4 Calcul matriciel

Soit le tableau rectangulaire A à n lignes et p colonnes où a_{ij} est l'élément à l'intersection de la $i$ème ligne et de la $j$ème colonne, c'est une *matrice*. On la représente par :

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix}$$

Chaque ligne i de A est un *vecteur ligne* :

$$\begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{ip} \end{bmatrix}$$

Chaque colonne j est un *vecteur colonne* :

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix}$$

4.1 Définitions

1- *Dimension* : Le nombre de lignes n et le nombre de colonnes p de la matrice (notation : $(n \times p)$).

2- *Matrice transposée* : Matrice obtenue en interchangeant les rôles des lignes et des colonnes (notation : A'). La dimension de A' est $(p \times n)$.

$$A' = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix}' = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1p} & a_{2p} & \cdots & a_{np} \end{bmatrix}$$

3- *Matrice carrée* : Matrice qui a autant de lignes que de colonnes ($n = p$).

4- *Matrices triangulaires* :

- *Triangulaire haute* : Matrice carrée dont les éléments au-dessous de la diagonale principale sont égaux à zéro.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & a_{nn} \end{bmatrix}$$

- *Triangulaire basse* : Matrice carrée dont les éléments au-dessus de la diagonale principale sont égaux à zéro.

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

5- *Matrice symétrique* : Matrice carrée égale à sa transposée ($a_{ij} = a_{ji}$ pour tout couple i, j). En général, on ne représente que la partie triangulaire haute (ou basse), c'est-à-dire les éléments situés sur la diagonale principale et au-dessus (au-dessous) de la diagonale principale.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \vdots & \vdots \\ & & & a_{nn} \end{bmatrix}$$

6- *Matrice diagonale* : Matrice carrée dont les seuls éléments éventuellement non nuls sont sur la diagonale principale (notation : $diag(A)$).

$$D = diag(A) = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & a_{nn} \end{bmatrix}$$

7- *Matrice identité* : Matrice diagonale dont les éléments diagonaux sont tous égaux à 1 (notation : I).

8- *Matrice nulle* : Matrice dont tous les éléments sont égaux à zéro.

9- *Matrice J* : Matrice dont tous les éléments sont égaux à 1.

4.2 Opérations élémentaires

1- *Addition* :

$$C = A + B \Longleftrightarrow c_{ij} = a_{ij} + b_{ij}$$

C, A et B sont de mêmes dimensions.

2- *Multiplication par un scalaire* :

$$C = \lambda A \Longleftrightarrow c_{ij} = \lambda a_{ij}$$

C et A sont de mêmes dimensions

3- *Produit de deux matrices* : Pour que le produit $C = AB$ de deux matrices A et B soit défini, il faut que le nombre de colonnes de A soit égal au nombre de lignes de B. Si A est de dimensions $(n \times p)$ et B de dimensions $(p \times m)$, le résultat C du produit de A par B est de dimensions $(n \times m)$.

$$C = AB \Longleftrightarrow c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}$$

En général, la multiplication n'est pas *commutative* c'est-à-dire que $AB \neq BA$. Pour cette raison, on dit parfois du produit AB que c'est la *prémultiplication* de B par A ou la *postmultiplication* de A par B.

4- *Matrice idempotente* : Une matrice $A(n \times n)$ est idempotente si : $AA = A$.

5- *Matrice inverse* : On appelle matrice inverse de A la matrice A^{-1} telle que :

$$AA^{-1} = A^{-1}A = I$$

Si A^{-1} existe, elle est *unique*, on dit alors que A est *non singulière*.

Si A n'a pas d'inverse elle est dite *singulière*.

- L'inverse d'une matrice symétrique est symétrique.
- L'inverse de la transposée de A est la transposée de A^{-1}.
- $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$
- $(\alpha A)^{-1} = \dfrac{1}{\alpha} A^{-1}$
- $D = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & a_{nn} \end{bmatrix} \Longrightarrow D^{-1} = \begin{bmatrix} a_{11}^{-1} & 0 & \cdots & 0 \\ 0 & a_{22}^{-1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & a_{nn}^{-1} \end{bmatrix}$

6- *Matrice orthogonale* : $A(n \times n)$ est orthogonale si : $AA' = I \Longleftrightarrow A' = A^{-1}$

7- *Rang* : Le nombre de vecteurs lignes linéairement indépendants si $n \leq p$, le nombre de vecteurs colonnes linéairement indépendants si $n > p$ (notation : $rg(A)$).

- $rg(A') = rg(A)$
- $rg(A'A) = rg(AA') = rg(A)$
- $rg(AB) \leq min(rg(A), rg(B))$

8- *Inverse généralisée* : Soit A une matrice $(n \times p)$ de rang r. $G(p \times n)$ est une *inverse généralisée* de A si :

G est de rang r et si elle satisfait à $AGA = A$.

Souvent en statistique, on utilise l'inverse généralisée de Penrose (1955) qui satisfait aussi aux conditions :

$$GAG = G \quad (GA)' = GA \quad (AG)' = AG$$

9- *Transposée d'un produit de matrices* : $(AB)' = B'A'$.

10- *Trace* : Soit A une matrice carrée. La somme des éléments diagonaux de A s'appelle la trace de A (notation : $tr(A)$).

- $tr(A) = tr(A')$

Soient $A(n \times p)$ et $B(p \times n)$ deux matrices, alors :

- $tr(AB) = tr(BA)$
- $tr(AA') = tr(A'A) = \sum_{i=1}^{n}\sum_{j=1}^{p} a_{ij}^2$

11- *Déterminant d'une matrice* :

Considérons le système d'équations :

$$\begin{cases} 3x_1 + 5x_2 + 1x_3 & = 7 \\ 2x_1 + 4x_2 + 5x_3 & = 4 \\ 1x_1 + 2x_2 + 2x_3 & = 6 \end{cases}$$

On l'écrit matriciellement :

$$\begin{bmatrix} 3 & 5 & 1 \\ 2 & 4 & 5 \\ 1 & 2 & 2 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ 6 \end{bmatrix}$$

On l'écrit symboliquement : $AX = B$.

Si la matrice A est inversible (A^{-1} existe) alors le système d'équations admet la solution : $X = A^{-1}B$ obtenue en prémultipliant $AX = B$ par A^{-1}.

Pour savoir si A^{-1} existe, on calcule un scalaire appelé le <u>*déterminant*</u> de A (noté $det(A)$).

On utilise la propriété : Si $det(A) \neq 0$ alors A^{-1} existe.

Comment calcule-t-on le déterminant de A ?

Pour préciser le mode de calcul notons :

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

avec $a_{11} = 3$, $a_{12} = 5$, $\cdots$, etc

Le déterminant de cette matrice est :

$$det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

En remplaçant les a_{ij} par leurs valeurs, on a :

$$det(A) = 3 \times 4 \times 2 + 5 \times 5 \times 1 + 1 \times 2 \times 2 - 1 \times 4 \times 1 - 3 \times 5 \times 2 - 5 \times 2 \times 2 = -1$$

En toute généralité, la définition d'un déterminant est la suivante :

Soit $A(n \times n)$ une matrice carrée, le déterminant de A est le scalaire :

$$det(A) = \sum \pm a_{1j(1)} a_{2j(2)} \cdots a_{nj(n)}$$

la sommation est sur les $n!$ permutations $j(1)$, $j(2)$, $\cdots$, $j(n)$ des entiers $1, 2, \cdots, n$. Les signes + et - sont associés aux nombres pair et impair d'interchangements de paires nécessaires pour modifier la permutation $1, 2, \cdots, n$ en $j(1), j(2), \cdots, j(n)$.

On peut aussi définir le déterminant d'une matrice carrée A à partir des <u>*mineurs*</u> et des <u>*cofacteurs*</u> de A.

Le mineur de l'élément a_{ij} de A est le déterminant obtenu en supprimant la i*ème* ligne et la j*ème* colonne de A.

Le cofacteur de a_{ij} est le mineur multiplié par $(-1)^{i+j}$, on le note A_{ij}.

En terme de cofacteurs, le déterminant de A s'écrit de la façon suivante :

$$det(A) = a_{i1}A_{i1} + \cdots + a_{in}A_{in} \qquad i = 1, \cdots, n$$

$$det(A) = a_{1j}A_{1j} + \cdots + a_{nj}A_{nj} \qquad j = 1, \cdots, n$$

- Si $A(n \times n)$ est une matrice triangulaire :

$$det(A) = \prod_{i=1}^{n} a_{ii}$$

- Si les éléments d'une ligne ou d'une colonne d'une matrice $A(n \times n)$ sont multipliés par un scalaire λ, le déterminant de la nouvelle matrice est égal à $\lambda det(A)$. Si chaque élément est multiplié par λ, alors $det(\lambda A) = \lambda^n det(A)$.

- Si deux lignes (ou deux colonnes) d'une matrice sont interchangées, le signe du déterminant est inversé.

- Si A et B sont deux matrices de dimension $(n \times n)$, alors :

$$det(AB) = det(A) \times det(B)$$

- $A(n \times n)$ est non singulière si et seulement si $det(A) \neq 0$.

- Si $A(n \times n)$ est non singulière, alors :

$$det(A^{-1}) = \frac{1}{det(A)}$$

12- *Matrice adjointe* :

C'est la matrice des cofacteurs, on la note $A^*(n \times n)$.

$$A^* = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \cdots & \cdots & \cdots \\ A_{n1} & \cdots & A_{nn} \end{bmatrix}$$

On montre que si $det(A) \neq 0$:

$$A^{-1} = \frac{1}{det(A)} A^{*'} = \frac{1}{det(A)} \begin{bmatrix} A_{11} & \cdots & A_{n1} \\ \cdots & \cdots & \cdots \\ A_{1n} & \cdots & A_{nn} \end{bmatrix}$$

4.2.1 Exemple 1 : écriture matricielle

Quand les grandeurs mesurées et les unités sont hétérogènes, on est conduit à transformer les données de telle sorte qu'elles aient une origine commune et s'expriment dans une même unité. Pour cela, on centre et on réduit ces grandeurs.

Soit $X(n \times p)$ la matrice des grandeurs mesurées. Comment s'écrit la matrice $Z(n \times p)$ des variables *centées réduites* ?

Réponse

Soient :

- x_{ij} la $j\grave{e}me$ variable observée sur le $i\grave{e}me$ individu
- $\bar{x}_j$ la moyenne de la $j\grave{e}me$ variable
- s_j l'écart type de la $j\grave{e}me$ variable
- $z_{ij} \; = \; \dfrac{x_{ij} - \bar{x}_j}{s_j}$ la variable centrée réduite

La matrice X des données s'écrit :

$$X \;=\; \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

On définit les matrices D et J de la façon suivante :

$$D \;=\; \begin{bmatrix} \frac{1}{s_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{s_2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \frac{1}{s_p} \end{bmatrix}$$

$$J \;=\; \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

L'écriture matricielle de la transformation "centrage, réduction" est alors :

$$Z \;=\; (I - \frac{1}{n}J)XD$$

avec I la matrice identité.

4.2.2 Exemple 2 : indépendance linéaire

Dans l'exemple des trois sociétés dont les actions sont cotées en bourse, on a vu que les sociétés d'investissements **CSI, CIS, SIC** ne forment pas une base de l'espace vectoriel des trois sociétés C.E.A, I.N.R.A et C.N.R.S. Calculons, par la méthode des cofacteurs, le déterminant de la matrice carrée A de dimension 3 formée par **CSI, CIS, SIC** :

$$A \;=\; \begin{bmatrix} 40 & 40 & 20 \\ 70 & 10 & 20 \\ 30 & 50 & 20 \end{bmatrix}$$

$$det(A) = 40(10 \times 20 - 20 \times 50) - 40(70 \times 20 - 20 \times 30) + 20(70 \times 50 - 10 \times 30) = 0$$

La matrice A est telle que :

- Une de ses lignes est combinaison linéaire des deux autres.
- Son déterminant est nul.

A partir de cette remarque, on admettra que :

Quand une matrice a une ligne (ou une colonne) qui est combinaison linéaire des autres lignes (ou colonnes) alors, son déterminant est égal à zéro. Ce résultat est intéressant pour vérifier l'indépendance linéaire de vecteurs.

4.3 Encore un peu de géométrie

Dans un espace vectoriel E de dimension p, on définit généralement le produit scalaire sur des vecteurs de base B_i, $(i = 1, \cdots, p)$. La base d'un espace vectoriel n'est pas unique (cf. § 3.9.2). Supposons que les vecteurs de la base choisie sont des vecteurs colonnes (cf. § 3.9.4) rangés dans une matrice. Soit :

$$B = [B_1 \quad B_2 \quad \cdots \quad B_p]$$

la matrice des vecteurs de base de dimension $(p \times p)$.

Dans E muni de la base B, deux vecteurs quelconques U et V sont de la forme :

$$U = \sum_{i=1}^{p} x_i B_i \text{ et } V = \sum_{i=1}^{p} y_i B_i$$

avec x_i et y_i des nombres réels, i variant de 1 à p.

Les x_i et les y_i, $(i = 1, \cdots, p)$ peuvent être considérés comme les composantes de deux vecteurs X et Y de dimension $(p \times 1)$, donc $U = BX$ et $V = BY$. On en déduit que :

$$< U, V > = < BX, BY >$$

Soit Q la matrice $(p \times p)$ dont l'élément (i, j) est $< B_i, B_j >$ alors :

$$< U, V > = < BX, BY > = X'B'BY = X'QY$$

On dit que Q est la *métrique* relative à la base B.

Puisqu'une métrique est liée à une base particulière, on déduit que tout changement de base définit un produit scalaire différent.

- Dans l'espace Euclidien E muni de la base B, il est clair que $X'QX > 0$ pour tout vecteur X de dimension $(p \times 1)$ différent de zéro. Dans ce cas, on dit que la matrice Q est *définie positive*. On peut vérifier par ailleurs que Q est *symétrique* : $Q = BB' = (BB')'$.
- $X'QX$ est désignée sous le nom de *forme quadratique*.
- D'un point de vue géométrique, il existe une correspondance étroite entre la classe des matrices symétriques définies positives et la classe des *ellipsoïdes* centrés à l'origine. Nous allons montrer cette correspondance dans le cas d'un espace de dimension $p = 2$.

Soit un nuage de points constitué par n observations de la variable X_1 et de la variable X_2. La forme du nuage de points, plus ou moins allongée dans une direction donnée, est une représentation de la relation qui lie les observations de X_1 aux observations de X_2 ; le coefficient de corrélation, la covariance sont des mesures de cette liaison. Le problème de la représentation géométrique des observations se pose de la façon suivante :

quelle est la figure géométrique qui enveloppe au mieux le nuage de points ?

La réponse à cette question est obtenue en écrivant que la distance d'un point au centre du nuage est inférieure ou égale à une constante positive. Comme la distance dérive du produit scalaire et que $X'QX$ est le produit scalaire associé à la métrique Q, on déduit que, les points à une distance du centre du nuage inférieure à C, sont tels que $X'QX \leq C^2$, soit:

$$X'QX = \begin{bmatrix} X_1 - \bar{X}_1 & X_2 - \bar{X}_2 \end{bmatrix} \begin{bmatrix} 1 & b \\ b & 1 \end{bmatrix} \begin{bmatrix} X_1 - \bar{X}_1 \\ X_2 - \bar{X}_2 \end{bmatrix} \leq C^2$$

avec X_1 et X_2 les coordonnées d'un point, $\bar{X}_1$ et $\bar{X}_2$ les coordonnés du point moyen et $Q = \begin{bmatrix} 1 & b \\ b & 1 \end{bmatrix}$ la métrique (matrice symétrique, définie positive). Si on développe cette expression, on trouve :

$$(X_1 - \bar{X}_1)^2 + 2b\,(X_1 - \bar{X}_1)(X_2 - \bar{X}_2) + (X_2 - \bar{X}_2)^2 \leq C^2$$

qui n'est autre que l'équation d'une surface limitée par une conique. Cette conique sera une ellipse si $b^2 \leq 1$.

On observe que, si le coefficient b est de signe opposé à celui du produit $(X_1 - \bar{X}_1)(X_2 - \bar{X}_2)$, l'ellipse est orientée dans le sens des variations conjointes des variables X_1 et X_2 (par exemple, si X_1 et X_2 augmentent simultanément le grand axe de l'ellipse est une droite de pente positive). Intuitivement, on dira que le terme $2b\ (X_1 - \bar{X}_1)(X_2 - \bar{X}_2)$ doit être négatif pour que l'ellipse et le nuage de points aient des orientations identiques. Dans ces conditions, quand X_1 et X_2 sont liées positivement (resp. négativement), le produit $(X_1 - \bar{X}_1)(X_2 - \bar{X}_2)$ est positif (resp. négatif) et donc b doit être négatif (resp. positif). En d'autres termes, si la mesure de liaison est le coefficient de corrélation, il interviendra avec le signe "-"(resp. "+") si X_1 et X_2 sont liées positivement (resp. négativement). Ainsi, la métrique Q associée à l'enveloppe elliptique du nuage de points ne peut pas être la matrice de corrélation $\begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$ car ρ et $(X_1 - \bar{X}_1)(X_2 - \bar{X}_2)$ sont de même signe, mais doit être l'inverse de la matrice de corrélation $\frac{1}{1-\rho^2} \begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}$ car dans ce cas $-\rho$ et $(X_1 - \bar{X}_1)(X_2 - \bar{X}_2)$ sont de signes opposés.

D'un point de vue général, on admettra que $X'Q^{-1}X \leq 1$ est l'ellipsoïde d'inertie associé à la forme quadratique $X'QX$. La justification de cette affirmation est donnée au paragraphe 4.4 (Q et Q^{-1} ont les mêmes vecteurs propres).

On va illustrer, sur des exemples, ce qui vient d'être dit sur les représentations géométriques qui prennent en compte une mesure de liaison entre X_1 et X_2.

Soient S_1^2 et S_2^2 des estimations des variances de X_1 et de X_2, $X_1^* = \frac{X_1 - \bar{X}_1}{S_1}$ et $X_2^* = \frac{X_2 - \bar{X}_2}{S_2}$ les variables X_1 et X_2 centrées réduites.

Plaçons nous d'abord dans l'espace des individus repérés par leurs coordonnées centrées réduites X_1^* et X_2^*. Considèrons les deux cas suivants :

1- Les variables X_1^* et X_2^* ne sont pas liées par une relation linéaire (*i.e.* le coefficient de corrélation $\rho = 0$). Le lieu des points M de coordonnées X_1^* et X_2^*, à une distance unité de l'origine est le cercle d'équation :

$$X_1^{*2} + X_2^{*2} = 1$$

Sous forme matricielle l'équation du cercle s'écrit $X^{*'} \, Q^{-1} \, X^* = 1$ avec $X^{*'}$ le vecteur transposé de X^* et

$$X^* = \begin{bmatrix} X_1^* \\ X_2^* \end{bmatrix}, \quad Q^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

2- Les variables X_1^* et X_2^* sont liées par une relation linéaire (*i.e.* le coefficient de corrélation $\rho \neq 0$). L'inverse de la matrice de corrélation est :

$$Q^{-1} = \frac{1}{(1-\rho^2)} \begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}$$

Après développement, l'équation de l'ellipse $X^{*'} \, Q^{-1} \, X^* = 1$ s'écrit :

$$X_1^{*2} - 2\rho \, X_1^* \, X_2^* + X_2^{*2} = 1 - \rho^2 \qquad (1)$$

On remarque que :

- Si $\rho = 0$ on retrouve l'équation du cercle.
- Si $\rho = \pm 1$, on obtient des équations de droites qui passent par le point moyen μ de coordonnées (X_1, X_2). Pour chacune d'elle, le signe de la pente est celui du coefficient de corrélation ($\rho = \pm 1 \Longrightarrow X_1^* = \pm X_2^*$).

Suivant les valeurs du coefficient de corrélation, (1) permet de passer de l'équation d'un cercle ($\rho = 0$) à celle d'une ellipse ($|\rho| < 1$) puis, à celle d'une droite ($\rho = \pm 1$). Cette représentation géométrique (figure 4.1) rend compte du degré de liaison entre X_1^* et X_2^* car la valeur du coefficient de corrélation ρ caractérise l'applatissement de l'ellipse c'est-à-dire *l'exentricité* de l'ellipse.

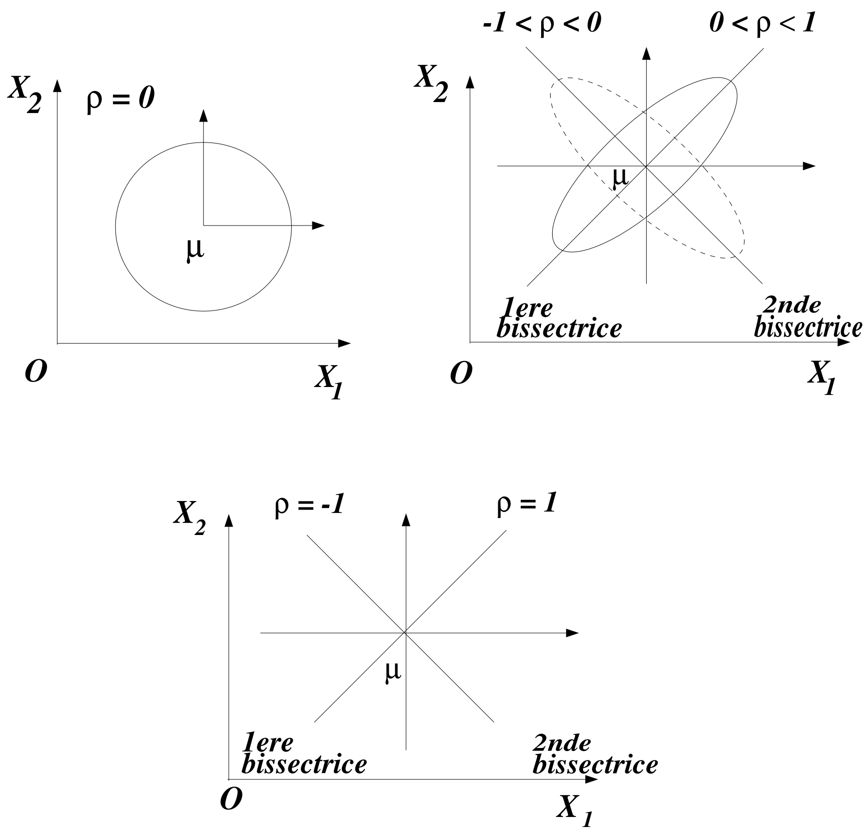

Figure 4.1 : Représentation dans l'espace des individus, repérés par leurs coordonnées centrées réduites, des ellipses d'inertie centrées au point moyen μ pour différentes valeurs du coefficient de corrélation (ρ).

Plaçons nous maintenant dans l'espace des individus repérés par leurs coordonnées X_1 et X_2. L'équation (1) s'écrit :

$$(\frac{X_1 - \bar{X}_1}{S_1})^2 \; - \; 2\rho(\frac{X_1 - \bar{X}_1}{S_1})(\frac{X_2 - \bar{X}_2}{S_2}) \; + \; (\frac{X_2 - \bar{X}_2}{S_2})^2 \; = \; 1 - \rho^2$$

Soit :

$$[S_2(X_1 - \bar{X}_1)]^2 - 2\rho[S_2(X_1 - \bar{X}_1)][S_1(X_2 - \bar{X}_2)] + [S_1(X_2 - \bar{X}_2)]^2 = S_1^2 S_2^2\,(1 - \rho^2) \quad (2)$$

C'est l'équation d'une ellipse, d'excentricité ρ, centrée au point moyen μ de coordonnées $(\bar{X}_1, \bar{X}_2)$. On considère les deux cas suivants :

1- Le coefficient de corrélation $\rho \; = \; 0$:

- Si $S_1 = S_2 = S$, l'équation (2) se réduit à l'équation d'un cercle.
- Si $S_1 \neq S_2$, l'équation (2) est soit l'équation d'une ellipse de grand axe parallèle à X_1 si $S_1 > S_2$ soit de grand axe parallèlle à X_2 si $S_2 > S_1$.

2- Le coefficient de corrélation $\rho \; \neq \; 0$:

- Si $\rho \; = \; \pm 1$, l'équation (2) se réduit à l' équation générale d'une droite passant par le point moyen μ de coordonnées $(\bar{X}_1, \bar{X}_2)$, de pente $\pm\frac{S_1}{S_2}$ selon que $\rho = \pm 1$.
- Si $|\rho| \; < 1$, l'équation (2) est l'équation d'une ellipse pour laquelle la pente du grand axe dépend du signe du coefficient de corrélation et du rapport des valeurs de S_1 et S_2.

Ces ellipses d'inertie, plus ou moins dégénérées en fonction des valeurs du coefficient de corrélation, dont les orientations dans le plan dépendent du signe du coefficient de corrélation et du rapport des variances, représentent bien l'allure générale des nuages de points (figure 4.2).

Quelques remarques s'imposent :

- Dans les deux cas qui viennent d'être présentés, les métriques utilisées sont différentes si on considère que sur l'espace des individus, les points individus sont repérés par :

  1- des coordonnées centrées réduites X_1^* et X_2^*. La métrique est alors la matrice de corrélation.

  2- des coordonnées centrées $X_1 - \bar{X}_1$ et $X_2 - \bar{X}_2$. La métrique est alors la matrice de variance-covariance.

- Il existe d'autres métriques qui permettent d'interpréter des représentations géométriques en fonction de la nature des données et de l'objectif fixé par l'analyse. On peut citer par exemple la métrique du χ^2 utilisée en analyse factorielle des correspondances (Caillez et Pages, 1976). En fait, il est presque toujours possible de travailler avec la métrique du "géomètre" à condition de faire les transformations idoines sur les données.

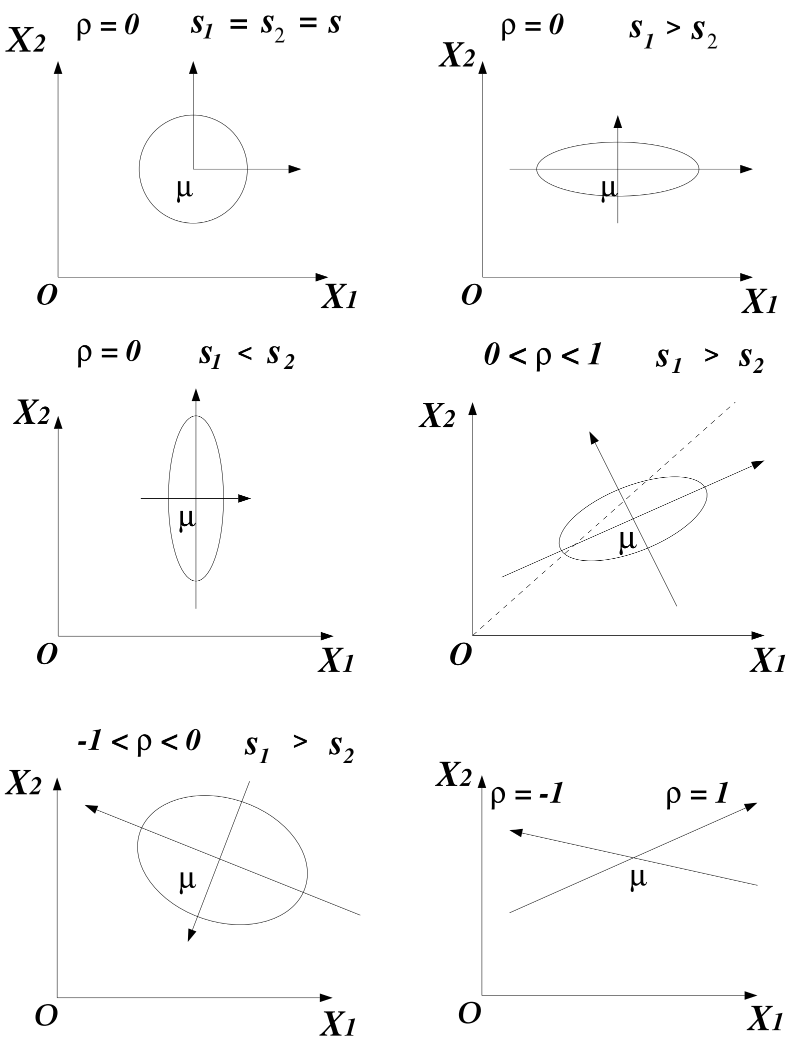

Figure 4.2 : Représentation dans l'espace des individus, repérés par leurs coordonnées centrées au point moyen μ, des ellipses d'inertie pour des valeurs différentes du coefficient de corrélation (ρ) et des écarts types (S_1 et S_2).

- Quand un produit scalaire est donné dans un espace, il est possible de considérer une gamme plus étendue de produits scalaires et la gamme correspondante d'ellipses associées dans laquelle le cercle ne représente qu'un cas particulier. Sur la figure 4.3 on a représenté un plan dans lequel les vecteurs OB_1 et OB_2 sont perpendiculaires. Si on prend comme produit scalaire dans ce plan celui qui est défini en choisissant pour base OB_1 et OB_2 alors, l'ensemble des points à une distance unité de l'origine O est le cercle de rayon unité centré en O. Supposons maintenant qu'un couple de vecteurs arbitraires OB_1^*, OB_2^* est choisi comme base orthonormale pour une <u>*autre*</u> définition du produit scalaire sur le <u>*même*</u> espace. Alors, on peut vérifier que l'ensemble des points à une distance unité de l'origine O est une ellipse. Cette ellipse est telle que la tangente à l'ellipse, au point d'intersection de l'ellipse et de OB_1^*, est parallèle à OB_2^* et que la tangente à l'ellipse, au point d'intersection de l'ellipse et de OB_2^*, est parallèle à OB_1^*, aucun autre axe de l'ellipse n'a cette propriété. On dit que OB_1 et OB_2 d'une part et OB_1^* et OB_2^* d'autre part sont les <u>*demi-axes conjugués*</u> du cercle et de l'ellipse respectivement.

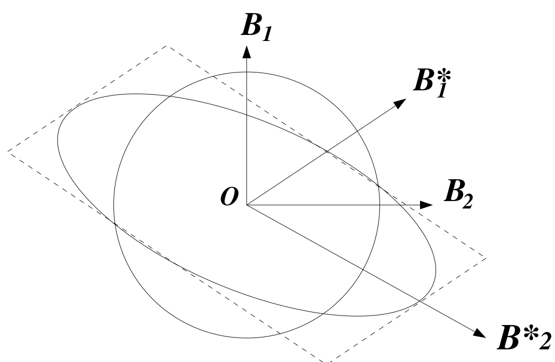

Figure 4.3 : Ellipses associées aux bases orthonormales OB_1, OB_2 et OB_1^*, OB_2^*.

Les résultats précédents se généralisent à des espaces de dimensions $p > 2$.

Conséquence

Un objectif de la statistique étant de décrire et de résumer des nuages de points, on utilise le concept d'ellipsoïde centré à l'origine pour rechercher des directions et des points particuliers qui caractérisent ces nuages de points. Ces directions et ces points particuliers sont : les axes principaux, les axes conjugués, les points moyens des ellipsoïdes. Pour déterminer ces axes on raisonne en termes de changement de base en utilisant les notions de <u>*valeurs propres*</u> et de <u>*vecteurs propres*</u>.

4.4 Valeurs et vecteurs propres

Soient D et $D^\perp$ deux droites orthogonales du plan passant par l'origine O. On note S la *symétrie* par rapport à D.

S est une application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}^2$.

Quels sont les points de $\mathbb{R}^2$ tels que $S(X) = \lambda X$ (Cailliez et Pages, 1976 : pp 100-101) **?**

$$V_1 \in D \Longrightarrow S(V_1) = \lambda_1 V_1 = 1 \times V_1$$

$$V_2 \in D^\perp \Longrightarrow S(V_2) = \lambda_2 V_2 = -1 \times V_2$$

Le point O joue un rôle particulier : $S(O) = \lambda O$ pour tout λ.

V_1 et V_2 sont dits *vecteurs propres* de S associés aux *valeurs propres* λ_1 et λ_2.

- Tout vecteur de D est un vecteur propre associé à la valeur propre $\lambda_1 = 1$.
- Tout vecteur de $D^\perp$ est un vecteur propre associé à la valeur propre $\lambda_2 = -1$.

Il y a donc *deux infinités* de vecteurs propres associés à l'application linéaire S, ils se répartissent sur les deux sous espaces vectoriels de dimension 1 : D et $D^\perp$.

On dit que :

- D est le *sous espace propre* associé à la valeur propre λ_1 =1.
- $D^\perp$ est le *sous espace propre* associéà la valeur propre $\lambda_2 = -1$.

En généralisant cet exemple de la symétrie par rapport à D, on définit les vecteurs propres et les valeurs propres de la façon suivante :

4.4.1 Définition

Soit E un espace vectoriel de dimension p, f une application linéaire de E dans E. $V \in E$ est vecteur propre de f associé à la valeur propre $\lambda \in \mathbb{R}$ si $f(V) = \lambda V$ $(V \neq O)$.

Si E est muni de la base $B = [B_1, B_2, \cdots, B_p]$,

$$V = \sum_{i=1}^{p} v_i B_i$$

Soit A la matrice associée à l'application linéaire f, la base B étant choisie dans E :

$$f(V) = \lambda V \Longleftrightarrow AV = \lambda V \Longleftrightarrow (A - \lambda I)V = 0$$

V est vecteur propre de la matrice A de valeur propre associée λ. Le système homogène $(A - \lambda I)V = 0$ admettra une solution non triviale si et seulement si $det(A - \lambda I) = 0$.

4.4.2 Résultats généraux

- A toute matrice carrée A d'ordre p sont associées au plus p valeurs propres réelles. Ces valeurs propres sont les racines d'un polynôme de degré p :

  $det(A - \lambda I) = 0$ appelé *polynôme caractéristique*.

  Tout vecteur propre V_i associé à la valeur propre λ_i de la matrice carrée A satisfait au système d'équations :

  $$(A - \lambda_i I)V_i = 0$$

- Si A est symétrique :

  1- Toutes les racines du polynôme caractéristique sont réelles.

  2- A la racine simple λ_i correspond un sous espace propre de dimension 1.

  3- L'ordre de multiplicité d'une valeur propre est la dimension du sous espace propre associé.

  4- Si λ_i et λ_j sont deux valeurs propres distinctes, les vecteurs propres associés sont orthogonaux. En effet :

  Soient λ_i, λ_j deux valeurs propres distinctes et V_i, V_j les vecteurs propres associés. On a :

  $$AV_i = \lambda_i V_i \text{ et } AV_j = \lambda_j V_j$$

  En prémultipliant la première de ces équations par V_j' (vecteur propre transposé de V_j) et la seconde par V_i' (vecteur propre transposé de V_i), on obtient :

  $$V_j' A V_i = \lambda_i V_j' V_i \text{ et } V_i' A V_j = \lambda_j V_i' V_j$$

  En faisant la différence de ces deux expressions, on déduit :

  $$\lambda_i V_j' V_i = \lambda_j V_i' V_j \Longleftrightarrow (\lambda_i - \lambda_j) V_j' V_i = 0$$

  car $V_j' V_i = V_i' V_j \in I\!R$ donc :

  $$\lambda_i \neq \lambda_j \Longrightarrow V_j' V_i = 0$$

  Le produit scalaire des vecteurs V_i et V_j est égal à zéro, ces deux vecteurs sont orthogonaux.

5- Il existe une matrice orthogonale P telle que :

$$A = P'DP$$

où D est une matrice diagonale. Les éléments situés sur la diagonale sont les valeurs propres de A. Les colonnes de la matrice P sont les vecteurs propres normalisés de A. Même si les valeurs propres ne sont pas distinctes, il est encore possible de choisir les éléments de leurs vecteurs propres permettant d'obtenir un ensemble de vecteurs propres mutuellement orthogonaux.

Ces propriétés sont importantes car si on applique la transformation orthogonale :

$$X = PY$$

aux p variables d'une forme quadratique $X'AX$, la forme quadratique devient:

$$X'AX = Y'P'APY = Y'DY = \lambda_1 y_1^2 + \cdots + \lambda_r y_r^2$$

les λ_i sont les valeurs propres de A et r est le <u>rang</u> de la forme quadratique. Toute forme quadratique peut être réduite à une somme de carrés pondérés en calculant les valeurs et les vecteurs propres de la matrice A.

- Si une matrice A d'ordre p admet p valeurs propres différentes λ_i, le système de vecteurs propres est une base de E.

- Si A est une matrice carrée inversible d'ordre p qui admet p valeurs propres non nulles λ_i et les vecteurs propres V_i. Alors, la matrice inverse A^{-1} admet pour valeurs propres les inverses des valeurs propres de A et a les mêmes vecteurs propres que A.

  En effet, les vecteurs propres V_i de A, associés aux valeurs propres λ_i, satisfont au système d'équations :

$$(A - \lambda_i I)V_i = 0$$

  et en prémultipliant par A^{-1} on obtient :

$$(I - \lambda_i A^{-1})V_i \iff (A^{-1} - \lambda_i^{-1} I)V_i = 0$$

4.4.3 Exemple 3 : recherche de valeurs et de vecteurs propres

Chercher les valeurs propres et les vecteurs propres des matrices :

$$A = \begin{bmatrix} 1 & 1 \\ 8 & 3 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

Que constate-t-on ?

Réponse

La matrice A a deux valeurs propres *distinctes* : $\lambda_1 = 5$ et $\lambda_2 = -1$ (solutions de $(1-\lambda)(3-\lambda) - 8 = 0$).

La matrice B a deux valeurs propres *confondues* : $\lambda_1 = \lambda_2 = 1$ (solutions de $(1-\lambda)^2 = 0$), on dit que 1 est valeur propre d'ordre 2.

La matrice C n'a qu'une valeur propre *réelle* : $\lambda = 1$ (solution de $(1-\lambda)(\lambda^2+1) = 0$).

4.4.4 Exemple 4 : vecteurs propres et droites de régression

On veut étudier la relation entre la longueur du corps x et la profondeur y de poitrine de 22 vaches laitières (Dagnélie, 1969 : pp 62). Les données sont présentées dans le tableau 4.1 :

x	y	x	y
168	71	169	68
150	65	148	67
154	67	145	66
165	69	163	69
148	68	161	69
151	70	176	74
159	70	159	73
151	69	155	71
169	74	158	70
157	71	161	73
146	71	150	65

Tableau 4.1 : Valeurs observées de la longueur du corps (x) et de la profondeur de la poitrine (y) de 22 vaches laitières.

Comme on veut étudier les variations simultanées des deux caractéristiques x et y il n'y a pas plus de raisons d'utiliser la droite de régression de y en x que la droite de régression de x en y. Dagnelie (1969 : pp 82) propose d'utiliser le critère des moindres rectangles (minimum de la somme des produits des écarts $|x_i - x(y_i)|$ et $|y_i - y(x_i)|$) pour obtenir une droite de régression dite droite des *moindres rectangles*. Cette dernière droite de régression est comprise entre les droites de régression de y en x et x en y. Par ailleurs, en minimisant perpendiculairement à la droite elle-même on obtiendrait une nouvelle droite de régression dite droite de *régression orthogonale*. Ces quatre droites de régression passent toutes par le point moyen $(\bar{x}, \bar{y})$, elles ont pour équations :

- régression de x en y : $y = 0,519x - 12,2$ (1)
- droite des moindres rectangles : $y = 0,310x + 20,8$ (2)
- droite de régression orthogonale : $y = 0,196x + 38,7$ (3)
- régression de y en x : $y = 0,185x + 40,5$ (4)

Sachant :

$\bar{x} = 157,41$, $\bar{y} = 69,55$

$var(x) = 68,95$, $var(y) = 6,614$, $cov(x,y) = 12,73$

1- Retrouver l'équation de la droite de régression orthogonale en calculant les valeurs et les vecteurs propres de la matrice de variance-covariance V.

2- Retrouver l'équation de la droite des moindres rectangles en calculant les valeurs propres et les vecteurs propres de la matrice de corrélation.

Réponse

1- La matrice de variance-covariance V représente (à une constante multiplicative près) la matrice d'inertie du nuage de points (x,y). Chercher à déterminer la droite de régression orthogonale est équivalent à chercher la direction de plus grande variabilité de ce nuage de points, c'est-à-dire le vecteur propre associé à la plus grande valeur propre de V. Pour cela, on résout :

$$det(V - \lambda I) = 0$$

avec :

$$V = \begin{bmatrix} 68,950 & 12,730 \\ 12,730 & 6,614 \end{bmatrix}$$

On déduit l'équation caractéristique :

$$\lambda^2 - 75,564\lambda + 293,982 = 0$$

qui admet deux racines : $\lambda_1 = 71,45$ et $\lambda_2 = 4,11$

Le vecteur propre de composantes v_1 et v_2 associé à la première valeur propre satisfait au système d'équations :

$$\begin{cases} -2,50v_1 + 12,73v_2 & = 0 \\ 12,73v_1 - 64,84v_2 & = 0 \end{cases}$$

Ces deux équations sont identiques. Si on choisit $v_1 = 1$, on déduit que $v_2 = 0,196$. La pente de ce vecteur propre est égale à 0,196 (rapport v_2 sur v_1). La droite de régression orthogonale $y = ax + b$ passe par le point moyen, la valeur du terme constant est donc : $b = 69,55 - 0,196 \times 157,41 = 38,70$.

2- La matrice de corrélation est :

$$C = \begin{bmatrix} 1 & 0,596 \\ 0,596 & 1 \end{bmatrix}$$

L'équation caractéristique est :

$$(1-\lambda)^2 - 0,355 = 0$$

qui admet deux racines $\lambda_1 = 1,596$ et $\lambda_2 = 0,404$

Le vecteur propre de composantes v_1 et v_2 associé à la première valeur propre satisfait au système d'équations :

$$\begin{cases} -0,596 v_1 + 0,596 v_2 & = 0 \\ 0,596 v_1 - 0,596 v_2 & = 0 \end{cases}$$

La pente du premier vecteur propre est égale à 1, c'est le rapport $\dfrac{v_2}{v_1} = \dfrac{0,596}{0,596}$. La matrice de corrélation n'est autre que la matrice de variance-covariance des variables centrées réduites. On déduit donc que :

$$\frac{\dfrac{y - 69,55}{2,572}}{\dfrac{x - 157,41}{8,304}} = 1$$

Soit : $y = 0,310x + 20,8$ qui est l'équation de la droite des moindres rectangles. Ainsi, la droite des moindres rectangles n'est autre que la droite de régression orthogonale calculée sur les variables centrées réduites.

Sur la figure 4.4 on a représenté, les données et les différentes droites de régression.

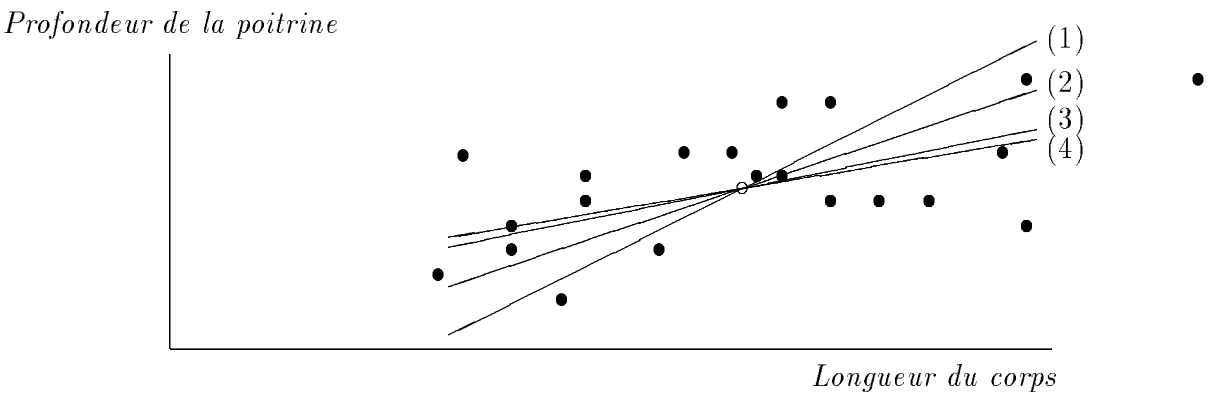

Figure 4.4 : Régression de x en y (1), droite des moindres rectangles (2), régression orthogonale (3), régression de y en x (4). La moyenne des observations est notée $\circ$.

4.5 Projecteurs associés à une décomposition en somme directe

Au paragraphe 3.8 nous avons vu la décomposition en somme directe d'un espace vectoriel E de dimension p. Supposons que l'on ait une décomposition en somme directe de E en k sous espaces $E_1, \cdots, E_k$ de dimensions $p_1, \cdots, p_k$ ($E = E_1 \oplus \cdots \oplus E_k$ et $p = p_1 + \cdots, p_k$).

Tout vecteur V de E s'écrit de façon unique $V = \sum_{i=1}^{k} V_i$ avec $V_i \in E_i$, $i = 1, \cdots, k$. Les k applications linéaires f_i de E dans E qui font correspondre $f_i(V) = V_i \in E_i \subset E$ à V sont appelées *projecteurs* associés à la somme directe $E = E_1 \oplus \cdots \oplus E_k$.

Ces applications linéaires sont telles que $f_i(f_i(V)) = f_i(V_i) = V_i$, elles sont dites *idempotentes*.

Les projecteurs sont des applications linéaires idempotentes et réciproquement toute application linéaire idempotente est un projecteur. Les matrices A_i associées aux projecteurs f_i sont idempotentes ($A_i A_i = A_i$), leurs valeurs propres sont toutes égales à 1 ou à 0.

4.5.1 Exemple 5 : projecteurs associés à des décompositions en somme directe

On considère dans $\mathbb{R}^3$ la base orthonormale

$$B = [\, B_1 \quad B_2 \quad B_3 \,] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

et les décompositions en somme directe suivantes :

- $\mathbb{R}^3 = D_1 \oplus P$.
- $\mathbb{R}^3 = D_2 \oplus P$.

avec

- P le sous espace vectoriel de $\mathbb{R}^3$ engendré par les vecteurs de base B_1 et B_2.

- D_1 le sous espace de $\mathbb{R}^3$ engendré par le vecteur B_3 orthogonal à P.

- D_2 le sous espace de $\mathbb{R}^3$ engendré par la bissectrice de l'angle (B_1, B_3).

La base B étant choisie :

1- Ecrire les matrices associées aux projecteurs sur D_1 et P d'une part et aux projecteurs D_2 et P d'autre part. Vérifier que ces matrices sont idempotentes.

2- Ecrire les matrices associées aux différents projecteurs dans la base

$$B' = [\,B_1 \quad B_2 \quad B_3'\,] = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Réponse

1- Sur la figure 4.5 on a représenté les sous espaces impliqués dans les deux décompositions en somme directe. Soit f_D le projecteur sur le sous espace vectoriel de dimension 1 (la droite D_1 ou la droite D_2 suivant le cas étudié) et f_P le projecteur sur le sous espace vectoriel de dimension 2 (le plan P). f_D et f_P sont des applications linéaires de $I\!R^3$ dans $I\!R^3$.

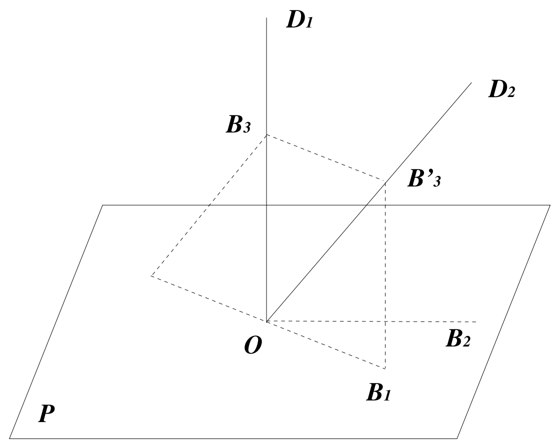

Figure 4.5 : Décomposition de $I\!R^3$ en une somme directe suivant 1) la droite D_1 et le plan P, 2) la droite D_2 et le plan P.

Dans la décomposition en somme directe $I\!R^3 = D_1 \oplus P$ les sous espaces vectoriels D_1 et P sont complémentaires orthogonaux. Les projections sur ces sous espaces sont telles que :

- A tout point V de coordonnées v_1, v_2 et v_3 dans $I\!R^3$, l'application linéaire f_{D_1} fait correspondre sa projection sur OB_3 c'est-à-dire le vecteur de composantes $(0,\ 0,\ v_3)$. La matrice associée à cette application linéaire est :

$$A_{D_1} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

car

$$A_{D_1}V = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ v_3 \end{bmatrix}$$

ce qui est le résultat attendu.

- A tout point V de coordonnées v_1, v_2 et v_3 dans $I\!R^3$, l'application linéaire f_P fait correspondre sa projection sur le plan P, c'est-à-dire le vecteur de composantes $(v_1,\ v_2,\ 0)$. La matrice associée à cette application linéaire est :

$$A_P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

car

$$A_PV = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ 0 \end{bmatrix}$$

ce qui est le résultat attendu.

- On vérifie, en effectuant les produits matriciels, que :

$$A_{D_1}A_{D_1} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = A_{D_1}$$

$$A_PA_P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = A_P$$

Les matrices A_{D_1} et A_P sont idempotentes.

- Soit I la matrice unité, on remarque que :

$$A_{D_1} = I - A_P \ \text{ et } \ A_p = I - A_{D_1}$$

Ce résultat est général.

Dans la décomposition en somme directe $I\!R^3 = D_2 \oplus P$ les sous espaces vectoriels D_2 et P sont complémentaires, ils ne sont pas orthogonaux. Les projections sur ces sous espaces sont telles que :

- A tout point V de coordonnées v_1, v_2 et v_3 dans $I\!R^3$, l'application linéaire f_{D_2} fait correspondre sa projection sur la bissectrice de l'angle (B_1, B_3). Dans la base B, cette bissectrice est de composantes $(1,\ 0,\ 1)$. La matrice associée à cette application linéaire est :

$$A_{D_2} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- En appliquant le résultat que $A_P = I - A_{D_2}$, on déduit que :

$$A_P = I - A_{D_2} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

- On vérifie, en effectuant les produits matriciels, que :

$$A_{D_2}A_{D_2} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = A_{D_2}$$

$$A_P A_P = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = A_P$$

Les matrices A_{D_2} et A_P sont idempotentes.

Dans les deux décompositions de $I\!\!R^3$ qui viennent d'être faites, on remarque que les matrices A_p associées aux projecteurs sur P, sont différentes.

2- Soit M la matrice des coordonnées des vecteurs B_1, B_2 et B_3 dans la base B' (M est dite matrice de changement de base) :

$$M = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Les matrices associées aux applications linéaires f_{D_1} et f_P d'une part, f_{D_2} et f_P d'autre part dans la nouvelle base sont :

$$A'_{D_1} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ et } A'_P = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$A'_{D_2} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ et } A'_P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Ces matrices sont idempotentes, ce sont des matrices associées à des projecteurs.