ -1 & si \quad i > j \\ 0 & Sinon \end{cases}$$

$$u_{ij} = \begin{cases} 1 & si\ i = j \neq n \\ 2^{i-1} & si \quad j = n \\ 0 & Sinon \end{cases}$$

3 Factorisation de Cholesky

Exercice 9. On considère la matrice A suivante

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

1. Justifier pourquoi il existe une factorisation de Cholesky pour A.
2. Calculer la matrice B tel que $A = BB^T$où Best une matrice triangulaire inférieure.
3. Résoudre, en utilisant la factorisation de Cholesky, le système $Ax = b$ où :

$$b = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

4. Calculer le coût de cette résolution en terme d'opérations.

Exercice 10. Effectuez la décomposition de Cholesky des matrices suivantes :

$$\begin{pmatrix} 1 & 2 & 1 \\ 2 & 5 & 4 \\ 1 & 4 & 6 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 1 \\ 1 & 5 & 5 \\ 1 & 5 & 14 \end{pmatrix}, \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

4 Intégration numérique

Exercice 11. Pou une intégration numérique par la méthode de quadrature, on choisit de mettre les nœuds $c_1 = \frac{1}{4}, c_2 = \frac{1}{2}$ et $c_3 = \frac{3}{4}$ avec des poids b_1, b_2 et b_3. On souhaite que la méthode soit d'ordre 3.

1. Trouver les poids b_i où $i \in \{1, 2, 3\}$.
2. Donner, donc les expressions, approchées par cette méthode, des intégrales :

$$\int_0^1 g(x)dx, \int_a^b f(x)dx.$$

3. Cette méthode est-elle d'ordre 4 ? Justifier votre réponse.

Exercice 12. Pour une intégration par des formules de quadratures, on considère les nœuds suivants :

— $(c_1, c_2, c_3, c_4) = (0, 1/3, 2/3, 1)$,

— $(c_1, c_2, c_3, c_4, c_5) = (0, 1/4, 1/2, 3/4, 1)$

1. Déterminer pour chaque cas les poids (b_1, b_2, b_3, b_4) des nœuds.
2. Déterminer l'ordre de ces formules de quadratures.

Exercice 13. Déterminer c_1, b_1, b_2 dans la formule de quadrature suivante afin que son ordre soit maximal.

$$\int_0^1 g(t) \sim b_1 g(0) + b_2 g(c_2)$$

Exercice 14. Montrer que si les nœuds d'une formule de quadrature satisfont $\forall i\ c_i = 1 - c_{s+1-i}$, et si la formule à un ordre $p \geq s$, alors on a nécessairement $b_i = b_{s+1-i}$, c'est à dire qu'elle est symétrique.

5 Méthodes d'Euler et méthode de Runge-Kutta

Exercice 15. Trouver une méthode numérique dans l'esprit de celle de Runge, mais basée sur la méthode des trapèzes.

Exercice 16. Écrire l'équation différentielle

$$y'' + y' = 0, \quad y(0) = 1, \quad y'(0) = 1$$

sous forme résolue du premier ordre. Calculer la solution exacte et la solution numérique avec la méthode de Runge sur $[0, 1]$, avec $h = \frac{1}{2}$.

6 Transformée de Fourier discrète

Exercice 17. Calculer à la main la transformée de Fourier discrète de la suite $(0, 1, 2, 3, 0, -3, -2, -1)$.

Exercice 18. Soit f la fonction 2π-périodique définie par

$$f(x) = \begin{cases} 4x/\pi & \text{si } |x| < \pi \\ 0 & \text{si } x = \pi. \end{cases}$$

1. Montrer que les coefficients de Fourier $\hat{f}(k)$ de f sont :

$$\hat{f}(0) = 0, \qquad \hat{f}(k) = \frac{4(-1)^{k+1}}{i\pi k} \quad k \neq 0.$$

2. Calculer la transformée de Fourier discrète pour y_j, $j = 0, \ldots, N-1$ avec $y_j = f(2\pi j/N)$.
3. Vérifier le résultat obtenu dans l'exercice précédent pour $N = 8$.
4. Estimer la différence $|z_k - \hat{f}(k)|$.