Exercice 6

1. Soit $\mathbf{A} \in \mathcal{M}_n(\mathbb{R})$ une matrice symétrique inversible admettant une factorisation $\mathbf{LU}$ où $\mathbf{L}$ est une matrice triangulaire inférieure à diagonale unité et $\mathbf{U}$ est une matrice triangulaire supérieure. Montrer que $\mathbf{A}$ admet une factorisation $\mathbf{A} = \mathbf{LDL}^t$ où $\mathbf{L}$ est une matrice triangulaire inférieure à diagonale unité et $\mathbf{D}$ est une matrice diagonale inversible.
2. Montrer qu'une matrice $\mathbf{A} \in \mathcal{M}_n(\mathbb{R})$ admet une factorisation $\mathbf{A} = \mathbf{LDL}^t$ où $\mathbf{L}$ est une matrice triangulaire inférieure à diagonale unité et $\mathbf{D}$ est une matrice diagonale dont tous les coefficients diagonaux sont strictement positifs si et seulement si $\mathbf{A}$ est symétrique définie positive. Montrer que cette factorisation est unique.
3. En déduire que $\mathbf{A}$ admet une unique factorisation de Cholesky $\mathbf{A} = \mathbf{BB}^t$ où $\mathbf{B}$ est une matrice triangulaire dont les coefficients diagonaux sont strictement positifs si et seulement si $\mathbf{A}$ est matrice symétrique définie positive.
4. On suppose $\mathbf{A}$ symétrique définie positive. Décrire une méthode permettant de calculer explicitement les coefficients de la matrice $\mathbf{B}$ précédente.
5. *(algo)* Ecrire la fonction CHOLESKY permettant de calculer la matrice $\mathbf{B}$ de la méthode précédente. Quel est le coût de la méthode ?

---

1. On suppose qu'il existe une matrice $\mathbf{U}$ triangulaire supérieure et une matrice $\mathbf{L}$ triangulaire inférieure à diagonale unité telle que $\mathbf{A} = \mathbf{LU}$. Comme de plus, $\mathbf{A}$ est inversible, $\mathbf{U}$ est inversible, et donc que les éléments diagonaux de $\mathbf{U}$ sont non nuls. On définit $\mathbf{D} \in \mathcal{M}_n(\mathbb{R})$ comme la matrice diagonale formée des éléments diagonaux de $\mathbf{U}$:

$$d_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ u_{ii} & \text{si } i = j. \end{cases} \qquad \mathbf{D} = \begin{pmatrix} u_{11} & 0 & \cdots & 0 \\ 0 & u_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & u_{nn} \end{pmatrix}$$

La matrice $\mathbf{D}$ est inversible. Il est clair que $\mathbf{U} = \mathbf{D}\tilde{\mathbf{U}}$ où $\tilde{\mathbf{U}} = \mathbf{D}^{-1}\mathbf{U}$. Les composantes de $\tilde{\mathbf{U}}$ sont

$$\tilde{u}_{ij} = \frac{u_{ij}}{u_{ii}} \quad \forall (i,j) \in [1:n]^2.$$

$\tilde{\mathbf{U}}$ est une matrice triangulaire supérieure à diagonale unité.

Puisque par ailleurs, $\mathbf{A} = \mathbf{LU}$, on a $\mathbf{A} = \mathbf{LD}\tilde{\mathbf{U}}$. Nous allons montrer que $\tilde{\mathbf{U}} = \mathbf{L}^t$, ce qui montrera que $\mathbf{A} = \mathbf{LDL}^t$. Comme $\mathbf{A}$ est symétrique, on a

$$\mathbf{LU} = \mathbf{A} = \mathbf{A}^t = (\mathbf{LD}\tilde{\mathbf{U}})^t = \tilde{\mathbf{U}}^t\mathbf{D}^t\mathbf{L}^t = \underbrace{(\tilde{\mathbf{U}}^t)}_{\text{tri. inf. diag. unité}} \underbrace{(\mathbf{D}^t\mathbf{L}^t)}_{\text{tri. sup}}$$

Par unicité de la factorisation $\mathbf{LU}$ (cf. exercice 4), on en déduit que $\tilde{\mathbf{U}}^t = \mathbf{L}$.

2. On procède par double implication :
   1. Étape 1 : on montre que si $\mathbf{A}$ est symétrique définie positive (SDP), alors il existe une matrice diagonale $\mathbf{D} > 0$ (i.e $d_{ii} > 0$ pour tout $i \in [1:n]$) et une matrice $\mathbf{L}$ triangulaire inférieure à diagonale unité telle que $\mathbf{A} = \mathbf{LDL}^t$.
   2. Étape 2 : On montre que s'il existe une matrice diagonale $\mathbf{D} > 0$ (i.e $d_{ii} > 0$ pour tout $i \in [1:n]$) et une matrice $\mathbf{L}$ triangulaire inférieure à diagonale unité telle que $\mathbf{A} = \mathbf{LDL}^t$, alors $\mathbf{A}$ est symétrique définie positive (SDP).

Enfin, on montre que l'unicité d'une telle factorisation a postériori.

Rappel sur les matrices symétriques définies positives Une matrice $\mathbf{A}$ est symétrique définie positive si

a- $\mathbf{A}$ est symétrique : $\mathbf{A}^t = \mathbf{A}$.

b- $\mathbf{A}$ est définie : s'il existe $x \in \mathbb{R}^n$ tel que $\langle \mathbf{A}x, x\rangle = 0$ alors $x = 0$ ($\langle \cdot, \cdot \rangle$ désigne le produit scalaire euclidien de $\mathbb{R}^n$).

c- $\mathbf{A}$ est positive : pour tout $x \in \mathbb{R}^n$, $\langle \mathbf{A}x, x\rangle \geq 0$.

Remarque. *une matrice SDP est inversible. En effet, si $\mathbf{A}$ est SDP et s'il existe $x \in \mathbb{R}^n$ tel que $\mathbf{A}x = 0$, alors $\langle \mathbf{A}x, x\rangle = 0$. D'après la propriété b, on en déduit que $x = 0$.*

Étape 1 (implication)

On montre que si $\mathbf{A}$ est symétrique définie positive (SDP), alors il existe une matrice diagonale $\mathbf{D} > 0$ (i.e $d_{ii} > 0$ pour tout $i \in [1:n]$) et une matrice $\mathbf{L}$ triangulaire inférieure à diagonale unité telle que $\mathbf{A} = \mathbf{LDL}^t$. Pour cela, on va utiliser la proposition suivante (preuve ci-dessous) :

Proposition 0.1. *Soit $\mathbf{A}$ une matrice symétrique définie positive. Alors tous les mineurs principaux de $\mathbf{A}$ sont non nuls.*

Soit $\mathbf{A}$ une matrice SDP. Alors, d'après la proposition précédente, ses mineurs principaux sont non nuls, et donc $\mathbf{A}$ admet une décomposition $\mathbf{LU}$. Par conséquent, en utilisant la question 1, il existe une matrice diagonale $\mathbf{D}$ inversible et une matrice triangulaire inférieure à diagonale unité $\mathbf{L}$ telle que

$$\mathbf{A} = \mathbf{LDL}^t.$$

Pour terminer l'étape 1, il suffit de montrer que $\mathbf{D} > 0$. Tout d'abord, puisque A est SDP,

$$\forall x \in \mathbb{R}^n,\ \langle Ax, x\rangle > 0 \quad \Leftrightarrow \quad \langle D(L^t x), L^t x\rangle > 0\ \forall x \in \mathbb{R}^n. \tag{1}$$

Par ailleurs, on remarque que pour tout $i \in [1:n]$, $d_{ii} = \langle De_i, e_i\rangle$. Ainsi, en choisissant $x = (L^t)^{-1} e_i$ dans (1) (L^t est inversible car L est une matrice triangulaire à diagonale unité), on a donc

$$d_{ii} = \langle De_i, e_i\rangle = \langle D(L^t x), L^t x\rangle > 0.$$

demonstration de la Proposition 0.1. Nous prouvons cette proposition par contradiction. Décomposons $\mathbf{A}$ par blocs : Soit $k \in \mathbb{N}$, $k \le n$, on a

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_k & \mathbf{B} \\ \mathbf{B}^t & \mathbf{C} \end{bmatrix}$$

où $\mathbf{A}_k$ est la sous matrice de $\mathbf{A}$ formée des k premières colonnes et k premières lignes de $\mathbf{A}$, $\mathbf{B} \in \mathcal{M}_{k,n-k}(\mathbb{R})$, $\mathbf{C} \in \mathcal{M}_{n-k,n-k}(\mathbb{R})$.

Supposons que le mineur principal d'ordre k de $\mathbf{A}$ soit nul. Cela signifie exactement que $\mathbf{A}_k$ n'est pas inversible. Donc il existe $x_k \in \mathbb{R}^k$ non nul tel que $\mathbf{A}_k x_k = 0$. Par conséquent, $(\mathbf{A}_k x_k, x_k) = 0$. On définit maintenant $x \in \mathbb{R}^n$ par

$$x = \begin{bmatrix} x_k \\ \mathbf{0} \end{bmatrix} \quad \text{i.e} \quad x_i = \begin{cases} (x_k)_i & \text{si } i \in [1:k], \\ 0 & \text{si } i \in [k+1:n], \end{cases}$$

où $\mathbf{0}$ désigne le vecteur nul de $\mathbb{R}^{n-k}$. On a

$$\langle \mathbf{A}x, x\rangle = \langle \mathbf{A}_k x_k + \mathbf{B}\,\mathbf{0}, x_k\rangle + \langle \mathbf{B}^t x_k + \mathbf{C}\mathbf{0}, \mathbf{0}\rangle = \langle \mathbf{A}_k x_k, x_k\rangle = 0.$$

Comme $\mathbf{A}$ est symétrique définie positive, on en déduit que $x = 0$ ce qui entraine que $x_k = 0$. On aboutit donc à une contradiction. Donc, pour tout $k \in [1:n]$, la matrice $\mathbf{A}_k$ est inversible. ☐

Étape 2 (implication réciproque)

On montre qu'une matrice de la forme $\mathbf{A} = \mathbf{LDL}^t$ où $\mathbf{L}$ est une matrice triangulaire inférieure à diagonale unité et $\mathbf{D} > 0$ est une matrice diagonale strictement positive, est symétrique définie positive.

- On montre que $\mathbf{A}$ est symétrique :

$$\mathbf{A}^t = (\mathbf{LDL}^t)^t = \underbrace{(\mathbf{L}^t)^t}_{=\mathbf{L}} \underbrace{\mathbf{D}^t}_{=\mathbf{D}} \mathbf{L}^t = \mathbf{LDL}^t = \mathbf{A}.$$

- On montre que $\mathbf{A}$ est positive : on considère $x \in \mathbb{R}^n$ et on pose $y = \mathbf{L}^t x$. Alors,

$$\langle \mathbf{A}x, x\rangle = \langle \mathbf{LDL}^t x, x\rangle = \langle \mathbf{DL}^t x, \mathbf{L}^t x\rangle = \langle \mathbf{D}y, y\rangle \ge 0 \tag{2}$$

  car $\mathbf{D}$ est une matrice positive.

- On montre que $\mathbf{A}$ est définie : soit $x \in \mathbb{R}^n$ tel que $\langle \mathbf{A}x, x\rangle = 0$. Posons $y = \mathbf{L}^t x$. D'après (2), on a

$$\langle \mathbf{D}y, y\rangle = \sum_{i=1}^{n} d_{ii} y_i^2 = 0.$$

  Comme, pour tout $i \in [1:n]$ $d_{ii} > 0$, on en déduit que $y_i = 0$ pour tout $i \in [1:n]$. Donc $y = 0$ et comme $= \mathbf{L}^t$ est inversible, on en déduit que $x = (\mathbf{L}^t)^{-1} y = 0$.

Pour terminer cette question, il reste à montrer l'unicité de la factorisation $\mathbf{LDL}^t$. Supposons que

$$\mathbf{A} = \mathbf{L}_1\mathbf{D}_1(\mathbf{L}_1)^t = \mathbf{L}_2\mathbf{D}_2(\mathbf{L}_2)^t$$

avec $\mathbf{L}_1$ et $\mathbf{L}_2$ triangulaires inférieures à diagonale unité et $\mathbf{D}_1$, $\mathbf{D}_2$ diagonales ayant des coefficients strictement positifs. On pose $\mathbf{U}_1 = \mathbf{D}_1(\mathbf{L}_1)^t$ et $\mathbf{U}_2 = \mathbf{D}_1(\mathbf{L}_2)^t$. Alors $\mathbf{U}_1$ et $\mathbf{U}_2$ sont triagulaires supérieures inversibles. Donc

$$\mathbf{A} = \mathbf{L}_1\mathbf{U}_1 = \mathbf{L}_2\mathbf{U}_2$$

par unicité de la factorisation $\mathbf{LU}$, on trouve que $\mathbf{L}_1 = \mathbf{L}_2$ et $\mathbf{U}_1 = \mathbf{U}_2$. On en déduit que $\mathbf{D}_1 = \mathbf{D}_2$, ce qui termine la preuve d'unicité.

3. Comme pour la question précédente, nous procédons par double implication :

Étape 1 (implication)

Supposons que $\mathbf{A}$ est SDP. Alors, d'après la question 2, il existe une matrice diagonale $\mathbf{D}$ ayant des coefficients diagonaux strictement positifs, et une matrice triangulaire inférieure à diagonale unité telle que $\mathbf{A} = \mathbf{LDL}^t$. On introduit la matrice diagonale $\mathbf{D}^{1/2}$ définie par

$$(d^{1/2})_{ij} = \begin{cases} \sqrt{d_{ii}} & \text{si } i = j, \\ 0 & \text{sinon.} \end{cases}$$

On peut vérifier que $\mathbf{D}^{1/2}\mathbf{D}^{1/2} = \mathbf{D}$. Alors, comme $\mathbf{D}^{1/2} = (\mathbf{D}^{1/2})^t$, on a

$$\mathbf{A} = \mathbf{LDL}^t = \mathbf{LD}^{1/2}\mathbf{D}^{1/2}\mathbf{L}^t = (\mathbf{LD}^{1/2})((\mathbf{D}^{1/2})^t\mathbf{L}^t) = (\mathbf{LD}^{1/2})(\mathbf{LD}^{1/2})^t.$$

Si on définit par $\mathbf{B} = \mathbf{LD}^{1/2}$, on a donc $\mathbf{A} = \mathbf{BB}^t$. Pour conclure, il suffit de vérifier que $\mathbf{B}$ est une matrice triangulaire inférieure dont les éléments diagonaux sont strictement positifs :
- $\mathbf{B}$ est le produit d'une matrice triangulaire inférieure et d'une matrice diagonale, donc $\mathbf{B}$ est triangulaire inférieure
- On a

$$b_{ii} = \sum_{k=1}^{n} \ell_{ik} \underbrace{(d^{1/2})_{ki}}_{=0 \text{ sauf si } k=i} = \underbrace{\ell_{ii}}_{=1}(d^{1/2})_{ii} = \sqrt{d_{ii}} > 0.$$

Étape 2 (implication réciproque)

Réciproquement, soit $\mathbf{A}$ une matrice se décomposant sous la forme $\mathbf{A} = \mathbf{BB}^t$ où $\mathbf{B}$ est une matrice triangulaire inférieure dont les éléments diagonaux sont strictement positifs. Vérifions que $\mathbf{A}$ est SDP.
- $\mathbf{A}$ est symétrique. En effet,
$$\mathbf{A}^t = (\mathbf{BB}^t)^t = (\mathbf{B}^t)^t\,\mathbf{B}^t = \mathbf{BB}^t = \mathbf{A}.$$
- $\mathbf{A}$ est positive. En effet, pour tout $x \in \mathbb{R}^n$,
$$\langle \mathbf{A}x, x\rangle = \langle \mathbf{BB}^t x, x\rangle = \langle \mathbf{B}^t x, \mathbf{B}^t x\rangle = \|\mathbf{B}^t x\|_2^2 \geq 0.$$
- $\mathbf{A}$ est définie. Supposons qu'il existe $x \in \mathbb{R}^n$ tel que $\langle \mathbf{A}x, x\rangle = 0$. Alors, d'après le calcul précédent,
$$\|\mathbf{B}^t x\|_2 = 0,$$
et donc $\mathbf{B}^t x = 0$. Comme $\mathbf{B}$ est inversible ($det(\mathbf{B}^t) = \prod_{i=1}^n b_{i,i} > 0$), on en déduit que $x = 0$.

4. Comme $\mathbf{B}$ est triangulaire inférieure $b_{mp} = 0$ pour tout $p > m$. Donc,

$$\mathbf{A} = \mathbf{BB}^t \Leftrightarrow a_{ij} = \sum_{k=1}^{n} b_{ik}{b^t}_{kj} = \sum_{k=1}^{n} b_{ik}b_{jk} = \sum_{k=1}^{\min(i,j)} b_{ik}b_{jk} \quad \forall (i,j) \in [1:n]^2. \tag{3}$$

Nous allons nous servir de l'égalité précédente pour calculer successivement les **colonnes** de $\mathbf{B}$:

Étape 1 Nous calculons la première colonne de $\mathbf{B}$ c'est à dire les coefficient b_{i1} pour tout $i \in [1:n]$. L'égalité (3) dans le cas $j = 1$ donne

$$a_{i1} = \sum_{k=1}^{1} b_{ik}b_{1k} = b_{11}b_{i1}.$$

On en déduit d'abord que

$$b_{11} = \sqrt{a_{11}},$$

puis, b_{11} étant désormais connu,

$$b_{i1} = \frac{a_{i1}}{a_{11}} \quad \forall i \in [2:n].$$

Étape j On suppose que les $j-1$ premières colonnes de $\mathbf{B}$ ont déjà été calculées lors des étapes précédentes (b_{ik} est connu pour tout $k \in [1:j-1]$ et pour tout $i \in [0:n]$). Nous allons calculer les coefficients de la colonne j de $\mathbf{B}$, c'est à dire b_{ij} pour $i \in [j:n]$ (en effet, on sait déjà que $b_{ij} = 0$ quand $i < j$). On remarque que pour les coefficients qui nous intéressent, $i \geq j$ $(\min(i,j) = j)$, si bien que L'égalité (3) devient

$$a_{ij} = \sum_{k=1}^{j} b_{ik} b_{jk} \quad \forall i \in [j:n].$$

D'après l'hypothèse de récurrence, pour $k \leq j-1$, les coefficients b_{ik} et b_{jk} sont connus. L'équation précédente devient donc

$$a_{ij} = \underbrace{\sum_{k=1}^{j-1} b_{ik} b_{jk}}_{\text{connu}} + b_{ij} b_{jj} \qquad \forall i \in [j:n].$$

L'égalité va nous permettre de calculer tous les termes b_{ij}, en commençant par le coefficient diagonal b_{jj} : on obtient

$$b_{jj} = \sqrt{a_{jj} - \sum_{k=1}^{j-1} b_{jk}^2}, \tag{4}$$

puis

$$b_{ij} = \frac{a_{ij} - \sum_{k=1}^{j-1} b_{ik} b_{jk}}{b_{jj}}. \tag{5}$$

Le fait que $b_{jj} > 0$ provient directement de la question précédente ($\mathbf{A}$ SDP, donc il existe $\mathbf{B}$ triangulaire inférieure à diagonale strictement positive telle que $\mathbf{A} = \mathbf{B}\mathbf{B}^t$). Puisque $b_{jj} \neq 0$, la formule (5) a bien un sens. Noter, que tout comme à l'étape 1, il faut d'abord calculer le coefficient diagonal b_{jj}. Par contre pour $i > j$, les calculs des coefficients b_{ij} peuvent être effectués en parallèle.

Remarque. *Pour construite $\mathbf{B}$, nous avons effectuer une identification colonne par colonne de (3). Nous aurions aussi pu faire cette identification ligne par ligne.*

5.

Fonction Cholesky(A) : B

[A est une matrice SDP,
B est un matrice carrée triangulaire inférieures à diagonale strictement positive
Ce programme calcule la matrice B telle que $A = BB^t$]

[Initialisations]
$n \leftarrow size(A, 1)$
$B \leftarrow zeros(n, n)$ *[Initialisation de B à la matrice nulle]*

[On calcule B colonne par colonne]
Pour j **de** 1 **à** n **faire**

[1- On calcule le terme diagonal b_{jj} par la formule (4)]
$B(j, j) \leftarrow A(j, j)$
Pour k **de** 1 **à** $j - 1$ **faire**
$B(j, j) \leftarrow B(j, j) - B(j, k)^2$
Fin Pour

[2- On calcule les termes extradiagonaux b_{ij}, $i \in [j + 1 : n]$ par la formule (5)]
Pour i **de** $j + 1$ **à** n **faire**
$B(i, j) \leftarrow A(i, j)$
Pour k **de** 1 **à** $j - 1$ **faire**
$B(i, j) \leftarrow B(i, j) - B(i, k) \times B(j, k)$
Fin Pour
$B(i, j) = B(i, j)/B(j, j)$
Fin Pour
Fin Pour
Retourner B ;
Fin

Algorithme 1: Algorithme pour la factorisation de Cholesky

On peut vérifier que le coût de calcul est d'environ $\frac{n^3}{3}$ opérations (multiplications, additions) plus n extractions de racines carrées. Quand la matrice est symétrique définie positive, il est donc plus avantageux d'utiliser la factorisation de Cholesky que la factorisation $\mathbf{LU}$ (dont le coût de calcul est de l'ordre de $\frac{2n^3}{3}$ opérations).