INTRODUCTION A L'ANALYSE DES DONNÉES

par

Claude DENIAU et Ludovic LEBART

Les pages qui suivent constituent un premier exposé élémentaire portant sur « l'analyse des données », et devraient être accessibles au plus grand nombre.

S'il est nécessaire d'avoir une certaine expérience statistique pour apprécier pleinement l'intérêt des méthodes employées, le langage mathématique d'un « bon bachelier » devrait, théoriquement, suffire puisque dans une première partie figurent les rappels nécessaires d'algèbre linéaire.

Les diverses méthodes d'analyse exposées ici, ont toutes le même but : réduire de façon synthétique et optimale une masse de données statistiques trop encombrante pour être exploitée telle quelle.

Les idées à la base de ces réductions sont anciennes : l'Analyse Factorielle des psychologues se proposait déjà de décrire de grands ensembles d'information en les condensant sous la forme d'un petit nombre de facteurs. C'est pourquoi l'on parlera indifféremment dans cet article d'Analyse Factorielle ou d'Analyse des données.

Les résultats présentés sont à mettre à l'actif de la Statistique descriptive, car nous ne poserons (sauf mention contraire) aucun modèle probabiliste « a priori ».

L'essentiel des idées exposées dans ce texte ont reçu un certain bain de jouvence (ou sont issues directement) du Laboratoire de M. le professeur J. P. Benzecri à la Faculté des Sciences de Paris.

Malgré les simplifications, les lourdeurs (en particulier sur le plan des notations) que nous imposait notre travail de vulgarisation, nous espérons n'avoir pas trop déformé la pensée des auteurs. Nous pensons que ce travail pourra inciter certains utilisateurs à approfondir les idées présentées ici.

SOMMAIRE

PREMIÈRE PARTIE

RAPPELS D'ALGÈBRE LINÉAIRE

1. ESPACES VECTORIELS

Définition (1,1)

Un ensemble E pour lequel on suppose données une application de $E \times E \to E((x, y) \to x + y)$ et une application de $\mathbf{R} \times E \to E((\lambda, x) \to \lambda x)$ s'appelle un espace vectoriel réel [1] si :

1. Pour tout couple (x, y) d'éléments de E, on a : $x + y = y + x$.
2. Pour tout triplet (x, y, z) d'éléments de E on a :

$$x + (y + z) = (x + y) + z.$$

3. Il existe $0 \in E$ tel que pour tout élément $x \in E$:

$$x + 0 = 0 + x = x.$$

4. Pour tout $x_1 \in E$, il existe $x_2 \in E$ tel que $x_1 + x_2 = 0$.
5. Pour tout couple (α, β) de nombres réels et tout $x \in E$ on a :

$$(\alpha + \beta)x = \alpha x + \beta x.$$

6. Pour tout $\alpha \in \mathbf{R}$ et tout couple (x, y) d'éléments de E on a :

$$\alpha(x + y) = \alpha x + \alpha y.$$

7. Pour tout couple (α, β) d'éléments de $\mathbf{R}$ et tout $x \in E$ on a :

$$\alpha(\beta x) = (\alpha\beta)x.$$

8. Pour tout $x \in E$, on a $1x = x$.

Propriétés (1,1) : (simples à démontrer)

— Pour tout $x, y, z, \in E$: $x + y = y + z$ entraîne $x = z$.
— Pour tout scalaire α et tout élément $x \in E$, on a $0x = \alpha 0 = 0$.
— La relation $\lambda y = 0$ est équivalente à $y = 0$ ou $\lambda = 0$.
— Pour tout $x \in E$, l'élément $x' = (-1)x$ est **l'unique** élément de E tel que $x + x' = 0$.

(1) Quand on écrira espace vectoriel cela sous-entendra « réel ».

Exemples (1,1) : **R** est un espace vectoriel pour les applications :

$$(\alpha, \beta) \to \alpha + \beta \qquad \text{où} \quad \alpha, \beta, \lambda, \mu \in \mathbf{R}$$
$$(\lambda, \mu) \to \lambda\mu$$

— soit $\mathbf{R}^n = \{ z \,;\, z = (x_1, ..., x_n), x_i \in \mathbf{R}\ 1 \leqslant i \leqslant n \}$

Si on se donne sur $\mathbf{R}^n$ les deux opérations suivantes :

$$(x_1, ..., x_n) + (x'_1, ..., x'_n) = (x_1 + x'_1, x_2 + x'_2, ..., x_n + x'_n)$$
$$\alpha(x_1, ..., x_n) = (\alpha x_1, ..., \alpha x_n)$$

$\mathbf{R}^n$ est un espace vectoriel, **produit des** n **espaces R.**

Proposition (1,1) :

Si $E_1, ..., E_n$ sont n espaces vectoriels et si on se donne sur $E_1 x ... x E_n$ les opérations définies ci-dessus, $E_1 x ... x E_n$ est un espace vectoriel.

Définition (1,2)

On dit qu'une partie non vide F de E est un sous-espace vectoriel si pour tout couple (x, y) d'éléments de F et tout couple (λ, μ) d'éléments de **R**, on a :

$$\lambda x + \mu y \in F$$

Proposition (1,2) : (très simple à démontrer)

Si E_1 et E_2 sont deux sous-espaces vectoriels de E il en est de même de $E_1 \cap E_2$ (la proposition s'étend à une famille quelconque de sous-espaces vectoriels).

Si E_1 et E_2 sont deux sous-espaces vectoriels de E

$S = \{ x \in E \,;\, x = x_1 + x_2\ x_1 \in E_1, x_2 \in E_2 \}$ est un sous-espace vectoriel de E.

On le note $E_1 + E_2$ et on l'appelle espace vectoriel **somme** de E_1 et E_2.

Définition (1,3)

Soient deux sous-espaces vectoriels E_1 et E_2 d'un espace E. Les propriétés suivantes sont équivalentes :

i) $E_1 \cap E_2 = \{ 0 \}$

ii) Tout élément de $E_1 + E_2$ s'écrit d'une manière et d'une seule sous la forme $x_1 + x_2$ où $x_1 \in E$ et $x_2 \in E_2$.

On dit alors que la somme $E_1 + E_2$ est **directe** et on la note $E_1 \oplus E_2$ (on dit aussi que E_1 et E_2 sont supplémentaires dans E).

Variété linéaire affine

Soit E un espace vectoriel réel et $a \in E$ et soit l'application :

$$t_a : E \to E$$
$$t_a : x \to a + x.$$

Définition (1,4)

Soient E_1 un sous-espace de E et $a \in E$. L'image de E_1 par t_a (notée $a + E_1$) s'appelle une **variété linéaire affine.**

2. APPLICATIONS LINÉAIRES ET BILINÉAIRES

Définition (2,1)

Soient E et F deux espaces vectoriels. On dit qu'une application f de E dans F est linéaire si :

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$$

quels que soient

$$x, y \in E, \lambda, \mu \in \mathbf{R}$$

(Montrer que $f(0) = 0$)

Exemple (2,1)

Soient E un espace vectoriel et $a \in \mathbf{R}$. On définit une application h_a de E dans E par :

$$h_a : x \rightarrow ax$$

C'est une application linéaire.

Définition (2,2)

Soient E et F deux espaces vectoriels et f une application linéaire de E dans F.

i) On appelle image de f et on note f(E) l'ensemble des $y \in F$ tels qu'il existe $x \in E$ vérifiant $y = f(x)$.

ii) On appelle noyau de f et on note Ker(f) l'ensemble des $x \in E$ tel que $f(x) = 0$.

Propriétés (2,2)

a) f(E) est un sous-espace vectoriel de F et Ker (f) un sous-espace vectoriel de E (facile à démontrer).

b) Une application linéaire bijective est appelée **isomorphisme** (si f est un isomorphisme, f^{-1} est aussi un isomorphisme).

Proposition (2,2)

Soient trois espaces vectoriels E, F, G, f une application linéaire de E dans F et g une application linéaire de F dans G.

Alors $h = g o f$ est une application linéaire de E dans G.

3. COMBINAISONS LINÉAIRES, BASES, DIMENSION

Définition (3,1)

Soient $x_1, ..., x_n$ des éléments d'un espace vectoriel E ; on appelle **combinaison linéaire** de $x_1, ..., x_n$, tout élément $x \in E$ possédant la propriété suivante :

Il existe des scalaires $\lambda_1, ..., \lambda_n \in \mathbf{R}$ tels que

$$x = \lambda_1 x_1 + ... + \lambda_n x_n$$

On notera $x = \sum_{1 \leqslant i \leqslant n} \lambda_i x_i$

Exemple (3,1) soit $a \in \mathbf{R}^n$ $\quad a = (\lambda_1, ..., \lambda_n)$ $\lambda_i \in \mathbf{R}$ $1 \leqslant i \leqslant n$

considérons

$$\begin{aligned} e_1 &= (1, 0 ..., 0) \\ &\cdots\cdots\cdots \\ e_i &= (0, ..., 1, 0 ... 0) \\ &\cdots\cdots\cdots \\ e_n &= (0, ..., 0, 1) \end{aligned} \qquad a = \sum_{1 \leqslant i \leqslant n} \lambda_i e_i$$

tout élément de $\mathbf{R}^n$ s'écrit comme combinaison linéaire de $e_1, ... e_n$. (On peut montrer que l'écriture est unique en remarquant que

$$(\xi_1, ..., \xi_n) = (\lambda_1, ..., \lambda_n)$$

si et seulement si $\xi_i = \lambda_i$ $1 \leqslant i \leqslant n$.)

Définition (3, 2)

Soit E un espace vectoriel et $(x_i)_{1 \leqslant i \leqslant n}$ une famille d'éléments de E.

i) On dit que la famille $(x_i)_{1 \leqslant i \leqslant n}$ est **libre** si et seulement si :

$$\sum \lambda_i x_i = 0 \quad \text{entraîne} \quad \lambda_i = 0 \qquad 1 \leqslant i \leqslant n$$

ii) On dit que la famille $(x_i)_{1 \leqslant i \leqslant n}$ **engendre** E (ou est une **famille génératrice** de E) si tout élément $x \in E$ s'écrit comme combinaison linéaire d'éléments de la famille $(x_i)_{1 \leqslant i \leqslant n}$.

iii) On dit que la famille $(x_i)_{1 \leqslant i \leqslant n}$ est une **base** de E si tout élément $x \in E$ s'écrit de manière unique comme combinaison linéaire de la famille $(x_i)_{1 \leqslant i \leqslant n}$.

Remarque (3, 2)

On montrera à titre d'exercice que (i/, ii/) est équivalent à (iii/). On admettra que « tout espace vectoriel possède une base » et que si un espace vectoriel possède deux bases, elles ont même « nombre » d'éléments.

Exemple (3, 2)

Montrer que la famille $(e_i)_{1 \leqslant i \leqslant n}$ d'éléments de $\mathbf{R}^n$, définie dans l'exemple (3, 1 page 62) est une base de $\mathbf{R}^n$.

Définition (3, 3)

« Le nombre » d'éléments d'une base d'un espace vectoriel est appelé **dimension** de E. On notera dim(E). (Nous ne nous intéresserons qu'à des espaces vectoriels de dimension finie.)

Définition (3, 4)

Soient E et F deux espaces vectoriels et f une application linéaire de E dans F. On appelle **rang de** f la dimension de f(E).

4. ESPACE $\mathscr{L}(E, F)$

Proposition (4, 1)

Soient E et F deux espaces vectoriels. L'ensemble des applications linéaires de E dans F pour lequel on suppose données les opérations :

(f, g) $f + g$ [définie par $(f + g)(x) = f(x) + g(x)$ pour tout $x \in E$]

(λ, f) λf [définie par $(\lambda f)(x) = \lambda f(x)$ pour tout $x \in E$]

est un espace vectoriel. On le notera $\mathscr{L}(E, F)$. (Démonstration triviale).

Définition (4, 1)

L'ensemble des applications linéaires de E dans **R** est appelé l'espace dual de E et se note $E^* = \mathscr{L}(E, \mathbf{R})$.

(Un élément de E* est appelé : **forme linéaire** sur E).

Proposition (4, 2)

Il existe une application linéaire unique f d'un espace vectoriel E dans un espace vectoriel F, telle que :

$$\text{quel que soit } 1 \leqslant i \leqslant n \qquad f(a_i) = b_i \qquad (1)$$

où $(b_i)_{1 \leqslant i \leqslant n}$ est une famille quelconque d'éléments de F et $(a_i)_{1 \leqslant i \leqslant n}$ une base de E.

Preuve :

- **Unicité :** supposons que f existe vérifiant (1) $x \in E$ s'écrit de manière unique :

$$x = \sum_{1 \leqslant i \leqslant n} \alpha_i a_i$$

$$(a_i)_{1 \leqslant i \leqslant n} \text{ base de E, donc } f(x) = \sum \alpha_i f(a_i) = \sum \alpha_i b_i$$

donc si f existe elle est unique.

- **Existence :** l'application $x \to \sum_{1 \leqslant i \leqslant n} \alpha_i b_i$ de E dans F est linéaire. (Donc f est déterminée par ses valeurs sur les éléments de la base).

Proposition (4, 3)

Soit E un espace vectoriel de dimension finie, alors $\dim(E) = \dim(E^*)$.

Preuve :

Soit $(e_i)_{1 \leqslant i \leqslant n}$ une base de E ; considérons la famille $(e_i^*)_{1 \leqslant i \leqslant n}$ d'éléments de E* tels que :

soient $x^* \in E^*$ et $x \in E$ $\qquad x = \sum_{1 \leqslant i \leqslant n} \lambda_i e_i$

$$x^*(x) = \sum \lambda_i x^*(e_i) \tag{1}$$

Posons $x^*(e_i) = \lambda_i^* \qquad 1 \leqslant i \leqslant n$.
D'après la proposition (4, 2), x^* est parfaitement déterminé.
Définissons une famille $(e^*)_{1 \leqslant i \leqslant n}$ d'éléments de E^* par

$$e_j^*(x) = \lambda_j$$

ou ce qui revient au même, par $e_j^*(e_i) = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$

(1) s'écrit alors $x^*(x) = \sum_{1 \leqslant i \leqslant n} \lambda_i^* e_i^*(x)$

d'où $x^* = \sum_{1 \leqslant i \leqslant n} \lambda_i^* e_i^* \qquad$ quel que soit $\qquad x^* \in E^*$

donc $(e_i^*)_{1 \leqslant i \leqslant n}$ engendre E^*.
Montrons que la famille $(e_i^*)_{1 \leqslant i \leqslant n}$ est libre.

$$\sum_{1 \leqslant i \leqslant n} \alpha_i e_i^* = 0 \Rightarrow \text{quel que soit } x \in E \sum_{1 \leqslant i \leqslant n} \alpha_i e_i^*(x) = 0$$

en particulier pour $x = e_j$

$$\sum \alpha_i e_i^*(e_j) = \alpha_i = 0 \quad \text{quel que soit} \quad 1 \leqslant i \leqslant n$$

ce qui achève la démonstration.

5. MATRICES

Soit f une application de E, dont une base est $(a_i)_{1 \leqslant i \leqslant m}$, dans F, dont une base est $(b_j)_{1 \leqslant j \leqslant n}$

$$f : E \rightarrow F$$

$x \in E$ s'écrit $x = \sum_{1 \leqslant i \leqslant m} \lambda_i a_i$

$$f(x) = \sum_{1 \leqslant i \leqslant m} \lambda_i f(a_i) \tag{1}$$

mais $f(x) \in F$, donc $f(x) = \sum_{1 \leqslant j \leqslant n} \mu_j b_j \qquad (1')$

Or, quel que soit i $1 \leqslant i \leqslant n$ les $f(a_i) \in F$ donc :

$$\left.\begin{aligned} f(a_1) &= b_1\alpha_{11} + \dots + b_n\alpha_{1n} \\ &\ \vdots \\ f(a_m) &= b_1\alpha_{m1} + \dots + b_n\alpha_{mn} \end{aligned}\right\} \tag{2}$$

Remplaçons dans (1) les $f(a_i)$ $1 \leqslant i \leqslant m$ par leurs expressions (2). Alors $f(x) = (b_1\alpha_{11} + ... + b_n\alpha_{1n})\lambda_1 + ... + (b_1\alpha_{m1} + ... + b_n\alpha_{mn})\lambda_m$ (3)

Écrivons que les seconds membres de (1′) et (3) sont égaux et identifiés. On en déduit :

$$\left.\begin{array}{l} \mu_1 = \alpha_{11}\lambda_1 + ... + \alpha_{m1}\lambda_m \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \mu_n = \alpha_{1n}\lambda_1 + ... + \alpha_{mn}\lambda_m \end{array}\right\} \qquad (4)$$

L'expression (4) détermine d'une manière et d'une seule l'application linéaire f par rapport aux bases $(a_i)_{1\leqslant i\leqslant m}$ et $(b_j)_{1\leqslant j\leqslant n}$ de E et F. Pour connaître f il suffit donc de connaître le tableau (5)

$$A = \begin{pmatrix} \alpha_{11} & \cdots\cdots\cdots & \alpha_{m1} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \alpha_{1n} & \cdots\cdots\cdots & \alpha_{mn} \end{pmatrix} \qquad (5)$$

Un tel tableau s'appelle une **matrice** à m colonnes et n lignes. La $i^{\text{ème}}$ colonne de la matrice A est formée avec les composantes de $f(a_i)$ dans (2).

Écriture matricielle d'une application linéaire.

La relation (4) du paragraphe précédent :

$$\begin{array}{l} \mu_1 = \alpha_{11}\lambda_1 + ... + \alpha_{m1}\lambda_m \\ \vdots \\ \mu_n = \alpha_{1n}\lambda_1 + ... + \alpha_{mn}\lambda_m \end{array}$$

où f(x) a pour composantes $(\mu_1, ..., \mu_n)$

où x a pour composantes $(\lambda_1, ..., \lambda_m)$

s'écrit :

$$\begin{pmatrix} \mu_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \mu_n \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \cdots\cdots\cdots & \alpha_{m1} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \alpha_{1n} & \cdots\cdots\cdots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \lambda_m \end{pmatrix} \qquad (6)$$

Remarque (5, 1)

A est la matrice de f **associée** aux bases $(a_i)_{1\leqslant i\leqslant m}$ de E et $(b_j)_{1\leqslant j\leqslant n}$ de F ; elle dépend du choix de ces bases.

Notations :

— Les coordonnées par rapport à la base $(a_i)_{1\leqslant i\leqslant m}$ d'un élément $x \in E$

$$x = \sum_{1\leqslant i\leqslant m} \lambda_i a_i$$

se disposent en une matrice du type « **colonne** » notée X (ou M(x)) avec

$$X = \begin{pmatrix} \lambda_1 \\ \cdot \\ \cdot \\ \cdot \\ \lambda_m \end{pmatrix}$$

— Une forme linéaire f est déterminée par les $(f(a_i))_{1 \leq i \leq m}$ où $(a_i)_{1 \leq i \leq m}$ est une base de E. Les $f(a_i) = \alpha_i$ sont disposés en une matrice du type « **ligne** » $M(f) = (\alpha_1 \dots \alpha_n)$.

— Enfin, posons

$$Y = \begin{pmatrix} \mu_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \mu_n \end{pmatrix}$$

l'expression (6) s'écrit

$$\boxed{Y = AX}$$

Transposée d'une application linéaire

Soient E et F deux espaces vectoriels et $f : E \rightarrow F$ une application linéaire. Soit u une forme linéaire sur F,($u \in F^*$) ; alors, l'application composée $u o f$ est une forme linéaire sur E.

On peut définir une application linéaire ${}^t f : F^* \rightarrow E^*$ en posant :

$${}^t f(u) = u o f$$

quel que soit $u \in F^*$.

Définition

L'application ${}^t f$ s'appelle la **transposée** de l'application linéaire f.

Proposition (à vérifier)

i) Soient E, F deux espaces vectoriels et f, g deux applications linéaires de E dans F, alors ${}^t(f + g) = {}^t f + {}^t g$.

ii) Soit G un autre espace vectoriel et $f : E \rightarrow F$, $g : F \rightarrow G$ deux applications linéaires :

$${}^t(f o g) = {}^t g o {}^t f$$

Quelques éléments de calcul matriciel

Une matrice A à m colonnes et n lignes est appelée matrice de type (m, n) et notée :

$$A = (a_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \qquad \text{ou} \qquad A(m, n)$$

(Si m = n la matrice est **carrée**).

Définition (5, 1)

On appelle transposée de la matrice $A = (a_{ij})$ la matrice $B = (b_{ij})$ définie par $b_{ij} = a_{ji}$; on la notera $B = {}^t A$.

Exemple :

$$ {}^t(x_1, \ldots, x_n) = \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix} $$

(le vérifier à partir de la définition de la transposée d'une application linéaire).

Définition (5, 2)

i) On dit qu'une matrice A est symétrique si ${}^tA = A$.

ii) On dit qu'une matrice D est diagonale si $a_{ij} = 0$ pour $i \neq j$.

iii) On dit qu'une matrice A est antisymétrique si ${}^tA = -A$

$$\left(\text{où } -A = (-a_{ij})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}}\right)$$

(On vérifiera que ces matrices sont nécessairement carrées).

Addition des matrices (5, 3)

Soient E et F deux espaces vectoriels réels, et f, g deux applications linéaires de E dans F. Choisissons $(a_i)_{1 \leqslant i \leqslant m}$ et $(b_j)_{1 \leqslant j \leqslant n}$ pour bases respectives de E et F

$A = (\alpha_{ij})$ matrice de f

$B = (\beta_{ij})$ matrice de g

$f(a_i) = b_1\alpha_{i1} + \ldots + b_n\alpha_{in}$

Posons $h = f + g$

$g(a_i) = b_1\beta_{i1} + \ldots + b_n\beta_{in}$

Les éléments de la ième colonne de la matrice associée à h seront les composantes de $h(a_i)$, donc si on appelle $C = (\gamma_{ij})$ la matrice associée à h :

$$\gamma_{ij} = \alpha_{ij} + \beta_{ij} \qquad 1 \leqslant i, j \leqslant n$$

et on écrit $C = A + B$.

Définition (5, 3)

Soient E, F et G trois espaces vectoriels ayant pour base respective $(a_i)_{1 \leqslant i \leqslant m}$, $(b_j)_{1 \leqslant j \leqslant n}$, $(c_k)_{1 \leqslant k \leqslant p}$ et deux applications linéaires f et g

$$g : E \to F \qquad f : F \to G$$

soit $h = f \circ g$.

Désignons par $A = (\alpha_{jk})_{\substack{1 \leqslant j \leqslant n \\ 1 \leqslant k \leqslant p}}$ $B = (\beta_{ij})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}}$ et $C = (\gamma_{ik})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant p}}$

les matrices associées à f, g et h. Le lecteur vérifiera que :

$$\gamma_{ik} = \sum_{1 \leqslant j \leqslant n} \beta_{ij}\alpha_{jk}$$

On dit alors que C est la **matrice produit** de A et B et on note $C = AB$.

Remarque (5, 4)

Le produit n'est défini que si le nombre de colonnes de A est égal au nombre de lignes de B (le vérifier).

Propriétés (5,5)

Soient A, B, C des matrices ; lorsque ces opérations sont définies :

$$\begin{aligned}
&(AB)C = A(BC) = ABC \\
&A(B + C) = AB + AC \\
&A(\lambda B) \quad = \lambda AB \quad \lambda \in \mathbf{R} \\
&AB \neq BA \\
&{}^t(A + B) = {}^tA + {}^tB \\
&{}^t(AB) \quad = {}^tB\,{}^tA \\
&{}^t({}^tA) = A.
\end{aligned}$$

Définition (5,4)

Soit A une matrice carrée, on appelle **trace** de $A = (\alpha_{ij})_{1 \leqslant i,j \leqslant m}$ et on note tr (A) la somme des éléments de la forme α_{ii} (éléments diagonaux de A)

$$\mathrm{tr}\, A = \sum_{1 \leqslant i \leqslant m} a_{ii}$$

Proposition (5,6)

$$\begin{aligned}
\mathrm{tr}(A + B) &= \mathrm{tr}(A) + \mathrm{tr}(B) \\
\mathrm{tr}({}^tA) &= \mathrm{tr}(A) \\
\mathrm{tr}(\lambda A) &= \lambda\, \mathrm{tr}(A) \\
\mathrm{tr}(AB) &= \mathrm{tr}(BA) = \sum_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}} a_{ij} b_{ji}
\end{aligned}$$

(la démonstration est laissée aux soins du lecteur)

La matrice carrée (m, m) (matrice carrée d'ordre m) dont tous les éléments diagonaux sont égaux à 1 et dont tous les autres sont nuls est notée

$$I_m = \begin{pmatrix} 1 & & & & 0 \\ & \cdot & & & \\ & & \cdot & & \\ & & & \cdot & \\ 0 & & & & 1 \end{pmatrix}$$

On a $I_m A = A I_m = A$ et $\operatorname{tr}(I_m) = m$

Définition (5,5)

On dit qu'une matrice carrée $A(m, m)$ est **inversible** si et seulement si il existe $B(m, m)$ telle que

$$AB = BA = I_m$$

B est alors appelée matrice inverse de A et notée A^{-1}.

Définition (5,6)

On appelle **rang d'une matrice** le rang de l'application linéaire qu'elle représente.

6. FORMES BILINÉAIRES, FORMES QUADRATIQUES

Définition (6,1)

Soient E, F deux espaces vectoriels et f une application de E dans F ; f est **bilinéaire** si elle est linéaire par rapport à chacune des variables [c'est-à-dire, si $x \to f(x, y)$ et $y \to f(x, y)$ sont linéaires].

Matrice associée à une forme bilinéaire

$$(a_i) \text{ base de E} \qquad x = \sum_{1 \leqslant i \leqslant m} a_i x_i$$

$$(b_j) \text{ base de F} \qquad y = \sum_{1 \leqslant j \leqslant n} b_j x_j$$

$$f(x, y) = \sum_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}} f(a_i b_j) x_i y_j$$

[On laisse le soin au lecteur de démontrer que les $f(a_i b_j) = \alpha_{ij}$ définissent complètement f.]

$$\text{On écrira : } f(x, y) = (y_1, \ldots, y_n) \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{m1} \\ \vdots & & \vdots \\ \alpha_{1n} & \cdots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$$

soit en écriture matricielle $F = {}^tYAX$.

Mais ${}^tYAX$ est un scalaire, donc ${}^t({}^tYAX) = {}^tYAX$

donc $$\boxed{{}^tX{}^tAY = {}^tYAX}$$

si le rang de A **est égal à** r, f **est dite de rang** r.

Définition (6,2)

On dit qu'une forme bilinéaire f, définie sur $E \times E$ est **symétrique** si quel que soit $x, y \in E$ $f(x, y) = f(y, x)$.

Remarque (6,2)

f symétrique entraîne : quel que soit $1 \leqslant i, j \leqslant n$:

$$\alpha_{ij} = f(a_i, a_j) = f(a_j, a_i) = \alpha_{ji}$$

donc la matrice associée à f est **symétrique.**

Définition (6,3)

f étant une forme bilinéaire symétrique on appelle forme **quadratique** q associée, l'application de E dans **R** définie par :

$$q(x) = f(x, x) \qquad \text{quel que soit} \qquad x \in E$$

(f est appelée la **forme polaire** de q)

Écriture matricielle : $Q(X) = {}^tXAX$ où A est symétrique et Q la matrice associée à q.

Action d'un changement de base

1) **Sur un vecteur** ; soit E un espace vectoriel de dimension n, $B = (a_i)_{1 \leqslant i \leqslant n}$ et $B' = (a'_i)_{1 \leqslant i \leqslant n}$ deux bases de E.

Considérons la matrice $S = (s_{ij})$ ayant pour colonne i les coordonnées de a'_i sur B :

$$a'_i = \sum_{1 \leqslant i \leqslant n} s_{ij} a_j \qquad (1)$$

S est inversible. (En effet, on sait qu'il existe $s \in \mathscr{L}(E, E)$ unique tel que $a'_i = s(a_i)$ (propriétés 4,2) et $s(a_i)$ est une famille libre, s est de rang n, donc s est inversible).

Soit $x \in F$,

si on pose $$x = \sum_{1 \leqslant i \leqslant n} a_i x_i = \sum_{1 \leqslant i \leqslant n} a'_i x'_i$$

en tenant compte de (1) $x_j = \sum_{1 \leqslant i \leqslant n} s_{ij} x' \qquad i \leqslant j \leqslant n$

que l'on peut écrire :

$$X = SX' \qquad \text{et} \qquad X' = S^{-1}X$$

La matrice S de passage donne les anciennes coordonnées $(x_1, \ldots, x_n)$ **en fonction des nouvelles** $(x'_1, \ldots, x'_n)$.

2) **Sur les formes polaires et quadratiques**

Soient $x \in E$ et X sa représentation dans B et Y sa représentation dans B'.

Soient $x_1 \in E$ et X_1 sa représentation dans B et Y_1 sa représentation dans B'.

On a :

$$X = SY \qquad Y = S^{-1}X$$
$$X_1 = SY_1 \qquad Y_1 = S^{-1}X_1$$

Soit q la forme quadratique et Q sa représentation dans B :

$$Q(X) = {}^tXAX$$

Exprimons X et ${}^tX$ en fonction de Y :

$$X = SY \quad \text{et} \quad {}^tX = {}^tY{}^tS$$

La représentation de q dans la base B′ est $Q(Y) = {}^tY{}^tSASY$

$$= {}^tYBY \quad \text{où} \quad \boxed{{}^tSAS = B}$$

La représentation de la forme polaire f dans B et B′ sera :

$$F(X, X_1) = {}^tXAX_1$$
$$F(Y, Y_1) = {}^tY{}^tSASY_1$$
$$= {}^tYBY_1$$

Proposition (6,1)

Étant donné E espace vectoriel, $(a_i)_{1\leqslant i\leqslant n}$ $(a'_i)_{1\leqslant i\leqslant n}$ deux bases de E et S la matrice de passage de $(a_i)_{1\leqslant i\leqslant n}$ à $(a'_i)_{1\leqslant i\leqslant n}$. Pour toute application bilinéaire f de E dans E : $B = {}^tSAS$ où A est la matrice associée à f par rapport à $(a_i)_{1\leqslant i\leqslant n}$ et B la matrice associée à f par rapport à $(a'_i)_{1\leqslant i\leqslant n}$.

Définition (6,4)

On dit qu'une forme quadratique q sur un espace vectoriel E est **positive** si $q(x, x) \geqslant 0$ pour tout $x \in E$.

Définition (6,5)

Pour qu'une forme quadratique positive q sur un espace vectoriel E soit non **dégénérée (ou définie)** il faut et il suffit que $q(x, x) > 0$ pour tout $x \neq 0$ dans E.

7. ESPACE VECTORIEL NORMÉ, DISTANCE

Définition (7,1)

E étant un espace vectoriel, toute application $\| \ \|$ de E dans $\mathbf{R}^+$ vérifiant :

$$\begin{cases} \|x\| = 0 & \text{si et seulement si} \quad x = 0 \\ \|x + y\| \leqslant \|x\| + \|y\| & x, y \in E \\ \|\lambda x\| = |\lambda| \, \|x\| & \lambda \in R \end{cases}$$

est appelée une **norme** sur E (E est alors un espace **vectoriel normé**).

Norme associée à une forme quadratique positive non dégénérée

Soit q une forme quadratique positive non dégénérée alors :

quel que soit $x \in \mathbf{R}$

$$q(\lambda x) = \lambda^2 q(x)$$

donc

$$\sqrt{q(\lambda x)} = |\lambda| \sqrt{q(x)}$$

d'autre part $q(x) = 0$ si et seulement si $x = 0$ (q est définie positive) enfin vérifier que

$$\sqrt{q(x + y)} \leqslant \sqrt{q(x)} + \sqrt{q(y)}$$

Proposition (7, 1)

L'application de E dans $\mathbf{R}^+$ définie par $x \rightarrow \sqrt{q(x)}$ est une norme sur E.

Définition (7, 2)

Soit E un espace vectoriel, muni d'une norme $x \rightarrow q(x)$ q étant une forme quadratique définie positive, E est alors appelé **espace euclidien.**

Définition (7, 3)

On appelle **distance** sur un ensemble E, une application d de E dans $\mathbf{R}^+$ telle que :

i) $d(x, y) = 0$ si et seulement si $x = y$.

ii) $d(x, y) = d(y, x)$ quels que soient $x, y \in E$.

iii) $d(x, z) \leqslant d(x, y) + d(y, z)$ quels que soient $x, y, z \in E$.

Proposition (7, 2)

Si $x \rightarrow \|x\|$ est une norme sur un espace vectoriel E, alors

$$d(x, y) = \|x - y\|$$

est une distance sur E telle que :

$$d(\lambda x, \lambda y) = |\lambda|\, d(x, y) \qquad \text{et} \qquad d(x + z, y + z) = d(x, y)$$

i) $d(x, y) = 0$ si et seulement si $\|x - y\| = 0$

$\|x - y\| = 0$ si et seulement si $x = y$

ii)

$$\begin{aligned} d(x, y) &= \|x - y\| \\ &= |-1|\, \|y - x\| \\ &= \|y - x\| \\ &= d(y, x) \end{aligned}$$

iii) $$\begin{aligned} d(x, z) &= \|x - z\| \\ &\leqslant \|x - y\| + \|y - z\| \\ &\leqslant d(x, y) + d(y, z) \end{aligned}$$

Les deux autres propriétés se démontrent facilement.

8. DÉTERMINANT

Définition (8, 1)

Étant donné $(E_i)_{1 \leqslant i \leqslant n}$ des espaces vectoriels on dit qu'une application de $E_1 x \dots x E_n$ dans **R** est une **forme multilinéaire,** si chaque application partielle $x_i \to f(x_1, \dots x_i \dots x_n)$ est linéaire. (On dit aussi « forme n-linéaire »).

Définition (8, 2)

Une forme n-linéaire définie sur E^n $(\overbrace{Ex \dots xE}^{n})$ où E est un espace vectoriel est **alternée** si :

$$f(x_1, \dots, x_i, \dots x_n) = 0 \qquad (x_1, \dots, x_i, \dots x_n) \in E^n$$

pour tout élément ayant deux coordonnées égales.

Proposition (8, 1)

i) Une forme multilinéaire alternée f change de signe lorsque l'on transpose deux composantes.

En effet :

$$\begin{aligned} f(x_1, \dots, x_i + x_j, \dots, x_j + x_i, \dots x_n) &= 0 \\ &+ f(x_1, \dots, x_i, \dots x_j, \dots x_n) \\ &+ f(x_1, \dots, x_j, \dots x_i, \dots x_n) \end{aligned}$$

donc :

$$f(x_1, \dots, x_i, \dots, x_j, \dots x_n) = - f(x_1, \dots, x_j, \dots, x_i, \dots, x_n)$$

ii) Si σ est une permutation et k le nombre de transpositions de σ

$$\sigma : i \to \sigma(i)$$

alors

$$f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = (-1)^k f(x_1, \dots, x_n) = \varepsilon_\sigma f(x_1, \dots, x_n)$$

où $\varepsilon_\sigma = \pm 1$ (suivant la parité de k).

Formes n-linéaires alternées sur un espace vectoriel de dimension n

Étudions les formes n-linéaires alternées sur E^n :

— soient E un espace vectoriel, $(a_i)_{1 \leqslant i \leqslant n}$ une base et $(x_i)_{1 \leqslant i \leqslant n}$ n éléments de E

$$x_i = \sum_{1 \leqslant j \leqslant n} \lambda_{ij} a_j$$

$$f(x_1, \dots, x_n) = f\left(\sum_j \lambda 1_{ij} a_j, \dots, \sum_j \lambda_{nj} a_j\right) \qquad 1 \leqslant j \leqslant n$$

$$f(x_1, \dots, x_n) = \sum_{\substack{1 \leqslant j_1 \leqslant n \\ \dots\dots\dots\dots \\ 1 \leqslant j_n \leqslant n}} \lambda_{1j_1} \dots \lambda_{nj_n} f(a_{j_1}, \dots, a_{j_n})$$

Les termes de la forme $f(a_{j_1}, \dots, a_{j_p}, \dots, a_{j_q}, \dots, a_{j_n})$ sont nuls si $j_q = j_p$.

Il ne reste que n! termes (nombre de permutations d'un ensemble à n éléments ; ensemble noté $\mathfrak{S}_n$).

Soit la permutation $\sigma : (1, \dots, n) \to (j_1, \dots, j_n)$

$$f(x_1, \dots, x_n) = \sum_{\sigma \in \mathfrak{S}_n} \lambda_{1,\sigma(1)} \dots \lambda_{n\sigma(n)} f(a_{\sigma(1)}, \dots, a_{\sigma(n)})$$

$$f(x_1, \dots, x_n) \sum_{\sigma \in \mathfrak{S}_n} \varepsilon_\sigma \lambda_{1\sigma(1)} \dots \lambda_{n\sigma(n)} f(a_1, \dots a_n) \qquad (1)$$

Dans l'expression (1) les $f(a_1, \dots, a_n)$ sont indépendants du n-uplet d'éléments de E.

Montrons que les formes construites répondent à la question.

— multilinéaire : évident

— alterné : supposons que $x_i = x_{i'}$ (c'est-à-dire $\lambda_{ij} = \lambda_{i'j}$ $1 \leqslant j \leqslant n$) à tout terme

$$\varepsilon_\sigma \lambda_{1\sigma(1)} \dots \lambda_{i\sigma(i)} \dots \lambda_{i'\sigma(i')} \dots \lambda_{n\sigma(n)}$$

correspond

$$\varepsilon_{\sigma'} \lambda_{1\sigma(1)} \dots \lambda_{i'\sigma(i')} \dots \lambda_{i\sigma(i)} \dots \lambda_{n\sigma(n)}$$

ces deux termes se diffèrent d'une transposition donc : $\varepsilon_\sigma = -\varepsilon_{\sigma'}$ donc les termes de la somme s'éliminent deux à deux.

La condition (1) est une condition nécessaire et suffisante.

Donc : **sur un espace de dimension** n, **il existe une forme n-linéaire alternée déterminée à une constante multiplicative près.**

Définition (8, 3)

Le déterminant, par rapport à une base ordonnée, d'un espace vectoriel est celle des formes n-linéaires alternées qui prend la valeur 1 pour les n éléments de la base.

$$\det [x_1, \dots, x_n] = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon_\sigma \lambda_{1\sigma(1)} \dots \lambda_{n\sigma(n)}$$

on le notera :

$$\det [x_1, \dots, x_n] = \begin{vmatrix} \lambda_{11} & \dots\dots\dots\dots & \lambda_{n1} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \lambda_{1n} & \dots\dots\dots\dots & \lambda_{nn} \end{vmatrix} \qquad (1')$$

(déterminant d'ordre n).

Le déterminant est linéaire par rapport aux éléments d'une ligne ou d'une colonne (multilinéarité).

$$\det [x_1, ..., x_n] = \lambda_{i1}A_{i1} + ... + \lambda_{in}A_{in} \tag{2}$$

Calcul de A_{11} (par exemple) coefficient de λ_{11} dans une expression du type (2). Si nous mettons λ_{11} en facteur dans l'expression (1), ce nombre est facteur d'une somme qui n'est autre que le déterminant d'ordre $(n - 1)$ obtenu en supprimant du tableau (1′) la première ligne et la première colonne.

Pour déterminer A_{ij} on amène la $i^{\text{ème}}$ ligne à la première [ce qui revient à changer $(i - 1)$ fois le signe du déterminant].

Amenons ensuite la $j^{\text{ème}}$ ligne à la première [ce qui revient à changer $(j - 1)$ fois de signe].

Conclusion : $A_{i,j}$ est égal au produit par $(-1)^{i-1+j-1} = (-1)^{i+j}$ du déterminant obtenu à partir de (1′) en supprimant la $i^{\text{ème}}$ ligne et la $j^{\text{ème}}$ colonne.

$A_{i,j}$ est appelé le **mineur** associé à λ_{ij} et $(-1)^{i+j}A_{ij}$ est appelé le **cofacteur** associé à λ_{ij}.

Déterminant associé à une application linéaire

$$(x_1, ..., x_n) \in E^n \qquad \text{et} \qquad f, g \in \mathscr{L}(E, E)$$

$$(x_1, ..., x_n) \to \det [g(x_1), ..., g(x_n)]$$

C'est une forme multilinéaire : (g linéaire, déterminant multilinéaire).

C'est une forme alternée : (trivial).

Cette forme multilinéaire alternée est proportionnelle à $\det [x_1, ... x_n]$

$$\det [g(x_1), ..., g(x_n)] = h \det [x_1, ..., x_n] \qquad h \in \mathbf{R}$$

$$\det [g(a_1), ..., g(a_n)] = h$$

$$\det [g(x_1), ..., g(x_n)] = \det [g(a_1) ... g(a_n)] \det [x_{11}, ..., x_n]$$

Mais se donner n éléments $(x_1, ..., x_n)$ c'est se donner une application linéaire qui aux éléments $(a_1, ..., a_n)$ associe $(x_1, ..., x_n)$ donc :

$$\det [g \circ f(a_1), ..., g \circ f(a_n)] = \det [g(a_1), ..., g(a_n)] \det [f(a_1), ..., f(a_n)]$$

en introduisant les matrices

$$f \to A \qquad g \to B \qquad \det [BA] = \det B \det A$$

Proposition (8, 2)

i) $\det A \cdot \det A^{-1} = 1$.

ii) $\det A = \det S^{-1}AS$ ou S matrice de changement de base (le déterminant d'une application linéaire de E dans E est indépendant de la base à laquelle E est rapportée).

(La démonstration est laissée aux soins du lecteur).

Application

Proposition (8, 3)

Soit $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$ une matrice carrée. A est inversible si et seulement si $\det A \neq 0$.

C.N. Supposons A inversible et soit A^{-1} son inverse, alors

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

donc $\det A \neq 0$.

C.S. Supposons $\det A \geqslant 0$. Si $A = (a)$, $B = (a^{-1})$ est telle que

$$AB = BA = I$$

Soit $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$ $n \geqslant 2$; considérons $B = (b_{ij})_{1 \leqslant ij \leqslant n}$ où $b_{ij} = c_{ji\,/\det A}$ et c_{ji} le cofacteur de a_{ij}.

L'élément (i, j) de AB est :

$$\sum a_{ik}b_{kj} = \left(\sum_{1 \leqslant k \leqslant n} a_{ik}c_{jk} \right) / \det A$$

pour $(1 \leqslant i, j \leqslant n)$.

Or

$$\sum_{1 \leqslant k \leqslant n} a_{ik}c_{jk} = \det A \qquad (\text{si } i = j)$$

et

$$\sum_{1 \leqslant k \leqslant n} a_{ik}c_{jk} = 0 \qquad (\text{si } i \neq j).$$

Définition (8, 4)

Soit $A = (a_{ij})_{1 \leqslant ij \leqslant n}$ une matrice carrée. On appelle **adjoint** de A et on note $\mathrm{Adj}(A) = (d_{ij})_{1 \leqslant ij \leqslant n}$ la matrice transposée des cofacteurs de $(a_{ij})_{1 \leqslant ij \leqslant n}$ donc :

$$A^{-1} = \mathrm{Adj}(A)_{/\det A}$$

9. VECTEURS PROPRES D'UNE MATRICE

Définition (9, 1)

Soit $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$ une matrice carrée ; on appelle **valeur propre** de A (respectivement **vecteur propre** de A) tout élément $\lambda_i \in \mathbf{R}$ $(1 \leqslant i \leqslant n)$ (respectivement tout élément $x_i \neq 0$) tel que :

$$Ax_i = \lambda_i x_i$$

où

$$(A - \lambda_i)x_i = 0$$

Or un tel système homogène a une solution non triviale si et seulement si $\det (A - \lambda_i) = 0$.

Proposition (9, 1)

Soit $A = (a_{ij})_{1 \leqslant ij \leqslant n}$ une matrice. Les valeurs propres de A sont les n racines (distinctes ou non) de $\det (A - \lambda) = 0$.

Le polynôme $\det(A - \lambda)$ est appelé « **polynôme caractéristique** » de A, (il est de degré n par rapport à λ).

Proposition (9, 2)

> Soient $A = (a_{ij})_{1 \leqslant ij \leqslant n}$ une matrice, P une matrice régulière et $B = PAP^{-1}$; les valeurs propres de A sont identiques à celles de B ; d'autre part $\mathrm{tr}A = \mathrm{tr}B$ et rang A = rang B.

En effet,

- $\det(B - \lambda I) = \det(PAP^{-1} - \lambda I) = \det(PAP^{-1} - \lambda PP^{-1})$

$$\det[P(A - \lambda)P^{-1}] = \det(A - \lambda)$$

donc les valeurs propres de A sont égales à celles de B.

- $\mathrm{tr}B = \mathrm{tr}PAP^{-1} = \mathrm{tr}(P^{-1}PA) = \mathrm{tr}(A)$
- P et P^{-1} régulières donc rang (A) = rang (B).

Proposition (9, 3)

Si D est une matrice diagonale ses valeurs propres sont ses éléments diagonaux (trivial).

Proposition (9, 4)

Si λ est une valeur propre de A et si $r \in \mathbf{N}^*$, alors λ^r est une valeur propre de A^r ; si A est inversible, alors λ^{-r} est valeur propre de A^{-r}.

- Si $AX = \lambda X$, alors $A(AX) = A(\lambda X) = \lambda AX = \lambda^2 X$ d'où par récurrence sur r $A^r X = \lambda^r X$.
- Si $AX = \lambda X$, alors $\lambda^{-1}A^{-1}AX = \lambda^{-1}A^{-1}\lambda X$ et $\lambda^{-1}X = A^{-1}X$ d'où par récurrence, sur r $A^{-r}X = \lambda^{-r}X$.

Proposition (9, 5)

Si $(A - \lambda_1 I)X_1 = 0$ et $(A - \lambda_2 I)X_2 = 0$ où A est une matrice symétrique et $\lambda_1 \neq \lambda_2$, alors ${}^tX_1X_2 = 0$.

Par extension si A et B sont symétriques et si $(B - \lambda_1 A)X_1 = 0$ et $(B - \lambda_2 A)X_2 = 0$ alors ${}^tX_2AX_1 = 0$.

- Si $\lambda_1 \neq \lambda_2$

$AX_1 = \lambda_1 X_1$ en multipliant à gauche par ${}^tX_2$, on obtient :

$$ {}^tX_2AX_1 = \lambda_1 {}^tX_2X_1 \tag{1}$$

$AX_2 = \lambda_2 X_2$ on obtient de même

$$ {}^tX_1AX_2 = \lambda_2 {}^tX_1X_2 $$

mais A est symétrique donc :

$$ {}^tX_2AX_1 = \lambda_2 {}^tX_2X_1 \tag{2}$$

de (1) et (2) on tire

$$(\lambda_1 - \lambda_2){}^tX_2X_1 = 0 \qquad \text{avec} \qquad \lambda_1 \neq \lambda_2$$

donc ${}^tX_2X_1 = 0$.

• $BX_1 = \lambda_1 AX_1$ en multipliant à gauche par ${}^tX_2$, on obtient

$${}^tX_2BX_1 = \lambda_1 {}^tX_2AX_1 \qquad (3)$$

$BX_2 = \lambda_2 AX_2$ en multipliant à droite par ${}^tX_1$, on obtient

$${}^tX_1BX_2 = \lambda_2 {}^tX_1AX_2$$

mais A et B sont symétriques, donc :

$${}^tX_2BX_1 = \lambda_2 {}^tX_2AX_1 \qquad (4)$$

de (3) et (4) on tire

$$(\lambda_1 - \lambda_2){}^tX_2AX_1 = 0 \qquad \text{avec} \qquad \lambda_1 \neq \lambda_2$$

donc ${}^tX_2AX_1 = 0$.

Matrices orthogonales

Définition (9, 2)

On appelle matrice orthogonale une matrice $P = (m_{ij})_{1 \leqslant i,j \leqslant n}$ telle que

$${}^tPP = 1.$$

On remarque que si P_i et P_j sont deux colonnes de P ${}^tP_jP_i = 0$ si $i \neq j$ et ${}^tP_iP_i = 1$.

Proposition (9, 6)

Si P est orthogonale, $|\det P| = 1$.
En effet, $\det I = \det {}^tP \det P$ et $\det {}^tP = \det P$.

Exemple :

$$\begin{pmatrix} \operatorname{Cos}\theta & -\operatorname{Sin}\theta & 0 \\ \operatorname{Sin}\theta & \operatorname{Cos}\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ est orthogonale}$$

Proposition (9, 7)

Si A est une matrice symétrique, et λ une valeur propre de A, alors $\lambda \in \mathbf{R}$.

• Soit X le vecteur propre de A associé à λ

$$AX = \lambda X$$

$${}^tXA = \lambda {}^tX \qquad (1)$$

Le conjugué de (1) est :

$${}^tXA = \bar{\lambda} {}^tX$$

(nous rappelons que nous considérons des espaces vectoriels réels)

$${}^tXAX = \bar{\lambda}{}^tXX = \lambda {}^tXX$$

donc $\lambda = \bar{\lambda}$.

Diagonalisation d'une matrice symétrique

Proposition (9,8)

Si $A\,_i(a)_{\leqslant t,j \leqslant n}$ est une matrice symétrique, il existe une matrice orthogonale $P = (p_{ij})_{1 \leqslant ij \leqslant n}$ telle que ${}^tPAP$ soit diagonale (la démonstration est laissée de côté).

Proposition (9,9)

Soient A une matrice symétrique et P la matrice orthogonale qui diagonalise A. Alors les valeurs propres de A sont les éléments diagonaux de $D = {}^tPAP$. Le rang de A est le nombre d'éléments diagonaux non nuls de D.

D matrice diagonale, donc ses valeurs propres sont ses éléments diagonaux et son rang le nombre des éléments diagonaux non nuls.

Puisque P est orthogonale $({}^tP)^{-1} = P$, $D = {}^tPA({}^tP)^{-1}$ a même rang et les mêmes valeurs propres que A.

Matrices idempotentes

Définition (9,3)

A est idempotente si et seulement si $A^2 = A$.

Proposition (9,10)

Si A est idempotente, les valeurs propres de A sont soit 0, soit 1.

Soit λ valeur propre de A, $AX = \lambda X$ où $X \neq 0$ vecteur propre

$$A(AX) = \lambda AX = \lambda^2 X$$

or $AA = A$ donc $AX = \lambda^2 X$

$$\lambda^2 = \lambda \text{ donc } \lambda = 0 \text{ ou } 1$$

Remarque : Le rang de A est le nombre valeurs propres non nulles de $D = {}^tPAP$ où P est la matrice orthogonale qui diagonalise A.

Propriétés des matrices définies positives

Proposition (9,11)

Une matrice symétrique A est définie positive si et seulement si les valeurs propres de A sont positives.

- Soit P la matrice orthogonale qui diagonalise A :

$$D = {}^tPAP = \begin{pmatrix} \lambda_1 & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

où les λ_i sont les valeurs propres de A.

Soit $Y = {}^tPX$ telle que $X = ({}^tP)^{-1} Y = PY$.

Alors ${}^tXAX = {}^t(PY)A(PY) = {}^tYDY = \sum_{1 \leqslant i \leqslant n} \lambda_i y_i^2$

Or, $\Sigma \lambda_i y_i^2 = 0$ entraîne $y_i = 0 \quad 1 \leqslant i \leqslant n$

$PY = X = 0$; donc A est définie positive.

• Réciproquement si A est définie positive ; supposons qu'une valeur propre λ_1 soit non positive :

$$\text{soit } Y' = \begin{pmatrix} 1 \\ 0 \\ . \\ . \\ . \\ 0 \end{pmatrix} \quad \text{et} \quad X' = PY' \text{ donc } X' \neq 0$$

mais ${}^tX'AX = {}^tY'{}^tPAPY' = {}^tY'DY = \lambda_1 \leqslant 0$

(contraire à l'hypothèse).

Proposition (9,12)

Si A est une matrice symétrique définie positive. Alors $\det A > 0$, le rang de A est égal à n (évident).

Proposition (9,13)

Si $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$ est une matrice définie positive et $C = (c_{ij})_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant n}}$ une matrice telle que rang $C = n$, alors ${}^tCAC$ est définie positive.

Proposition (9,14)

Si A est définie positive et C régulière, alors ${}^tCAC$ est définie positive.

- ${}^tCAC$ est symétrique.
- $Y \neq 0$, ${}^tY({}^tCAC)Y = {}^tXAX$ où $X = CY$. Puisque A est définie positive et $X \neq 0$, ${}^tXAX > 0$.

Mais ${}^tY({}^tCAC)Y$ est donc positif pour $Y \neq 0$ et on a gagné !

Proposition (9,15)

Si A est définie positive, il en est de même de A^{-1}.

(On pose $C = A^{-1}$ dans la démonstration de Proposition 8,12).

Proposition (9,16)

Soit A une matrice symétrique définie positive, il existe une matrice régulière C telle que $CA{}^tC = I$ et ${}^tCC = A^{-1}$.

Soit P une matrice orthogonale telle que

$$D = {}^tPAP = \begin{pmatrix} \lambda_1 & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \qquad \lambda_i > 0 \qquad 1 \leqslant i \leqslant n$$

et soit
$$M = \begin{pmatrix} 1/\sqrt{\lambda_1} & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & 1/\sqrt{\lambda_n} \end{pmatrix}$$

Alors $C = {}^tM{}^tP$ produit de matrices régulières est régulière et $CA{}^tC = {}^tM{}^tPAPM = {}^tMDM = I$.

Enfin, si $CA{}^tC = I$ alors ${}^tC(CA{}^tC)C = {}^tCIC = {}^tCC$ d'où ${}^tCCA = I$ et ${}^tCC = A^{-1}$. On a gagné !

DEUXIÈME PARTIE

ANALYSE DES DONNÉES

1. ANALYSE EN COMPOSANTES PRINCIPALES

1.1. LE PROBLÈME POUR L'UTILISATEUR

Le problème posé par l'analyse factorielle peut schématiquement se résumer ainsi :

A) **On dispose de données statistiques qui ne peuvent se représenter que dans un espace de dimension élevé** ; ce qui signifie le plus couramment que l'on dispose d'un nombre assez important de variables et d'observations, par exemple 60 caractères (variables) mesurés sur 1 000 individus (observations), ou même 8 caractéristiques économiques (variables) mesurées sur 21 régions (observations). Quelquefois, la distinction entre variables et observations est artificielle ou conventionnelle ; en principe, on réservera le nom d'observation à tout ce qui a un caractère répétitif, à tout ce qui peut être considéré comme une réalisation d'un vecteur aléatoire qui aurait autant de composantes qu'il existe de variables.

Pratiquement donc, on dispose d'un tableau rectangulaire de valeurs numériques dont les dimensions sont telles qu'elles sont un obstacle à l'assimilation rapide par le statisticien de l'information contenue dans cet ensemble de nombres.

B) **On désire représenter, dans la mesure du possible, ces données statistiques dans un espace de faible dimension, avec le minimum de perte d'information** ; ainsi, par exemple, on résumerait les 60 variables mesurées sur les 1 000 individus par cinq variables (les données de ces cinq variables permettant de reconstituer approximativement les 60).

On passe ainsi de 1 000 points dans un espace à 60 dimensions, à 1 000 points dans un espace à 5 dimensions.

En fait, cinq dimensions sont encore beaucoup pour notre « rétine».

Ainsi, nous chercherons d'abord des espaces à une ou deux ou trois dimensions, en les complétant éventuellement par quelques dimensions supplémentaires.

C) **On ne fait aucune hypothèse particulière préalable ;** ceci pour insister sur le côté purement descriptif de l'opération, à l'opposé des démarches des pionniers de l'analyse factorielle, cherchant un petit nombre de facteurs interprétables, à partir de modèles assouplis et rendus plus dociles par l'introduction de paramètres supplémentaires.

1.2. LE PROBLÈME POUR LE STATISTICIEN

Le statisticien a le choix entre plusieurs espaces de départ, pour raisonner : l'espace dit « des observations » (où les données de base seront par exemple représentées par 60 points variables dans un espace à 1 000 dimensions) ou « l'espace des variables » (à 60 dimensions, avec 1 000 points observations dans le cas de notre exemple, chaque point observation ayant 60 coordonnées qui sont les valeurs des 60 variables lui correspondant).

Dans ce dernier espace, par exemple, la distance entre deux points constitue un nombre caractéristique de la ressemblance des deux observations correspondantes. Soit x_1 et x_2 ces observations (x_1 et x_2 sont donc des vecteurs à 60 composantes que nous désignerons par x_{1i} et x_{2i} l'indice i variant de 1 à 60).

La distance $d(x_1, x_2)$ s'écrit :

$$d^2(x_1, x_2) = \sum_{i=1}^{n} (x_{1i} - x_{2i})^2$$

Elle est d'autant plus faible que les composantes x_{1i} et x_{2i} sont proches pour toutes les valeurs de i.

Dire que l'observation x_3 est à la même distance de x_1 que x_2 pose néanmoins, en général, beaucoup de problèmes d'interprétation car l'égalité :

$$\sum_{i=1}^{n} (x_{1i} - x_{2i})^2 = \sum_{i=1}^{n} (x_{1i} - x_{3i})^2$$

ne renseigne pas beaucoup sur les composantes responsables de cette égalité... (cette difficulté pourra être levée par la « représentation simultanée »).

Notons qu'intervient ici un important problème d'unité et de pondération entre les différentes variables : en effet, si la $j^{ème}$ composante a des valeurs relativement élevées, le terme $(x_{1j} - x_{2j})^2$ jouera un rôle prépondérant dans la somme :

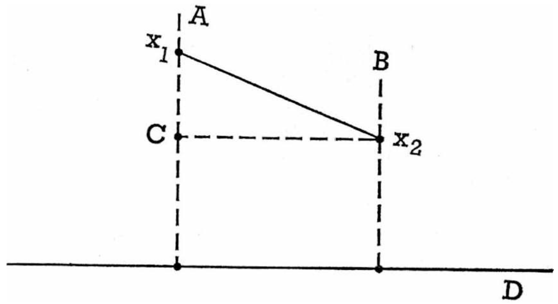
$$\sum_{i=1}^{n} (x_{1i} - x_{2i})^2,$$

par suite, la $j^{ème}$ variable aura un poids supérieur aux autres dans la définition des proximités entre observations.

Plaçons-nous donc dans l'espace à p dimensions des variables, $\mathbf{R}^p$, où se trouvent les n points observations.

A) **Nous allons chercher le sous-espace à une dimension** D **(la droite)** tel que les distances entre les projections des points-observations mesurées sur cette droite soient le plus proche possible des distances définies plus haut, dans l'espace à p dimensions.

D'après le théorème de Pythagore, la distance entre deux points dans l'espace $\mathbf{R}^p$: AB^2 se décompose entre distance BC^2 de leur projection sur D, et distance CA^2, CA étant un vecteur de l'espace à $p - 1$ dimension, orthogonal à D.

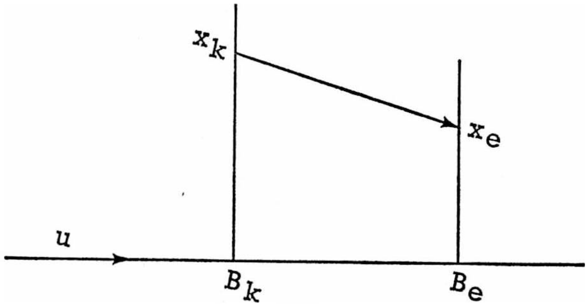


Il est donc naturel de chercher un sous-espace D qui maximise les distances projetées du type BC^2 ou, ce qui revient au même qui minimise les distances « résiduelles » telles que la distance AC^2.

Soient x_k, x_e deux points observations du nuage.

Soient $u_1, u_2, \dots, u_p$ les composantes d'un vecteur unitaire u porté par la droite cherchée D.

La grandeur de la projection B_kB_e est le produit scalaire du vecteur $x_k - x_e$, par le vecteur u.



$$\overline{B_kB_e} = {}^t u(x_k - x_e) = \sum_{i=1}^{p} u_i(x_{ki} - x_{ei})$$

Si l'on veut que, en moyenne, les longueurs des projections soient maxima, de façon à ce que la déformation du nuage soit minimum, il faut donc que en sommant pour tous les couples (k, e) de points-observations :

$$S = \sum_{k,e} \overline{B_k B_e^2} = \sum_{k,e} [{}^t u(x_k - x_e)]^2 \qquad \text{soit maximum,}$$

avec la condition :

$$\sum_{i=1}^{n} u_i^2 = {}^t u \cdot u = 1 \text{ puisque u est supposé unitaire.}$$

$$\text{Or} \qquad [{}^t u(x_k - x_e)]^2 = \left[\sum_{i=1}^{n} u_i(x_{ki} - x_{ei}) \right]^2$$

$$= \sum_{1 \leqslant i,j \leqslant p} u_i u_j (x_{ki} - x_{ei})(x_{kj} - x_{ej})$$

Par suite, en intervertissant les indices (i, j) et (k, e) :

$$S = \sum_{1 \leqslant i,j \leqslant p} u_i u_j \left(\sum_{k,e} (x_{ki} - x_{ei})(x_{kj} - x_{ej}) \right)$$

La quantité entre parenthèse est le terme général v_{ij} d'une matrice V ;

On peut donc écrire

$$S = \sum_{1 \leqslant i,j \leqslant p} u_i u_j v_{ij}.$$

En notation matricielle, il faut donc maximiser la forme quadratique :

$$S = {}^t u V u \qquad \text{avec la condition} \qquad {}^t u u = 1.$$

Précisons quelle est cette matrice V :

Son terme général s'écrit :

$$v_{ij} = \sum_{1 \leqslant k,e \leqslant n} (x_{ki} - x_{ei})(x_{kj} - x_{ej}) \tag{1}$$

C'est, à un coefficient près, la covariance des variables x_i et x_j. Vérifions-le :

$$v_{ij} = \sum_{1 \leqslant k,e \leqslant n} (x_{ki}x_{kj} - x_{ei}x_{kj} - x_{ki}x_{ej} + x_{ei}x_{ej})$$

$$v_{ij} = n \sum_{k=1}^{n} x_{ki}x_{kj} + n \sum_{e=1}^{n} x_{ei}x_{ej} - 2 \sum_{e=1}^{n} x_{ei} \sum_{k=1}^{n} x_{kj}$$

$$v_{ij} = 2n \left(\sum_{k=1}^{n} x_{ki}x_{kj} - \frac{1}{n} \sum_{e=1}^{n} x_{ei} \cdot \sum_{k=1}^{n} x_{kj} \right)$$

On reconnaît là le terme général de la covariance des variables x_i et x_j, au facteur $2n^2$ près.

Ainsi, les composantes d'un vecteur u unitaire porté par la droite D sur laquelle se projette le nuage en se déformant le moins possible maximisent ${}^t uVu$, (où V est maintenant la matrice des covariances des variables 2 à 2) avec la contrainte ${}^t uu = 1$.

Il s'agit là d'un problème classique en algèbre et en géométrie analytique. (Recherche des axes principaux d'une quadrique.)

La matrice des covariances V est symétrique et définie positive. Nous savons (cf. rappel) qu'une telle matrice admet des valeurs propres positives correspondant à des vecteurs propres r_i orthogonaux. Soit R la matrice ayant les vecteurs propres r_i en colonne ; on a donc ici :

$$VR = R\Delta$$

Δ étant une matrice diagonale ayant les valeurs propres δi sur sa diagonale.

Multipliant par la transposée ${}^t R$ (qui est aussi, ici, l'inverse de R) :

$$V = R\Delta{}^t R$$

Notre forme quadratique s'écrit : ${}^t uVu = {}^t uR\Delta{}^t Ru$.

Posons :

$${}^t y = {}^t uR.$$

Et notons que :

$$y = {}^t({}^t uR) = {}^t Ru \qquad \text{et que} \qquad Ry = u$$

On a alors :

$${}^t uVu = {}^t y\Delta y = \sum_{i=1}^{p} \delta_i y_i^2.$$

La condition ${}^t uu = 1$ devient, puisque $u = Ry$:

$${}^t y{}^t RRy = 1 = {}^t yy \qquad ({}^t RR = I)$$

On est donc ramené au problème suivant :

chercher le maximum de $\sum_{i=1}^{p} \delta_i y_i^2$ sachant que $\sum_{i=1}^{p} y_i^2 = 1$.

C'est, si l'on veut, un petit problème de programmation linéaire : la solution se trouve nécessairement en un sommet du domaine défini par la relation :

$\sum_{i=1}^{p} y_i^2 = 1$, donc pour une valeur $y_i^2 = 1$, $y_j^2 = 0$, $j \neq i$

La somme $\sum_{i=1}^{p} \delta_i y_i^2$ devant être rendue maximum, l'indice i choisi sera celui correspondant à la plus grande valeur propre δ_i, qui sera la valeur effective du maximum.

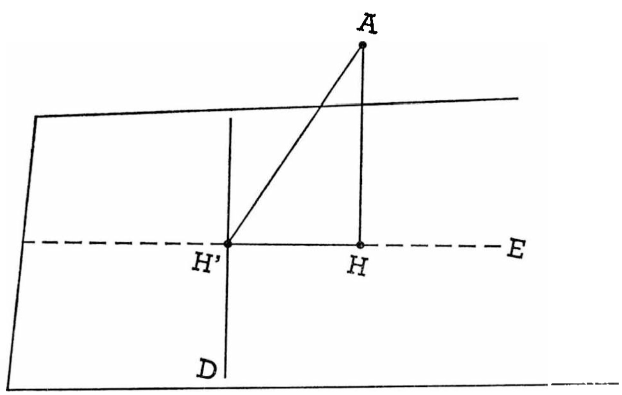
Comme $u = Ry$, le vecteur u cherché sera le vecteur propre correspondant à cette plus grande valeur propre.

Ce vecteur u s'appelle la première composante principale du nuage de points.

On peut donc énoncer le résultat suivant :

« les cosinus directeurs de l'espace à une dimension sur laquelle se projette le nuage des points-observations en se déformant le moins possible (avec notre définition de la non-déformation) sont les composantes du vecteur propre correspondant à la plus grande valeur propre de la matrice des covariances associée au nuage. »

B) Si l'on cherche maintenant le sous-espace à deux dimensions (par conséquent représentable sur un graphique) ayant toujours ces propriétés de déformation minimale, l'on s'aperçoit que ce sous-espace contient D, et qu'il contient le sous-espace à une dimension E ayant également la propriété de « déformation minimale » tout en étant assujetti à être orthogonal à D.



On voit facilement qu'une base de ce sous-espace est constituée par les deux vecteurs propres de la matrice V correspondant aux deux plus grandes valeurs propres. Il en serait de même pour l'espace à p dimensions de déformation minimale, dont une base serait constituée par les p premiers vecteurs propres de V.

Nous avons ainsi trouvé des axes factoriels qui ont la propriété d'extraire progressivement le plus d'information possible concernant les proximités entre les points.

La qualité de cette information peut être mesurée par l'importance de la, ou des plus grandes valeurs propres, selon que l'on utilise un ou plusieurs facteurs pour la représentation.

1.3. PRÉSENTATION DES RÉSULTATS (par le statisticien, pour l'utilisateur)

Le plus souvent, et particulièrement lorsqu'on dispose d'un ensemble hétérogène de variables, on utilise des transformations simples sur ces variables afin de les rendre maniables et comparables.

Ainsi, la transformation la plus couramment utilisée, héritée de la statistique classique, consiste à centrer et réduire les variables, c'est-à-dire à leur retrancher leurs moyennes et à les diviser par leurs écarts-types.

L'analyse précédente revient alors à chercher les vecteurs propres de la matrice des corrélations des variables.

Toutefois, lorsque l'ensemble des variables est homogène, et lorsque les différences entre les échelles des différentes variables ont une signification, il est plus raisonnable de s'en tenir aux données initiales brutes.

La représentation la plus courante consistera à faire figurer les proximités entre points-observations sur un graphique-plan dont les axes sont les deux premiers axes factoriels.

Si ceux-ci s'avèrent insuffisants pour résumer ces proximités, on adjoindra, sur un deuxième graphique les positions de ces points dans le système des troisième et quatrième axes factoriels, etc...

Soit x_{ik} la $k^{\text{ème}}$ observation de la variable i.

Les coordonnées du $k^{\text{ème}}$ point-observation sur les deux premiers axes factoriels (de cosinus directeurs u et s) sont les produits scalaires ${}^t u x_k$ avec :

$$ {}^t u x_k = \sum_i u_i x_{ik} $$

et ${}^t s x_k$ avec :

$$ {}^t s x_k = \sum_i s_i x_{ik} $$

Ainsi, ces nouvelles coordonnées apparaissent comme des combinaisons linéaires des variables initiales.

Corrélations variables-composantes :

Supposons maintenant les variables centrées et réduites : calculons le coefficient de corrélation entre la $i^{\text{ème}}$ variable x_i et le $k^{\text{ème}}$ facteur u_k dont la $j^{\text{ème}}$ observation s'écrit :

$$ \sum_{e=1}^{p} u_{ke} x_{ej} $$

$$ c(i, k) = \frac{\frac{1}{n} \sum_{j=1}^{n} x_{ij} \left(\sum_{e}^{p} u_{ke} x_{ej} \right)}{\sqrt{\operatorname{var}(x_i) \cdot \operatorname{var}(u_k)}} $$

$\operatorname{var}(x_i) = 1$ puisque la variable x_i est réduite.

$\operatorname{var}(u_k) = \lambda_k$ ($k^{\text{ème}}$ valeur propre $= \max({}^t u_k V u_k)$, (maximum lié)).

D'autre part la somme

$$\frac{1}{n}\sum_{j=1}^{n} x_{ij}\left(\sum_{e=1}^{p} u_{ke}x_{ej}\right)$$

s'écrit :

$$\frac{1}{n}\sum_{e=1}^{p}\left(\sum_{j=1}^{n} x_{ij}x_{ej}\right)u_{ke} = \sum_{e=1}^{p} v_{ie}u_{ke}$$

Comme u_k est vecteur propre de V :

$$\sum_{e=1}^{p} v_{ik}u_{ke} = \lambda_k u_i$$

Ainsi, le coefficient c(i, k) s'écrit : $c(i, k) = u_{ki} \cdot \sqrt{\lambda_k}$.

On a donc le résultat suivant :

— **Dans le cas où l'analyse est effectuée sur des variables centrées réduites, les coefficients de corrélation entre une variable « i » et un facteur « k » sont proportionnels à la** i$^{\text{ème}}$ **composante de l'axe factoriel** « k ».

Ainsi, si l'on porte sur deux axes les p points ayant pour coordonnées (u_i, s_i), i$^{\text{èmes}}$ composantes des premier et second axes factoriels, on obtiendra dans le plan p points-variables, dont les proximités peuvent s'interpréter en terme de corrélation.

Notons que si l'on porte ces p points-variables sur le même graphique que les n-points-observations, on enrichit considérablement sa lecture, à condition d'être prudent, car il existe un arbitraire dans le choix des échelles des deux représentations.

En effet, au vu des coordonnées des points-variables, on a aussitôt une idée de la signification de facteurs, et l'on sait quelles variables sont responsables de la proximité entre telle ou telle observation.

Il se dégage ainsi une notion de proximité entre variables et observations, confirmée par la remarque suivante, le mot « barycentre » y étant pris dans un sens large.

— les points variables sont des « barycentres » des points-observations, chaque point-observation étant affecté du poids « valeur de la variable pour cette observation ». De la même façon, les points-observations sont des « barycentres » des points-variables.

En effet, nous avons vu que la coordonnée de la j$^{\text{ème}}$ observation sur le k$^{\text{ème}}$ axe factoriel s'écrivait :

$$c_{kj} = \sum_{i=1}^{p} u_{ki}x_{ij} \qquad \text{(pour } k = 1, 2, \ldots)$$

(ce qui prouve que le point c_j est « barycentre » des i points-variables u_i avec les « poids » x_{ij}).

Comme u_k est vecteur propre de $V = (v_{ij}) = \left(\frac{1}{n}\sum_{e=1}^{n} x_{ie}x_{je}\right)$

On a de même

$$Vu_k = \lambda_k u_k$$

Soit :

$$\sum_{i=1}^{p} \left(\frac{1}{n} \sum_{e=1}^{n} x_{ie} x_{je} \right) u_{ki} = \lambda_k u_{kj}$$

Que l'on peut écrire :

$$\frac{1}{n} \sum_{e=1}^{n} \left(\sum_{i=1}^{p} u_{ki} x_{ie} \right) x_{je} = \lambda_k u_{kj}$$

On reconnaît c_{ke} défini plus haut dans la parenthèse :

$$\frac{1}{n} \sum_{e=1}^{n} c_{ke} x_{je} = \lambda_k u_{kj}$$

Ainsi, en remplaçant u_k par $u_k/\sqrt{n\lambda_k}$ on a les deux relations :

$$\left\{ \begin{aligned} \sqrt{n\lambda_k}\, c_{kj} &= \sum_{e=1}^{p} u_{ke} x_{ej} \\ \sqrt{n\lambda_k}\, u_{kj} &= \sum_{e=1}^{n} c_{ke} x_j \end{aligned} \right.$$

1.4. EXEMPLE D'APPLICATION

1) **Variables :**

A. Catégories socio-professionnelles (8 variables exprimées en pourcentage de chaque catégorie).

B. Consommations médicales. 11 variables : consultations, visites, actes en B, K, R pour ville et hôpital, pharmacie, soins dentaires et fréquentation hospitalière, par personne protégée.
(B = analyses, K = actes techniques, R = radiologie).

2) **Observations** : les départements français (repérés sur la figure 1 par leurs numéros minéralogiques).

3) **Remarques générales :**

Le premier facteur correspond à une valeur propre représentant 40 % de la somme des valeurs propres ; le second est relatif à une valeur propre représentant 20 % de cette somme (trace de la matrice des corrélations).

La figure 1, qui donne les positions des points-variables et des points-observations résume donc 60 % de la dispersion totale.

Mais cela ne signifie pas que nous n'avons que 60 % de l'information disponible, car les 40 % restant ne sont pas « intéressants » du point de vue de l'interprétation. En effet, si nous ajoutons d'autres variables sans aucune corrélation avec les précédentes, ni entre elles, celles-ci augmenteront la trace (somme des éléments diagonaux de la matrice des corrélations) sans augmenter les premières valeurs propres. Tout en ayant la même information, le « pourcentage d'explication » diminuera.

4) **Lecture de la figure 1 :** (voir page 92).

Le premier axe factoriel, porté en abscisse, peut être qualifié « d'axe d'urbanisation » au vu des variables qui y participent (abscisses, sur cet axe, des divers points-variables).

Il oppose en effet le pourcentage d'agriculteurs, à gauche, aux pourcentages de professions libérales, de cadres et d'employés, à droite.

Les abscisses des points-départements sur cet axe sont également significatives : à gauche la Creuse, la Mayenne et la Vendée, à droite la Seine. Notons que la qualification de ce premier axe factoriel « d'axe d'urbanisation » n'est qu'une appréciation qui a un caractère approché.

Si cet axe n'avait aucune interprétation évidente, l'analyse n'en aurait pas été moins intéressante.

Le deuxième axe sépare principalement les consommations de ville des consommations d'hôpital.

Dans la partie supérieure du graphique, se trouvent essentiellement des départements du Midi.

Une proximité entre deux points-variables (respectivement deux points-départements) s'interprète en terme de corrélation (respectivement de comportement similaire vis-à-vis des variables), et ceci d'autant plus que ces points occupent des positions excentriques (ce qui signifie alors qu'ils sont bien corrélés, dans un sens ou dans un autre, avec les facteurs).

S'il n'est pas très important de connaître les échelles choisies sur les axes, pour interpréter les diverses proximités, la position de l'origine, elle, doit donc être connue.

2. ANALYSE DANS UNE MÉTRIQUE ASSOCIÉE A UNE FORME QUADRATIQUE DÉFINIE POSITIVE

2.1. MÉTHODE

Soient $x_1, ..., x_n$ des éléments de $\mathbf{R}^p$. On définit la distance entre deux éléments x_i et x_j de $\mathbf{R}^p$, par :

$$d^2(i, j) = {}^t(x_i - x_j)Q(x_i - x_j)$$

où Q est une matrice associée à une forme quadratique définie positive. Or, on sait qu'il existe une matrice régulière C telle que :

$$Q = {}^tCC$$

donc $d^2(i, j) = {}^t(x_i - x_j){}^tCC(x_i - x_j)$.

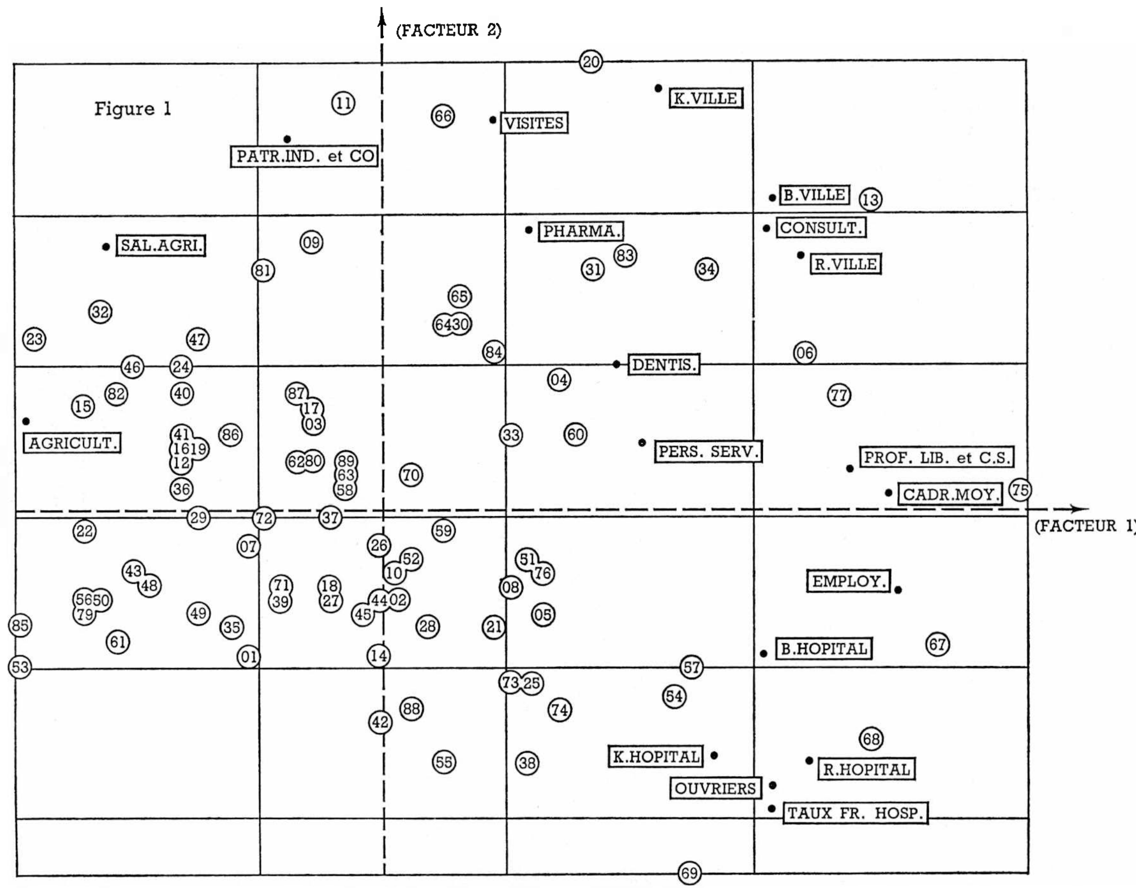

(FACTEUR 2)
(FACTEUR 1)
K.VILLE
VISITES
PATR.IND. et CO
B.VILLE
CONSULT.
R.VILLE
PHARMA.
SAL.AGRI.
DENTIS.
AGRICULT.
PERS. SERV.
PROF. LIB. et C.S.
CADR.MOY.
EMPLOY.
B.HOPITAL
K.HOPITAL
R.HOPITAL
OUVRIERS
TAUX FR. HOSP.

Figure 1

Posons : $y_i = Cx_i$ $1 \leqslant i \leqslant n$ (représentant x_i dans la nouvelle base)

$$d^2(i, j) = {}^t(y_i - y_j)(y_i - y_j)$$

Nous cherchons une forme linéaire v, telle :

$$v(x_j) = \sum_{1 \leqslant i \leqslant p} v_i x_{ij}$$

et avec $v_i = v(a_i)$ $1 \leqslant i \leqslant p$ où les a_i sont les éléments de base initiale. Cette forme linéaire v étant astreinte à « résumer » l'information contenue dans l'observation j. C'est l'application linéaire « projection sur **R** ». Pour avoir cette bonne approximation de l'information contenue il faut que les distances entre les projections soient maxima (c'est-à-dire telles que les ${}^tv(x_i - x_j)$ soient maxima) dans la nouvelle base :

$$ {}^tvx = ({}^tvC^{-1})y$$

posons ${}^tu = {}^tv\, C^{-1}$

Comme au (1.2, page 85) dire que la déformation est minimum c'est dire que :

$${}^tu\, \mathscr{V}u \text{ est maximum avec } {}^tuu = 1$$

où $\mathscr{V}$ est à un coefficient près la matrice de covariance des y

$$\mathscr{V} = (v_{ij}) \qquad \text{avec} \qquad v_{ij} = \sum_{1 \leqslant k,e \leqslant p} (y_{ik} - y_{ie})(y_{jk} - y_{je})$$

comme $y = Cx$, on voit facilement que : $\mathscr{V} = CV^tC$ où V est définie par l'expression (1) (page 85) donc que la forme linéaire u cherchée est solution de (cf. 1.2).

$$\begin{cases} CV^tCu = \lambda u \\ {}^tuu = 1 \end{cases} \qquad (1)$$

Nous savons résoudre ce problème.

En effet, la forme linéaire v cherchée initialement est telle que :

$${}^tu = {}^tvC^{-1}$$

donc

$$u = {}^t(C^{-1})v$$

et

$$v = {}^tCu$$

Multiplions (1) à gauche par ${}^tC$

$${}^tCCV({}^tCu) = \lambda({}^tCu)$$

c'est-à-dire

$$\boxed{(QV)(v) = \lambda v}$$

On voit que la forme linéaire cherchée est le vecteur propre de QV relatif à la plus grande valeur propre. (Nous notons au passage que CV^tC est symétrique et a les mêmes valeurs propres que QV).

On constate alors que l'on a les normalisations suivantes :

i) **Norme de v :**

$(QV)(v) = \lambda v$ en multipliant à gauche par V $(VQ)(Vv) = \lambda Vv$.

Alors que v est appelé **facteur** du nuage de points, $Vv = x$ sera appelé **axe factoriel** du nuage.

On avait la relation :

$$v = {}^tCu$$

$$ {}^tC^{-1}v = u \qquad \text{or} \qquad {}^tuu = 1 \qquad \text{donc} \qquad {}^tvC^{-1}{}^tC^{-1}v = 1$$

soit

$$\boxed{{}^tvQ^{-1}v = 1}$$

ii) **Norme de** $x = Vv$

$$QV(v) = \lambda v \qquad \text{donc} \qquad V(v) = \lambda Q^{-1}v$$

$$ {}^tvVv = \lambda{}^tvQ^{-1}v$$

d'où :

$$ {}^tvVv = \lambda$$

$$x = Vv \qquad {}^tv = {}^txV^{-1}$$

alors

$$VQx = \lambda x$$

Multiplions à gauche par ${}^tv$

$$ {}^txV^{-1}VQx = \lambda{}^txV^{-1}x$$

$$\begin{aligned} {}^txQx &= \lambda{}^txV^{-1}x \\ &= \lambda{}^tvV(V^{-1}V)v \\ &= \lambda{}^tvVv \end{aligned}$$

$$\boxed{{}^txQx = \lambda^2}$$

2.2. EXEMPLE D'APPLICATION : ANALYSE DISCRIMINANTE

Le problème général de l'analyse discriminante est le suivant : il existe plusieurs populations multidimensionnelles $P_1, P_2, \ldots P_q$.

Les observations statistiques forment donc q « paquets » d'effectifs divers, le « paquet » i contenant par exemple N_i observations du vecteur x.

[**Exemple :** on mesure 15 variables sur 1 000 ménages répartis en 8 catégories socio-professionnelles (qui constituent une partition des 1 000 ménages].

Le problème est le suivant : disposant d'une observation multidimensionnelle supplémentaire (c'est-à-dire : un vecteur observation), nous nous posons la question :

De quelle population cette observation provient-elle le plus vraisemblablement ?

[Pour notre exemple : peut-on prédire la catégorie socio-professionnelle d'un nouveau ménage caractérisé par ses 15 variables ?]

Le problème prend tout son sens lorsque les valeurs des variables sont les seuls indices permettant de classer les individus. Si les populations P_1, P_2, ... P_q correspondent à des maladies, et les variables mesurées sur les individus à des résultats d'analyses et à des symptômes, l'analyse discriminante permettra de réaliser un diagnostic. (D'où le nom de « fonction diagnostic » donné parfois aux facteurs extraits).

La figure 2 représente le cas où l'on mesure seulement deux variables par individu, avec quatre populations.

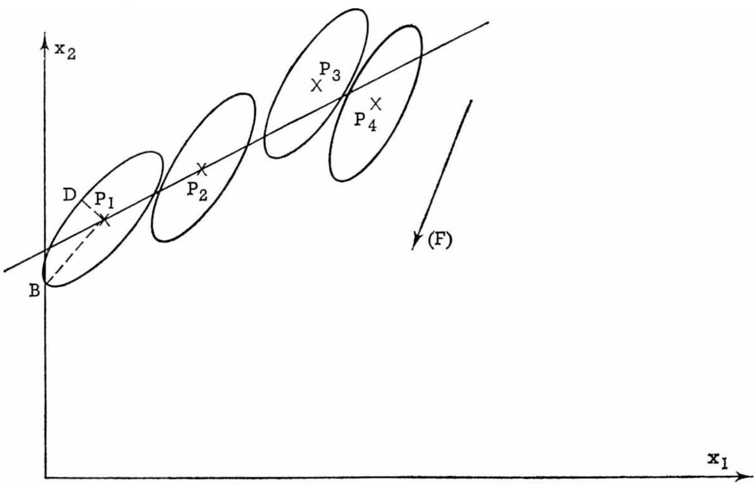


figure 2

Nous supposerons (ce qui est une hypothèse très forte) que les matrices des covariances expérimentales sont voisines dans les diverses populations. Leur valeur moyenne est notée « V ».

Nous voulons donc trouver des facteurs, c'est-à-dire des opérateurs-projection sur des espaces à une dimension, qui permettent de représenter les points-observations en accentuant l'appartenance à telle ou telle population.

La nécessité d'une métrique particulière s'impose comme on le voit sur la figure A : les points B et D sont à des distances différentes du point moyen P_1, et pourtant ils ont visiblement la même probabilité d'appartenir à cette population. L'équation des divers ellipsoïdes d'égale densité étant de la forme

$$ {}^t x V^{-1} x = k \qquad (k = \text{constante positive}), $$

nous choisirons une distance de la forme :

$$ d^2(x_i, x_j) = {}^t(x_i - x_j)V^{-1}(x_i - x_j) $$

c'est-à-dire une distance qui fait en sorte que B et D soient équidistants de P_1.

Avec cette métrique, nous analyserons le nuage des points moyens, dont la matrice des covariances sera notée M.

On doit donc chercher des fonctions u telles que :

$$V^{-1}Mu = \lambda u$$

Les fonctions u sont appelées les **fonctions discriminantes.**

Remarquons que s'il y a q populations, le rang de M est inférieur ou égal à q — 1. Il y a donc moins de q — 1 fonctions discriminantes.

Dans le cas de la figure 2, la première fonction discriminante serait l'opérateur : « projection sur Δ parallèlement à F ».

S'il n'y a que **deux populations,** la matrice M (de rang 1) n'est que le produit du vecteur « différence des vecteurs moyens » par son transposé (à un coefficient près).

$$M = (\bar{x}_1 - \bar{x}_2)^t(\bar{x}_1 - \bar{x}_2)$$

On a donc :

$$V^{-1}(\bar{x}_1 - \bar{x}_2)^t(\bar{x}_1 - \bar{x}_2)u = \lambda u \qquad (1)$$

Multipliant à gauche par $^t(\bar{x}_1 - \bar{x}_2)$

$$[^t(\bar{x}_1 - \bar{x}_2)V^{-1}(\bar{x}_1 - \bar{x}_2)]\ \{\ ^t(\bar{x}_1 - \bar{x}_2)u\ \} = \lambda(^t(\bar{x}_1 - \bar{x}_2)u)$$

La quantité entre crochet est un nombre réel (on a une forme bilinéaire) ; elle est égale à λ, valeur propre cherchée.

On l'appelle **distance généralisée** des deux populations multidimensionnelles.

On voit facilement que $u = V^{-1}(\bar{x}_1 - \bar{x}_2)$ est vecteur propre de $V^{-1}M$, pour la valeur propre λ précédente.

Cette fonction qui s'exprime simplement en fonction de V, de $\bar{x}_1$, et $\bar{x}_2$ (respectivement points moyens des premières et secondes populations) est la fonction discriminante des deux populations (la matrice des covariances interne V commune aux diverses populations se calcule, en pratique, en faisant une moyenne pondérée des diverses matrices de covariances internes expérimentales).

(à suivre)

BIBLIOGRAPHIE

1) **Publication du Laboratoire de J.-P. Benzécri** (Fac. Sciences Paris).

J.-P. BENZECRI, **Leçon sur l'analyse statistique des données dimensionnelles** (ronéoté).

J.-P. BENZECRI, **Distance distributionnelle et métrique du χ^2 en analyse factorielle des correspondances** (ronéoté).

B. CORDIER, **Analyse factorielle des correspondances** (thèse).

J.-P. BENZECRI, **Réduction d'un élément du produit tensoriel de deux espaces euclidiens** (ronéoté).

2) **Divers**

MC KENDALL and A. STUART, **The advanced theory of statistic** (Griffin).

Donald F. MORRISSON, **Multivariate statistical methods** (Mc Graw Hill).
Algèbre (ouvrages de base).

GODEMENT, **Algèbre** (Hermann).

QUEYSANNE, **Algèbre** (A. Colin).

BARBUT, **Mathématiques des Sciences Humaines,** Tome I et II (PUF collection Sup) (à recommander aux psychologues et sociologues).

BIBLIOGRAPHIE

VINCENT (A.). — La mesure de la productivité. — Paris, Dunod, 1968, 303 p. (Collection « Sigma » dirigée par Henri HIERCHE, n° 12.)

Le concept de productivité doit être démystifié. Ce n'est pas un état d'esprit : c'est une constatation : le rapport entre une production et les facteurs qui ont permis de l'obtenir (ou certains de ces facteurs). A priori, ce concept paraît donc simple ; mais les voies pour le cerner sont nombreuses et parfois arides. André L. A. VINCENT, dans son ouvrage **La mesure de la productivité** cherche à atteindre un double objectif : présenter et classer méthodiquement les différentes manières de mesurer la productivité ; ce premier effort essentiel consistant à montrer que l'on a véritablement à faire à un instrument de mesure même s'il est protéiforme, trouve sa justification dans les multiples applications qui peuvent en être faites, qu'il s'agisse d'économie théorique ou pratique.

La plus grande partie du livre est consacrée à l'éclaircissement des approches quantitatives du concept de productivité. Reprenons sommairement le cheminement suivi par M. Vincent.

La multiplicité des formules de productivité résulte de la variété des éléments pris en considération. Cette variété tient à la période de temps, à l'unité de production, aux quantités produites, aux facteurs de production, aux relations choisies.

En tout état de cause, une formule de productivité se présente toujours avec une production au numérateur et un facteur de production ou un ensemble de facteurs au dénominateur.

La spécificité de la mesure de productivité dépend donc de la période de temps pendant laquelle les quantités ont été mesurées, de la manière dont elles sont mesurées, du niveau de production pris en considération et enfin de la méthode retenue pour réaliser la synthèse de produits différents. L'hétérogénéité des produits pose le problème crucial d'une mesure des quantités produites, indépendante du niveau des prix. La dissociation entre volume et valeur par l'utilisation d'indices synthétiques donne un élément de réponse. Les indices les plus souvent utilisés sont ceux de Laspeyres et de Paasche. On peut reprocher à l'auteur de ne pas insister sur les fondements théoriques de ceux-ci. Pourquoi pondérer des prix par des quantités de l'année de référence (et vice versa) plutôt que par celles de l'année observée ?

Le schéma des comptes d'exploitation, appliqué aux différents niveaux de production, entreprise, branche, nation, permet de servir de support à la définition des différentes mesures de la productivité : **productivité globale des facteurs, productivité nette du travail, productivité brute du travail.**

Parmi les différentes définitions de la **productivité globale des facteurs,** celle de **productivité globale exhaustive** incorpore tous les éléments du compte d'exploitation, mais suppose que ceux-ci peuvent être dissociés en volume et indice de prix.

La **productivité nette du travail** correspond à une mesure de la production au numérateur où l'on a retranché de la production brute les consommations courantes et les amortissements.

La **productivité brute du travail** correspond à une production brute au numérateur. Enfin la définition de la **productivité intégrale du travail** comprend au dénominateur l'ensemble des facteurs de production (les éléments financiers du compte d'exploitation mis à part) mesurés à partir du travail qu'ils incorporent ; à ce sujet, il est regrettable qu'aucune allusion n'ait été faite à la théorie de Marx sur la valeur travail.

Dans tous les cas, deux méthodes de mesure peuvent être utilisées : la **méthode des volumes** et la **méthode des prix.** Dans le premier cas, on divise un indice de volume de la production par un indice de volume des facteurs ; dans le deuxième cas, on divise un indice de prix des facteurs, par un indice de prix de la production.

***

Ces définitions permettent une comparaison dans le temps des productivités aux différents niveaux de production. Celles-ci ne seraient que partielles si on ne pouvait effectuer de comparaison dans l'espace.

L'agrégation des indices de productivité permet la comparaison des productivités de différentes branches de production. La méthode des valeurs fictives permet la comparaison des productivités entre économies nationales. Cette méthode suppose toutefois que les structures des économies nationales soient comparables puisque l'on pondère les quantités d'un pays par les prix de l'autre (méthode utilisée par Gilbert et Kravis).

Le problème de changement de structure est alors abordé à partir d'un exemple très simple, expliquant comment un indice moyen des prix peut différer notablement de la moyenne arithmétique des indices partiels (le fameux paradoxe) : l'application est rapidement faite à l'effet qualité.

***

Ces mesures de la productivité permettant des comparaisons dans le temps et dans l'espace, trouvent toute leur utilité dans leur application à des problèmes d'économie théorique ou pratique.

A des liaisons comptables, on peut substituer des relations causales et déterminer dans quelle mesure, ressources naturelles et population agissent sur la productivité. A ces liaisons causales, s'ajoutent des liaisons d'objectifs : déterminer comment se répartissent les fruits de la productivité, réduction des durées de travail, accroissement de la consommation des ménages, rapport population totale/population active... Le concept de productivité trouve encore un domaine d'application dans la théorie des échanges internationaux (coûts comparés). Enfin la notion de productivité est au cœur même de tous les problèmes de gestion économique, qu'il s'agisse de l'entreprise privée ou de la planification nationale.

La question se pose alors de savoir quel crédit accorder à une mesure qui peut prendre les formes les plus diverses et aboutir à des résultats numériques très différents, alors que le phénomène à mesurer est unique. De plus, certains biens paraissent hors valeur (ou du moins, leur valorisation suppose l'utilisation de techniques particulières) : la santé, la vie humaine, les services gratuits. Que devient alors le concept de productivité ?

G. MOUTET

VICTOROFF (D.). — Publicité et communications de masse.

Coopération, n° 1, 1969, pp. 21-24.

Dans le monde contemporain la publicité et les communications de masse prennent une place qui devient prépondérante de jour en jour. Les moyens les plus divers sont utilisés pour faire passer les messages publicitaires ; tous n'ont pas le même impact : la radio, la télévision et le cinéma sont plus persuasifs que la presse mais c'est le contact personnel qui est de loin le plus efficace.

Quelle que soit la nature du véhicule de diffusion collective employé, le message qu'il transmet n'agit en général qu'indirectement. Il faut tenir compte des prédispositions psychiques du public, d'autant plus importantes qu'elles sont totalement inconscientes, toute personne ayant tendance à choisir le message dont le contenu est conforme à ses opinions. (On a pu constater aux U.S.A., en 1959, lors d'une campagne télévisée, patronnée par le parti républicain, qu'elle fut suivie par un public de télespectateurs comportant deux fois plus de républicains que de démocrates.)

Simultanément un autre facteur médiatisant joue : ce sont les valeurs propres aux groupes. Le message diffusé par ces média a pour effet de renforcer les attitudes et opinions préexistantes plutôt que d'en créer de nouvelles ; les cas de conversion existent néanmoins lorsque les facteurs médiatisants ne jouent plus ou jouent à contre-temps.

Les contacts personnels se combinent avec les média pour renforcer la persuasion. Qui crée ces contacts personnels ? certains individus plus exposés que d'autres aux techniques de la diffusion collective jouent le rôle de guide d'opinion et répandent autour d'eux le message transmis par les média. Ce ne sont pas d'ailleurs les personnes les plus éminentes du groupe, mais celles qui sont les plus typiques. Une seule condition à leur réussite : agir très discrètement car si le public s'aperçoit de leur emprise il se produit alors une réaction très forte qui fait échouer toute tentative de persuasion.

D. Victoroff conclut en disant quelques mots de l'évolution récente des recherches concernant les communications de masse. Aux U.S.A., en G.B., les chercheurs se préoccupent de savoir ce que les individus récepteurs font des mass-media et ils se sont plus précisément attachés à l'étude de la télévision.

Il est frappant de voir que le public américain ne considère pas le message publicitaire comme une entité isolée ; au contraire, il ne fait pas de différence entre les « commercials » et l'ensemble du programme, ce qui d'ailleurs, ôte beaucoup de forces au message.

On est ainsi très loin de l'image d'une publicité agissant selon le modèle de conditionnement pur et simple, et il faut détruire le mythe concédant à la publicité moderne un pouvoir illimité.

Sylvie GUIRAUD

Trésor et pouvoir économique par **Pierre TRAIMOND,**

Collection « Connaissances Économiques » dirigée par Jean MARCHAL, librairie Cujas 1968 (320 p.).

Le pouvoir étant capacité de dominer et de manœuvrer, le pouvoir monétaire du Trésor est analysé dans les relations institutionnelles de cet intermédiaire financier privilégié avec la Banque de France. C'est sans conteste la meilleure méthode qui puisse permettre de montrer comment ces relations sont la source de création et de diffusion d'actifs et comment les circuits du Trésor sont intégrés au système monétaire. Il restait ensuite à l'auteur à opposer le pouvoir de régulation par l'action des dépenses publiques au pouvoir par l'action des stratégies monétaires (action par les taux d'intérêt, par le multiplicateur du crédit, par le contrôle de la liquidité).

Le grand mérite de cet ouvrage est dans l'intérêt et la précision de la documentation qui vient à l'appui de la thèse. Il s'agit à l'évidence de la documentation d'un spécialiste très au courant des pratiques du Trésor et qui possède de larges ouvertures sur le système monétaire dans son ensemble. Certains contesteront les conclusions de la première partie : il n'est pas prouvé que le pouvoir du Trésor « trouve sa raison d'être dans les carences du système bancaire ». Prudence n'est pas forcément carence et ce n'est pas le rôle des banques d'aventurer l'épargne des déposants ; la rentabilité est un des grands critères du crédit mais le risque n'est pas le moins important. Le Trésor n'a pas le même souci. D'aillleurs l'auteur a parfaitement montré comment le crédit à court terme accordé par les banques doit rester sous l'emprise de la Banque de France qui organise et contrôle efficacement le système bancaire, alors que le Trésor assure la « transformaion » des actifs en prêts à long terme, c'est-à-dire en investissement.

Cette critique n'enlève rien à la valeur d'un ouvrage qui devrait figurer dans la bibliothèque de chaque économiste, tant il est riche d'enseignement sur les circuits du Trésor et sur leur intégration au système monétaire.

Roger COSTE

Le directeur de la publication : G. DUNOD.

Dépôt légal : 4e trimestre 1969. Numéro 6169. Imprimé en France.

Imprimerie Nouvelle, Orléans. — N° 6041.

CONSOMMATION (ANNALES DU C. R. E. D. O. C.)

1965

Nº 1. — Quelle est la rentabilité des capitaux investis dans les logements en location ? — Analyse des phénomènes d'induction (Évolution de l'emploi dans le commerce par région entre 1954 et 1962). — Quelques réactions des ménages à l'égard de leur logement. — Un modèle des dépenses médicales. — La consommation en France de 1963 à 1964.

Nº 2. — Analyse économique et planification urbaine. — Louer ou acheter son logement. — Réflexions sur le rôle de l'avenir dans ce choix. — Les produits surgelés. — La consommation des boissons de 1960 à 1963. — La fréquentation des colonies de vacances jusqu'en 1964.

Nº 3. — Les études d'armature urbaine régionale. — Quelques problèmes posés par la prévision de la demande en services collectifs. — Conditions de logement et insatisfaction des ménages en 1961. — Les dépenses de location de voitures sans chauffeur.

Nº 4. — Le Plan, accélérateur de croissance. — L'ajustement de l'offre de viande à la demande. — Étude de la série épargne des ménages (1950-1964).

1966

Nº 1. — Recherche et aménagements urbains.

Nº 2. — La consommation des Français en 1964. — Étude bibliographique sur l'utilisation des services collectifs. — L'influence des facteurs économiques sur la consommation médicale. — L'influence de la Sécurité Sociale sur les dépenses médicales des exploitants agricoles.

Nº 3. — Les conditions du marché du logement et le comportement des ménages. — La consommation pharmaceutique des Français. — Les loisirs aux U.S.A. — Les jeunes ménages et leurs conditions de logement en 1963. — La consommation en France en 1964-1965.

Nº 4. — Une méthode pour étudier la solvabilité de la demande de logement. — La loi et les travaux d'Engel. — Le « Federal Reserve Board » et les recherches sur l'épargne.

1967

Nº 1. — Une étude économétrique de la demande de viande. — La consommation des Français en 1965. — Intégration des méthodes d'approche psycho-sociologiques à l'étude de l'épargne.

Nº 2. — Un indicateur de la morbidité appliqué aux données d'une enquête sur la consommation médicale. — La diffusion des services collectifs : phénomène économique ou social ? — Les travaux de préparation du Ve Plan et l'élaboration d'un modèle national de fonctionnement du marché du logement. — Les conditions de vie des familles.

Nº 3. — L'épargne des exploitants agricoles. — Structure et équilibre du marché du textile. — Les dépenses touristiques.

1968

Nº 1. — Étude critique de méthodes d'enquête. — Étude sur l'offre et la demande de créance.

Nº 2. — Théorie et politique de l'épargne. — Un modèle prévisionnel de la demande de logements. — L'évolution de la consommation de viande.

Nº 3. — La consommation et la demande de monnaie. — Valeur prédictive des intentions d'achats au niveau du ménage pris individuellement.

Nº 4. — Quelques éléments sur le comportement des propriétaires vis-à-vis du prix du logement acheté et de la mise de fonds versée. — Facteurs « irrationnels » de l'offre d'épargne (recherches allemandes).

1969

Nº 1. — L'offre de monnaie par les banques commerciales. — L'économie des services de soins médicaux en France. — L'évolution de la consommation de produits laitiers de 1950 à 1966.

Nº 2. — L'économie des services de soins médicaux en France. — La formation de l'épargne liquide (l'exemple du Crédit Mutuel). — Consommation individuelle et consommation collective. — Étude sur la demande en logement des ménages.

SOMMAIRE DES PROCHAINS NUMÉROS

Durée d'observation et précision dans les enquêtes de consommation. — Influence des caractéristiques de l'offre sur le recours aux services collectifs. — La consommation en France de 1959 à 1968. — Un essai de classification de titres boursiers, fondée sur l'analyse factorielle.

sommaire

ÉTUDES

MÉTHODOLOGIE

BIBLIOGRAPHIE

CENTRE DE RECHERCHES ET DE DOCUMENTATION SUR LA CONSOMMATION

45, boulevard de la Gare, PARIS-13e

Tél. POR. 97-59

1969 nº 3

juillet septembre