MAT7381 Chapitre 1
Rappels sur les matrices

1.1 Rang d'une matrice et espace vectoriel

Notation Soit **A** une matrice réelle. Nous utiliserons la notation suivante :

$\mathcal{C}(\mathbf{A})$ = L'espace vectoriel engendré par les colonnes de **A**.

$\mathcal{N}(\mathbf{A})$ = le noyau de $\mathbf{A} = \{\boldsymbol{x} \mid \mathbf{A}\boldsymbol{x} = \mathbf{0}\}$

$n(\mathbf{A})$ = la dimension de $\mathcal{N}(\mathbf{A})$

$\mathbf{A}'$ = la transposée de **A**.

$\boldsymbol{e}$ = un vecteur colonne dont toutes les composantes sont égales à 1.

$\boldsymbol{e}_i$ = un vecteur colonne dont toutes les composantes sont nulles sauf la i^e qui est égale à 1.

Définition* 1.1.1** Une matrice symétrique **A** ($n \times n$) est dite ***définie positive si et seulement si $\boldsymbol{x}'\mathbf{A}\boldsymbol{x} > 0$ pour tout vecteur non nul $\boldsymbol{x} \in \mathbb{R}^n$. Elle est ***semi définie positive*** si et seulement si $\boldsymbol{x}'\mathbf{A}\boldsymbol{x} \geq 0$ pour tout vecteur $\boldsymbol{x} \in \mathbb{R}^n$. ■

Remarque *Toute matrice* **A** *qui peut s'écrire sous la forme* $\mathbf{A} = \mathbf{B}'\mathbf{B}$ *est semi-définie positive*: $\boldsymbol{x}'\mathbf{A}\boldsymbol{x} = \boldsymbol{x}'\mathbf{B}'\mathbf{B}\boldsymbol{x} \geq 0$ *pour tout* $\boldsymbol{x}$, *puisque* $\boldsymbol{x}'\mathbf{B}'\mathbf{B}\boldsymbol{x} = (\mathbf{B}\boldsymbol{x})'(\mathbf{B}\boldsymbol{x})$ *est la somme des carrés des composantes de* $\mathbf{B}\boldsymbol{x}$ ■

***Définition* 1.1.2.** La trace d'une matrice carrée $\mathbf{A} = [a_{ij}]$ est la somme des éléments de la diagonale :

$$tr(\mathbf{A}) = \sum_i a_{ii}$$

Théorème 1.1.1

i) $tr(\mathbf{AB}) = tr(\mathbf{BA})$

ii) $tr(\mathbf{A}'\mathbf{A}) \geq 0$

iii) $tr(\mathbf{A}'\mathbf{A}) = 0 \Leftrightarrow \mathbf{A}'\mathbf{A} = \mathbf{0} \Leftrightarrow \mathbf{A} = \mathbf{0}$.

Démonstration Exercice ■

***Définition* 1.1.3** Le rang d'une matrice **A**, noté $r(\mathbf{A})$, est la dimension de $\mathcal{C}(\mathbf{A})$ [qui est égale aussi à la dimension de $\mathcal{C}(\mathbf{A}')$].

Théorème 1.1.2. $r(\mathbf{AB}) \leq r(\mathbf{A})$, $r(\mathbf{AB}) \leq r(\mathbf{B})$.

Démonstration. Les colonnes de **AB** sont des combinaisons linéaires des colonnes de **A**. Donc $\mathcal{C}(\mathbf{AB}) \subset \mathcal{C}(\mathbf{A}) \Rightarrow dim\{\mathcal{C}(\mathbf{AB})\} \leq dim\{\mathcal{C}(\mathbf{A})\} \Rightarrow r(\mathbf{AB}) \leq r(\mathbf{A})$. Par le même argument, $r(\mathbf{B}'\mathbf{A}') \leq r(\mathbf{B}')$, ce qui est équivalent à $r(\mathbf{AB}) \leq r(\mathbf{B})$, puisque $r(\mathbf{B}'\mathbf{A}') = r(\mathbf{AB})$ et $r(\mathbf{B}') = r(\mathbf{B})$ ■

Définition* 1.1.4** Si **A** est une matrice $n \times n$ et $\mathbf{A}\boldsymbol{x} = \lambda\boldsymbol{x}$, $\boldsymbol{x} \neq \mathbf{0}$, où λ est un scalaire, on dit que λ est une ***valeur propre de **A** et $\boldsymbol{x}$ un ***vecteur propre*** correspondant à la valeur propre λ. ■

Chaque valeur propre est solution de l'équation $|\mathbf{A}-\lambda\mathbf{I}_n| = 0$. Il y a donc au plus n valeurs propres distinctes de **A**, puisque $|\mathbf{A} - \lambda\mathbf{I}_n|$ est un polynôme de degré n en λ.

***Définition* 1.1.5** Si $\mathcal{A}$ est un sous-espace vectoriel de $\mathbb{R}^n$, alors le complément orthogonal $\mathcal{A}^{\perp}$ de $\mathcal{A}$ dans $\mathbb{R}^n$ est défini comme l'ensemble des vecteurs $\boldsymbol{x} \in \mathbb{R}^n$ qui sont orthogonaux à tous les vecteurs dans $\mathcal{A}$:

$$\mathcal{A}^{\perp} = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x}'\boldsymbol{y} = \mathbf{0} \quad \forall \boldsymbol{y} \in \mathcal{A}\}$$

***Définition* 1.1.6** Si $\mathcal{A}$ et $\mathcal{B}$ sont deux espaces vectoriels, alors la somme $\mathcal{A} + \mathcal{B}$ est définie comme l'ensemble de tous les vecteurs de la forme $\boldsymbol{a} + \boldsymbol{b}$, $\boldsymbol{a} \in \mathcal{A}$, $\boldsymbol{b} \in \mathcal{B}$:

$$\mathcal{A} + \mathcal{B} = \{\boldsymbol{a} + \boldsymbol{b} \mid \boldsymbol{a} \in \mathcal{A}\ , \boldsymbol{b} \in \mathcal{B}\}$$

La notation $\oplus$ remplace + lorsque $\mathcal{A}$ et $\mathcal{B}$ satisfont $\mathcal{A} \cap \mathcal{B} = \{\mathbf{0}\}$. Dans ce cas, $\dim(\mathcal{A} \oplus \mathcal{B}) = \dim(\mathcal{A}) + \dim(\mathcal{B})$.

Remarques.

i) Étant donné un espace vectoriel $\mathcal{A} \subset \mathbb{R}^n$ de dimension $q < n$, il existe un espace $\mathcal{B}$ de dimension n-q tel que $\mathcal{A} \cap \mathcal{B} = \{\mathbf{0}\}$ et $\dim(\mathcal{A} \oplus \mathcal{B}) = \dim(\mathcal{A}) + \dim(\mathcal{B}) = n$. L'espace $\mathcal{B}$ est dit *complémentaire* à $\mathcal{A}$. Pour le montrer, il suffit de noter que, puisque $q < n$, il existe un vecteur $\boldsymbol{b}_1 \in \mathbb{R}^n$, $\boldsymbol{b}_1 \notin \mathcal{A}$, de sorte que $\mathcal{A} \oplus \mathcal{C}(\boldsymbol{b}_1)$ est un espace de dimension q+1. Si $q+1 < n$, il existe un deuxième vecteur $\boldsymbol{b}_2 \in \mathbb{R}^n$, $\boldsymbol{b}_2 \notin \mathcal{A} \oplus \mathcal{C}(\boldsymbol{b}_1)$ tel que $\mathcal{A} \oplus \mathcal{C}(\boldsymbol{b}_1) \oplus \mathcal{C}(\boldsymbol{b}_2)$ est un espace de dimension q+2. On peut ainsi définir n-q vecteurs $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_{n-q}$ indépendants qui alors constituent une base d'un espace $\mathcal{B}$ de dimension n-q tel que $\mathcal{A} \cap \mathcal{B} = \{\mathbf{0}\}$ et $\dim(\mathcal{A} \oplus \mathcal{B}) = \dim(\mathcal{A}) + \dim(\mathcal{B}) = n$.

ii) Si $\mathcal{A}$ est un sous-espace vectoriel de $\mathbb{R}^n$, alors $\mathbb{R}^n = \mathcal{A} \oplus \mathcal{A}^{\perp}$. Il suffit dans le développement en i) choisir des vecteurs $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_{n-q}$ orthogonaux à $\mathcal{A}$.

iii) Le complément orthogonal de $\mathcal{C}(\mathbf{A})$ est $\mathcal{N}(\mathbf{A}')$: $\mathcal{C}(\mathbf{A})^{\perp} = \mathcal{N}(\mathbf{A}')$. Donc si $\mathbf{A}$ a n lignes, alors $\mathbb{R}^n = \mathcal{C}(\mathbf{A}) \oplus \mathcal{C}(\mathbf{A})^{\perp} = \mathcal{C}(\mathbf{A}) \oplus \quad (\mathbf{A}')$.

iiv) Si $\mathbf{A}$ est $n \times q$, alors $n = dim\{\mathcal{C}(\mathbf{A})\} + dim\{\mathcal{N}(\mathbf{A}')\} = r(\mathbf{A}) + n(\mathbf{A}')$.

Théorème 1.1.3 *Soit* $\mathbf{A}$ *une matrice* $n \times q$. *Alors* $r(\mathbf{A}) = r(\mathbf{AA}') = r(\mathbf{A}'\mathbf{A}) = r(\mathbf{A}')$.

Démonstration de $r(\mathbf{A}) = r(\mathbf{AA}')$ Selon le théorème 1.1.2, $r(\mathbf{A}) \geq r(\mathbf{AA}')$ et il reste donc à montrer que $r(\mathbf{A}) \leq r(\mathbf{AA}')$, ce qui est équivalent à $n(\mathbf{AA}') \leq n(\mathbf{A}')$ puisque, d'après la dernière remarque, $n = r(\mathbf{A}) + n(\mathbf{A}') = r(\mathbf{AA}') + n(\mathbf{AA}')$. Pour montrer que $n(\mathbf{AA}') \leq n(\mathbf{A}')$, nous montrerons que $\mathcal{N}(\mathbf{AA}') \subset \mathcal{N}(\mathbf{A}')$, c'est-à-dire, que $\mathbf{A}'\boldsymbol{x} = \mathbf{0}$ pour tout $\boldsymbol{x} \in \mathcal{N}(\mathbf{AA}')$. Soit $\boldsymbol{x} \in \mathcal{N}(\mathbf{AA}') \Rightarrow \mathbf{AA}'\boldsymbol{x} = \mathbf{0} \Rightarrow \boldsymbol{x}'\mathbf{AA}'\boldsymbol{x} = \mathbf{0} \Rightarrow \mathbf{A}'\boldsymbol{x} = \mathbf{0}$ $\Rightarrow \boldsymbol{x} \in \mathcal{N}(\mathbf{A}')$. Donc $\mathcal{N}(\mathbf{AA}') \subset \mathcal{N}(\mathbf{A}')$, et c'est ce qu'il fallait démontrer. ■

***Définition* 1.1.7** Une matrice $n \times r$ $\mathbf{A}$ est dite *de plein rang* si et seulement si $r(\mathbf{A}) = r$, c'est-à-dire, si et seulement si ses colonnes sont linéairement indépendantes. ■

Théorème 1.1.4 *Si une matrice* $n \times r$ $\mathbf{A}$ *est de plein rang, alors elle a une inverse à gauche, c'est-à-dire, il existe une matrice* $n \times r$ $\mathbf{B}$ telle que $\mathbf{I}_r = \mathbf{B}'\mathbf{A}$, *l'identité d'ordre* r.

Démonstration. Les lignes de $\mathbf{A}$ sont des vecteurs à r composantes qui engendrent un espace dans $\mathbb{R}^r$ de dimension $r(\mathbf{A}') = r(\mathbf{A}) = r$. Donc les lignes de $\mathbf{A}$ engendrent $\mathbb{R}^r$. Pour $i = 1, \ldots, r$, soit $\boldsymbol{e}_i' = (0, 0, \ldots, 1, 0, \ldots, 0)$ un vecteur $1 \times r$ dont les composantes sont toutes nulles sauf la i^e, qui est 1. Alors $\boldsymbol{e}_i$ est combinaison linéaire des lignes de **A**, c'est-à-dire, il existe un vecteur ligne $\boldsymbol{b}_i$' tel que $\boldsymbol{e}_i' = \boldsymbol{b}_i'\mathbf{A}$. En prenant pour **B** la matrice dont les colonnes sont les $\boldsymbol{b}_i$, le système d'équations $\boldsymbol{e}_i' = \boldsymbol{b}_i'\mathbf{A}$ s'écrit, sous forme matricielle, comme $\mathbf{B'A} = \mathbf{I}_r$, ce qu'il fallait démontrer. ■

Théorème 1.1.5 *Si* **A** *est de plein rang, alors* $r(\mathbf{AC}) = r(\mathbf{C})$.

Démonstration. Puisque **A** est de plein rang, il existe par le théorème 1.1.4, une matrice **B** telle que $\mathbf{B'A} = \mathbf{I}_r$, où $r = r(\mathbf{A})$. Utilisant le théorème 1.1.2, nous avons $r(\mathbf{AC}) \geq r(\mathbf{B'AC}) = r(\mathbf{C}) \geq r(\mathbf{AC}) \Rightarrow r(\mathbf{AC}) = r(\mathbf{C})$. En particulier, si **A** est une matrice carrée non-singulière, $r(\mathbf{AC}) = r(\mathbf{C})$. ■

Théorème 1.1.6 *Soit* **A** *une matrice* $m \times n$ *de rang* r. *Alors il existe des matrices de plein rang* **B** $(m \times r)$ *et* **C** $(n \times r)$ *telles que* $\mathbf{A} = \mathbf{BC}'$. ■

Démonstration . Soit **B** une matrice $m \times r$ dont les colonnes forment une base pour $\mathcal{C}(\mathbf{A})$. Alors chaque colonne de **A** peut être exprimée comme une combinaison linéaire des colonnes de **B**. Si $\boldsymbol{a}_i$ est la i^e colonne de **A**, alors il existe un vecteur $\boldsymbol{c}_i$ tel que $\boldsymbol{a}_i = \mathbf{B}\boldsymbol{c}_i$. Ces équations peuvent s'écrire sous la forme $\mathbf{A} = \mathbf{BC}'$, où **C**' est la matrice dont les colonnes sont les $\mathbf{c}_i$. Puisque **B** et **C** ont chacune r colonnes, elles ne peuvent être de rang supérieur à r. Elles ne peuvent être de rang inférieur à r non plus, puisque d'après le théorème 1.1.2, $r = r(\mathbf{A}) = r(\mathbf{BC}') \leq r(\mathbf{B})$ et $r = r(\mathbf{A}) \leq r(\mathbf{C}')$. Donc **B** et **C** sont toutes deux de rang r. ■

Théorème 1.1.7 *Toute valeur propre non nulle de* **AB** *est une valeur propre de* **BA** (*en d'autres termes*, **AB** et **BA** *ont les mêmes valeurs propres non nulles*.)

Démonstration Soit λ une valeur propre non nulle de **AB** et $\boldsymbol{x}$ un vecteur propre correspondant. On montre que λ est alors une valeur propre de **BA**. On a $\mathbf{AB}\boldsymbol{x} = \lambda\boldsymbol{x} \Rightarrow \mathbf{BAB}\boldsymbol{x} = \lambda\mathbf{B}\boldsymbol{x}$, ce qui signifie que $\mathbf{B}\boldsymbol{x}$ est un vecteur propre de **BA** correspondant à la valeur propre λ. ■

Définition* 1.1.8** Les ***valeurs singulières d'une matrice **A** sont les racines carrées des valeurs propres non nulles de **A**'**A** (et donc aussi les valeurs propres non nulles de **AA**', d'après le théorème 1.1.7)

1.2 Matrices idempotentes et projecteurs

***Définition* 1.2.1**
Une matrice **A** est dite idempotente si $\mathbf{A}^2 = \mathbf{A}$. ■

Théorème 1.2.1 *Soit* **A** *est une matrice* ***idempotente***, $\mathbf{A}^2 = \mathbf{A}$, *de rang* r, *et soit* $\mathbf{A} = \mathbf{BC}'$, *où* **B** *et* **C** *sont de plein rang. Alors* $\mathbf{B'C} = \mathbf{C'B} = \mathbf{I}_r$

Démonstration. Puisque **B** et **C** sont de plein rang, par le théorème 1.4, il existe une matrice **U** et une matrice **V** telles que $\mathbf{U'B} = \mathbf{I}_r$ et $\mathbf{V'C} = \mathbf{I}_r$. Alors $\mathbf{A}^2 = \mathbf{A} \Rightarrow \mathbf{BC'BC'} = \mathbf{BC'} \Rightarrow \mathbf{U'BC'BC'V} = \mathbf{U'BC'V} \Rightarrow \mathbf{I}_r\mathbf{C'B}\mathbf{I}_r = \mathbf{I}_r\mathbf{I}_r \Rightarrow \mathbf{C'B} = \mathbf{I}_r$. En transposant les deux membres de la dernière égalité, on obtient le résultat $\mathbf{B'C} = \mathbf{I}_r$. ■

Théorème 1.2.2 *Si* **A** *est une matrice idempotente de rang* r, *alors* $r = tr(\mathbf{A})$.

Démonstration . Selon le théorème 1.1.6 il existe des matrices de plein rang **B** et **C** telles que $\mathbf{A} = \mathbf{BC}'$ et $\mathbf{B}'\mathbf{C} = \mathbf{C}'\mathbf{B} = \mathrm{I}_r$. Alors $tr(\mathbf{A}) = tr(\mathbf{A}^2) = tr(\mathbf{BC}'\mathbf{BC}') = tr(\mathbf{C}'\mathbf{BC}'\mathbf{B}) = tr(\mathbf{I}_r\mathbf{I}_r) = tr(\mathbf{I}_r) = r = r(\mathbf{A})$ ∎

Théorème 1.2.3 *Si* **A** *est une matrice idempotente, alors ses valeurs propres sont* 1 *ou* 0.

Démonstration . Soit λ une valeur propre non nulle de **A** et $\boldsymbol{x}$ un vecteur propre correspondant. Alors $\mathbf{A}\boldsymbol{x} = \lambda\boldsymbol{x} \Rightarrow \mathbf{A}^2\boldsymbol{x} = \lambda\mathbf{A}\boldsymbol{x} \Rightarrow \mathbf{A}\boldsymbol{x} = \lambda(\lambda\boldsymbol{x}) \Rightarrow \lambda\boldsymbol{x} = \lambda^2\boldsymbol{x} \Rightarrow \lambda(1-\lambda)\boldsymbol{x} = 0 \Rightarrow \lambda = 0$ ou 1. ∎

Théorème 1.2.4 (*Théorème de Cochran*) . *Soit* $\mathbf{A} = \mathbf{A}_1 + \ldots + \mathbf{A}_k$. *Alors les deux énoncés suivants sont équivalents* :

i) $\mathbf{A}^2 = \mathbf{A}$ *et* $r(\mathbf{A}) = \sum_i r(\mathbf{A}_i)$

ii) $\mathbf{A}_i^2 = \mathbf{A}_i$ *pour tout i et* $\mathbf{A}_i\mathbf{A}_j = \mathbf{0}$ *pour* $i \neq j$.

Démonstration Par le théorème 1.1.6, il existe des matrices $\mathbf{B}_i$ et $\mathbf{C}_i$ telles que $\mathbf{A}_i = \mathbf{B}_i\mathbf{C}_i'$, où $\mathbf{B}_i$ et $\mathbf{C}_i$ ont $r_i = \mathrm{r}(\mathbf{A}_i)$ colonnes et sont de rang r_i. Soit $\mathbf{B} = [\mathbf{B}_1, \ldots, \mathbf{B}_k]$, et $\mathbf{C} = [\mathbf{C}_1, \ldots, \mathbf{C}_k]$. Alors $\mathbf{A} = \Sigma\mathbf{A}_i = \Sigma\mathbf{B}_i\mathbf{C}_i' = \mathbf{BC}'$ où **B** et **C** ont Σr_i colonnes.

Nous démontrons d'abord que i) ⇒ ii). Supposons que i) est vrai. Alors $r(\mathbf{B}) = r(\mathbf{A})$ car d'une part $r(\mathbf{B}) \leq \Sigma r_i$ et $\Sigma_i r_i = r(\mathbf{A})$ par hypothèse, et d'autre part $r(\mathbf{A}) \leq r(\mathbf{B})$ par le théorème 1.1.2. De même $r(\mathbf{C}) = r(\mathbf{A})$. Donc **B** et **C** sont de plein rang. Par le théorème 1.7, $\mathbf{C}'\mathbf{B} = \mathbf{I}$, c'est-à-dire,

$$\begin{pmatrix} \mathbf{C}_1' \\ \mathbf{C}_2' \\ \vdots \\ \mathbf{C}_k' \end{pmatrix} \begin{pmatrix} \mathbf{B}_1 & \mathbf{B}_2 & \cdots & \mathbf{B}_k \end{pmatrix} = \begin{pmatrix} \mathbf{C}_1'\mathbf{B}_1 & \cdots & \mathbf{C}_1'\mathbf{B}_k \\ \vdots & \ddots & \vdots \\ \mathbf{C}_k'\mathbf{B}_1 & \cdots & \mathbf{C}_k'\mathbf{B}_k \end{pmatrix} = \mathbf{I}$$

d'où $\mathbf{C}_i'\mathbf{B}_i = \mathbf{I}$ et $\mathbf{C}_i'\mathbf{B}_j = \mathbf{0} \Rightarrow \mathbf{A}_i^2 = \mathbf{B}_i\mathbf{C}_i'\mathbf{B}_i\mathbf{C}_i' = \mathbf{B}_i\mathbf{C}_i' = \mathbf{A}_i$ et $\mathbf{A}_i\mathbf{A}_j = \mathbf{B}_i\mathbf{C}_i'\mathbf{B}_i\mathbf{C}_i' = \mathbf{0}.$

Réciproquement, ii) $\Rightarrow \mathbf{A}^2 = (\Sigma_i\mathbf{A}_i)^2 = \Sigma_i\mathbf{A}_i^2 + \Sigma\Sigma_{i\neq j}\mathbf{A}_i\mathbf{A}_i = \Sigma_i\mathbf{A}_i^2 = \Sigma_i\mathbf{A}_i = \mathbf{A}$ et

$r(\mathbf{A}) = tr(\mathbf{A}) = tr(\Sigma_i\mathbf{A}_i) = \Sigma_i tr(\mathbf{A}_i) = \Sigma_i r(\mathbf{A}_i)$ ∎

1.3 Triangularisation et diagonalisation

Transformée de Householder

***Définition* 1.3.1** Une matrice réelle carrée **P** est dite *orthogonale* si et seulement si $\mathbf{P}'\mathbf{P} = \mathbf{I}$. Ceci entraîne que $\mathbf{PP}' = \mathbf{I}$.

Remarque . *Une matrice orthogonale est une transformation rigide, c'est-à-dire, elle préserve les longueurs et les angles:* $\|\mathbf{P}\boldsymbol{x}\| = \|\boldsymbol{x}\|$ *et le produit scalaire entre* $\mathbf{P}\boldsymbol{x}$ *et* $\mathbf{P}\boldsymbol{y}$ *est égal au produit scalaire entre* $\boldsymbol{x}$ et $\boldsymbol{y}$. ∎

Définition* 1.3.2** Soit $\boldsymbol{u}$ un vecteur tel que $\boldsymbol{u}'\boldsymbol{u} = 1$. La matrice $\mathbf{H} = \mathbf{I} - 2\boldsymbol{uu}'$ est appelée ***transformée de Householder.

Remarque La transformée de Householder est orthogonale et symétrique. $\mathbf{H} = \mathbf{I} - 2\boldsymbol{uu}'$ a l'effet d'une réflexion par rapport à l'espace des vecteurs orthogonaux à $\boldsymbol{u}$:
si $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$, où $\boldsymbol{x}_2 = \alpha\boldsymbol{u}$ et $\boldsymbol{x}_1 \perp \boldsymbol{u}$, alors $\mathbf{H}\boldsymbol{x} = \boldsymbol{x}_1 - \boldsymbol{x}_2$. ∎

Théorème 1.3.1 *Soit* $\boldsymbol{x} \in \mathbb{R}^n$ *un vecteur tel que* $\boldsymbol{x}'\boldsymbol{x} = 1$. *Alors il existe une transformée de* Householder $\mathbf{H} = \mathbf{I} - 2\boldsymbol{uu}'$ *dont la première colonne est* $\boldsymbol{x}$.

Démonstration. Il faut que la première colonne de $\mathbf{H} = \mathbf{I} - 2\boldsymbol{uu}'$ soit égale à $\boldsymbol{x} = [x_1, \ldots, x_n]$

$$\begin{pmatrix} 1-u_1^2 \\ -2u_1u_2 \\ \vdots \\ -2u_1u_k \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix}$$

Les solutions sont

$$\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_k \end{pmatrix} = \begin{pmatrix} \sqrt{u_1}(1-x_1)/2 \\ -x_2/2u_1 \\ \vdots \\ -x_k/2u_1 \end{pmatrix}$$

Alors il est aisé de montrer que $\boldsymbol{u}'\boldsymbol{u} = 1$ et donc que $\mathbf{H} = \mathbf{I} - 2\boldsymbol{uu}'$ avec $\boldsymbol{u}' = (u_1, \ldots, u_n)$ est une transformée de *Householder* avec première colonne $\boldsymbol{x}$. ■

Théorème 1.3.2 *Soit* $\mathbf{A}$ *une matrice réelle* $n \times n$. *Alors il existe une matrice orthogonale* $\mathbf{P}$ *telle que* $\mathbf{P}'\mathbf{AP} = \mathbf{T}$ *où* $\mathbf{T}$ *est une matrice triangulaire* :

$$\mathbf{T} = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ 0 & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & t_{nn} \end{pmatrix}$$

Démonstration. Soit λ une valeur propre de $\mathbf{A}$ et $\boldsymbol{x}$ un vecteur propre correspondant tel que $\boldsymbol{x}'\boldsymbol{x} = 1$: $\mathbf{A}\boldsymbol{x} = \lambda\boldsymbol{x}$. Par le théorème 1.3.1, il existe une transformation de Householder $\mathbf{H}$ telle que $\boldsymbol{x} = \mathbf{H}\boldsymbol{e}_1$, où $\boldsymbol{e}_1' = (1, 0, \ldots, 0)$. Alors $\mathbf{A}\boldsymbol{x} = \lambda\boldsymbol{x} \Rightarrow \mathbf{AH}\boldsymbol{e}_1 = \lambda\mathbf{H}\boldsymbol{e}_1 \Rightarrow \mathbf{H}'\mathbf{AH}\boldsymbol{e}_1 = \lambda\mathbf{H}'\mathbf{H}\boldsymbol{e}_1 = \lambda\boldsymbol{e}_1$. Ce dernier énoncé montre que $\mathbf{H}'\mathbf{AH}$ peut s'écrire sous la forme suivante:

$$\mathbf{H}'\mathbf{AH} = \begin{pmatrix} \lambda & \boldsymbol{a}_1' \\ \boldsymbol{0} & \mathbf{A}_1 \end{pmatrix}$$

Par le même argument, il existe une transformée de Householder $\mathbf{H}_1$ telle que

$$\mathbf{H}_1'\mathbf{AH}_1 = \begin{pmatrix} \lambda & \boldsymbol{a}_2' \\ \boldsymbol{0} & \mathbf{A}_2 \end{pmatrix}$$

Soit

$$\mathbf{G}' = \begin{pmatrix} 1 & \boldsymbol{0}' \\ \boldsymbol{0} & \mathbf{H}_1' \end{pmatrix}$$

Alors $\mathbf{G}'\mathbf{H}'\mathbf{AHG}$ est de la forme

$$\mathbf{G}'\mathbf{H}'\mathbf{AHG} = \begin{pmatrix} \lambda & * & * \\ 0 & \lambda_1 & \boldsymbol{a}_3' \\ 0 & 0 & \mathbf{A}_2 \end{pmatrix}$$

où * désigne une composante quelconque (nulle ou non). On vérifie aisément que la matrice $\mathbf{G}'\mathbf{H}'$ est orthogonale. Ce procédé peut être appliqué jusqu'à ce que la matrice soit triangularisée ■

Théorème 1.3.3 *Si* **A** *est une matrice réelle symétrique, alors il existe une matrice orthogonale* **P** *telle que*

$$\mathbf{P}'\mathbf{AP} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

Démonstration. Par le théorème 1.3.2, il existe une matrice orthogonale **P** telle que $\mathbf{P'AP}$ est triangulaire. Mais si **A** est symétrique, alors $\mathbf{P'AP}$ l'est aussi, ce qui entraîne que $\mathbf{P'AP}$ est en fait diagonale ■

Théorème 1.3.4 *Si* **A** *est une matrice symétrique et* **P** *est une matrice orthogonale telle que* $\mathbf{D} = \mathbf{P'AP}$ *est diagonale, alors les valeurs propres de* **A** *sont les éléments de la diagonale de* **D** *et les vecteurs propres correspondants sont les colonnes de* **P.**

Démonstration. Exercice ■

Théorème 1.3.5 *Une matrice réelle symétrique est définie positive si et seulement si ses valeurs propres sont toutes positives ; et elle est semi définie positive si et seulement si ses valeurs propres sont toutes positives ou nulles.*

Démonstration. Soit **A** une matrice réelle symétrique. Alors il existe une matrice **P** orthogonale telle que $\mathbf{P}'\mathbf{AP} = \mathbf{D}$, où $\mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$. Nous allons démontrer la première affirmation. Supposons que **A** est définie positive. Alors $\boldsymbol{x}'\mathbf{A}\boldsymbol{x} > 0$ pour tout $\boldsymbol{x}$. Soit $\boldsymbol{x} = \mathbf{P}\boldsymbol{e}_i$. Alors $\boldsymbol{x}'\mathbf{A}\boldsymbol{x} = \boldsymbol{e}_i'\mathbf{D}\boldsymbol{e}_i = \lambda_i > 0$. Puisque ceci est vrai pour tout i, nous avons montré que si **A** est définie positive, alors $\lambda_i > 0$ pour i. Inversement, supposons que $\lambda_i > 0$ pour tout i. Alors

$$\boldsymbol{x}'\mathbf{A}\boldsymbol{x} = \boldsymbol{x}'\mathbf{PP}'\mathbf{APP}'\boldsymbol{x} = \boldsymbol{y}'\mathbf{D}\boldsymbol{y} \text{ (où } \boldsymbol{y} = \mathbf{P}'\boldsymbol{x}) = \textstyle\sum_{i=1}^{n} \lambda_i y_i^2 > 0.$$ ■

Remarque *Si* **A** *est une matrice réelle symétrique de rang* $r \le n$, *alors* $\mathbf{A} = \mathbf{PDP}'$, *où* **P** *est une matrice* $n \times r$ *telle que* $\mathbf{P}'\mathbf{P} = \mathbf{I}_r$ *et* **D** *est une matrice* $r \times r$ *diagonale dont les éléments de la diagonale sont les valeurs propres non nulles de* **A**. ■

Théorème 1.3.6 *Si* **A** *est une matrice réelle symétrique et définie positive, alors il existe une matrice non-singulière et symétrique* **E** *telle que* $\mathbf{A} = \mathbf{E}^2$; *et il existe une matrice non-singulière et symétrique* **C** *telle que* $\mathbf{C}'\mathbf{AC} = \mathbf{I}$.

Démonstration. Dans la décomposition $\mathbf{P'AP} = \mathbf{D} \Rightarrow \mathbf{A} = \mathbf{PDP}'$, définissons la matrice diagonale $\mathbf{D}^{½}$ dont les éléments de la diagonale sont les racines carrées des éléments de **D.** On a alors $\mathbf{A} = \mathbf{PDP}' = \mathbf{PD}^{½}\mathbf{D}^{½}\mathbf{P}' = \mathbf{PD}^{½}\mathbf{P}'\mathbf{PD}^{½}\mathbf{P}' = \mathbf{E}^2$ où $\mathbf{E} = \mathbf{PD}^{½}\mathbf{P}'$. Ensuite, nous avons $\mathbf{C}'\mathbf{AC} = \mathrm{I}$ avec $\mathbf{C} = \mathbf{C}' = \mathbf{PD}^{-½}\mathbf{P}'$. ■

Théorème 1.3.7 *Soit* $\mathbf{A}_{n\times n}$ *une matrice réelle symétrique et* $\lambda_1 \le \lambda_2 \le \ldots \le \lambda_n$ *ses valeurs propres. Alors pour tout* $\boldsymbol{x} \in \mathbb{R}^n$, $\lambda_1 \le \dfrac{\boldsymbol{x}'\mathbf{A}\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{x}} \le \lambda_n$.

Démonstration. Étant donné que **A** est réelle et symétrique, il existe une matrice orthogonale **P** telle que $\mathbf{P'AP} = \mathbf{D} \Rightarrow \mathbf{A} = \mathbf{PDP}'$, où **D** est une matrice diagonale dont les éléments de la diagonale sont $\lambda_1, \lambda_2, \ldots, \lambda_n$. Soit $\boldsymbol{x}$ un vecteur quelconque dans $\mathbb{R}^n$. On a alors $\boldsymbol{x}'\mathbf{A}\boldsymbol{x} = \boldsymbol{x}'\mathbf{PP}'\mathbf{APP}'\boldsymbol{x} = \boldsymbol{x}'\mathbf{PDP}'\boldsymbol{x} = \boldsymbol{z}'\mathbf{D}\boldsymbol{z}$, où $\boldsymbol{z} = [z_1; z_2;\ldots,z_n]' = \mathbf{P}'\boldsymbol{x}$; et $\boldsymbol{x}'\boldsymbol{x} = \boldsymbol{x}'\mathbf{PP}'\boldsymbol{x} = \boldsymbol{z}'\boldsymbol{z}$. Alors $\dfrac{\boldsymbol{x}'\mathbf{A}\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{x}} = \dfrac{\mathbf{z}'\mathbf{D}\mathbf{z}}{\mathbf{z}'\mathbf{z}} = \dfrac{\sum_{i=1}^n \lambda_i z_i^2}{\sum_{i=1}^n z_i^2}$.

Or $\lambda_1 = \dfrac{\lambda_1 \sum_{i=1}^n z_i^2}{\sum_{i=1}^n z_i^2} \le \dfrac{\sum_{i=1}^n \lambda_i z_i^2}{\sum_{i=1}^n z_i^2} \le \dfrac{\lambda_n \sum_{i=1}^n z_i^2}{\sum_{i=1}^n z_i^2} = \lambda_n$, ce qui complète la démonstration. ■

1.4 Dérivées de vecteurs et de formes quadratiques

Théorème 1.4.1 *Soit $\mathbf{A}_{n\times q}$ une matrice réelle et $\boldsymbol{x}$ un vecteur dans $\mathbb{R}^q$. Alors le vecteur des dérivées de $\mathbf{A}\boldsymbol{x}$ par rapport à la j^e composante x_j de $\boldsymbol{x}$ est*

$$\frac{\partial}{\partial x_j}\mathbf{A}\boldsymbol{x} = \mathbf{A}\boldsymbol{e}_j$$

où $\boldsymbol{e}_j = [0 ; 0; \ldots;1;\ldots;0]'$, la valeur « 1 » étant en j^e position.

Démonstration Soient $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_q$ les colonnes de $\mathbf{A}$ et $\boldsymbol{x} = [x_1; x_2;\ldots,x_q]'$. Alors

$$\mathbf{A}\boldsymbol{x} = [a_1 | \ldots | a_j | \ldots | a_q]\begin{bmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_q \end{bmatrix} = x_1\boldsymbol{a}_1 + \ldots + x_j\boldsymbol{a}_j + \ldots + x_q\boldsymbol{a}_q .$$

Donc $\dfrac{\partial}{\partial x_j}\mathbf{A}\boldsymbol{x} = \dfrac{\partial}{\partial x_j}x_1\boldsymbol{a}_1 + \ldots + \dfrac{\partial}{\partial x_j}x_j\boldsymbol{a}_j + \ldots + \dfrac{\partial}{\partial x_j}x_q\boldsymbol{a}_q = 0 + \ldots + \dfrac{\partial}{\partial x_j}x_j\boldsymbol{a}_j + \ldots + 0 = \boldsymbol{a}_j$ ■

***Corollaire* 1.4.1** S*i* $\boldsymbol{u}(x_1;\ldots; x_j;\ldots; x_q) = \begin{bmatrix} u_1 \\ \vdots \\ u_j \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} u_1(x_1;\ldots;x_q) \\ \vdots \\ u_j(x_1;\ldots;x_q) \\ \vdots \\ u_n(x_1;\ldots;x_q) \end{bmatrix}$, et $\dfrac{\partial \boldsymbol{u}}{\partial \boldsymbol{x}}$ la matrice dont la $(i ; j)^e$ composante est $\left[\dfrac{\partial u_i}{\partial x_j}\right]$. Lorsque $\boldsymbol{u} = \mathbf{A}\boldsymbol{x}$, alors $\dfrac{\partial \boldsymbol{u}}{\partial \boldsymbol{x}} = \mathbf{A}$. ■

Théorème 1.4.2 *Soit $\mathbf{A}_{n\times n}$ une matrice réelle symétrique et $\boldsymbol{x}$ un vecteur dans $\mathbb{R}^q$. Alors* $\dfrac{\partial}{\partial \boldsymbol{x}}\boldsymbol{x}'\mathbf{A}\boldsymbol{x} = 2\boldsymbol{x}'\mathbf{A}$.

Démonstration $\dfrac{\partial}{\partial x_j}\boldsymbol{x}'\mathbf{A}\boldsymbol{x} = \lim\limits_{\delta\to 0}\dfrac{(\boldsymbol{x}+\delta\boldsymbol{e}_j)'\mathbf{A}(\boldsymbol{x}+\delta\boldsymbol{e}_j) - \boldsymbol{x}'\mathbf{A}\boldsymbol{x}}{\delta}$, où $\boldsymbol{e}_j = [0 ; 0; \ldots;1;\ldots;0]'$, la valeur « 1 » étant en j^e position.

$$\lim_{\delta\to 0}\frac{(\boldsymbol{x}+\delta\boldsymbol{e}_j)'\mathbf{A}(\boldsymbol{x}+\delta\boldsymbol{e}_j) - \boldsymbol{x}'\mathbf{A}\boldsymbol{x}}{\delta} = \lim_{\delta\to 0}\frac{\delta\boldsymbol{e}_j'\mathbf{A}\boldsymbol{x} + \delta\boldsymbol{x}'\mathbf{A}\boldsymbol{e}_j + \delta^2\boldsymbol{e}_j'\mathbf{A}\boldsymbol{e}_j}{\delta}$$

$= \boldsymbol{x}'\mathbf{A}\boldsymbol{e}_j + \boldsymbol{x}'\mathbf{A}\boldsymbol{e}_j = 2\boldsymbol{x}'\mathbf{A}\boldsymbol{e}_j = 2\boldsymbol{x}'\boldsymbol{a}_j$ où $\boldsymbol{a}_j$ est la j^e colonne de $\mathbf{A}$. Le résultat suit. ■

1.5 Matrices idempotentes et projecteurs

Théorème 1.5.1 *Soit*

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}$$

où $\mathbf{A}_{11}$ *et* $\mathbf{A}_{22}$ *sont carrées. Le déterminant de* $\mathbf{A}$ *est alors donné par* :

$$|\mathbf{A}| = |\mathbf{A}_{11}|\,|\mathbf{A}_{22}-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}| = |\mathbf{A}_{22}|\,|\mathbf{A}_{11}-\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}|$$

Démonstration. Nous utiliserons le fait que $\begin{vmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{B} & \mathbf{C} \end{vmatrix} = |\mathbf{A}|\,|\mathbf{C}|$ et que le déterminant d'un produit est égal au produit des déterminants.

Puisque $\begin{vmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{I} \end{vmatrix} = 1$, on a $|\mathbf{A}| = \begin{vmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{I} \end{vmatrix} \begin{vmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{vmatrix} =$

$$\left|\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{I} \end{pmatrix}\begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}\right| = \begin{vmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ 0 & \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \end{vmatrix} = |\mathbf{A}_{11}|\,|\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}|$$

On démontre la deuxième inégalité en multipliant $\mathbf{A}$ à droite par $\begin{pmatrix} \mathbf{I} & -\mathbf{A}_{12}\mathbf{A}_{11}^{-1} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}$ ■

Théorème 1.5.2 *Soit*

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}$$

où $\mathbf{A}_{11}$ *et* $\mathbf{A}_{22}$ *sont carrées. Alors*

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{A}_{22.1}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{A}_{22.1}^{-1} \\ -\mathbf{A}_{22.1}^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{A}_{22.1}^{-1} \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{A}_{11.2}^{-1} & -\mathbf{A}_{11.2}^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{A}_{11.2}^{-1} & \mathbf{A}_{22}^{-1} + \mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{A}_{11.2}^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \end{pmatrix}$$

où $\mathbf{A}_{11.2} = \mathbf{A}_{11}-\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}$ et $\mathbf{A}_{22.1} = \mathbf{A}_{22}-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}$

Démonstration. Il suffit de vérifier que le produit $\mathbf{A}^{-1}\mathbf{A}$ donne bien la matrice identité. ■

Théorème 1.5.3 *Soit* $\mathbf{A}$ *une matrice réelle* $n\times n$ *nonsingulière et* $\boldsymbol{u}$ *un vecteur* $n\times 1$ *tels que* $\mathbf{A}$-$\boldsymbol{uu}$' *soit non singulière. Alors* $(\mathbf{A}-\boldsymbol{uu}')^{-1} = \mathbf{A}^{-1} + \dfrac{\mathbf{A}^{-1}\boldsymbol{uu}'\mathbf{A}^{-1}}{1-\boldsymbol{u}'\mathbf{A}^{-1}\boldsymbol{u}}$ ■

Démonstration. Il suffit d'effectuer le produit $(\mathbf{A}-\boldsymbol{uu}')\left(\mathbf{A}^{-1} + \dfrac{\mathbf{A}^{-1}\boldsymbol{uu}'\mathbf{A}^{-1}}{1-\boldsymbol{u}'\mathbf{A}^{-1}\boldsymbol{u}}\right)$ pour voir qu'il donne bien la matrice identité. ■