Rappels de calcul matriciel

T. Moreau et M. Chavance

Octobre 2006

Chapitre 1

LES MATRICES. PREMIERES DEFINITIONS

1.1 Définition

Une *matrice* $A(n,p)$ est un tableau rectangulaire de nombres comprenant n lignes et p colonnes. Un nombre de ce tableau est un *élément* de A ; n et p sont les *dimensions* de A. On note

$$A = \begin{pmatrix} a_{11} & a_{12} & & \dots & & a_{1p} \\ a_{21} & a_{22} & & \dots & & a_{2p} \\ \vdots & \vdots & & & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{ip} \\ \vdots & \vdots & & & & \vdots \\ a_{n1} & a_{n2} & & \dots & & a_{np} \end{pmatrix}$$

ou $A = \|a_{ij}\|$. Les p éléments $a_{i1}, \dots, a_{ip}$ constituent la ligne i et les n éléments $a_{1j}, \dots, a_{nj}$ constituent la colonne j.

1.2 Exemples

1.2.1 Les matrices les plus simples

Un nombre est une matrice (1, 1). La matrice $V = (x_1, x_2, x_3)$ à 1 ligne et 3 colonnes est appelée *vecteur ligne*. La matrice $W = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ à 3 lignes et

1 colonne est appelée *vecteur colonne*. Dans l'espace à 3 dimensions $\Re^3$, un point peut être représenté par ses 3 coordonnées (x_1, x_2, x_3) et V, ou W, peut ainsi désigner un point de $\Re^3$.

1.2.2 Exercice

Les tensions systolique et diastolique suivantes ont été mesurées sur 10 sujets:

$$\begin{array}{lll} sujet1 : (180, 112) & sujet2 : (152, 82) & sujet3 : (167, 80) \\ sujet4 : (154, 106) & sujet5 : (148, 80) & sujet6 : (164, 98) \\ sujet7 : (156, 98) & sujet8 : (171, 96) & sujet9 : (150, 106) \\ sujet10 : 160, 111) & & \end{array}$$

Représenter le plus simplement possible les données sous forme de matrice de deux façons différentes.

1.3 Matrice transposée

1.3.1 Définition

La transposée de la matrice $A\,(n, p)$ est une matrice (p, n) dont la ième ligne est la ième colonne de A (i= 1, ..., p) et dont la jème colonne est la jème ligne de A (j=1, ..., n). Cette matrice est notée A^t ou A'.

1.3.2 Exercices

Ecrire la matrice transposée de $A = \begin{pmatrix} 1 & 4 & 5 \\ -1 & 8 & 0 \\ 3 & 1 & -3 \\ 2 & 0 & 1 \\ 7 & 2 & 1 \end{pmatrix}$, puis la matrice transposée de la matrice obtenue.

Ecrire la matrice transposée du vecteur ligne $V = \begin{pmatrix} 1 & 0 & 3 \end{pmatrix}$, puis la transposée de la matrice obtenue.

1.3.3 Propriété

La transposée de la transposée d'une matrice est la matrice elle même

$$(A')' = A$$

.

1.4 Matrices carrées

1.4.1 Définitions

Une matrice est dite carrée si son nombre de lignes est égal a son nombre de colonnes. Ce nombre est appelé *ordre* de la matrice. Les éléments de la forme a_{ii} constituent la *diagonale* de A, appelée parfois *diagonale principale* pour la distinguer de la *seconde diagonale*.

Exemple: une matrice carrée est dite magique si la somme des éléments de chacune de ses lignes est égal à la somme des éléments de chacune de ses colonnes. Soit la matrice magique :

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

Que peut on dire de A' ?

1.4.2 Matrices carrées particulières

Matrice *diagonale* : tous les éléments sont nuls sauf ceux de la diagonale, $\forall i$ $\forall j$, $i \neq j \Rightarrow a_{ij} = 0$.

Matrice *scalaire* : c'est une matrice diagonale dont tous les éléments diagonaux sont égaux a un même nombre a.

Matrice *unité* : c'est une matrice scalaire avec a=1. On la note I ou I_n, si l'on veut préciser son ordre.

Matrice *symétrique* : c'est une matrice A telle que $\forall i$ $\forall j$, $a_{ij} = a_{ji}$. Remarque : A symétrique $\Leftrightarrow A' = A$.

Matrice *antisymétrique* : c'est une matrice A telle que $\forall i$ $\forall j$ $a_{ij} = -a_{ji}$.

Remarque 1 : la diagonale d'une matrice antisymétrique est nulle car $\forall i$ $a_{ii} = -a_{ii} = 0$.

Remarque 2 : A antisymétrique $\Leftrightarrow A' = -A$.

1.4.3 Matrices de covariance et de corrélation

Soient $(X_1, X_2, \ldots, X_p)$ p variables aléatoires pour lesquelles $var(X_i) = \sigma_i^2$ et $cov(X_i, X_j) = \sigma_{ij}$. La matrice de covariance du p-uple $(X_1, X_2, \ldots, X_p)$

est une matrice carrée symétrique d'ordre p

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2p} \\ \vdots & \vdots & & \vdots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_p^2 \end{pmatrix}$$

On définit de même la matrice de correlation du p-uple $(X_1, X_2, \dots, X_p)$. C'est une matrice carrée symétrique dont les éléments sont les coefficients de corrélation des variables entre elles: $\rho_{ij} = corr(X_i, X_j)$

$$\Omega = \begin{pmatrix} 1 & \rho_{12} & \dots & \rho_{1p} \\ \rho_{21} & 1 & \dots & \rho_{2p} \\ \vdots & \vdots & & \vdots \\ \rho_{p1} & \rho_{p2} & \dots & 1 \end{pmatrix}$$

Si toutes les variables X_i sont deux à deux indépendantes, $i \neq j \Rightarrow \sigma_{ij} = 0$. Il s'en suit que Σ est diagonale, ce qui peut se noter $\Sigma = Diag(\sigma_1^2, \sigma_2^2, \dots, \sigma_p^2)$, et que Ω est égale à la matrice unité. La réciproque n'est vraie que si les X_i sont gaussiennes.

1.4.4 ATTENTION

Toutes les matrices carrées symétriques ayant des éléments positifs sur la diagonale ne sont pas forcément des matrices de covariance. De même toutes les matrices carrées symétriques ayant des 1 sur la diagonale et des éléments plus petits que 1 ailleurs ne sont pas nécessairement des matrices de corrélation. Exemple $\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$

1.5 Matrices définies par blocs

Une notation souvent utilisée consiste à écrire une matrice $A(n,p)$ comme une juxtaposition de sous-matrices ou *blocs*. On dit alors que A est *partitionnée*. Il faut bien entendu que les dimensions des blocs soient compatibles. Exemple: A, matrice à n lignes et p colonnes peut s'écrire $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$

$$= \begin{pmatrix} a_{11} & \dots & a_{1l} & a_{1(l+1)} & \dots & a_{1p} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{k1} & \dots & a_{kl} & a_{k(l+1)} & \dots & a_{kp} \\ a_{(k+1)1} & \dots & a_{(k+1)l} & a_{(k+1)(l+1)} & \dots & a_{(k+1)p} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{nl} & a_{n(l+1)} & \dots & a_{np} \end{pmatrix}$$

où A_{11} et A_{12} ont k lignes, A_{21} et A_{22} n-k lignes, A_{11} et A_{21} l colonnes, A_{12} et A_{22} p-l colonnes.

1.5.1 Transposition

On vérifie facilement que la transposée de A, partitionnée comme ci-dessus, vaut : $A' = \begin{pmatrix} A'_{11} & A'_{21} \\ A'_{12} & A'_{22} \end{pmatrix}$

1.5.2 Exercices

1) Montrer que $A = (P \ \ Q) \Rightarrow A' = \begin{pmatrix} P' \\ Q' \end{pmatrix}$

2) Soit $A = \begin{pmatrix} P & Q & R \\ S & T & U \end{pmatrix}$. Calculer A'.

Chapitre 2

Opérations sur les matrices

2.1 Multiplication d'une matrice par un nombre

Le produit d'une matrice A et d'un nombre ou *scalaire* λ est une matrice $B = \lambda \times A$ obtenue en multipliant par λ tous les éléments de A.

$$b_{ij} = \lambda \times a_{ij}$$

B a donc les mêmes dimensions que A.

2.2 Somme de deux matrices de mêmes dimensions

La somme $C = A + B$ de deux matrices de *mêmes dimensions* s'obtient en additionnant les éléments de mêmes indices.

$$c_{ij} = a_{ij} + b_{ij}$$

La différence se définit de façon évidente

$$A - B = A + (-1) \times B$$

L'addition étant associative, il est possible d'effectuer la somme de plusieurs matrices de mêmes dimensions.

2.3 Exercices

2.3.1

Soit la matrice magique $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 1 & 1 \\ 1 & 3 & 2 \end{pmatrix}$. Calculer $B = \frac{A+A'}{2}$ et $C = \frac{A-A'}{2}$. Que peut on dire de B et de C, de $B + C$? Remarque: De façon générale, toute matrice carrée A peut s'écrire de manière unique comme la somme d'une matrice symétrique et d'une matrice antisymétrique.

2.3.2

On s'intéresse à la diminution de la tension systolique et de la tension diastolique de 10 sujets soumis á un traitement antihypertenseur. Les matrices des observations avant et après traitement sont respectivement

$$A = \begin{pmatrix} 180 & 112 \\ 152 & 82 \\ 167 & 80 \\ 154 & 106 \\ 148 & 80 \\ 164 & 98 \\ 156 & 98 \\ 171 & 96 \\ 150 & 106 \\ 160 & 104 \end{pmatrix} \text{ et } B = \begin{pmatrix} 175 & 100 \\ 160 & 90 \\ 150 & 60 \\ 140 & 83 \\ 141 & 72 \\ 149 & 80 \\ 154 & 96 \\ 150 & 84 \\ 137 & 76 \\ 145 & 77 \end{pmatrix}$$

calculer la matrice des diminutions.

2.3.3

Soient les matrices: $A = \begin{pmatrix} -1 & 2 & -1 \\ 0 & 0 & 0 \\ 1 & -2 & 1 \end{pmatrix}, B = A', C = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$. Calculer la matrice $A = 5A + 2B + C$. Quelle propriété présente M ? vérifier que cette propriété est conservée pour une combinaison linéaire de votre choix.

2.4 Produit de deux matrices

2.4.1 Définition

Soient les deux matrices A de dimensions (n, p) et B de dimensions (p, q), le produit $A \times B$ est une matrice C de dimensions (n, q) dont le terme c_{ij} vaut

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ip}b_{pj}$$

$$\begin{array}{ccccc} C & = & A & \times & B \\ (\text{n, q}) & & (\text{n, p}) & & (\text{p, q}) \end{array}$$

2.4.2 Méthode pratique

Le terme c_{ij} est obtenu en faisant la somme des produits deux à deux des termes reliés par des flèches.

$$\begin{array}{cccccc}
 & & & & & \textit{colonne } j \\
 & & & & & b_{1j} \\
 & & & & \nearrow & b_{2j} \\
 & & & & \nearrow & \vdots \\
 & & & & & b_{pj} \\
 & & \nearrow & \nearrow & \nearrow & \\
\textit{ligne } i & a_{i1} & a_{i2} & \ldots & a_{ip} &
\end{array}$$

2.4.3 Exemples

1) Effectuer le produit $A \times B$ avec $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix}$, puis effectuer $A' \times B$.

2) Exprimer sous forme de produit matriciel la somme puis la moyenne des éléments du vecteur colonne $V = \begin{pmatrix} 1 \\ 0 \\ 4 \\ 6 \\ 4 \end{pmatrix}$.

3) Calculer $M = \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} \times \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$ puis $N = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \times \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$

Quelle est la forme des matrices M et N ? Généraliser à des matrices $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ et $\begin{pmatrix} a & b \\ b & a \end{pmatrix}$.

4) Calculer $C = \begin{pmatrix} 2 & 0 & 1 \\ -1 & 1 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ -1 & 0 \end{pmatrix}$

2.5 Propriétés du produit

2.5.1 Non commutativité

Le produit de deux matrices n'est pas, en général, commutatif. Si les matrices ne sont pas carrées et de même ordre, les deux produits $A \times B$ et $B \times A$ ne peuvent pas être simultanément définis. Si les matrices sont carrées et de même ordre, on a en général $A \times B \neq B \times A$. Il peut pourtant arriver que $A \times B = B \times A$. On dit alors que A et B commutent.

2.5.2 Associativité

Le produit de plusieurs matrices est associatif :

$$A_1 \times (A_2 \times A_3) = (A_1 \times A_2) \times A_3 = A_1 \times A_2 \times A_3$$

Remarque 1: Bien entendu, le produit des n matrices $A_1, A_2, \ldots, A_n$ n'est défini que si $\forall i \; i = 1, \ldots, n-1$, le nombre de colonnes de A_i est égal au nombre de lignes de A_{i+1}.

Remarque 2 : Puisque le produit est associatif, il est possible de définir la puissance kième d'une matrice carrée d'ordre p A. On la note A^k. C'est évidemment une matrice carré d'ordre p.

Définition : Une matrice carrée telle que $A^2 = A$, et donc $A^n = A$ est dite *idempotente.*

2.5.3 distributivité

Le produit est distributif par rapport à la somme :

$$A \times (B + C) = A \times B + A \times C$$

2.5.4 Multiplication par un nombre λ

On vérifie facilement que $(\lambda A) \times B = \lambda(A \times B) = A \times (\lambda B)$

2.5.5 Produit transposé

Soient deux matrices A et B de dimensions respectives (n, p) et (p, q). Leur produit C est une matrice (n, q) et l'on peut vérifier que

$$C' = B' \times A'$$

Plus généralement, si le produit des matrices $A_1, A_2, \ldots, A_n$ est défini, il vérifie

$$(A_1 \times A_2 \ldots \times A_n)' = A'_n \times A'_{n-1} \ldots \times A'_1$$

2.5.6 matrice unité

Soit A une matrice (n, p), rappelons que la matrice unité d'ordre n se note I_n. On peut vérifier que

$$I_n \times A = A \times I_p = A$$

Si $n = p$, A est carrée et $A \times I_p = I_p \times A = A$. Les matrices unités doivent leur nom au fait qu'elles jouent le rôle d'élément neutre pour le produit matriciel.

2.6 Exercices

2.6.1

Reprenons l'exemple des tensions systolique (TS) et diastolique (TD) avant et après traitement. Calculer la matrice ligne $M = (m_1, m_2)$ où m_1 est la moyenne des diminutions des TS et m_2 la moyenne des diminutions des TD en utilisant la matrice ligne (1,p) $C = \begin{pmatrix} 1 & 1 & \ldots & 1 \end{pmatrix}$. Comment effectuer le calcul pour obtenir les deux moyennes sous forme d'une matrice colonne ?

2.6.2

Soit la matrice $A = \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}$. Calculer le plus simplement possible $C = A^8$.

2.6.3

Soient les matrices

$$J = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} K = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} L = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$

Calculer J^2, K^2, L^2 et $J \times K \times L$

2.6.4

Soient les matrices $A = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \end{pmatrix}$ $B = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ $C = \begin{pmatrix} 0 \\ 5 \\ 0 \end{pmatrix}$. Calculer $A \times B$ et $A \times C$. Conclusion ?

2.7 Somme et produit de matrices partitionnées

Soient $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ et $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$. Si les blocs A_{ij} et B_{ij} ont même dimension, la somme $A + B$ s'écrit de façon évidente

$$A + B = \begin{pmatrix} A_{11} + B_{11} & A_{12} + B_{12} \\ A_{21} + B_{21} & A_{22} + B_{22} \end{pmatrix}$$

Le produit de deux matrices partitionnées est effectué selon la règle énoncée en 4, en traitant chaque bloc comme un élément. Il faut que les dimensions coincident, c'est-à-dire que les colonnes de la première matrice et les lignes de la seconde soient partitionnées de la même manière.

Exercice : Vérifier la règle précédente avec $A = \begin{pmatrix} B & C \\ D & E \end{pmatrix}$ et $N = \begin{pmatrix} P \\ Q \end{pmatrix}$, où $B = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$, $C = D = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, $E = \begin{pmatrix} 4 & 1 \\ 2 & 1 \end{pmatrix}$, $P = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$, $Q = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

2.8 Trace d'une matrice

Définition : La trace d'une matrice carrée A, notée $tr(A)$ est la somme de ses éléments diagonaux

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

2.8.1 Propriétés

$\forall \alpha$ réel $tr(\alpha A) = \alpha tr(A)$;
$\forall A$ et B de mêmes dimensions $tr(A + B) = tr(A) + tr(B)$ et $tr(A - B) = tr(A) - tr(B)$;
Si C est une matrice (n, p) et D une matrice (p, n)

$$tr(CD) = tr(DC) = \sum_{ij} c_{ij} d_{ji}$$

en particulier, pour toute matrice carrée C, $tr(C'C) = tr(CC') = \sum_{ij} c_{ij}^2$

2.8.2 Exercice

Calculer $tr(AB)$ avec $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 4 & 0 & 2 \end{pmatrix}$

2.9 Problème 1 : calcul de la matrice de covariance

Soit la bivariable (X_1, X_2) dont on possède un échantillon de taille n. Les valeurs sont disposées dans une matrice $X(n, 2)$: $X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ \vdots & \vdots \\ x_{n1} & x_{n2} \end{pmatrix}$ 1)

Trouver la matrice A qui vérifie $A \times X = \begin{pmatrix} \sum_i x_{i1}^2 & \sum_i x_{i1} x_{i2} \\ \sum_i x_{i1} x_{i2} & \sum_i x_{i2}^2 \end{pmatrix}$

2) Soit la matrice $(2, n)$ $B = \begin{pmatrix} \sum_i x_{i1} & \sum_i x_{i1} & \dots & \sum_i x_{i1} \\ \sum_i x_{i2} & \sum_i x_{i2} & \dots & \sum_i x_{i2} \end{pmatrix}$. Ecrire l'opération matricielle qui permet d'obtenir B en utilisant la matrice (n, n) $E = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix}$

3) Effectuer le produit $B \times X$
4) Conclure en donnant une formule simple permettant de calculer l'estimation de la matrice de covariance de (X_1, X_2) à partir de la matrice X des données.
5) Qu'obtient on si l'on commence par centrer et réduire les observations X ?
6) Estimer les matrices de covariance et de corrélation entre la tension systolique et la tension diastolique à partir d'un échantillon de 10 observations donné par $X' = \begin{pmatrix} 174 & 178 & 182 & 178 & 182 & 162 & 158 & 162 & 154 & 170 \\ 86 & 94 & 98 & 106 & 90 & 86 & 94 & 74 & 86 & 86 \end{pmatrix}$.

Chapitre 3

DETERMINANT D'UNE MATRICE CARREE

3.1 Définition

Soit une matrice carrée X d'ordre n. Notons C_i le vecteur colonne constitué par la ième colonne de X. On démontre qu'il existe un nombre Δ et un seul qui soit une fonction f des C_i et qui vérifie

a) f se change en $-f$ quand on transpose C_i et C_j (i.e. quand on echange leurs places).

b) S'il existe des indices i, i_1, i_2 tels que $C_i = C_{i1} + C_{i2}$, alors

$$\begin{aligned} f(C_1, C_2, \ldots, C_i, \ldots, C_n) &= f(C_1, C_2, \ldots, C_{i1} + C_{i2}, \ldots, C_n) \\ &= f(C_1, C_2, \ldots, C_{i1}, \ldots, C_n) + f(C_1, C_2, \ldots, C_{i2}, \ldots, C_n) \end{aligned}$$

c) Si $C_i = \lambda V_i$, on a

$$\begin{aligned} f(C_1, C_2, \ldots, C_i, \ldots, C_n) &= f(C_1, C_2, \ldots, \lambda V_i, \ldots, C_n) \\ &= \lambda f(C_1, C_2, \ldots, V_i, \ldots, C_n) \end{aligned}$$

Δ est appelé le *déterminant* de X. Ce qui précède reste valable si les vecteur-lignes L_i constitués par les lignes de X remplacent les vecteurs colonnes dans ces énoncés.

Notations : le déterminant s'écrit comme la matrice elle-même, mais entre des traits parallèles au lieu de parenthèses : $\begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix}$ est un déterminant.

Si X est une matrice (n, n), on dit que Δ est un déterminant d'ordre n.

3.2 Calcul pratique d'un déterminant

Nous allons donner l'expression du déterminant des matrices les plus simples, puis donner une règle permettant d'obtenir le déterminant d'une matrice carrée quelconque.

3.2.1 Déterminant d'ordre 1

C'est facile : $|a| = a$

3.2.2 Déterminant d'ordre 2

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

Il est facile de vérifier les propriétés a), b) et c) de la définition.

3.2.3 Déterminant d'ordre 3

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33}+a_{21}a_{32}a_{13}+a_{31}a_{12}a_{23}-a_{13}a_{22}a_{31}-a_{23}a_{32}a_{11}-a_{33}a_{12}a_{21}$$

Il est très facile d'obtenir l'expression du déterminant d'une matrice d'ordre 3 en recopiant ses deux premières lignes sous la dernière et en reliant les éléments situés sur une même diagonale.

$$\begin{array}{ccccc} a_{11} & & a_{12} & & a_{13} \\ & \searrow & & & \\ a_{21} & & a_{22} & & a_{23} \\ & \searrow & & \searrow & \\ a_{31} & & a_{32} & & a_{33} \\ & \searrow & & \searrow & \\ a_{11} & & a_{12} & & a_{13} \\ & & & \searrow & \\ a_{21} & & a_{22} & & a_{23} \end{array} \text{ ainsi que } \begin{array}{ccccc} a_{11} & & a_{12} & & a_{13} \\ & & & \swarrow & \\ a_{21} & & a_{22} & & a_{23} \\ & \swarrow & & \swarrow & \\ a_{31} & & a_{32} & & a_{33} \\ & \swarrow & & \swarrow & \\ a_{11} & & a_{12} & & a_{13} \\ & \swarrow & & & \\ a_{21} & & a_{22} & & a_{23} \end{array}$$

Les produits d'éléments orientés selon la diagonale principale sont affectés du signe "+" et ceux d'él'ements orientés selon la seconde diagonale du signe "-". Le déterminant est la somme de ces produits. Ce procédé de calcul est connu sous le nom de *règle de Sarrus*.

3.2.4 Déterminant d'une matrice carrée quelconque

Il est facile de vérifier qu'un déterminant d'ordre 3 peut s'exprimer sous la forme

$$\Delta = a_{11}\begin{vmatrix} a_{21} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$
$$= a_{11}\Delta_{11} - a_{12}\Delta 12 + a_{13}\Delta_{13}$$

Δ_{11}, $-\Delta_{12}$, Δ_{13} sont les *cofacteurs* respectifs de a_{11}, a_{12} et a_{13}. On dit que Δ a été développé selon la première ligne. On peut aussi écrire

$$\Delta = -a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{22}\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} - a_{32}\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix}$$
$$= -a_{12}\Delta_{12} + a_{22}\Delta_{22} - a_{32}\Delta_{32}$$

On dit dans ce cas que Δ a été développé selon la deuxième colonne.

De façon générale, le déterminant d'une matrice A d'ordre n peut s'obtenir en choisissant une ligne ou une colonne quelconque et en faisant la somme du produit de ses éléments par leur cofacteur. Le cofacteur de a_{ij} est égal au produit par $(-1)^{i+j}$ du déterminant de la matrice obtenue en supprimant la ligne i et la colonne j de A.

3.3 Propriétés du déterminant

Soit $\Delta = |X|$ le déterminant de la matrice X d'ordre n. On pourrait montrer facilement à partir de la définition 1 que
1) Δ change de signe quand on permute deux lignes ou deux colonnes de X;
2) le déterminant de λX est $|\lambda X| = \lambda^n \Delta$;
3) si une colonne est nulle, $\Delta = 0$;
4) si une ligne est nulle, $\Delta = 0$;
5) si deux colonnes, C_i et C_j, ou deux lignes, L_i et L_j sont proportionnelles, $\Delta = 0$;
$C_i = \lambda C_j \Longrightarrow \Delta = 0$
6) plus généralement, si une colonne (respectivement une ligne) peut s'exprimer comme une combinaison linéaire des autres colonnes (respectivement des autres lignes), $\Delta = 0$

$$C_i = \sum_{j \neq i} \lambda_j C_j \Longrightarrow \Delta = 0$$

7) Δ ne change pas si l'on ajoute à l'une des colonnes (respectivement à l'une des lignes) une combinaison linéaire des autres colonnes (respectivement des autres lignes).

3.4 Exemples

1) Calculer $\Delta_1 = \begin{vmatrix} 1 & 5 & 2 \\ 2 & 0 & 4 \\ 3 & 7 & 6 \end{vmatrix}$ puis $\Delta_2 = \begin{vmatrix} 3 & 2 & 1 \\ 4 & 3 & 2 \\ 5 & 4 & 3 \end{vmatrix}$;

2) Calculer $\Delta_1 = \begin{vmatrix} 1 & a \\ 1 & b \end{vmatrix}$ puis $\Delta_2 = \begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix}$ (déterminants de Vandermonde)

3.5 Opérations sur les déterminants

1) Soient A et B deux matrices de même ordre, on a

$$|A \times B| = |A| \times |B|$$

2) Une conséquence immédiate est que

$$|A \times B| = |B \times A| = |A| \times |B| = |B| \times |A|$$

3) Le produit de plusieurs déterminants est associatif

$$|A \times B \times C| = |A \times B| \times |C| = |A| \times |B \times C| = |A| \times |B| \times |C|$$

4) La matrice A et sa transposée A' ont même déterminant

$$|A'| = |A|$$

3.6 Définition

Une matrice carrée d'ordre n est dite *régulière* si $|X| \neq 0$.

X régulière $\iff$ aucun vecteur colonne ne peut s'exprimer comme une combinaison linéaire des autres colonnes
$\iff$ aucun vecteur ligne ne peut s'exprimer comme une combinaison linéaire des autres lignes.

On dit aussi que la matrice X est de *rang* n .

3.7 Exemples

1) Les matrices suivantes sont elles régulières ? (calculer leur déterminant) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & -1 \\ 0 & 0 & -2 \end{pmatrix}$. Ce résultat se généraliseà toute matrice diagonale ou triangulaire d'ordre n quelconque, dont les éléments diagonaux sont non nuls.
2) Monter, sans calcul, que le déterminant des matrices suivantes est nul: $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} -2 & -10 & 1 \\ 0 & 0 & 2 \\ 1 & 5 & 3 \end{pmatrix}$.

3.8 déterminant d'une matrice partitionnée

Vérifier que $\begin{vmatrix} 2 & 1 & 3 & 0 \\ 1 & 1 & 2 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 3 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} \times \begin{vmatrix} 1 & 3 \\ 3 & 1 \end{vmatrix}$. De façon générale, on peut montrer que si A, B et C sont des blocs carrés, $\begin{vmatrix} A & C \\ 0 & B \end{vmatrix} = |A| \times |B|$. De façon encore plus générale, on a $\begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} = |A_{11}| \times \left|A_{22} - A_{21}A_{11}^{-1}A_{12}\right| = |A_{22}| \times \left|A_{11} - A_{12}A_{22}^{-1}A_{21}\right|$

Chapitre 4

INVERSE D'UNE MATRICE CARREE REGULIERE

4.1 Définition

Soit une matrice carrée A d'ordre n. Son *inverse* est une matrice de mêmes dimensions, notée A^{-1} et définie par

$$A \times A^{-1} = A^{-1} \times A = I_n$$

où I_n désigne la matrice identité d'ordre n.

4.2 Conditions d'existence

L'inverse A^{-1} n'est définie que si A est régulière: $|A| \neq 0$.

4.3 Propriétés

1) Soient p matrices carrées régulières d'ordre n $A_1, A_2, \ldots, A_p$, alors, en raison de l'associativité du produit:

$$(A_1 \times A_2 \times \ldots A_p)^{-1} = A_p^{-1} \times A_{p-1}^{-1} \times \ldots \times A_1^{-1}$$

2) L'inverse de la transposée de A est la transposée de l'inverse de A:

$$(A')^{-1} = \left(A^{-1}\right)'$$

En effet, $A \times A^{-1} = I \Longrightarrow (A \times A^{-1})' = I' = I = (A^{-1})' \times A'$
Par conséquent, l'inverse d'une matrice symétrique est symétrique:

$$A = A' \Longrightarrow A^{-1} = (A')^{-1} = \left(A^{-1}\right)'$$

3) Le déterminant de l'inverse est l'inverse du déterminant:

$$\left|A^{-1}\right| = \frac{1}{|A|}$$

car $A \times A^{-1} = I \Longrightarrow |A| \times |A^{-1}| = 1$

4.4 Calcul de la matrice inverse

Une manière d'obtenir l'inverse d'une matrice A est d'appliquer la formule

$$A^{-1} = \frac{1}{|A|} \times C'$$

où C est la matrice obtenue en remplaçant chaque élément de A par son cofacteur. On remarque sur cette formule que A^{-1} ne peut exister que si $|A| \neq 0$.

4.5 Exemples

1) Calculer A_1^{-1} et A_2^{-1} avec $A_1 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ et $A_2 = \begin{pmatrix} a_1 & 0 & 0 & \dots & 0 \\ 0 & a_2 & 0 & \dots & 0 \\ 0 & 0 & a_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_n \end{pmatrix}$.

2) (suite de l'exemple 4 du 2.4.3) Calculer les inverses de $A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ et $B = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$

3) Plus généralement, montrer que l'inverse de $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est $A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

4) Montrer que l'inverse de la matrice magique $M = \begin{pmatrix} -1 & 3 & 1 \\ 3 & 1 & -1 \\ 1 & -1 & 3 \end{pmatrix}$ est la matrice magique $N = \frac{1}{36}\begin{pmatrix} -2 & 10 & 4 \\ 10 & 4 & -2 \\ 4 & -2 & 10 \end{pmatrix}$ En pratique, il existe de nombreux programmes d'inversion de matrices et vous n'aurez plus jamais besoin d'effectuer vous même ce calcul.

4.6 Application: résolution d'un système linéaire de n équations à n inconnues (système de Cramer)

Soit à résoudre le système

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ &\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= b_n \end{aligned}$$

où les x_i sont les inconnues. Ce système s'écrit sous forme matricielle

$$A \times X = B$$

avec $A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$, $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ et $B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$. Alors, si $|A| \neq 0$, la solution est immédiate. Elle est donnée par

$$X = A^{-1} \times B$$

Exemple : utiliser l'exemple 4) du paragraphe précédent pour résoudre le système

$$\begin{aligned} -x_1 + 3x_2 + x_3 &= 1 \\ 3x_1 + x_2 - x_3 &= 2 \\ x_1 - x_2 + 3x_3 &= 3 \end{aligned}$$

4.7 Matrices orthogonales

Définition : Une matrice A, régulière, est dite *orthogonale* si $A^{-1} = A'$. Propriétés

1) $|A| = \pm 1$
car $|A \times A^{-1}| = |I| = 1 = |A| \times |A'| = |A|^2$
2) Soient V_i et V_j deux vecteurs colonnes (ou lignes) de A.
Alors, puisque $A \times A' = I$
$i \neq j \Longrightarrow V_i' \times V_j = 0$
$i = j \Longrightarrow V_i' \times V_j = 1$

Exemple : Vérifier que $M = \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}$ est orthogonale.

4.8 Inverse généralisée

Soit $X = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$, calculer $M = X'X$. La matrice M est elle régulière ? Soit $M^- = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, vérifier que l'on a $MM^-M = M$.

Ce résultat est général. Pour toute matrice M, il existe une inverse généralisée, notée M^- qui vérifie $MM^-M = M$. Si M est régulière, M^- n'est autre que l'invers M^{-1}, sinon, l'inverse généralisée n'est pas unique.

Application : Vérifier que $\hat{\theta} = (A'A)^-A'Y$ est solution de l'équation $A'Y = A'A\theta$ et que $\hat{\theta} + [(A'A)^-(A'A) - I]V$ où V est un vecteur quelconque de dimension n, est également solution.

Mode de calcul. Soit M une matrice carrée, régulière, d'ordre n et de rang q. Après réarrangement éventuel des lignes et des colonnes, on peut écrire

$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}$, où M_{11} est régulière, d'ordre q maximal. On montre qu'alors $M_{22} = M_{21}M_{11}^{-1}M_{12}$. En effet, puisque les colonnes formées par les blocs M_{12} et M_{22} sont fonctions linéaires des colonnes formées par M_{11} et M_{21}, on a

$$M_{12} = M_{11}A \Rightarrow A = M_{11}^{-1}M_{12}$$

$$M_{22} = M_{21}A \Rightarrow M_{22} = M_{21}M_{11}^{-1}M_{12}$$

où A est la matrice des coefficients de ces combinaisons linéaires.

Vérifions que $M^- = \begin{pmatrix} M_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix}$.

$$\begin{aligned} MM^- &= \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} M_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix} &&= \begin{pmatrix} I & 0 \\ M_{21}M_{11}^{-1} & 0 \end{pmatrix} \\ MM^-M &= \begin{pmatrix} I & 0 \\ M_{21}M_{11}^{-1} & 0 \end{pmatrix} \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} &&= \begin{pmatrix} M_{11} & M_{12} \\ M_{21}M_{11}^{-1}M_{11} & M_{21}M_{11}^{-1}M_{12} \end{pmatrix} \\ &= \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} &&= M \end{aligned}$$

4.9 Exercices

Soient les matrices $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$, $X = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $Y = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$

1) Calculer AX et AY.
2) A est elle régulière ?
3) Soit B une matrice régulière. Que peut on dire de deux matrices U et V telles que $BU = BV$?

4.10 Problème: régression et estimateur des moindres carrés

Soient une variable explicative X, supposée fixée, et une variable aléatoire à expliquer Y (soit X est effectivement sous le contrôle de l'expérimentateur, soit on raisonne conditionnellement aux valeurs observées de X). *Régresser* Y sur X revient à chercher a et b tels que

$$Y_i = a + bX_i + \epsilon_i$$

où X_i et Y_i sont les valeurs correspondant au sujet i et ϵ_i une variable aléatoire, mesurant la variabilité individuelle de ce sujet, supposée d'espérance nulle et de variance identique pour tous les sujets.

Ce système d'équations peut s'écrire sous forme matricielle

$$\begin{aligned}
\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} &= \begin{pmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \\ 1 & X_n \end{pmatrix} \times \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix} \\
Y &= X \times \theta + \epsilon \qquad (4.1) \\
(n,1) &\ \ (n,2)(2,1)(n,1)
\end{aligned}$$

Une des façons d'obtenir des estimateurs â et $\hat{\text{b}}$ de a et b consiste à utiliser la méthode des moindres carrés. On cherche â et $\hat{\text{b}}$ qui minimisent

$$S = \sum_{i=1}^{n} [y_i - E(Y_i)]^2 = \sum_{i=1}^{n} [y_i - a - bx_i]^2$$

Ce minimum correspond à un point où les dérivées s'annulent:

$$\begin{aligned}
\frac{\partial S}{\partial a} &= -2\sum_1^n (y_i - a - bx_i) &= 0 \\
\frac{\partial S}{\partial b} &= -2\sum_1^n x_i(y_i - a - bx_i) &= 0
\end{aligned}$$

Il s'en suit que â et $\hat{\text{b}}$ doivent être solutions de

$$\begin{aligned}
na + b\sum_1^n x_i &= \sum_1^n y_i \\
a\sum_1^n x_i + b\sum_1^n x_i^2 &= \sum_1^n x_i y_i
\end{aligned}$$

1) Calculer $X'X$ et $X'Y$. En déduire l'expression matricielle simple donnant les estimateurs cherchés.

2) Application : reprendre les données du problème 1 et calculer l'équation de régression linéaire de la tension systolique en fonction de la tension diastolique.

Remarque 1) Ce procédé de calcul matriciel se généralise à la régression sur un nombre quelconque de variables.

Remarque 2) L'équation (4.1) et l'hypothèse sur l'espérance des ϵ_i conduisent à $E(Y) = X\theta$, équation d'où vient le nom de modèle linéaire, car

$E(Y)$ est une fonction linéaire des paramètres inconnus. Le modèle linéaire recouvre à la fois la régression linéaire et l'analyse de variance. Les estimateurs des moindres carrés de ses paramètres sont toujours donnés par la même équation matricielle.

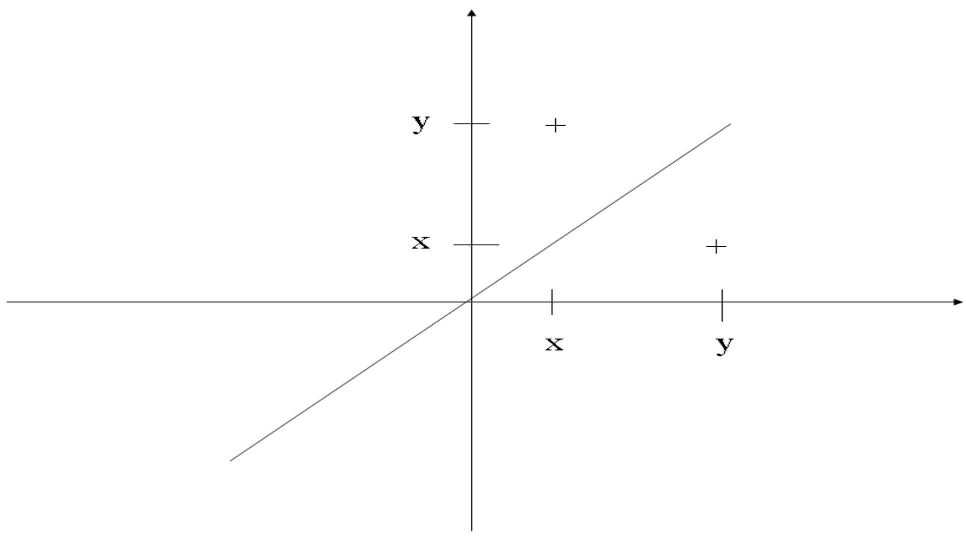
Chapitre 5

INTERPRETATION GEOMETRIQUE

5.1 Un exemple d'espace vectoriel: $\Re^n$

5.1.1 Le plan

Soit l'espace à deux dimensions $\Re^2$, muni d'un système d'axes orthonormés. Chaque point de $\Re^2$ est caractérisé par ses deux coordonnées, x_1 et x_2, qui sont les deux composantes du vecteur X joignant l'origine à ce point. Sous forme matricielle: $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.



L'ensemble des points de $\Re^2$ forment un ensemble de vecteurs sur lesquels on définit une *addition*:

$$X + Y = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix}$$

(identique à la somme vectorielle au sens géométrique) et une *multiplication par un scalaire*:

$$\lambda X = \lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \end{pmatrix}$$

Les propriés habituelles de l'addition et de la multiplication par un nombre confèrent à $\Re^2$ une structure d'*espace vectoriel*.

Remarques: En appelant e_1 et e_2 les vecteurs unités sur chacun des deux axes, $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ on a pour X quelconque de $\Re^2$

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 e_1 + x_2 e_2$$

D'autre part il est impossible de trouver deux valeurs λ_1 et λ_2 non nulles telles que $\lambda_1 e_1 + \lambda_2 e_2 = 0$, ou encore il n'existe aucune valeur λ telle que $e_1 = \lambda e_2$

5.1.2 Généralisation à $\Re^n$

Le système de vecteurs $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $\ldots$, $e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$ est tel qu'il n'existe aucune possibilité d'exprimer l'un de ceux-ci comme une combinaison linéaire des autres:

$$\lambda_1 e_1 + \lambda_2 e_2 + \ldots + \lambda_n e_n = 0 \Rightarrow \lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$$

On dit que les e_i sont des vecteurs linéairement indépendants.

D'autre part, tout vecteur de $\Re^n$ peut s'écrire

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n$$

On dit que les e_i forment une base de $\Re^n$; $x_1, x_2, \ldots, x_n$ sont les coordonnées de X dans la base $(e_1, e_2, \ldots, e_n)$, appelée base canonique.

De façon générale, tout système de n vecteurs linéairement indépendants constitue une base de $\Re^n$ et tout vecteur X de $\Re^n$ peut se mettre sous la forme d'une combinaison linéaire, et d'une seule, de vecteurs constituant une base donnée.

5.2 Sous-espace vectoriel de $\Re^n$

Soient k vecteurs linéairement indépendants. L'ensemble des combinaisons linéaires de ces k vecteurs forme un *sous-espace vectoriel* de dimension k.

Par exemple, si $(e_1, e_2, \ldots, e_n)$ est la base canonique de $\Re^n$, l'ensemble des vecteurs $X = \lambda_1 e_1 + \lambda_2 e_2$ forment un sous-espace vectoriel de $\Re^n$ de dimension 2 ; on dit aussi que ce sous-espace vectoriel est *engendré* par e_1 et e_2.

5.3 Applications linéaires et matrices

Soit X (p, 1), un vecteur colonne de $\Re^p$ et A une matrice (n, p). Le produit AX est un vecteur colonne de $\Re^n$; chacune de ses coordonnées dans $\Re^n$ est une combinaison linéaire des coordonnées de X. Pour cette raison, on appelle *application linéaire* la transformation qui à X fait correspondre $Y = AX$.

Propriétés: Posons $f(X) = AX$. L'application linéaire f vérifie les propriétés suivantes:

$$f(X_1 + X_2) = f(X_1) + f(X_2) \Longleftrightarrow A(X_1 + X_2) = AX_1 + AX_2$$

$$f(\lambda X) = \lambda f(X) \Longleftrightarrow A(\lambda X) = \lambda AX$$

La matrice A permet ainsi de définir une application de $\Re^p$ dans $\Re n$. Si $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $\ldots$, $e_p = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$ est la base canonique de $\Re^p$, on voit que
Ae_1 = premier vecteur colonne de A
Ae_2 = second vecteur colonne de A
$\ldots$

Ae_p = pième vecteur colonne de A
Les colonnes de A apparaissent ainsi comme les transformées des vecteurs de base de $\Re^p$ par l'application linéaire A.

5.4 Exercices

1) Soit l'application f de $\Re^2$ dans $\Re^2$ qui à tout point du plan, de coordonnées $\begin{pmatrix} x \\ y \end{pmatrix}$ fait correspondre son symétrique par rapport à la première bissectrice, de coordonnées $\begin{pmatrix} y \\ x \end{pmatrix}$.

Quelle est la matrice A associé à f ?

Quelle est l'application inverse f^{-1} ?

Quelle est sa matrice associée ?

2) Soit g l'application de $\Re^n$ dans $\Re^n$ qui à tout point $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ fait correspondre son symétrique par rapport à l'origine. Quelle est la matrice B associée à g ? Quelle est l'application inverse g^{-1} ? Quelle est sa matrice associée ?

3) Soit h l'application de $\Re^2$ dans $\Re^2$ définie par la matrice $C = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$. Représenter dans le plan les points $X_1 = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$, $X_2 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$, $X_3 = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$, et leurs transformées $Y_i = CX_i$. Que remarque-t-on ?

5.5 Rang d'une matrice et de l'application associée

Le rang de A est le nombre maximum de vecteurs colonnes de A linéairement indépendants. On démontre que le rang de A est aussi égal au nombre de vecteurs lignes de A indépendants. Autrement dit : $\text{rang}(A) = \text{rang}(A')$.

5.6 Noyau d'une application linéaire

Le noyau d'une application linéaire f de matrice A (n, p) est l'ensemble des vecteurs X de $\Re^p$ tels que

$$f(X) = AX = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = 0$$

Le noyau de f est un sous-espace vectoriel de $\Re^p$ qui contient le vecteur nul de $\Re^p$.

Si la matrice A (n, n) est régulière, son rang est n et on voit que le noyau de l'application linéaire associée f est réduit au vecteur nul de $\Re^p$. En effet, soit $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum x_i e_i$ un élément du noyau ; on a

$$AX = 0 \Longleftrightarrow \sum x_i A e_i = \sum x_i c_i = 0$$

où $Ae_i = c_i$ est le ième vecteur colonne de A. Or, puisque A est régulière, $\sum x_i c_i$ ne peut être nulle que si tous les x_i sont nuls. On a donc bien $X = 0$.

Remarque: une matrice régulière est une matrice inversible. On pourra vérifier que dans les exercices 5.4, les matrices A et B sont régulières et que les matrices associées aux applications inverses f^{-1} et g^{-1} sont leurs inverses A^{-1} et B^{-1}. En revanche, la matrice associée à l'application h n'est pas régulière. C'est ce qui explique la disposition des points Y_i.

5.7 Produit scalaire

Dans $\Re^2$ muni de la base orthonormée, soient deux vecteurs $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, le *produit scalaire* de X et Y est défini par

$$xy \cos \alpha$$

où x est la longueur de X, y, la longueur de Y et α l'angle formé par les deux vecteurs. On sait que le produit scalaire est aussi égal à

$$x_1 y_1 + x_2 y_2 = X'Y$$

5.7.1 Définitions

1) le *produit scalaire* de deux vecteurs $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ de $\Re^n$ est égal au produit matriciel $X'Y$;

2) deux vecteurs X et Y tels que $X'Y = 0$ sont dits *orthogonaux* ;

3) un vecteur X tel que $X'X = 1$ est dit *normé* ;

Exemple: $X = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}$ est normé et orthogonal à $Y = \begin{pmatrix} \frac{-\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}$ qui est aussi normé.

5.8 Changement de base dans un espace vectoriel

Dans $\Re 2$ muni de la base (e_1, e_2), un vecteur X a comme coordonnées x_1 et x_2 : $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ Une rotation des axes de $\pi/4$ définit une nouvelle base orthonormée dans laquelle les coordonnées de X deviennent

$$Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \frac{x_1+x_2}{\sqrt{2}} \\ \frac{-x_1+x_2}{\sqrt{2}} \end{pmatrix}$$

En effet $x_1 = OX\cos\alpha$, $x_2 = OX\sin\alpha$, $y_1 = OX\cos(\alpha - \frac{\pi}{4}) = \frac{\sqrt{2}}{2}(x_1 + x_2)$ $y_2 = OX\sin(\alpha - \frac{\pi}{4}) = \frac{\sqrt{2}}{2}(-x_1 + x_2)$

et $\sin(\alpha + \frac{\pi}{4} + \cos(\alpha + \frac{\pi}{4}) = \sqrt{2}\cos\alpha = \frac{x_1+x_2}{x_1^2+x_2^2} \ldots$

Plus généralement, soit un vecteur dont les coordonnées dans une base orthonormée sont données par $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ Considérons une autre base, non nécessairement orthonormée ; le même vecteur aura pour coordonnées dans cette base $Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ et on démontre que

$$X = PY$$

où P *matrice de passage* de l'ancienne base à la nouvelle est telle que sa ième colonne est constituée des coordonnées du ième nouveau vecteur de base dans l'ancienne.

Remarque: P est une matrice carrée dont les colonnes sont indépendantes, puisque constituées par les coordonnées des vecteurs de base, P est donc de rang n (régulière) ; il s'en suit que

$$Y = P^{-1}X$$

Exemple: Reprenant l'exemple du début de ce paragraphe, la matrice de passage de la base (e_1, e_2) à la base (v_1, v_2) vaut $P = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$ Soit $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ dans la base (e_1, e_2); dans (v_1, v_2) ce vecteur devient

$$Y = P^{-1}X = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{x_1+x_2}{\sqrt{2}} \\ \frac{-x_1+x_2}{\sqrt{2}} \end{pmatrix}$$

ce qui est bien conforme aux résultats trouvés.

Remarque: La matrice de passage de la base (v_1, v_2) à la base (e_1, e_2) est $\begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} = P^{-1}$

5.9 Propriétés des matrices de passage

La matrice de passage d'une base orthonormée à une autre base orthonormée est une matrice P orthogonale.

Exemple: la matrice P de l'exemple précédent est orthogonale : les bases (e_1, e_2) et (v_1, v_2) sont orthonormales.

Chapitre 6

VALEURS PROPRES ET VECTEURS PROPRES D'UNE MATRICE CARREE

6.1 Définitions

Soit une matrice carrée d'ordre n A, On appelle *vecteur propre* de A tout vecteur $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \neq 0$ tel qu'il existe un nombre λ vérifiant :

$$AX = \lambda X \tag{6.1}$$

On appelle *valeur propre* de A tout nombre λ tel qu'il existe un vecteur non nul X vérifiant (1).

6.2 Solutions

L'équation (1) est équivalente à $AX - \lambda X = 0$ ou encore

$$\begin{aligned} (A - \lambda I)X &= B(\lambda)X \\ &= 0 \end{aligned} \tag{6.2}$$

avec $A - \lambda I = B(\lambda)$. C'est l'écriture matricielle d'un système linéaire de n équations à n inconnues, x_1, x_2 , ..., x_n.

6.2.1 Retour sur les systèmes linéaires

Soit à résoudre

$$\begin{aligned} ax + by &= 0 \\ cx + dy &= 0 \end{aligned} \tag{6.3}$$

où x et y sont les inconnues. Ce système s'écrit sous forme matricielle $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ ou $BX = 0$ avec $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, et $X = \begin{pmatrix} x \\ y \end{pmatrix}$

La première équation donne $y = -\frac{ax}{b}$ que l'on reporte dans la seconde pour obtenir $x\frac{bc-ad}{b} = 0$.

Deux cas sont à distinguer:

1) $|B| = ad - bc \neq 0 \Longrightarrow \begin{pmatrix} x \\ y \end{pmatrix} = 0$

2) $|B| = ad - bc = 0 \Longrightarrow y = -\frac{ax}{b} = -\frac{cx}{d}$

La solution n'est pas unique, elle est constituée par l'ensemble des vecteurs $X = \begin{pmatrix} x \\ -\frac{a}{b}x \end{pmatrix}$ qui forment un sous espace vectoriel de dimension 1 dans R^2.

Plus généralement, si B est une matrice carrée d'ordre n et $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, un vecteur de R^n, le système de n équations à n inconnues x_1, x_2 ,..., x_n

$$BX = 0$$

n'admet de solutions différentes de $X = 0$ que si

$$|B| = 0$$

Un tel systéme d'équations est dit *homogène*.

6.2.2 Existence de valeurs et de vecteurs propres

Pour que le système (6.3) admette des solutions autres que $X = 0$, il faut d'après ce qui vient dêtre vu que

$$|A - \lambda I| = 0 \tag{6.4}$$

Les nombres λ vérifiant (6.4) sont les valeurs propres de A. A chaque valeur propre correspond un ensemble de vecteurs propres X satisfaisant (6.1), ensemble qui constitue un sous-ensemble vectoriel de R^n. On voit que (6.4) est un polynome de degré n en λ qui admet un certain nombre de racines, distinctes ou confondues.

6.2.3 Exemple 1

Soit la matrice diagonale, d'ordre n, $A = \begin{pmatrix} a_1 & 0 & 0 & \dots & 0 \\ 0 & a_2 & 0 & \dots & 0 \\ 0 & 0 & a_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_n \end{pmatrix}$ avec $\forall i \neq j$, $a_i \neq a_j$. L'équation caractéristique de A est

$$(a_1 - \lambda)(a_2 - \lambda) \dots (a_n - \lambda) = 0$$

La matrice A admet donc comme valeurs propres ses éléments diagonaux. Les vecteurs propres de A associés à la valeur propre a_i sont donnés par

$$AX = a_i X \Longleftrightarrow \begin{pmatrix} a_1 x_1 \\ a_2 x_2 \\ \vdots \\ a_i x_i \\ \vdots \\ a_n x_n \end{pmatrix} = \begin{pmatrix} a_i x_1 \\ a_i x_2 \\ \vdots \\ a_i x_i \\ \vdots \\ a_i x_n \end{pmatrix}$$

On voit que x_i peut être quelconque, mais il faut $x_1 = x_2 = \dots = x_{i-1} = x_{i+1} = \dots = x_n = 0$, soit

$$X = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ x_i \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

L'ensemble des vecteurs propres associés à la valeur propre a_i est donc un sous espace vectoriel de dimension 1 engendré par

$$e_i = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

il existe ainsi n vecteurs propres de A qui constituent la *base canonique* de R^n

6.2.4 Exemple 2

Reprenons la matrice de l'exemple précédent avec $a_1 = a_2 = a$. A admet maintenant $n-1$ valeurs propres distinctes. Les vecteurs propres associés à la valeur propre a sont donnés par

$$AX = aX \Longrightarrow \begin{pmatrix} ax_1 \\ ax_2 \\ a_3x_3 \\ \vdots \\ a_nx_n \end{pmatrix} = \begin{pmatrix} ax_1 \\ ax_2 \\ ax_3 \\ \vdots \\ ax_n \end{pmatrix} \Longrightarrow x_3 = x_4 = \ldots = x_n = 0$$

et $X = \begin{pmatrix} x_1 \\ x_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ Ainsi l'ensemble des vecteurs propres associés à la valeur propre a est un sous espace vectoriel de dimension 2 engendré par $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ et $e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$

De façon générale, à une valeur propre donnée peut correspondre un sous espace vectoriel de vecteurs propres de dimension quelconque, inférieure à n.

6.3 Exercice

Trouver les valeurs propres de $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, puis les vecteurs propres associés.

Chapitre 7

DIAGONALISATION D'UNE MATRICE CARREE

7.1 Définition

Une matrice carrée $A(n,n)$ est dite diagonalisable s'il existe une matrice $P(n,n)$, régulière et une matrice diagonale $D(n,n)$ telles que

$$D = P^{-1}AP$$

7.2 Propriétés

1) Les éléments composant la matrice D sont les valeurs propres de A, et les colonnes de la matrice P sont les vecteurs propres de A. En effet, posons $D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$ et $P = (V_1 V_2 \dots V_n)$ où V_i, ième colonne de P est un vecteur (n, 1) et où certains λ_i peuvent être égaux ou nuls

$$D = P^{-1}AP \Rightarrow PD = AP$$

On voit que la ième colonne de la matrice AP est donnée par AV_i, et la ième colonne de la matrice PD par $\lambda_i V_i$. Donc

$$\forall i \ AV_i = \lambda_i V_i$$

ce qui signifie que V_i est le vecteur propre associé à la valeur propre λ_i.

On voit que A diagonalisable $\Rightarrow$ P inversible $\Rightarrow$ les n vecteurs propres forment une base de $\Re^n$. En outre, P s'interprète comme la matrice de passage de la base de départ à la base formée par les vecteurs propres de A.

7.3 Exercice

Sachant que $A = \begin{pmatrix} 3/2 & 0 & 1/2 \\ 0 & 2 & 0 \\ 1/2 & 0 & 3/2 \end{pmatrix}$ se diagonalise en $D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, trouver une base orthonormée de vecteurs propres de A.

7.4 Diagonalisation d'une matrice symétrique

On démontre qu'une matrice symétrique est toujours diagonalisable et qu'il existe une base orthonormée formée de ses vecteurs propres.

7.5 Diagonalisation d'une matrice inverse

Soit A une matrice (n, n) régulière. De

$$D = P^{-1}AP \Longrightarrow D^{-1} = P^{-1}A^{-1}P$$

on déduit que

1) A diagonalisable $\Longrightarrow$ A^{-1} diagonalisable

2) les valeurs propres de A^{-1} sont les inverses de celles de A

3) les vecteurs propres de A^{-1} sont les vecteurs propres de A

7.6 Puissance nième d'une matrice carrée

$$D = P^{-1}AP \Longrightarrow D^n = P^{-1}A^nP$$

On en déduit

1) les valeurs propres de A^n sont les puissances nièmes de celles de A;

2) les vecteurs propres de A^n sont les vecteurs propres de A;

3) $A^n = PD^nP^{-1}$, ce qui peut fournir un procédé commode de calcul de A^n

7.7 Rang d'une matrice diagonalisable

$$D = P^{-1}AP \Longrightarrow |D| = \left|P^{-1}\right| |A| \, |P| = |A|$$

On en déduit que

rang(A) = n $\Longleftrightarrow$ rang(D) = n $\Longleftrightarrow$ 0 n'est pas valeur propre de A.

De façon générale, on démontre que le rang de A, diagonalisable, est égal au nombre de ses valeurs propres son nulles (en tenant compte de leur ordre de multiplicité)

7.8 Trace d'une matrice carrée diagonalisable

Soit A une matrice matrice carrée (n, n), diagonalisable en $D = P^{-1}AP$. Alors, $tr(A) = tr(D) = \sum$ (valeurs propres de A).

Comme nous l'avons vu en 2.8.1, $tr(AB) = tr(BA)$; on peut donc écrire

$$tr(A) = tr(PDP^{-1}) = tr(DP^{-1}P) = tr(D)$$

7.9 Résumé

- D diagonalisable $\Longleftrightarrow$ $D = P^{-1}AP$,avec D =diag(λ_i), λ_i ième valeur propre de A (certaines valeurs propres peuvent être multiples ou nulles).

- A diagonalisable régulière (= inversible = de rang n = de déterminant non nul) $\Longleftrightarrow$ toutes les valeurs propres sont différentes de 0.

- A symétrique $\Longrightarrow$ A est diagonalisable et il existe une base de vecteurs propres orthonormés.

Chapitre 8

FORMES QUADRATIQUES, APPLICATIONS

8.1 Définition

Soit une matrice A (n, n), symétrique, et X un vecteur colonne de $\Re^n$. Le produit $X'AX$, qui est un scalaire est appelé *forme quadratique* associée à A.

Exemples. Ecrire les formes quadratiques associées à $A = \begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}$ et à $B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

8.2 Matrices (semi) définies positives

Définition: Soit A (n, n), symétrique. Si la forme quadratique $X'AX$ associée à A est positive quel que soit X, A est dite *définie positive* (dp); si elle est positive ou nulle quel que soit X, A est dite semi-définie positive (sdp).

Propriété 1: Une matrice A sdp a ses éléments diagonaux positifs ou nuls.

Ce résultat se démontre en prenant successivement pour X les vecteurs de la base canonique $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$

Propriété 2: [A (n, n) dp et B (n, q),de rang $q \leq$ n] $\Rightarrow B'AB$ dp.

Démonstration: Soit Y (q, 1) avec $Y \neq 0$ et $X = BY$, avec X (n, 1) et $X \neq 0$ puisque B est de rang q. On a:

$$Y'(B'AB)Y = X'AX > 0$$

puisque A est dp.

Remarque 1: si B est de rang inférieur à q, $B'AB$ est seulement sdp.

Remarque 2: si A est sdp et B de rang inférieur ou égal q, $B'AB$ est sdp.

8.3 Exemples

1) Soient les matrices (n, n) $I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$ et $E_n = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix}$
Montrer que $I_n - \frac{1}{n}E_n$ est sdp.

2) En déduire en s'aidant du problème 1) (calcul de la matrice de covariance) qu'une matrice de covariance estimée, calculée à partir d'un tableau de données est sdp.

8.4 Application 1: équation d'une ellipse

Ellipse canonique: une ellipse est dite en position *canonique* si son grand axe et son petit axe coincident avec les axes du repère orthonormé. Dans ce repère orthonormé, son équation est alors

$$x_1^2\lambda_1 + x_2^2\lambda_2 = a \tag{8.1}$$

où λ_1 et λ_2 sont des nombres positifs donnés. Notons que (8.1) peut se mettre sous la forme

$$X'DX = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = a \tag{8.2}$$

Ellipse quelconque

Soient $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ un vecteur (2, 1), Σ une matrice (2, 2), symétrique, sdp et a un nombre strictement positif. Alors

$$X'\Sigma X = a \tag{8.3}$$

est l'équation d'une ellipse. Cette ellipse n'est en position canonique que si Σ est diagonale, comme le montre l'équation (8.2). Or il apparait graphiquement que dans le système d'axes orthonormés formés par le grand et le petit axe de l'ellipse, celle-ci est évidemment en position canonique. Dans ce système (Oy_1, Oy_2), son équation sera du type

$$Y'DY = a \tag{8.4}$$

où D est diagonale et $Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$.

Pour trouver les vecteurs définissant les directions Oy_1 et Oy_2, il faut ainsi passer de (8.3) à (8.4), c'est-à-dire diagonaliser Σ. Plus précisément, puisque Σ est symétrique, elle est toujours diagonalisable en une matrice diagonale D à l'aide d'une matrice de passage orthogonale P

$$D = P^{-1}\Sigma P = P'\Sigma P$$

P est la matrice de passage de la base orthonormée de départ à la base orthonormée des vecteurs propres de Σ (qui existe puisque P est symétrique). Un vecteur dont les coordonnées sont $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ dans la base de départ aura pour coordonnées $Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ dans la base des vecteurs propres selon:

$$X = PY \Leftrightarrow X' = Y'P'$$

L'équation (8.3) peut donc s'écrire

$$X'\Sigma X = Y'P'\Sigma PY = Y'DY = a$$

Ainsi

1) le grand et le petit axe de l'ellipse sont portés par les vecteurs propres de Σ;

2) les longueurs des axes de l'ellipse sont inversement proportionnelles aux valeurs propres λ_1 et λ_2 de Σ.

Généralisation à $\Re^n$: équation d'un ellipsoïde

Soit Σ (n, n), symétrique, sdp et X un vecteur de R^n, $X'\Sigma X = a > 0$ est l'équation d'un *ellipsoïde* de R^n (généralisation d'une ellipse). Il est encore possible de diagonaliser Σ et les calculs ci-dessus restent valables.

8.5 Application 2: lois normales

1) Densité d'un couple de variables normales indépendantes

Soient X et Y deux gaussiennes indépendantes:

$$X \sim N(\mu_1, \sigma_1^2),\ Y \sim N(\mu_2, \sigma_2^2)$$

En raison de l'indépendance, la densité de probabilité du couple (X, Y) est une fonction de deux variables qui est le produit des densités de probabilité de X et de Y. Elle s'écrit

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2} exp[-\frac{1}{2}(x-\mu_1\ y-\mu_2)\Sigma^{-1}\begin{pmatrix} x-\mu_1 \\ y-\mu_2 \end{pmatrix}$$

où $\Sigma = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}$ est la matrice de covariance du couple (X, Y).

2) loi multinormale

Soient p variables aléatoires gaussiennes $X_1, X_2, \ldots, X_i, \ldots, X_p)$ non indépendantes, de matrice de covariance Σ telles que

$$X_i \sim N(\mu_i, \sigma_i^2)$$

On montre que la densité du p-uple $(X_1, \ldots, X_i, \ldots, X_p)$ est une fonction de p variables qui a pour écriture matricielle

$$f(x_1, \ldots, x_i, \ldots, x_p) = \frac{1}{(2\pi)^{p/2}\,|\Sigma|^{1/2}} exp[\frac{-1}{2}(X-\mu)'\Sigma^{-1}(X-\mu)]$$

avec $(X-\mu) = \begin{pmatrix} X_1-\mu_1 \\ \vdots \\ X_i-\mu_i \\ \vdots \\ X_p-\mu_p \end{pmatrix}$.

3) Courbes d'isodensité

Une variable normale

Soit $f(x) = \frac{1}{\sqrt{2\pi}} exp[-\frac{(x-\mu)^2}{2\sigma^2}]$ la densité de probabilité d'une variable normale X. Les *courbes d'isodensité* sont formées de l'ensemble des points racines d'une équation de la forme $f(x) = c$ (deux points au plus, symétriques par rapport à μ). L'équation $f(x) = 0.058$ admet deux racines $x_1 = \mu - 1.96\sigma$ et $x_2 = \mu + 1.96\sigma$ telles que

$$P[\mu - 1.96\sigma < X < \mu + 1.96\sigma] = 0.95$$

Plus généralement
$\forall \alpha \in [0, 1], \exists c_\alpha$ tel que l'équation d'isodensité $f(x) = c_\alpha$ admet deux racines x_1 et x_2 vérifiant $P[x_1 < X < x_2] = 1 - \alpha$

couple de variables binormales

La densité de probabilité du couple (X, Y) est une fonction de deux variables à valeurs dans $\Re$. C'est l'équation d'une surface en forme de cloche $z = f(x, y)$. Les courbes d'isodensité, constituées des points solutions de $f(x, y) = c$ correspondent à l'intersection de cette surface avec les plans d'équation $z = c$. Pour tout $\alpha \in [0, 1]$, on peut calculer la valeur c_α telle que $P[(X, Y) \in U_\alpha] = 1 - \alpha$ où U_α désigne l'espace intérieur délimité par la courbe d'isodensité $f(x, y) = c_\alpha$ Il existe ainsi une famille de courbes d'isodensité dont chacune correspond à une valeur et une seule de α, donc de c_α. L'équation des courbes d'isodensité est donnée par

$$(X - \mu)'\Sigma^{-1}(X - \mu) = k_\alpha$$

avec $(X - \mu) = \begin{pmatrix} X - \mu_1 \\ X - \mu_2 \end{pmatrix}$ et $k_\alpha = -2log(2\pi\, |\Sigma|^{1/2}\, c_\alpha)$. Cette équation est celle d'une ellipse dont les axes coincident avec les directions des vecteurs propres de Σ^{-1}, qui sont les mêmes que les vecteurs propres de Σ.

Ainsi, dans le plan (Ox, Oy), les courbes d'isodensité sont des ellipses telles que, pour k_α donné: $P[(X,Y) \in U_\alpha] = 1 - \alpha$. L'ellipse n'est en position canonique que si X et Y sont indépendantes (Σ diagonale).

variables multinormales

Ce qui précède se généralise à $\Re^p$ où les surfaces d'isodensité de p variables normales sont des ellipsoïdes qui délimitent un domaine où les p variables ont une probabilité $1 - \alpha$ de se réaliser.

Chapitre 9

DERIVEES MATRICIELLES

9.1 Définition

Soit le vecteur $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et une fonction à valeurs dans $\Re$, $f(X) = f(x_1, \ldots, x_n)$. Le vecteur (n, 1)

$$\frac{\partial f(X)}{\partial X} = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

est appelé *dérivée matricielle* de f(X) par rapport à X.

9.2 Exemples

1) Prendre la fonction $A'X = \sum a_i x_i$ où A est un vecteur (n, 1). On a

$$\frac{\partial A'X}{\partial X} = A$$

car $\frac{\partial \sum a_i x_i}{\partial x_i} = a_i$

2) Considérer la fonction $X'X = \sum x_i^2$, alors

$$\frac{\partial X'X}{\partial X} = 2X$$

car $\frac{\partial \sum x_i^2}{\partial x_i} = 2x_i$.

3) Vérifier également que si A est une matrice carrée (n, n)

$$\frac{\partial X'AX}{\partial X} = 2AX$$

et que pour A matrice (n, p) et Y (p, 1)

$$\frac{\partial X'AY}{\partial X} = AY$$

9.3 Application: encore les moindres carrés

On connait Y (n, 1) et A (n, p) et on veut estimer θ qui minimise

$f(\theta) = (Y-A\theta)'(Y-A\theta) = Y'Y-Y'A\theta-\theta'A'Y+\theta'A'A\theta = Y'Y-2\theta'A'Y+\theta'A'A\theta$

La solution doit annuler la dérivé, or $\frac{\partial Y'Y}{\partial\theta} = 0$, $\frac{\partial\theta'A'Y}{\partial\theta} = A'Y$ et $\frac{\partial\theta'A'A\theta}{\partial\theta} = 2A'A\theta$ Il vient

$$\frac{\partial(Y-A\theta)'(Y-A\theta)}{\partial\theta} = -2A'Y + 2A'A\theta = 0$$

d'où $A'A\theta = A'Y$ et si $A'A$ est régulière, la solution est donnée par $\hat{\theta} = (A'A)^{-1}A'Y$.

Chapitre 10

MATRICES DE COVARIANCES

10.1 Matrice de covariance exacte

10.1.1 Définition

Soit le p-uple de variables aléatoires $X' = (X_1, X_2, \ldots, X_p)$. On pose:

$E(X)$ est le vecteur (p, 1) dont les éléments sont les espérances des X_i

$E(XX')$ est la matrice (p, p) dont les éléments sont les espérances des éléments de la matrice XX'

Alors, la matrice de covariance du vecteur X est la matrice (p, p), symétrique, semi-définie positive Σ définie par

$$\Sigma = E(XX') - E(X)E(X)' = \begin{pmatrix} Var(X_1) & cov(X_1X_2) & \ldots & cov(X_1X_p) \\ cov(X_2X_1) & var(X_2) & \ldots & cov(X_2X_p) \\ \vdots & \vdots & \ddots & \vdots \\ cov(X_pX_1) & cov(X_pX_2) & \ldots & var(X_p) \end{pmatrix}$$

10.1.2 Changement de variables linéaire

Soit A une matrice(q, p) de coefficients donnés et soit $Y' = (Y_1, Y_2, \ldots, Y_q)$ le n-uple de variables aléatoires défini par

$$Y = AX$$

Chaque Y_i est une combinaison linéaire des X_j dont les coefficients constituent la ième ligne de A et on a les résultats suivants

$$E(Y) = AE(X) \tag{10.1}$$

$$\Gamma = A\Sigma A' \tag{10.2}$$

où Γ désigne la matrice de covariance de Y de dimensions (q, q). (1) résulte de la linéarité de l'espérance. (2) peut se démontrer comme suit:

$$\begin{aligned}
\Gamma &= E(YY') - E(Y)E(Y)' \\
&= E(AXX'A') - AE(X)E(X)'A' \\
&= AE(XX')A' - AE(X)E(X)'A' \\
&= A[E(XX') - E(X)E(X)']A' \\
&= A\Sigma A'
\end{aligned}$$

10.1.3 Cas particulier: A vecteur ligne

Si A est de dimensions (1, p), $Y = AX$ est une variable aléatoire, combinaison linéaire des X_i dont la variance est donnée par

$$var(Y) = A\Sigma A'$$

10.2 Matrice de covariance estimée

10.2.1 Définition

Soit le p-uple aléatoire $(X_1, X_2, \ldots, X_p)$ dont on possède un échantillon de taille n et soit X le tableau (n, p) des observations

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

où la ième colonne est un échantillon de taille n de la variable aléatoire X_i. Par définition, l'estimation de la matrice de covariance Σ du p-uple aléatoire $(X_1, X_2, \ldots, X_p)$ est la matrice S dont les éléments sont les estimations des éléments de Σ. On a (cf problème 1)

$$(n-1)S = X'X - \frac{1}{n}X'J_nX$$

où J_n est la matrice (n, n) composée uniquement de '1'.

10.2.2 Changement de variables linéaire

Soit A une matrice (q, p) de coefficients connus et soit $Y' = AX'$ ou $Y = XA'$

$$Y' = \begin{pmatrix} y_{11} & y_{21} & \dots & y_{n1} \\ y_{12} & y_{22} & \dots & y_{n2} \\ \vdots & \vdots & \dots & \vdots \\ y_{1q} & y_{2q} & \dots & y_{nq} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & \dots & \vdots \\ a_{q1} & a_{q2} & \dots & a_{qp} \end{pmatrix} \begin{pmatrix} x_{11} & x_{21} & \dots & x_{n1} \\ x_{12} & x_{22} & \dots & x_{n2} \\ \vdots & \vdots & \dots & \vdots \\ x_{1p} & x_{2p} & \dots & x_{np} \end{pmatrix}$$

autrement dit la ième colonne de Y (=la ième ligne de Y') est un échantillon de taille n de la variable aléatoire $Y_i = a_{i1}x_1 + a_{i2}x_2 + \dots + a_{ip}x_p$

On a alors les résultats suivants:

1) $\begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \vdots \\ \bar{y}_q \end{pmatrix} = A \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_p \end{pmatrix}$ où $\bar{y}_j = \frac{1}{n}\sum_{k=1}^{n} y_{kj}$ et $\bar{x}_i = \frac{1}{n}\sum_{k=1}^{n} x_{ki}$

2) $W = ASA'$ où W est la matrice de covariance estimée du q-uple $Y_1, \dots, Y_q$.

En effet, d'après 2.1,

$$\begin{aligned} (n-1)W &= Y'Y - \frac{1}{n}Y'J_nY = AX'XA' - \frac{1}{n}AX'J_nXA' \\ &= A(X'X - \frac{1}{n}X'J_nX)a' \\ &= (n-1)ASA' \end{aligned}$$

10.2.3 Cas particulier : A vecteur ligne

Si $A = (a_1a_2 \dots a_p)$, alors $Y' = AX'$ est un vecteur ligne, échantillon de taille n de la variable aléatoire $Y = a_1X_1 + a_2X_2 + \dots + a_pX_p$, et en notant s^2 l'estimation de la variance de Y, on a

$$s^2 = ASA'$$