ANNEXE 5

CALCUL MATRICIEL ET GÉOMÉTRIE DANS $\mathbb{R}^K$

A.5.1 ÉLÉMENTS DE CALCUL MATRICIEL

A.5.1.1 Définitions

Une ***matrice*** est un tableau rectangulaire de nombres comportant par exemple K *lignes et* L *colonnes*, où K et L sont des entiers positifs. Elle est généralement représentée par une lettre majuscule comme, par exemple :

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1l} & \cdots & a_{1L} \\ a_{21} & a_{22} & \cdots & a_{2l} & \cdots & a_{2L} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kl} & \cdots & a_{kL} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{K1} & a_{K2} & \cdots & a_{Kl} & \cdots & a_{KL} \end{pmatrix} = (a_{kl})$$

où a_{kl} représente l'élément situé en ligne k et en colonne l ($k = 1, \ldots, K; l = 1, \ldots, L$). Les valeurs de K et L définissent la *dimension* de $\mathbf{A}$; cette matrice est dite « $K \times L$ », et notée parfois $\mathbf{A}_{K\times L}$.

Une matrice est ***carrée*** si $K = L$, c'est-à-dire si elle compte autant de lignes que de colonnes.

Un ***vecteur-colonne*** est une matrice constituée d'une seule colonne ($L = 1$) ; un ***vecteur-ligne*** est une matrice comptant une seule ligne ($K = 1$).

EXEMPLE A.5.1

Un tableau individus $\times$ caractères contenant les valeurs x_{ik} de p variables auprès de n individus ($i = 1, \ldots, n; k = 1, \ldots, p$) peut être représenté par une matrice $\mathbf{X}$ de dimension $n \times p$.

A.5.1.2 Opérations élémentaires

a) Transposée d'une matrice

Transposer une matrice consiste à permuter le rôle des lignes et des colonnes. La ***transposée*** de $\mathbf{A}$ est notée $\mathbf{A}'$.

EXEMPLE A.5.2

Si $\mathbf{A} = \begin{pmatrix} 1 & 3 & -1 \\ 4 & 1 & 2 \end{pmatrix}$, alors $\mathbf{A}' = \begin{pmatrix} 1 & 4 \\ 3 & 1 \\ -1 & 2 \end{pmatrix}$.

Si $\mathbf{A}$ est de dimension $K \times L$, sa transposée $\mathbf{A}'$ est de dimension $L \times K$. Observons aussi que $(\mathbf{A}')' = \mathbf{A}$.

La matrice carrée $\mathbf{A}_{K \times K}$ est ***symétrique*** si $\mathbf{A}' = \mathbf{A}$, c'est-à-dire encore si $a_{kl} = a_{lk}$ pour tout $k, l \in \{1, \ldots, K\}$.

EXEMPLE A.5.3

La matrice $\mathbf{A} = \begin{pmatrix} 1 & 2 & 5 \\ 2 & 0 & 3 \\ 5 & 3 & -1 \end{pmatrix}$ est symétrique.

b) Addition de deux matrices

On peut *sommer* deux matrices $\mathbf{A} = (a_{kl})$ et $\mathbf{B} = (b_{kl})$ de mêmes dimensions : $\mathbf{A}+\mathbf{B} = (a_{kl}+b_{kl})$. Il suffit de sommer les éléments appartenant à la même ligne et à la même colonne.

EXEMPLE A.5.2 (suite 1)

Si $\mathbf{A} = \begin{pmatrix} 1 & 3 & -1 \\ 4 & 1 & 2 \end{pmatrix}$ et $\mathbf{B} = \begin{pmatrix} -1 & -2 & 4 \\ 2 & 1 & 2 \end{pmatrix}$,

$$\mathbf{A} + \mathbf{B} = \begin{pmatrix} 0 & 1 & 3 \\ 6 & 2 & 4 \end{pmatrix}.$$

Insistons sur le fait qu'on ne peut additionner que des matrices ayant les mêmes dimensions. De plus, si $\mathbf{A}$ et $\mathbf{B}$ sont toutes deux de dimension $K \times L$, leur somme $\mathbf{A} + \mathbf{B}$ sera elle aussi une matrice de dimension $K \times L$.

Citons enfin deux propriétés élémentaires liées à la somme de matrices.

Propriété A.5.1 (commutativité) Soient $\mathbf{A}$ et $\mathbf{B}$ deux matrices de mêmes dimensions : $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$.

Propriété A.5.2 (associativité) Soient $\mathbf{A}$, $\mathbf{B}$ et $\mathbf{C}$ trois matrices de mêmes dimensions : $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$.

c) Multiplication d'une matrice par un scalaire

Multiplier une matrice $\mathbf{A} = (a_{kl})$ par un ***scalaire*** λ (c'est-à-dire un nombre réel) consiste à multiplier tous les éléments de $\mathbf{A}$ par λ :

$$\lambda\mathbf{A} = (\lambda a_{kl}).$$

EXEMPLE A.5.2 (suite 2)

Si $\mathbf{A} = \begin{pmatrix} 1 & 3 & -1 \\ 4 & 1 & 2 \end{pmatrix}$, alors $2\mathbf{A} = \begin{pmatrix} 2 & 6 & -2 \\ 8 & 2 & 4 \end{pmatrix}$.

d) Produit de deux matrices

La *multiplication* de deux matrices $\mathbf{A}$ et $\mathbf{B}$ ne se conçoit que si le nombre de colonnes de $\mathbf{A}$ est égal au nombre de lignes de $\mathbf{B}$. Ainsi, le produit de $\mathbf{A}_{K\times L} = (a_{kl})$ par $\mathbf{B}_{L\times H} = (b_{lh})$ nous donne une matrice $\mathbf{C}_{K\times H} = (c_{kh})$ où

$$c_{kh} = \sum_{l=1}^{L} a_{kl}b_{lh} \qquad k = 1,\ldots,K; h = 1,\ldots,H.$$

EXEMPLE A.5.2 (suite 3)

Si $\mathbf{A} = \begin{pmatrix} 1 & 3 & -1 \\ 4 & 1 & 2 \end{pmatrix}$ et $\mathbf{B} = \begin{pmatrix} -1 & 2 \\ 0 & 1 \\ -2 & 1 \end{pmatrix}$, nous avons :

$$\mathbf{C} = \mathbf{A}\times\mathbf{B} = \begin{pmatrix} 1 & 3 & -1 \\ 4 & 1 & 2 \end{pmatrix} \times \begin{pmatrix} -1 & 2 \\ 0 & 1 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ -8 & 11 \end{pmatrix}.$$

Ainsi, l'élément c_{11} (situé sur la *première ligne* et la *première colonne* de la matrice $\mathbf{C}$) est obtenu en sommant les produits des éléments de la *première ligne* de $\mathbf{A}$ par ceux de la *première colonne* de $\mathbf{B}$ qui occupent « la même position » :

$$1 = 1\times(-1) + 3\times 0 + (-1)\times(-2).$$

Dans la suite, le produit de $\mathbf{A}$ par $\mathbf{B}$ sera désigné indifféremment par $\mathbf{A}\times\mathbf{B}$ ou par $\mathbf{AB}$.

Les propriétés suivantes se démontrent aisément.

Propriété A.5.3 (associativité) Soient les matrices $\mathbf{A}_{K\times L}$, $\mathbf{B}_{L\times H}$ et $\mathbf{C}_{H\times Q}$:

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}).$$

Propriété A.5.4 (distributivité) Soient les matrices $\mathbf{A}_{K\times L}$, $\mathbf{B}_{L\times H}$ et $\mathbf{D}_{L\times H}$:

$$\mathbf{A}(\mathbf{B}+\mathbf{D}) = \mathbf{AB}+\mathbf{AD}.$$

De même, si l'on considère les matrices $\mathbf{B}_{L\times H}$, $\mathbf{D}_{L\times H}$ et $\mathbf{C}_{H\times Q}$:

$$(\mathbf{B}+\mathbf{D})\mathbf{C} = \mathbf{BC}+\mathbf{DC}.$$

Considérons enfin les matrices carrées $\mathbf{E}$ et $\mathbf{F}$, de mêmes dimensions $K \times K$. Nous pouvons facilement vérifier que $\mathbf{EF}$ n'est généralement pas égal à $\mathbf{FE}$. Contrairement à l'opération d'addition, celle de la multiplication de deux matrices ne jouit pas de la propriété de commutativité.

A.5.1.3 Utilisation de la notation matricielle dans l'analyse statistique

Le recours à la notation matricielle permet de simplifier l'écriture. Prenons l'exemple suivant.

EXEMPLE A.5.4

Considérons la matrice $\mathbf{X}$ représentant le tableau individus $\times$ caractères relatif à une série statistique bivariée de taille $n = 5$, où les éléments de la première colonne correspondent aux valeurs x_i de la première variable et ceux de la seconde colonne aux valeurs y_i observées pour l'autre variable :

$$\mathbf{X} = \begin{pmatrix} 2 & 7 \\ 4 & 10 \\ 5 & 9 \\ 6 & 13 \\ 8 & 11 \end{pmatrix}.$$

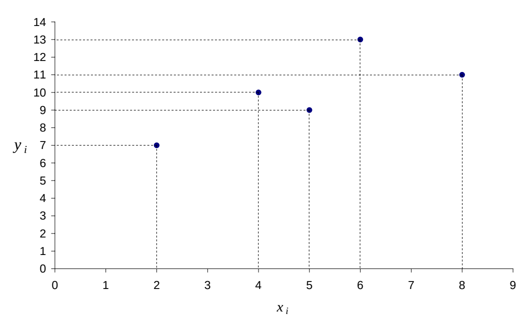

FIG. 5.1 – Série bivariée

Les moyennes marginales valent :

$$\begin{cases} \overline{x} &= \dfrac{2+4+5+6+8}{5} = 5; \\ \overline{y} &= \dfrac{7+10+9+13+11}{5} = 10. \end{cases}$$

D'autre part, pour calculer les variances marginales, nous pouvons *centrer* les séries marginales en remplaçant les valeurs x_i par $w_i = x_i - \overline{x}$ et y_i par $z_i = y_i - \overline{y}$. Nous obtenons alors la ***matrice des valeurs centrées***

$$\mathbf{X}_\mathrm{c} = \begin{pmatrix} -3 & -3 \\ -1 & 0 \\ 0 & -1 \\ 1 & 3 \\ 3 & 1 \end{pmatrix}$$

où les valeurs w_i et z_i sont contenues dans les colonnes de $\mathbf{X}_\mathrm{c}$. Les variances marginales sont égales à

$$\begin{cases} s_x^2 &= \dfrac{1}{n}\sum_i w_i^2 = \dfrac{9+1+0+1+9}{5} = 4; \\ s_y^2 &= \dfrac{1}{n}\sum_i z_i^2 = \dfrac{9+0+1+9+1}{5} = 4. \end{cases}$$

Par ailleurs, la covariance vaut

$$s_{xy} = \frac{1}{n}\sum_i w_i z_i = \frac{9+0+0+3+3}{5} = 3.$$

Le coefficient de corrélation vaut, quant à lui :

$$r = \frac{s_{xy}}{s_x s_y} = \frac{3}{2 \times 2} = 0.75.$$

Tous ces résultats peuvent aussi être aisément obtenus par calcul matriciel. En effet, la matrice $\mathbf{V}$ donnée par la relation

$$\mathbf{V} = \frac{1}{n}\mathbf{X}_\mathrm{c}'\mathbf{X}_\mathrm{c}$$

est la ***matrice de variance-covariance*** de la série observée :

$$\begin{aligned} \mathbf{V} &= \frac{1}{5}\begin{pmatrix} -3 & -1 & 0 & 1 & 3 \\ -3 & 0 & -1 & 3 & 1 \end{pmatrix} \times \begin{pmatrix} -3 & -3 \\ -1 & 0 \\ 0 & -1 \\ 1 & 3 \\ 3 & 1 \end{pmatrix} \\ &= \frac{1}{5}\begin{pmatrix} 20 & 15 \\ 15 & 20 \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} s_x^2 & s_{xy} \\ s_{xy} & s_y^2 \end{pmatrix}. \end{aligned}$$

Par ailleurs, en divisant les éléments $(x_i - \overline{x})$ de la première colonne de $\mathbf{X}_\mathrm{c}$ par $s_x = 2$ et ceux de la seconde, $(y_i - \overline{y})$, par $s_y = 2$, nous obtenons la ***matrice des valeurs centrées réduites*** :

$$\mathbf{X}_\mathrm{cr} = \begin{pmatrix} -3/2 & -3/2 \\ -1/2 & 0 \\ 0 & -1/2 \\ 1/2 & 3/2 \\ 3/2 & 1/2 \end{pmatrix}.$$

La matrice

$$\mathbf{R} = \frac{1}{n}\mathbf{X}'_\mathrm{cr}\mathbf{X}_\mathrm{cr} = \begin{pmatrix} 1 & 0.75 \\ 0.75 & 1 \end{pmatrix} = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix}$$

est alors la ***matrice de corrélation***.

Observons que les matrices $\mathbf{V}$ et $\mathbf{R}$ sont toutes deux symétriques.

L'intérêt de cette écriture est de pouvoir l'utiliser *quels que soient les nombres de variables et d'individus*.

A.5.1.4 Rang d'une matrice

Considérons Q vecteurs $\mathbf{y}_1, \ldots, \mathbf{y}_Q$ de mêmes dimensions[1]. Ces vecteurs sont ***linéairement indépendants*** si l'égalité[2]

$$\sum_{q=1}^{Q} \alpha_q \mathbf{y}_q = \mathbf{0}$$

n'est réalisée que pour $\alpha_1 = \alpha_2 = \ldots = \alpha_Q = 0$. Ceci revient encore à dire que $\mathbf{y}_1, \ldots, \mathbf{y}_Q$ sont linéairement indépendants si aucun de ces vecteurs n'est combinaison linéaire des autres.

EXEMPLE A.5.5

Considérons les vecteurs $\mathbf{y}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $\mathbf{y}_2 = \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}$ et $\mathbf{y}_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$. Sont-ils linéairement indépendants ? Pour répondre à cette question, recherchons les scalaires α_1, α_2 et α_3 tels que

(1) Il peut s'agir aussi bien de vecteurs-lignes que de vecteurs-colonnes.

(2) Le vecteur $\mathbf{0}$ est le vecteur nul (vecteur dont toutes les composantes sont nulles) de même dimension que les vecteurs $\mathbf{y}_1, \ldots, \mathbf{y}_Q$.

$\alpha_1\mathbf{y}_1 + \alpha_2\mathbf{y}_2 + \alpha_3\mathbf{y}_3 = \mathbf{0}$. Nous devons résoudre le système d'équations :

$$\begin{cases} \alpha_1 + 2\alpha_2 + \alpha_3 = 0 \\ -\alpha_2 = 0 \\ \alpha_1 + 2\alpha_2 - \alpha_3 = 0. \end{cases}$$

La deuxième équation nous donne $\alpha_2 = 0$ et nous permet de réécrire le système sous la forme suivante :

$$\begin{cases} \alpha_1 + \alpha_3 = 0 \\ \alpha_2 = 0 \\ \alpha_1 - \alpha_3 = 0. \end{cases}$$

En sommant membre à membre les première et troisième équations, nous constatons que $2\alpha_1 = 0$ et, par conséquent, que $\alpha_1 = \alpha_2 = \alpha_3 = 0$. L'égalité $\alpha_1\mathbf{y}_1 + \alpha_2\mathbf{y}_2 + \alpha_3\mathbf{y}_3 = \mathbf{0}$ n'est donc vérifiée que si α_1, α_2 et α_3 sont tous les trois nuls ; nous pouvons en conclure que les vecteurs $\mathbf{y}_1$, $\mathbf{y}_2$ et $\mathbf{y}_3$ sont linéairement indépendants.

En revanche, les vecteurs $\mathbf{y}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $\mathbf{y}_2 = \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}$ et $\mathbf{y}_4 = \begin{pmatrix} 4 \\ -1 \\ 4 \end{pmatrix}$ ne sont pas linéairement indépendants. En effet, $\mathbf{y}_4$ est combinaison linéaire de $\mathbf{y}_1$ et $\mathbf{y}_2$: $\mathbf{y}_4 = 2\mathbf{y}_1 + \mathbf{y}_2$.

Considérons à présent une matrice $\mathbf{A}$ de dimension $K \times L$, avec $K \geqslant L$: la matrice $\mathbf{A}$ compte donc autant ou davantage de lignes que de colonnes. Dans ce cas le ***rang*** de $\mathbf{A}$ est le nombre maximum de colonnes de $\mathbf{A}$ qui sont linéairement indépendantes. Dans le cas où $K < L$, le rang de $\mathbf{A}$ correspond au nombre maximum de lignes de $\mathbf{A}$ linéairement indépendantes. Nous avons donc nécessairement

$$\operatorname{rang}(\mathbf{A}) \in \{1, 2, \ldots, \min(K, L)\}.$$

On vérifie par ailleurs que

$$\operatorname{rang}(\mathbf{A}) = \operatorname{rang}(\mathbf{A}') = \operatorname{rang}(\mathbf{A}'\mathbf{A}).$$

EXEMPLE A.5.5 (suite 1)

Soit la matrice $\mathbf{A}_{3\times 3}$ dont les colonnes sont les vecteurs $\mathbf{y}_1$, $\mathbf{y}_2$ et $\mathbf{y}_3$:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 0 \\ 1 & 2 & -1 \end{pmatrix}.$$

Nous avons déjà vérifié ci-dessus que les trois colonnes de $\mathbf{A}$ étaient linéairement indépendantes. On peut, suivant une démarche analogue, montrer que les trois lignes de $\mathbf{A}$ sont également linéairement indépendantes. La matrice $\mathbf{A}$ est de rang égal à 3 : $\mathrm{rang}(\mathbf{A}) = 3$.

Déterminons à présent le rang de la matrice $\mathbf{B}_{3\times 3}$ dont les colonnes sont les vecteurs $\mathbf{y}_1$, $\mathbf{y}_2$ et $\mathbf{y}_4$:

$$\mathbf{B} = \begin{pmatrix} 1 & 2 & 4 \\ 0 & -1 & -1 \\ 1 & 2 & 4 \end{pmatrix}.$$

Nous l'avons vu, les trois colonnes de $\mathbf{B}$ ne sont pas linéairement indépendantes : $\mathrm{rang}(\mathbf{B}) < 3$. Puisque deux colonnes quelconques de $\mathbf{B}$ sont toujours linéairement indépendantes (aucune colonne de $\mathbf{B}$ n'est multiple d'une autre colonne), la matrice $\mathbf{B}$ est de rang égal à 2.

A.5.1.5 Déterminant d'une matrice carrée $\mathbf{A}_{K\times K}$

a) Définition de $|\mathbf{A}|$

Le déterminant d'une matrice carrée $\mathbf{A}_{K\times K} = (a_{kl})$ est un nombre réel, noté $|\mathbf{A}|$, qui se calcule comme suit :

- $K = 1$: si $\mathbf{A} = a$, alors $|\mathbf{A}| = a$;
- $K = 2$: si $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, alors $|\mathbf{A}| = a_{11}a_{22} - a_{21}a_{12}$;
- $K = 3$: si $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, alors

$$\begin{aligned} |\mathbf{A}| &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\quad -a_{11}a_{23}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33}; \end{aligned}$$

- Quand $K > 3$, la définition d'un déterminant n'est plus utilisable en pratique pour son calcul ; on recourt plutôt au résultat suivant.

b) Détermination de $|\mathbf{A}|$

Désignons par $\mathbf{M}_{kl}$ la sous-matrice carrée de $\mathbf{A}$ obtenue en *supprimant* la k-ème ligne et la l-ème colonne de $\mathbf{A}$. Le ***cofacteur*** de a_{kl}, noté A_{kl}, est défini par

$$A_{kl} = (-1)^{k+l}|\mathbf{M}_{kl}|.$$

Dans ce cas, $|\mathbf{A}|$ peut être calculé par l'une quelconque des deux relations suivantes :

$$\begin{aligned} |\mathbf{A}| &= \sum_{l=1}^{K} a_{kl} A_{kl} \qquad k \in \{1, \ldots, K\} \\ |\mathbf{A}| &= \sum_{k=1}^{K} a_{kl} A_{kl} \qquad l \in \{1, \ldots, K\}. \end{aligned}$$

EXEMPLE A.5.6

$$\begin{vmatrix} -3 & 2 \\ 1 & -3 \end{vmatrix} = 9 - 2 = 7;$$

$$\begin{aligned} \begin{vmatrix} -1 & 1 & 2 \\ -1 & -2 & 1 \\ 2 & 1 & -3 \end{vmatrix} &= (-1) \begin{vmatrix} -2 & 1 \\ 1 & -3 \end{vmatrix} - \begin{vmatrix} -1 & 1 \\ 2 & -3 \end{vmatrix} + 2 \begin{vmatrix} -1 & -2 \\ 2 & 1 \end{vmatrix} \\ &= (-1) \times 5 - 1 \times 1 + 2 \times 3 = 0. \end{aligned}$$

c) Propriétés

Les propriétés essentielles du déterminant $|\mathbf{A}|$ d'une matrice carrée $\mathbf{A}_{K \times K} = (a_{kl})$ sont les suivantes :

1. $|\mathbf{A}| = |\mathbf{A}'|$.
2. Si $\mathbf{A}$ possède une ligne ou une colonne nulle, son déterminant est nul.
3. Si $\mathbf{A}$ possède deux lignes ou deux colonnes égales, son déterminant est nul.
4. Si $\mathbf{A}$ n'est pas de rang maximal K — au moins une de ses lignes (colonnes) est combinaison linéaire des autres lignes (colonnes) —, son déterminant est nul.
5. Si $\mathbf{A}$ est diagonale — $a_{kl} = 0$ pour tout $k, l \in \{1, \ldots, K\}$ avec $k \neq l$ —, $|\mathbf{A}| = a_{11} a_{22} \ldots a_{KK}$.
6. Si $\mathbf{A}$ et $\mathbf{B}$ sont de même dimension, $|\mathbf{AB}| = |\mathbf{A}| \times |\mathbf{B}|$.
7. Si on multiplie les éléments d'une ligne (colonne) de $\mathbf{A}$ par $c \in \mathbb{R}$, $|\mathbf{A}|$ est multiplié par c.

A.5.1.6 Inverse d'une matrice carrée

a) Matrice adjointe

Soit $\mathbf{A}$ une matrice carrée de dimension $K \times K$. On appelle ***matrice adjointe*** de $\mathbf{A}$ — et on la note $\text{adj}\mathbf{A}$ — la transposée de la matrice des cofacteurs des éléments a_{kl} de $\mathbf{A}$:

$$\text{adj}\mathbf{A} = (A_{kl})'.$$

EXEMPLE A.5.6 (suite 1)

$$\text{adj}\begin{pmatrix} -3 & 2 \\ 1 & -3 \end{pmatrix} = \begin{pmatrix} -3 & -1 \\ -2 & -3 \end{pmatrix}' = \begin{pmatrix} -3 & -2 \\ -1 & -3 \end{pmatrix}$$

$$\text{adj}\begin{pmatrix} -1 & 1 & 2 \\ -1 & -2 & 1 \\ 2 & 1 & -3 \end{pmatrix} = \begin{pmatrix} 5 & -1 & 3 \\ 5 & -1 & 3 \\ 5 & -1 & 3 \end{pmatrix}' = \begin{pmatrix} 5 & 5 & 5 \\ -1 & -1 & -1 \\ 3 & 3 & 3 \end{pmatrix}$$

b) Matrice inverse

On peut montrer qu'une matrice carrée $\mathbf{A}$ est ***inversible*** si et seulement si $|\mathbf{A}| \neq 0$. Elle est dite dans ce cas ***non singulière*** et sa ***matrice inverse*** $\mathbf{A}^{-1}$ est telle que

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\text{adj}\mathbf{A}.$$

Notons que $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$ où $\mathbf{I}$ est la matrice identité[(3)] de dimension $K \times K$.

Remarquons que, pour pouvoir être inversible, la matrice $\mathbf{A}$ doit nécessairement être de rang maximal K, puisque $|\mathbf{A}|$ serait nul si le rang de $\mathbf{A}$ était strictement inférieur à K.

EXEMPLE A.5.6 (suite 2)

$$\begin{pmatrix} -3 & 2 \\ 1 & -3 \end{pmatrix}^{-1} = \frac{1}{7}\begin{pmatrix} -3 & -2 \\ -1 & -3 \end{pmatrix} = \begin{pmatrix} -3/7 & -2/7 \\ -1/7 & -3/7 \end{pmatrix}.$$

En revanche, la matrice

$$\mathbf{A}_{3\times 3} = \begin{pmatrix} -1 & 1 & 2 \\ -1 & -2 & 1 \\ 2 & 1 & -3 \end{pmatrix}$$

(3) $\mathbf{I} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$

n'est pas inversible puisque $|\mathbf{A}| = 0$. Elle est dite *singulière*. On vérifie par ailleurs aisément que cette matrice n'est pas de rang égal à 3 :

la 3^e ligne $= (-1) \times$ la 1^re ligne $+ (-1) \times$ la 2^e ligne,

c'est-à-dire

$$a_{3l} = (-1) \times a_{1l} + (-1) \times a_{2l}, \qquad l = 1, 2, 3.$$

---

c) Propriétés

Si $\mathbf{A}$ et $\mathbf{B}$ sont des matrices non singulières de mêmes dimensions, alors

1. $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$;
2. $(\mathbf{A}')^{-1} = (\mathbf{A}^{-1})'$;
3. $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

A.5.1.7 Formes quadratiques

a) Forme quadratique en K variables

Une forme quadratique en K variables $x_1, x_2, \ldots, x_K$ est une expression du type :

$$F(x_1, x_2, \ldots, x_K) = \sum_{k=1}^{K} \sum_{l=1}^{K} a_{kl} x_k x_l.$$

Si on pose

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_K \end{pmatrix} \quad \text{et} \quad \mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1K} \\ a_{21} & a_{22} & \cdots & a_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ a_{K1} & a_{K2} & \cdots & a_{KK} \end{pmatrix},$$

on peut vérifier, en notant $F(\mathbf{x}) = F(x_1, x_2, \ldots, x_K)$, que

$$F(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x}.$$

De plus, on peut supposer que $\mathbf{A}$ est symétrique. En effet, si $a_{kl} \neq a_{lk}$, nous pouvons remplacer ces deux valeurs par

$$b_{kl} = b_{lk} = \frac{a_{kl} + a_{lk}}{2} \cdot$$

EXEMPLE A.5.7

La forme quadratique

$$F(\mathbf{x}) = F(x_1, x_2, x_3) = 2x_1^2 + 2x_2^2 - 2x_3^3 + 4x_1x_3 + 2x_2x_3$$

peut encore s'écrire :

$$F(\mathbf{x}) = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 2 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

b) Formes quadratiques définies et semi-définies positives

Une forme quadratique $\mathbf{x}'\mathbf{A}\mathbf{x}$ — ou encore, par extension, la matrice carrée symétrique $\mathbf{A}$ définissant cette forme quadratique — est ***définie positive*** si $\mathbf{x}'\mathbf{A}\mathbf{x} > 0$ pour tout $\mathbf{x}$ non nul. Elle est ***semi-définie positive*** si $\mathbf{x}'\mathbf{A}\mathbf{x} \geqslant 0$ pour tout $\mathbf{x}$ non nul et s'il existe au moins un $\mathbf{x}$ non nul tel que $\mathbf{x}'\mathbf{A}\mathbf{x} = 0$.

c) Propriété

La matrice $\mathbf{A}$ est définie positive si et seulement si

$$a_{11} > 0\ ;\ \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0\ ;\ \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0\ ;\ \ldots\ ;\ |\mathbf{A}| > 0.$$

EXEMPLE A.5.7 (suite 1)

La matrice $\mathbf{A}$ de la forme quadratique est telle que

$$2 > 0\ ;\ \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = 4 > 0\ ;\ \begin{vmatrix} 2 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 1 & -2 \end{vmatrix} = -18 < 0.$$

Elle n'est donc pas définie positive.

A.5.2 NOTIONS DE GÉOMÉTRIE DANS $\mathbb{R}^K$

A.5.2.1 Vecteurs

Considérons le vecteur-colonne

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_K \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & \cdots & a_K \end{pmatrix}'.$$

Nous dirons que a_k est la k-ème composante de $\mathbf{a}$ ($k = 1, \ldots, K$). Nous pouvons représenter $\mathbf{a}$ *géométriquement*, dans l'espace euclidien $\mathbb{R}^K$ à K dimensions, par le segment de droite orienté dont l'extrémité initiale est l'origine O des axes de $\mathbb{R}^K$ et dont l'extrémité finale est le point A dont les coordonnées sont les composantes du vecteur $\mathbf{a}$. Nous désignerons ce segment de droite par $\overrightarrow{OA}$.

EXEMPLE A.5.8

Plaçons-nous dans $\mathbb{R}^2$ ($K = 2$). Les axes de $\mathbb{R}^2$ sont sous-tendus par les vecteurs $\overrightarrow{OE_1}$ et $\overrightarrow{OE_2}$ (voir figure A.5.2.1) ; les coordonnées des extrémités E_1 et E_2 de ces vecteurs correspondent aux composantes des vecteurs $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $\mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, respectivement.

Soit A le point dont la coordonnée sur le premier axe est égale à 3 et celle sur le second axe est égale à 2. Ce point A est l'extrémité du vecteur « géométrique » $\overrightarrow{OA}$ auquel nous pouvons associer le vecteur « algébrique » $\mathbf{a} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$. Il est clair que $\overrightarrow{OA} = 3\ \overrightarrow{OE_1}\ + 2\ \overrightarrow{OE_2}$ (vision géométrique) ou encore, de manière équivalente, que $\mathbf{a} = 3\mathbf{e}_1 + 2\mathbf{e}_2$ (vision algébrique).

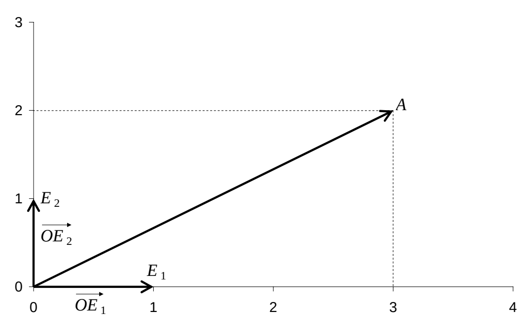

FIG. 5.2 – Correspondance entre $\overrightarrow{OA}$ et $\mathbf{a}$

Nous dirons encore que 3 et 2 sont les coordonnées de $\overrightarrow{OA}$ dans le repère $\{\overrightarrow{OE_1}, \overrightarrow{OE_2}\}$.

De manière générale, nous désignerons par $\overrightarrow{OE_1}, \overrightarrow{OE_2}, \ldots, \overrightarrow{OE_K}$ les vecteurs définissant les axes de $\mathbb{R}^K$. Nous pouvons leur associer respectivement les vecteurs « algébriques »

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \ldots, \mathbf{e}_K = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Si le point A de $\mathbb{R}^K$ a pour coordonnées a_1 sur le 1[er] axe, a_2 sur le 2[e] axe, ..., a_K sur le K^{e} axe, nous pouvons associer à $\overrightarrow{OA}$ le vecteur $\mathbf{a} = \begin{pmatrix} a_1 & a_2 & \cdots & a_K \end{pmatrix}'$. Nous avons, d'un point de vue géométrique, $\overrightarrow{OA} = a_1 \overrightarrow{OE_1} + a_2 \overrightarrow{OE_2} + \ldots + a_K \overrightarrow{OE_K}$ ou encore, d'un point de vue algébrique, $\mathbf{a} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + \ldots + a_K\mathbf{e}_K$.

A.5.2.2 Distance et norme euclidienne

Considérons dans $\mathbb{R}^K$ les points A et B — ou, de manière équivalente, les vecteurs $\overrightarrow{OA}$ et $\overrightarrow{OB}$ — dont les coordonnées sont les composantes de

$$\mathbf{a} = \begin{pmatrix} a_1 & \cdots & a_k & \cdots & a_K \end{pmatrix}' \quad \text{et} \quad \mathbf{b} = \begin{pmatrix} b_1 & \cdots & b_k & \cdots & b_K \end{pmatrix}'.$$

La ***distance euclidienne*** $d(A, B)$ entre les points A et B est donnée par

$$d(A, B) = \sqrt{\sum_{k=1}^{K} (a_k - b_k)^2}.$$

La ***norme*** (euclidienne) du vecteur $\overrightarrow{OA}$, notée $\| \overrightarrow{OA} \|$, correspond à la *longueur* (euclidienne) de ce vecteur ou encore, en d'autres termes, à la distance $d(O, A)$ entre l'origine O de $\mathbb{R}^K$ et le point A :

$$\| \overrightarrow{OA} \| = d(O, A) = \sqrt{\sum_{k=1}^{K} a_k^2}.$$

Un vecteur dont la norme est égale à 1 est dit ***normé*** ou ***unitaire***. Diviser un vecteur par sa norme donne un vecteur unitaire : $\overrightarrow{OA} / \| \overrightarrow{OA} \|$ est de norme égale à 1.

EXEMPLE A.5.9

Prenons $\mathbf{a} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$ et $\mathbf{b} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$ (voir figure 5.3).

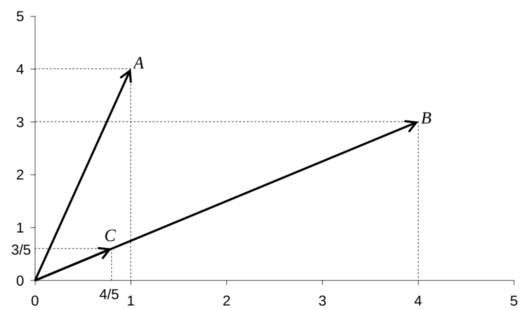

FIG. 5.3 – Distance et norme

La distance euclidienne entre les points A et B est égale à

$$d(A,B) = \sqrt{(1-4)^2 + (4-3)^2} = \sqrt{9+1} = \sqrt{10} = 3.16.$$

La norme de $\overrightarrow{OB}$ vaut :

$$\| \overrightarrow{OB} \| = \sqrt{4^2 + 3^2} = \sqrt{25} = 5.$$

Le vecteur $\overrightarrow{OC} = \dfrac{\overrightarrow{OB}}{\| \overrightarrow{OB} \|}$ est de norme égale à 1 ; son extrémité C a pour coordonnées les composantes de $\mathbf{c} = \frac{1}{5}\mathbf{b} = \begin{pmatrix} 4/5 \\ 3/5 \end{pmatrix}$ (voir figure 5.3).

Notons encore que les vecteurs $\overrightarrow{OE_1}, \ldots, \overrightarrow{OE_K}$ qui définissent les axes de $\mathbb{R}^K$ sont unitaires : $\| \overrightarrow{OE_k} \| = 1$ pour tout $k = 1, \ldots, K$.

A.5.2.3 Produit scalaire

Le ***produit scalaire*** de $\overrightarrow{OA}$ et $\overrightarrow{OB}$ — nous le désignerons par $\left\langle \overrightarrow{OA}, \overrightarrow{OB} \right\rangle$ — est le *nombre réel* défini par

$$\left\langle \overrightarrow{OA}, \overrightarrow{OB} \right\rangle = \mathbf{a}'\mathbf{b} = \begin{pmatrix} a_1 & \cdots & a_k & \cdots & a_K \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_k \\ \vdots \\ b_K \end{pmatrix} = \sum_{k=1}^{K} a_k b_k.$$

EXEMPLE A.5.9 (suite 1)

Le produit scalaire de $\overrightarrow{OA}$ et $\overrightarrow{OB}$ est égal à

$$\left\langle \overrightarrow{OA}, \overrightarrow{OB} \right\rangle = \mathbf{a}'\mathbf{b} = \begin{pmatrix} 1 & 4 \end{pmatrix} \begin{pmatrix} 4 \\ 3 \end{pmatrix} = 4 + 12 = 16.$$

On vérifie aisément que

$$\left\langle \overrightarrow{OA}, \overrightarrow{OE_1} \right\rangle = a_1, \left\langle \overrightarrow{OA}, \overrightarrow{OE_2} \right\rangle = a_2, \ldots, \left\langle \overrightarrow{OA}, \overrightarrow{OE_K} \right\rangle = a_K :$$

la coordonnée du point A sur un axe de $\mathbb{R}^K$ est égale au produit scalaire de $\overrightarrow{OA}$ et du vecteur unitaire qui définit l'axe.

Observons aussi que

$$\| \overrightarrow{OA} \| = \sqrt{\sum_{k=1}^{K} a_k^2} = \sqrt{\left\langle \overrightarrow{OA}, \overrightarrow{OA} \right\rangle}.$$

A.5.2.4 Cosinus d'un angle

La figure 5.4 permet de visualiser le ***cosinus*** d'un angle α ($\alpha \in [0°, 360°]$).

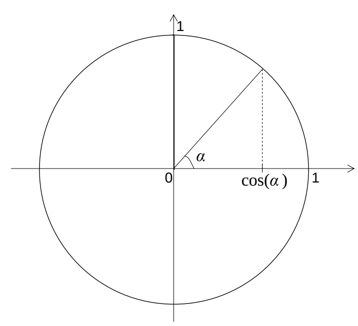

FIG. 5.4 – Cosinus de α

Nous avons, pour tout angle α :

$$-1 \leqslant \cos(\alpha) \leqslant 1 \quad \text{et} \quad 0 \leqslant \cos^2(\alpha) \leqslant 1.$$

Le tableau suivant indique la valeur de $\cos(\alpha)$ et $\cos^2(\alpha)$ pour différents angles particuliers α :

α	0°	45°	90°	180° (-180°)	270° (-90°)	360° (0°)
$\cos(\alpha)$	1	$\sqrt{2}/2$	0	-1	0	1
$\cos^2(\alpha)$	1	0.5	0	1	0	1

Désignons par α_{AB} l'angle entre les vecteurs $\overrightarrow{OA}$ et $\overrightarrow{OB}$. On peut montrer que son cosinus est égal à

$$\cos(\alpha_{AB}) = \frac{\left\langle \overrightarrow{OA}, \overrightarrow{OB} \right\rangle}{\| \overrightarrow{OA} \| \| \overrightarrow{OB} \|}\ .$$

EXEMPLE A.5.9 (suite 2)

Etant donné que $\left\langle \overrightarrow{OA}, \overrightarrow{OB} \right\rangle = 16$, $\| \overrightarrow{OB} \| = 5$ et $\| \overrightarrow{OA} \| = \sqrt{17} = 4.12$, nous obtenons :

$$\cos(\alpha_{AB}) = \frac{16}{(5)(4.12)} = 0.78.$$

Dans certaines situations, le cosinus de l'angle entre deux vecteurs possède une interprétation statistique intéressante. Replaçons-nous dans le cadre de travail présenté au paragraphe A.5.1.3. Soit $\mathbf{X}_c$ une matrice de valeurs centrées (de dimension $n \times 2$) :

$$\mathbf{X}_c = \begin{pmatrix} y_{11} & y_{12} \\ \vdots & \vdots \\ y_{i1} & y_{i2} \\ \vdots & \vdots \\ y_{n1} & y_{n2} \end{pmatrix}$$

avec $\overline{y}_1 = \frac{1}{n}\sum_{i=1}^{n} y_{i1} = 0$ et $\overline{y}_2 = \frac{1}{n}\sum_{i=1}^{n} y_{i2} = 0$. Considérons les vecteurs

$$\mathbf{y}_1 = \begin{pmatrix} y_{11} \\ \vdots \\ y_{i1} \\ \vdots \\ y_{n1} \end{pmatrix} \quad \text{et} \quad \mathbf{y}_2 = \begin{pmatrix} y_{12} \\ \vdots \\ y_{i2} \\ \vdots \\ y_{n2} \end{pmatrix};$$

notons $\overrightarrow{OY_1}$ et $\overrightarrow{OY_2}$ leurs représentations géométriques. Le cosinus de l'angle entre $\overrightarrow{OY_1}$ et $\overrightarrow{OY_2}$ est donné par

$$\cos\left(\alpha_{Y_1Y_2}\right) = \frac{\left\langle \overrightarrow{OY_1}, \overrightarrow{OY_2} \right\rangle}{\| \overrightarrow{OY_1} \| \| \overrightarrow{OY_2} \|} = \frac{\sum_{i=1}^{n} y_{i1} y_{i2}}{\sqrt{\left(\sum_{i=1}^{n} y_{i1}^2\right)\left(\sum_{i=1}^{n} y_{i2}^2\right)}} = \frac{s_{12}}{s_1 s_2} = r$$

où s_1, s_2, s_{12} et r sont respectivement l'écart-type de la série statistique $\{y_{i1}; i = 1, \ldots, n\}$, l'écart-type de la série $\{y_{i2}; i = 1, \ldots, n\}$, la covariance et le coefficient de corrélation de la série bivariée $\{(y_{i1}, y_{i2}); i = 1, \ldots, n\}$.

A.5.2.5 Vecteurs orthogonaux

Les vecteurs $\overrightarrow{OA}$ et $\overrightarrow{OB}$ sont ***orthogonaux*** — ou ***perpendiculaires*** — si et seulement si (*ssi*) l'angle α_{AB} est égal à 90° ou 270°$(= -90°)$, c'est-à-dire *ssi* $\cos(\alpha_{AB}) = 0$, autrement dit encore *ssi* $\left\langle \overrightarrow{OA}, \overrightarrow{OB} \right\rangle = 0$.

EXEMPLE A.5.9 (suite 3)

Soit $\overrightarrow{OD}$ le vecteur « géométrique » associé à $\mathbf{d} = \begin{pmatrix} 2 \\ -1/2 \end{pmatrix}$.

Puisque $\left\langle \overrightarrow{OA}, \overrightarrow{OD} \right\rangle = 2 - 2 = 0$, $\overrightarrow{OA}$ et $\overrightarrow{OD}$ sont des vecteurs orthogonaux dans $\mathbb{R}^2$.

Un ensemble de H vecteurs $\overrightarrow{OU_1}, \ldots, \overrightarrow{OU_H}$ est dit ***orthonormé*** lorsque les vecteurs $\overrightarrow{OU_h}$ sont normés — $\| \overrightarrow{OU_h} \| = 1$ pour tout $h = 1, \ldots, H$ — et orthogonaux deux à deux — $\left\langle \overrightarrow{OU_h}, \overrightarrow{OU_l} \right\rangle = 0$ pour tout $h \neq l \in \{1, \ldots, H\}$.

A.5.2.6 Projection orthogonale d'un point sur une droite passant par l'origine

Considérons la droite Δ passant par l'origine O de $\mathbb{R}^K$ et définie par le vecteur unitaire $\overrightarrow{OU}$ dont les coordonnées sont les composantes de $\mathbf{u} = \begin{pmatrix} u_1 & \cdots & u_K \end{pmatrix}'$ (voir figure 5.5).

La projection orthogonale du point A sur la droite Δ est le point de cette droite le plus proche de A. Désignons ce point par $P_\Delta(A)$. Le vecteur $\overrightarrow{AP_\Delta(A)}$ — vecteur dont l'origine est A et l'extrémité est $P_\Delta(A)$ — se caractérise par le fait qu'il est orthogonal à $\overrightarrow{OU}$ ou, autrement dit, perpendiculaire à la droite Δ.

On montre que

$$\overrightarrow{OP_\Delta(A)} = \left\langle \overrightarrow{OA}, \overrightarrow{OU} \right\rangle \overrightarrow{OU}$$

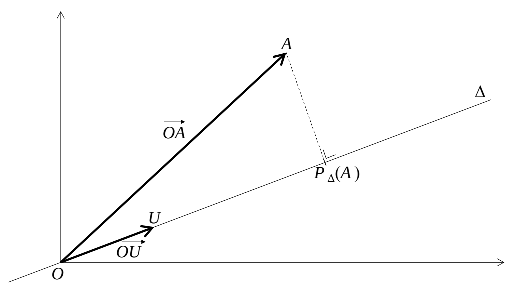

FIG. 5.5 – Projection orthogonale du point A sur la droite Δ

ce qui implique que

$$\| \overrightarrow{OP_\Delta(A)} \| = \left| \left\langle \overrightarrow{OA}, \overrightarrow{OU} \right\rangle \right| = \left| \sum_{k=1}^{K} a_k u_k \right| .$$

EXEMPLE A.5.10

Soient les points D, A et B dont les coordonnées dans $\mathbb{R}^2$ sont respectivement les composantes des vecteurs $\mathbf{d} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $\mathbf{a} = \begin{pmatrix} 5 \\ 1 \end{pmatrix}$ et $\mathbf{b} = \begin{pmatrix} 2 \\ -3 \end{pmatrix}$ (voir figure 5.6). Considérons la droite Δ ayant la même direction que $\overrightarrow{OD}$ et recherchons les coordonnées des projections orthogonales $P_\Delta(A)$ et $P_\Delta(B)$ de A et B sur cette droite.

Etant donné que $\| \overrightarrow{OD} \| = \sqrt{1+4} = \sqrt{5}$, le vecteur unitaire $\overrightarrow{OU}$ définissant la droite Δ a pour coordonnées les composantes de $\mathbf{u} = \frac{1}{\sqrt{5}}\mathbf{d} = \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}$.

Calculons $\left\langle \overrightarrow{OU}, \overrightarrow{OA} \right\rangle$ et $\left\langle \overrightarrow{OU}, \overrightarrow{OB} \right\rangle$:

$$\begin{aligned}
\left\langle \overrightarrow{OU}, \overrightarrow{OA} \right\rangle &= \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} 5 \\ 1 \end{pmatrix} = \frac{5}{\sqrt{5}} + \frac{2}{\sqrt{5}} = \frac{7}{\sqrt{5}}; \\
\left\langle \overrightarrow{OU}, \overrightarrow{OB} \right\rangle &= \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} 2 \\ -3 \end{pmatrix} = \frac{2}{\sqrt{5}} - \frac{6}{\sqrt{5}} = \frac{-4}{\sqrt{5}} .
\end{aligned}$$

Il s'ensuit que $\overrightarrow{OP_\Delta(A)} = \frac{7}{\sqrt{5}} \overrightarrow{OU}$; les coordonnées de $P_\Delta(A)$ dans $\mathbb{R}^2$ sont les composantes du vecteur $\frac{7}{\sqrt{5}}\mathbf{u} = \begin{pmatrix} 7/5 \\ 14/5 \end{pmatrix}$.

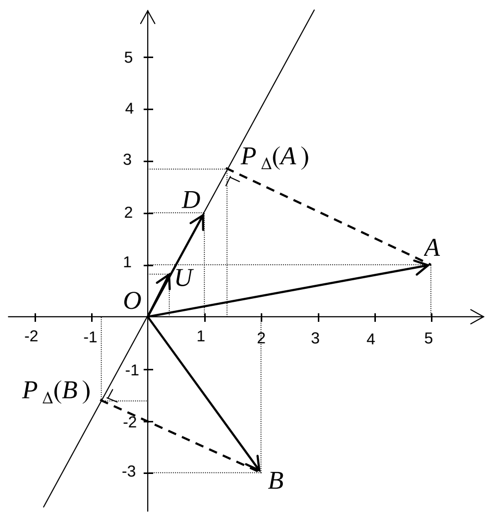

FIG. 5.6 – Projection orthogonale d'un point sur une droite

De même, $\overrightarrow{OP_\Delta(B)} = \frac{-4}{\sqrt{5}}\ \overrightarrow{OU}$; les coordonnées de $P_\Delta(B)$ dans $\mathbb{R}^2$ sont les composantes du vecteur $\dfrac{-4}{\sqrt{5}}\mathbf{u} = \begin{pmatrix} -4/5 \\ -8/5 \end{pmatrix}$.

A.5.2.7 Projection orthogonale d'un point sur un sous-espace de $\mathbb{R}^K$ engendré par H vecteurs orthonormés ($H < K$)

Soient $\overrightarrow{OU_1}, \ldots, \overrightarrow{OU_H}$ H vecteurs orthonormés dans $\mathbb{R}^K$. Ces vecteurs *définissent* — *engendrent* — le sous-espace $\mathcal{L}$ de $\mathbb{R}^K$ constitué de toutes les combinaisons linéaires $\sum_{h=1}^{H} \alpha_h\ \overrightarrow{OU_h}$ ($\alpha_h \in \mathbb{R}$). Ce sous-espace $\mathcal{L}$ contient l'origine O de $\mathbb{R}^K$ et est de dimension H.

Ainsi, par exemple, le vecteur normé $\overrightarrow{OU}$ de $\mathbb{R}^K$ définit la droite (sous-espace de dimension 1) passant par O dont les points sont les extrémités des vecteurs de la forme $\alpha\ \overrightarrow{OU}$ ($\alpha \in \mathbb{R}$) ; les vecteurs orthonormés $\overrightarrow{OU_1}$ et $\overrightarrow{OU_2}$ engendrent le plan (sous-espace de dimension 2) passant par O dont les points sont les extrémités des vecteurs de la forme $\alpha_1\ \overrightarrow{OU_1} + \alpha_2\ \overrightarrow{OU_2}$ ($\alpha_1, \alpha_2 \in \mathbb{R}$) ; etc.

La projection orthogonale $P_{\mathcal{L}}(A)$ du point A sur le sous-espace $\mathcal{L}$ est le point de $\mathcal{L}$ le plus proche de A. On montre que

$$\begin{aligned}\overrightarrow{OP_{\mathcal{L}}(A)} &= \sum_{h=1}^{H} \left\langle \overrightarrow{OU_h}, \overrightarrow{OA} \right\rangle \overrightarrow{OU_h} \\ &= \sum_{h=1}^{H} \overrightarrow{OP_{\Delta_h}(A)}\end{aligned}$$

où Δ_h est la droite engendrée par $\overrightarrow{OU_h}$ ($h = 1, \ldots, H$).

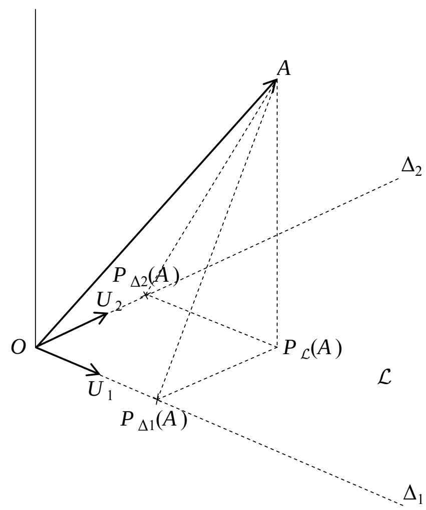

FIG. 5.7 – Projection orthogonale d'un point sur un plan

A.5.2.8 Valeurs et vecteurs propres d'une matrice carrée $\mathbf{A}_{K\times K}$

a) Définitions

Considérons une matrice $\mathbf{A}$ de dimension $K \times K$ et un vecteur-colonne $\mathbf{u}$ à K composantes. Le produit $\mathbf{Au}$, de dimension $K \times 1$, définit un nouveau vecteur à K composantes. La matrice $\mathbf{A}$ permet ainsi de définir une ***transformation*** de l'espace $\mathbb{R}^K$, c'est-à-dire une fonction de $\mathbb{R}^K$ dans $\mathbb{R}^K$: celle qui envoie le vecteur $\mathbf{u}$ de $\mathbb{R}^K$ sur le vecteur $\mathbf{v} = \mathbf{Au}$ de $\mathbb{R}^K$.

EXEMPLE A.5.11

Soit la matrice

$$\mathbf{A} = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}.$$

Elle définit une transformation de $\mathbb{R}^2$ qui envoie le vecteur $\mathbf{u}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ sur le vecteur $\mathbf{v}_1 = \mathbf{A}\mathbf{u}_1 = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 \\ 8 \end{pmatrix}$, le vecteur $\mathbf{u}_2 = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$ sur $\mathbf{v}_2 = \mathbf{A}\mathbf{u}_2 = \begin{pmatrix} 2 \\ 6 \end{pmatrix}$ et le vecteur $\mathbf{u}_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ sur $\mathbf{v}_3 = \begin{pmatrix} 5 \\ 5 \end{pmatrix} = 5\mathbf{u}_3$.

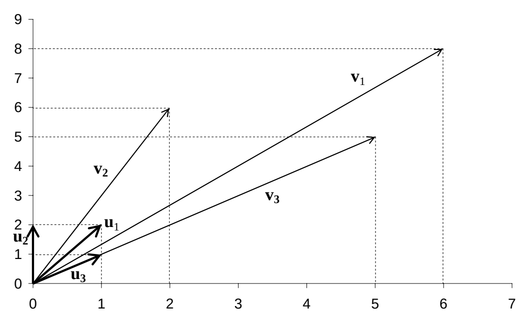

FIG. 5.8 – Transformation de $\mathbb{R}^2$ définie par $\mathbf{A}$

Nous pouvons visualiser l'effet de cette transformation sur la figure 5.8 dans laquelle, par souci de simplification, nous utilisons la même notation pour les vecteurs « algébriques » $\mathbf{u}_1$, $\mathbf{u}_2$ et $\mathbf{u}_3$, et pour leurs représentations géométriques.

Dans cet exemple, $\mathbf{u}_3$ est envoyé par $\mathbf{A}$ sur un multiple de lui-même.

Un vecteur $\mathbf{u}_{K\times 1}$ (non nul) et un scalaire λ sont respectivement ***vecteur propre*** et ***valeur propre*** (associée) de $\mathbf{A}$ si $\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$.

Si $\mathbf{u}$ est un vecteur propre de $\mathbf{A}$ associé à la valeur propre λ, alors $\alpha\mathbf{u}$ l'est aussi[4] pour tout $\alpha \in \mathbb{R}_0$. L'ensemble des vecteurs propres ainsi obtenus définit une ***direction propre*** associée à λ.

EXEMPLE A.5.11 (suite 1)

[4] En effet, $\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$ implique que $\mathbf{A}(\alpha\mathbf{u}) = \alpha\mathbf{A}\mathbf{u} = \alpha\lambda\mathbf{u} = \lambda(\alpha\mathbf{u})$.

Le vecteur $\mathbf{u}_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ est un vecteur propre de la matrice $\mathbf{A}$ associé à la valeur propre 5. Il en est de même pour tout multiple de $\mathbf{u}_3$. On vérifie par exemple que $\mathbf{A}\begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 10 \\ 10 \end{pmatrix} = 5\begin{pmatrix} 2 \\ 2 \end{pmatrix}$.

Le vecteur $\mathbf{u}_3$ définit ainsi dans $\mathbb{R}^2$ une direction propre pour $\mathbf{A}$, autrement dit une direction du plan qui n'est pas modifiée par la transformation définie par la matrice $\mathbf{A}$.

b) Détermination des valeurs et vecteurs propres de $\mathbf{A}$

On montre que λ est une valeur propre de $\mathbf{A}_{K\times K}$ si et seulement si λ annule le déterminant de la matrice $\mathbf{A} - \lambda\mathbf{I}$ où $\mathbf{I}$ est la matrice identité de dimension $K \times K$. Grâce à ce résultat, nous voyons que déterminer les valeurs propres de $\mathbf{A}$ revient à rechercher les solutions de l'équation $|\mathbf{A} - \lambda\mathbf{I}| = 0$.

EXEMPLE A.5.11 (suite 2)

Les valeurs propres de $\mathbf{A} = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$ sont les solutions λ de l'équation $|\mathbf{A} - \lambda\mathbf{I}| = 0$ où $\mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Nous pouvons écrire :

$$\begin{aligned} & |\mathbf{A} - \lambda\mathbf{I}| = 0 \\ \Leftrightarrow\quad & \begin{vmatrix} 4-\lambda & 1 \\ 2 & 3-\lambda \end{vmatrix} = 0 \\ \Leftrightarrow\quad & (4-\lambda)(3-\lambda) - 2 = 0 \\ \Leftrightarrow\quad & 12 - 4\lambda - 3\lambda + \lambda^2 - 2 = 0 \\ \Leftrightarrow\quad & \lambda^2 - 7\lambda + 10 = 0. \end{aligned}$$

Cette équation du second degré possède deux solutions réelles : $\lambda_1 = 5$ et $\lambda_2 = 2$.

Nous avons déjà vu que $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ était un vecteur propre de $\mathbf{A}$ associé à la valeur propre $\lambda_1 = 5$. Pour trouver un vecteur propre correspondant à $\lambda_2 = 2$, il suffit de chercher $\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$ vérifiant

$$\mathbf{A}\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = 2\begin{pmatrix} w_1 \\ w_2 \end{pmatrix},$$

c'est-à-dire

$$\begin{cases} 4w_1 + w_2 = 2w_1 \\ 2w_1 + 3w_2 = 2w_2 \end{cases}$$

ou encore

$$w_2 = -2w_1.$$

En prenant par exemple $w_1 = 1$, on obtient le vecteur propre $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$.

Tout multiple de $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$ est lui aussi vecteur propre de $\mathbf{A}$ associé à la valeur propre $\lambda_2 = 2$.

---

c) Remarques

Remarque A.5.1 Il arrive que l'équation $|\mathbf{A} - \lambda\mathbf{I}| = 0$ ne possède pas de solution réelle. Cela signifie que la matrice $\mathbf{A}$ ne possède aucune valeur propre (réelle), autrement dit encore que la transformation de $\mathbb{R}^K$ définie par $\mathbf{A}$ ne laisse aucune direction fixe.

EXEMPLE A.5.12

Considérons la matrice $\mathbf{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. L'équation $|\mathbf{A} - \lambda\mathbf{I}| = 0$ s'écrit

$$\begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = 0,$$

c'est-à-dire

$$\lambda^2 + 1 = 0,$$

ou encore $\lambda^2 = -1$. Cette équation ne possède pas de solution réelle. Il n'existe dans $\mathbb{R}^2$ aucune direction laissée fixe par la transformation définie par $\mathbf{A}$.

Remarque A.5.2 Toute matrice $\mathbf{A}_{K\times K}$ possède *au plus* K valeurs propres distinctes. Dans certains cas, une valeur propre est solution multiple de l'équation $|\mathbf{A} - \lambda\mathbf{I}| = 0$ — on parlera de *valeur propre multiple* — ; on peut alors lui associer un *sous-espace propre* de dimension égale à sa multiplicité. Illustrons cette situation au moyen de l'exemple suivant.

EXEMPLE A.5.13

Recherchons les valeurs propres de $\mathbf{A} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Pour ce faire, résolvons l'équation $|\mathbf{A} - \lambda\mathbf{I}| = 0$. Nous avons :

$$\begin{vmatrix} 4-\lambda & 0 & 0 \\ 0 & 4-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{vmatrix} = 0$$

$$\Leftrightarrow \quad (4-\lambda)(4-\lambda)(1-\lambda) = 0$$

$$\Leftrightarrow \quad (4-\lambda)^2(1-\lambda) = 0.$$

Cette équation possède deux solutions : $\lambda_1 = 4$, de multiplicité égale à 2, et $\lambda_2 = 1$.
Les vecteurs propres associés à $\lambda_1 = 4$ sont les vecteurs $\mathbf{u} = \begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix}'$ solutions de

$$\begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = 4 \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix},$$

c'est-à-dire tels que

$$\begin{cases} 4u_1 &= 4u_1 \\ 4u_2 &= 4u_2 \\ u_3 &= 4u_3 \end{cases}$$

et donc tels que $u_3 = 0$. Les vecteurs propres correspondant à $\lambda_1 = 4$ sont ainsi tous les vecteurs dont la 3[e] composante est nulle : ils définissent un *plan propre* (le plan horizontal engendré par les deux premiers axes de $\mathbb{R}^3$), c'est-à-dire un plan laissé fixe par la transformation définie par $\mathbf{A}$.
Les vecteurs propres associés à $\lambda_2 = 1$ sont les vecteurs $\mathbf{u} = \begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix}'$ solutions de

$$\begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix},$$

c'est-à-dire tels que

$$\begin{cases} 4u_1 &= u_1 \\ 4u_2 &= u_2 \\ u_3 &= u_3. \end{cases}$$

Il s'agit des vecteurs dont les deux premières composantes sont nulles, autrement dit des vecteurs multiples de $\begin{pmatrix} 0 & 0 & 1 \end{pmatrix}'$; ils engendrent une *direction propre* de la transformation définie par $\mathbf{A}$.
Observons que la direction propre associée à $\lambda_2 = 1$ est orthogonale au plan propre associé à $\lambda_1 = 4$.

Plus généralement, si l'équation $|\mathbf{A} - \lambda \mathbf{I}| = 0$ possède une solution de multiplicité k ($1 \leqslant k \leqslant K$), l'ensemble des vecteurs propres associés à cette valeur propre multiple constitue un sous-ensemble de $\mathbb{R}^K$ de dimension k et définit ainsi un *sous-espace propre à k dimensions* de la transformation définie par la matrice $\mathbf{A}$.

d) Valeurs et vecteurs propres d'une matrice symétrique

Si $\mathbf{A}_{K \times K}$ est une matrice *symétrique*, la somme des multiplicités de ses différentes valeurs propres est égale à K : si nous désignons par $\lambda_1, \ldots, \lambda_H$

les H ($H \leqslant K$) valeurs propres distinctes de $\mathbf{A}$ et par $k_1, \ldots, k_H$ leurs multiplicités respectives, nous avons $k_1 + \ldots + k_H = K$. En d'autres termes encore, les différents sous-espaces propres de la transformation de $\mathbb{R}^K$ définie par la matrice $\mathbf{A}$ — les sous-espaces définis par les vecteurs propres associés aux différentes valeurs propres de $\mathbf{A}$ — ont des dimensions dont la somme est égale à K.

EXEMPLE A.5.13 (suite 1)

$\mathbf{A} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ est une matrice symétrique. Nous l'avons vu, elle possède deux valeurs propres distinctes, $\lambda_1 = 4$ et $\lambda_2 = 1$, de multiplicités respectives 2 et 1. À $\lambda_1 = 4$ correspond un plan propre (sous-espace de $\mathbb{R}^3$ de dimension 2) et à $\lambda_2 = 1$ est associée une direction propre (sous-espace de $\mathbb{R}^3$ de dimension 1). Nous avons bien $2 + 1 = 3 = K$.

Les vecteurs propres associés à deux valeurs propres *distinctes* de $\mathbf{A}$ sont orthogonaux. Par ailleurs, si λ_h est une valeur propre de $\mathbf{A}$ de multiplicité $k_h > 1$, il lui correspond un sous-espace propre $\mathcal{L}_h$ de dimension k_h dans lequel on peut choisir k_h vecteurs propres (associés à λ_h) orthogonaux deux à deux, qui seront considérés par la suite comme les vecteurs engendrant le sous-espace propre $\mathcal{L}_h$. On peut ainsi définir un ensemble de K vecteurs propres $\mathbf{u}_1, \ldots, \mathbf{u}_K$ de $\mathbf{A}$, orthogonaux deux à deux et que l'on peut prendre, sans perte de généralité, de norme égale à 1. Ce système orthonormé de vecteurs propres de la matrice $\mathbf{A}$ peut être pris comme nouveau repère dans $\mathbb{R}^K$.

EXEMPLE A.5.13 (suite 2)

Le plan propre associé à $\lambda_1 = 4$ est le plan de $\mathbb{R}^3$ constitué de tous les vecteurs dont la 3e composante est nulle. Nous pouvons par exemple définir ce plan à partir des deux vecteurs propres orthonormés $\mathbf{u}_1 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}'$ et $\mathbf{u}_2 = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}'$. D'autres choix sont bien sûr possibles : par exemple, $\mathbf{u}_1 = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \end{pmatrix}'$ et $\mathbf{u}_2 = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{pmatrix}'$.
La direction propre associée à $\lambda_2 = 1$ est engendrée par le vecteur propre unitaire $\mathbf{u}_3 = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}'$; $\mathbf{u}_3$ est orthogonal à tout vecteur propre associé à $\lambda_1 = 4$, donc en particulier à $\mathbf{u}_1$ et $\mathbf{u}_2$.
Les vecteurs $\mathbf{u}_1$, $\mathbf{u}_2$ et $\mathbf{u}_3$ constituent un *repère* — une *base* — orthonormé(e) de $\mathbb{R}^3$.

On vérifie également que si la matrice symétrique $\mathbf{A}_{K \times K}$ possède les valeurs propres distinctes $\lambda_1, \ldots, \lambda_H$ de multiplicités respectives $k_1, \ldots, k_H$:

1. $(\lambda_1)^{k_1} \ldots (\lambda_H)^{k_H} = |\mathbf{A}|$;

2. $k_1\lambda_1 + \ldots + k_H\lambda_H = \mathrm{tr}(\mathbf{A})$, où $\mathrm{tr}(\mathbf{A})$ est la ***trace*** de la matrice $\mathbf{A}$, c'est-à-dire la somme de ses composantes sur la première diagonale ($\mathrm{tr}(\mathbf{A}) = \sum_{k=1}^{K} a_{kk}$).

Si la matrice symétrique $\mathbf{A}_{K\times K}$ est *définie positive*, toutes ses valeurs propres sont strictement positives.

EXEMPLE A.5.13 (suite 13)

$\mathbf{A} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ est définie positive. Ses deux valeurs propres sont strictement positives.

Si la matrice symétrique $\mathbf{A}_{K\times K}$ est *semi-définie positive*, ses valeurs propres sont positives ou nulles, et il y a au moins une valeur propre nulle. Le nombre de valeurs propres non nulles (strictement positives) est égal au rang de $\mathbf{A}$.

e) Le cas particulier de la matrice de corrélation

Considérons le tableau individus $\times$ caractères issu de l'observation de p variables quantitatives sur n individus. Ce tableau peut être représenté par la matrice

$$\mathbf{X}_{n\times p} = \begin{pmatrix} x_{11} & \cdots & x_{1k} & \cdots & x_{1p} \\ \vdots & & \vdots & & \vdots \\ x_{i1} & \cdots & x_{ik} & \cdots & x_{ip} \\ \vdots & & \vdots & & \vdots \\ x_{n1} & \cdots & x_{nk} & \cdots & x_{np} \end{pmatrix}.$$

Les p colonnes de $\mathbf{X}$ constituent p séries statistiques univariées d'effectif n que l'on peut centrer et réduire. Soit $\mathbf{X}_{\mathrm{cr}}$ la matrice contenant les données centrées réduites :

$$\mathbf{X}_{\mathrm{cr}} = \begin{pmatrix} z_{11} & \cdots & z_{1k} & \cdots & z_{1p} \\ \vdots & & \vdots & & \vdots \\ z_{i1} & \cdots & z_{ik} & \cdots & z_{ip} \\ \vdots & & \vdots & & \vdots \\ z_{n1} & \cdots & z_{nk} & \cdots & z_{np} \end{pmatrix}$$

où $z_{ik} = (x_{ik} - \overline{x}_k)/s_k$, avec $\overline{x}_k = \frac{1}{n}\sum_{i=1}^{n} x_{ik}$ et $s_k^2 = \frac{1}{n}\sum_{i=1}^{n}(x_{ik} - \overline{x}_k)^2$ [$i = 1, \ldots, n$ et $k = 1, \ldots, p$].

Nous l'avons vu dans l'exemple A.5.3, la matrice de corrélation $\mathbf{R}$ — matrice contenant les coefficients de corrélation de Bravais-Pearson entre les

p séries statistiques constituant les colonnes de $\mathbf{X}$ — peut s'obtenir facilement à partir de $\mathbf{X}_{\mathrm{cr}}$ grâce à la relation suivante :

$$\mathbf{R} = \frac{1}{n}\mathbf{X}_{\mathrm{cr}}'\mathbf{X}_{\mathrm{cr}}.$$

$\mathbf{R}$ est une matrice *symétrique* et (au moins) *semi-définie positive*. En effet, pour tout vecteur $\mathbf{x}$ non nul de $\mathbb{R}^p$:

$$\begin{aligned}
\mathbf{x}'\mathbf{R}\mathbf{x} &= \frac{1}{n}\mathbf{x}'\mathbf{X}_{\mathrm{cr}}'\mathbf{X}_{\mathrm{cr}}\mathbf{x} \\
&= \frac{1}{n}\left(\mathbf{X}_{\mathrm{cr}}\mathbf{x}\right)'\mathbf{X}_{\mathrm{cr}}\mathbf{x} \\
&= \frac{1}{n}\left\|\mathbf{X}_{\mathrm{cr}}\mathbf{x}\right\|^2 \geqslant 0.
\end{aligned}$$

On montre par ailleurs que la matrice $\mathbf{R}$ est *définie positive* si le rang de $\mathbf{R}$ est égal à p. Puisque le rang de $\mathbf{R}$ est égal[(5)] à celui de la matrice $\mathbf{X}_{\mathrm{cr}}$ ou encore de la matrice $\mathbf{X}$, nous pouvons affirmer que $\mathbf{R}$ est définie positive si les p colonnes de $\mathbf{X}$ (ou $\mathbf{X}_{\mathrm{cr}}$) sont linéairement indépendantes[(6)].

Les propriétés des valeurs et vecteurs propres d'une matrice symétrique présentées ci-avant s'appliquent donc aux valeurs et vecteurs propres de la matrice de corrélation $\mathbf{R}$. Ainsi, de manière générale, les valeurs propres de $\mathbf{R}$ sont positives ou nulles et leur somme est égale à $\mathrm{tr}(\mathbf{R}) = p$, c'est-à-dire au nombre de variables considérées. Le nombre de valeurs propres non nulles (strictement positives) est égal au $\mathrm{rang}(\mathbf{R}) = \mathrm{rang}(\mathbf{X})$, soit au nombre maximum de colonnes de $\mathbf{X}$ qui sont linéairement indépendantes.

(5) Voir paragraphe A.5.1.4.

(6) On suppose ici que $n \geqslant p$, ce qui est souvent le cas en pratique.