

# REVUE DE STATISTIQUE APPLIQUÉE

J. L. MALLET

## **Contribution à l'étude des facteurs non orthogonaux en analyse factorielle**

*Revue de statistique appliquée*, tome 19, n° 1 (1971), p. 57-76

[http://www.numdam.org/item?id=RSA\\_1971\\_\\_19\\_1\\_57\\_0](http://www.numdam.org/item?id=RSA_1971__19_1_57_0)

© Société française de statistique, 1971, tous droits réservés.

L'accès aux archives de la revue « *Revue de statistique appliquée* » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

*Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques*  
<http://www.numdam.org/>

# CONTRIBUTION A L'ÉTUDE DES FACTEURS NON ORTHOGONAUX EN ANALYSE FACTORIELLE

J. L. MALLET

Centre de Recherches Pétrographiques et Géochimiques (Nancy)

*Une analyse factorielle effectuée sur une matrice de corrélation  $C(m \times m)$  ayant fait ressortir  $r(r < m)$  facteurs orthogonaux et une matrice de saturation  $A(m \times r)$  telle que  $C = A A^T$ , l'auteur a déterminé un procédé qui permet d'obtenir des facteurs obliques plus facilement interprétables ; Etant entendu qu'un facteur est d'autant plus interprétable qu'il est lié fortement avec certaines variables et orthogonal aux autres.*

*La matrice de saturation  $V$  associée à de tels facteurs est caractérisée par des colonnes composées de 0 et de 1 (ou  $-1$ ), et des lignes ne comportant qu'un seul 1 (ou  $-1$ ). Si l'on considère les colonnes  $p$  et  $q$  de  $V$  ceci se traduit par la relation :*

$$\sum_{j=1}^m V_{jp} \cdot V_{jq} = 0$$

*Soit  $N_{pq}$  la quantité suivante associée aux colonnes  $p$  et  $q$  de  $V$  :*

$$N_{pq} = \sum_{j=1}^m V_{jp} \cdot V_{jq}$$

*L'auteur cherche une matrice de rotation  $\Lambda$  telle que  $V = A\Lambda$  et telle que  $\sum_{p,q} N_{pq}$  soit minimum.*

*A la suite de cet exposé il donne un exemple d'application de sa méthode ; celui-ci paraît être un peu trop typé pour être significatif.*

*Toutefois la méthode proposée est fort ingénieuse. Encore demande-t-elle les hypothèses de normalité, les facteurs ne pouvant être interprétés que dans une population homogène (ou normale).*

*Ce qui restreint peut-être quelque peu son champ d'application.*

## INTRODUCTION

### Position du problème

#### But de la méthode

Un des problèmes qui se posent en analyse factorielle est, à partir d'une solution orthogonale initiale, trouver une nouvelle solution plus facilement interprétable que la première.

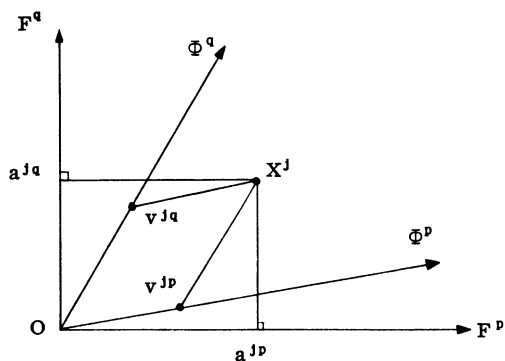


Figure 1

Afin de préciser cette notion de "solution plus facilement interprétable", considérons dans l'espace des facteurs orthogonaux contravariants initiaux  $F^q$  (de vecteurs unitaires covariants  $\vec{F}_q$ ), l'ensemble des projections des  $m$  variables contravariantes  $X^j$ .

Dans l'espace des  $r$  facteurs  $F^p$ , chaque variable  $X^j$  est repérée par ses  $r$  composantes (contravariantes)  $a^{jp}$ , et l'ensemble des  $a^{jp}$  constitue ce que l'on appelle la matrice factorielle (contravariante)  $A = (a^{jp})$  à  $m$  lignes et  $r$  colonnes pour laquelle on a toujours les relations  $|a^{jp}| < 1$ .

#### Solution plus facilement interprétable

Nous dirons que la matrice carrée  $\Lambda$  de rang  $r$  transforme la matrice  $A$  en une matrice  $V = [v^{jp}]$  plus facilement interprétable si les  $|v^{jp}|$  sont soit "voisins" de 1, soit "voisins" de 0. En d'autres termes, cela revient à dire que l'opérateur  $\Lambda$  fait passer du système de coordonnées (contravariantes)  $F^p$  au système de coordonnées (contravariantes)  $\Phi^p$  (de vecteurs unitaires  $\vec{\Phi}_p$ ) (cf. fig. 1) de telle façon que les points  $X^j$  soient "voisins" de l'un des  $\Phi^p$ .

#### Contrainte imposée à $\Lambda$

Désignons par  $\Lambda_p$  le  $p^{1\text{ème}}$  vecteur colonne de  $\Lambda$  : on remarque alors que dans le cas où  $\vec{\Phi}_p$  et  $\vec{\Phi}_q$  sont orthonormés, la matrice  $\Lambda$  doit elle-même être orthogonale, c'est-à-dire, entre autre, que ses vecteurs  $\Lambda_p$  et  $\Lambda_q$  doivent être chacun de module unité. Dans le cas général où  $\vec{\Phi}_p$  et  $\vec{\Phi}_p$  ne sont pas forcément orthogonaux, nous conserverons la contrainte :

$$||\Lambda_p|| = 1, \quad \forall p$$

En définitive, on a (1).

$$V = A \cdot \Lambda$$

$$\text{avec } \Lambda_p^T \cdot \Lambda_p = 1$$

-----  
 (1) dans ce qui suit on désignera par un caractère T en indice supérieur, la transposée d'une matrice :

$$\text{Transposée } (M) = M^T$$

### Remarque

Afin que l'on puisse trouver une matrice  $\Lambda$ , il faut s'imposer, sur la solution  $V$  des conditions traduisant le fait que les  $|v^{jp}|$  doivent être "voisins" (en un certain sens) de 0 ou de 1. La formulation de ces conditions constitue ce que l'on appelle un critère, et on dit alors que les  $|v^{jp}|$  sont "voisins" de 0 ou de 1 au sens du critère considéré.

### Critère utilisé

#### Introduction

Supposons qu'il n'y ait qu'un seul point  $X^j$ , il est évident, sur la figure 1, qu'une solution sera d'autant plus facilement interprétable que pour tout couple de facteurs  $\Phi^p$  et  $\Phi^q$ , le produit  $|v^{jp}| \cdot |v^{jq}|$  est plus petit (compte-tenu que  $0 \leq |v^{jp}| \leq 1$ ), ou encore que  $(v^{jp})^2 \cdot (v^{jq})^2$  est plus petit.

#### Critère

Dans le cas où il y a  $m$  variables, le critère utilisé consiste à chercher la matrice  $\Lambda$  telle que...

$$N_{pq} = \sum_{j=1}^m (v^{jp})^2 \cdot (v^{jq})^2$$

... soit minimum quel que soit  $p \neq q$ .

#### Remarque 1

La méthode quartimin due à Carroll consiste à minimiser la quantité :

$$N = \sum_{p < q}^r \sum_{j=1}^m (v^{jp})^2 \cdot (v^{jq})^2 = \sum_{p < q}^r N_{pq}$$

Remarquons alors que, contrairement à ce que semble suggérer Carroll, rien ne prouve que le fait de minimiser  $N$  globalement équivaut à minimiser séparément les diverses quantités  $N_{pq}$  (compte-tenu qu'elles ne sont pas indépendantes), et il n'est pas sûr du tout que ce faisant on obtienne une "solution plus facilement interprétable". La suggestion de Carroll serait vraie si les  $N_{pq}$  étaient indépendants, or ils sont deux à deux liés par des termes communs : ainsi  $N_{pq}$  et  $N_{qs}$  ont en commun les quantités  $(v^{jq})^2$  ce qui introduit fatalement un biais dans la méthode proposée par Carroll.

#### Remarque 2

Carroll visiblement non satisfait de sa méthode quartimin a proposé une nouvelle méthode, la méthode covarimin qui, elle, est manifestement fautive. L'idée de cette méthode est celle-ci : on cherche à rendre le plus disparate possible les quantités  $|v^{jp}|$  et  $|v^{jq}|$  ; pour ce faire, Carroll suggère de minimiser la covariance des quantités  $(v^{jp})^2$  et  $(v^{jq})^2$  et dans son esprit, "minimiser" signifie "rendre voisin de 0" ; or ceci est complètement faux car une covariance de quantités  $(v^{jp})^2$  et  $(v^{jq})^2$  positives ou nulles peut très bien être minimum tout en étant égale à un nombre négatif très grand, en valeur absolue, c'est-à-dire qu'il y aura alors un lien étroit entre les  $|v^{jp}|$  et  $|v^{jq}|$ . On

s'aperçoit en définitive que cette façon de faire risque de conduire au résultat inverse de celui escompté !

## RECHERCHE D'UN ALGORITHME PERMETTANT DE TROUVER $\Lambda$

### Introduction

Nous aurons dans l'étude qui va suivre besoin des quantités ci-dessous :

$$\left[ \begin{array}{l} A_j = j^{\text{ème}} \text{ vecteur ligne de } A \\ \Lambda_q = q^{\text{ème}} \text{ vecteur colonne de } \Lambda \\ H_q = \sum_{j=1}^n [(\Lambda_q^T \cdot A_j^T \cdot A_j \cdot \Lambda_q) \cdot A_j^T A_j] \end{array} \right.$$

Remarquons que  $H_q = H(A, \Lambda_q)$  est une matrice manifestement symétrique.

### Théorème 1

#### Théorème

$$\left[ \begin{array}{l} \text{On a la relation :} \\ N_{pq} = \Lambda_p^T \cdot H_q \cdot \Lambda_p \end{array} \right.$$

#### Démonstration

On a  $V = A \cdot \Lambda$  ; avec les notations adoptées en introduction, on peut écrire ceci sous la forme :

$$V = \begin{bmatrix} A_1 \\ \vdots \\ A_j \\ \vdots \\ A_n \end{bmatrix} \cdot [\Lambda_1 \cdots \Lambda_q \cdots \Lambda_r]$$

en faisant le produit ci-dessus blocs par blocs, on obtient : ...

$$v^{jp} = A_j \cdot \Lambda_p = \Lambda_p^T \cdot A_j^T$$

... car  $v^{jp}$  est un scalaire.

Il s'en suit que l'on peut écrire :

$$(v^{jp})^2 = \Lambda_p^T A_j^T A_j \Lambda_p$$

$$(v^{jq})^2 = \Lambda_q^T A_j^T A_j \Lambda_q$$

$$\iff (v^{jq})^2 \cdot (v^{jp})^2 = (\Lambda_q^T A_j^T A_j \Lambda_q) \cdot (\Lambda_p^T A_j^T A_j \Lambda_p)$$

$$\begin{aligned}
\iff \sum_{j=1}^m (v^{jq})^2 \cdot (v^{jp})^2 &= \sum_{j=1}^m (\Lambda_q^T A_j^T A_j \Lambda_q) \cdot (\Lambda_p^T A_j^T A_j \Lambda_p) \\
&= \Lambda_p^T \cdot \left\{ \sum_{j=1}^m [(\Lambda_q^T A_j^T A_j \Lambda_q) \cdot A_j^T A_j] \right\} \cdot \Lambda_p \\
&= \Lambda_p^T H_p \Lambda_p \qquad \text{c. q. f. d.}
\end{aligned}$$

**Théorème 2**

**Théorème**

$N_{pq} = \Lambda_p^T H_q \Lambda_p$  minimum compte tenu que  $\|\Lambda_p\| = 1 \iff \Lambda_p$  est le vecteur propre associé à la plus petite valeur propre  $\mu_q$  de  $H_q$  et dans ce cas  $\mu_q = N_{pq}$

**Démonstration**

posons  $\left\{ \begin{array}{l} H_q = H(A, \Lambda_q) \\ \Lambda_p = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_r \end{bmatrix}_p \end{array} \right.$

comme  $H_q$  est manifestement symétrique et indépendant de  $\Lambda_p$ , on peut écrire :

$$N_{pq} = \Lambda_p^T H_q \Lambda_p$$

par ailleurs, on doit avoir :

$$\Lambda_p^T \cdot \Lambda_p = 1$$

Le problème est donc de minimiser  $N_{pq} = \Lambda_p^T H_q \Lambda_p$  sous la condition  $\Lambda_p^T \Lambda_p = 1$ . Dans ce cas, la recherche d'une solution  $\Lambda_p$  se fait en utilisant les multiplicateurs de Lagrange  $\mu_q$  : pour celà on écrit...

$$T_{pq} = \Lambda_p^T H_q \Lambda_p - \mu_q \cdot (\Lambda_p^T \cdot \Lambda_p - 1)$$

... et on cherche le vecteur  $\Lambda_p$  tel que l'on ait  $\partial T_{pq} / \partial \Lambda_p = 0$

$$\frac{\partial T_{pq}}{\partial \Lambda_p} = 2 \cdot H_q \Lambda_p - 2 \mu_q \cdot \Lambda_p$$

$$\left\{ \begin{array}{l} \frac{\partial T_{pq}}{\partial \Lambda_p} = 0 \iff H_q \Lambda_p - \mu_q \cdot \Lambda_p = 0 \quad (I) \\ \iff H_q \Lambda_p = \mu_q \cdot \Lambda_p \\ \iff \Lambda_p \text{ est un vecteur propre de } H_q \text{ associé à la valeur propre } \mu_q. \end{array} \right.$$

Pour déterminer  $\mu_q$ , reprenons la relation (I) :

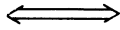
$$H_q \Lambda_p - \mu_q \Lambda_p = 0$$

multiplions à gauche cette équation par  $\Lambda_p^T$  :

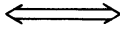
$$\Lambda_p^T H_q \Lambda_p - \mu_q \Lambda_p^T \Lambda_p = 0$$

mais

$$\Lambda_p^T \cdot \Lambda_p = 1$$



$$\mu_q = \Lambda_p^T H_q \Lambda_p$$



$$\mu_q = N_{pq} \quad \text{c. q. f. d.}$$

pour être minimum,  $N_{pq}$  doit donc bien être égal à la plus petite valeur propre  $\mu_q$  de  $H_q$ .

### Corollaire 1

[ La matrice  $H_q$  est non négative

Il est facile de voir que  $H_q$  est non négative car sa plus petite valeur propre  $\mu_q$  est égale à la quantité  $N_{pq}$  positive ou nulle.

### Corollaire 2

[ Pour  $\Lambda_q$  fixé et  $||\Lambda_p|| = 1$  fixé, la fonction  $N_{pq}(\Lambda_p, \Lambda_q)$  (fonction de  $\Lambda_p$ ) présente un minimum absolu  $N_{pq}$  et aucun minimum local. Le minimum absolu est obtenu deux fois pour deux valeurs opposées  $\Lambda_p = \pm \Lambda_p^*$

Supposons pour fixer les idées que  $\Lambda_q$  est fixé et que  $\Lambda$  est de taille  $(2 \times 2)$  ; posons :

$$\left\{ \begin{array}{l} \Lambda_p = \begin{bmatrix} x \\ y \end{bmatrix} \\ H_q = H(A, \Lambda_q) = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \\ z = \Lambda_p^T \cdot H_q \cdot \Lambda_p \end{array} \right.$$

On peut alors écrire :

$$z = ax^2 + 2 bxy + cy^2$$

Cette équation représente dans l'espace  $(x, y, z)$  soit un parabolofde dont l'axe de symétrie est parallèle à  $oz$ , soit un cylindre parabolique dont les génératrices sont parallèles au plan  $(x, y)$ . Ce parabolofde ou ce cylindre parabolique a forcément sa concavité tournée vers les  $z$  positifs car  $a$  et  $c$  sont toujours positifs ou nuls ; remarquons de plus que la surface  $z(x, y)$  passe par l'origine car :

$$\left. \begin{array}{l} x = 0 \\ y = 0 \end{array} \right\} \Rightarrow z = 0$$

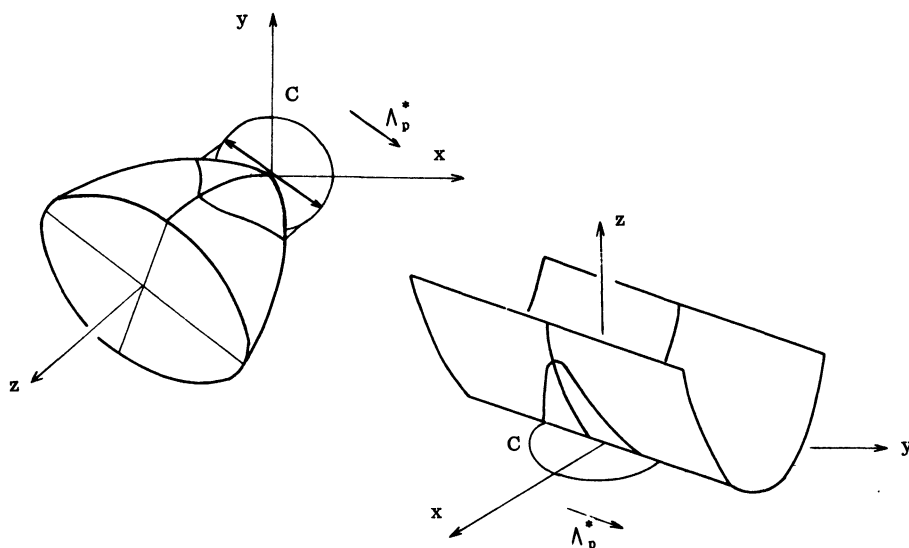


Figure 2

Comme les vecteurs  $\Lambda_p$  doivent être de module unité, leurs extrémités doivent se trouver sur le cercle C de rayon unité tracé dans le plan  $(x,y)$ . Construisons le cylindre D dont les génératrices parallèles à  $oz$  s'appuient sur le cercle C ; si nous traçons l'intersection (biquadratique) de ce cylindre avec la surface  $z(x,y)$ , nous voyons (cf. fig. 2) que cette intersection présente un minimum unique  $z^*$  de  $z$  obtenu pour  $(x = +x^*, y = +y^*)$  et  $(x = -x^*, y = -y^*)$ , c'est-à-dire pour deux vecteurs opposés  $\Lambda_p = \pm \Lambda_p^*$  avec :

$$\Lambda_p^* = \begin{bmatrix} x^* \\ y^* \end{bmatrix}$$

On voit sur la figure 2, et on peut démontrer qu'il n'y a pas de minimum local  $z^*$  le long de l'intersection de D avec la surface  $z(x,y)$ , ce qui démontre la proposition.

#### Remarque

La démonstration que nous venons de faire peut être extrapolée au cas où  $\Lambda$  est de taille supérieure à  $(2 \times 2)$ , mais cela nous intéresse peu car, comme nous le verrons par la suite, nous nous ramènerons toujours au cas où l'on a à traiter des matrices  $\Lambda$  de taille  $(2 \times 2)$ .

#### Détermination de $\Lambda_p$ et $\Lambda_q$ par un algorithme pour p et q fixés

##### Introduction

Désignons par...



$$\begin{cases} N_{pq} = f_1 [A, \Lambda_p, \Lambda_q] \\ W_q = f_2 [A, \Lambda_q] \end{cases}$$

... le procédé de calcul qui à tout vecteur  $\Lambda_q$  associe le vecteur propre  $W_q$  correspondant à la plus petite valeur propre  $N_{pq}$  de  $H_q = H(A, \Lambda_q)$

Algorithme

L'algorithme proposé est représenté par l'organigramme de la figure 3 :  
 on commence par se donner une valeur  $(\Lambda_q)_0$  initiale quelconque et on en déduit  $(\Lambda_p)_1 = f_2 [A, (\Lambda_q)_0]$   
 on calcule ensuite  $(\Lambda_q)_1 = f_2 [A, (\Lambda_p)_1]$  qui est une valeur plus proche de la solution finale  $\Lambda_q$  que  $(\Lambda_q)_0$   
 on calcule ensuite  $(\Lambda_p)_2 = f_2 [A, (\Lambda_q)_1]$  qui est une valeur plus proche de la solution finale  $\Lambda_p$  que  $(\Lambda_p)_1$   
 .....  
 .....

La convergence du procédé est testée sur l'évolution de la valeur du paramètre  $(N_{pq})_1 = f_1 [A, (\Lambda_p)_1, (\Lambda_q)_1]$

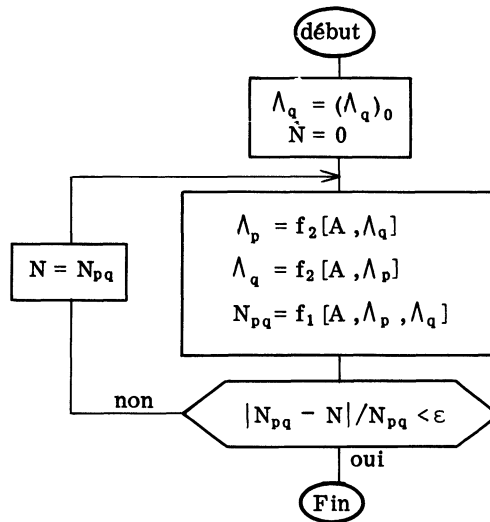


Figure 3

Convergence du procédé

La convergence du procédé est assurée par le fait que l'hypersurface  $N_{pq}(\Lambda_p, \Lambda_q)$  présente pour  $\Lambda_q$  (ou  $\Lambda_p$ ) fixé un minimum absolu unique ; pour s'en convaincre, il suffit de considérer la figure 4 :

Nous voyons que lorsqu'on se déplace sur la surface  $N_{pq}(\Lambda_p, \Lambda_q)$  suivant le procédé décrit au paragraphe précédent, la quantité  $N_{pq}$  ne peut pas augmenter, ce qui démontre la convergence du procédé.

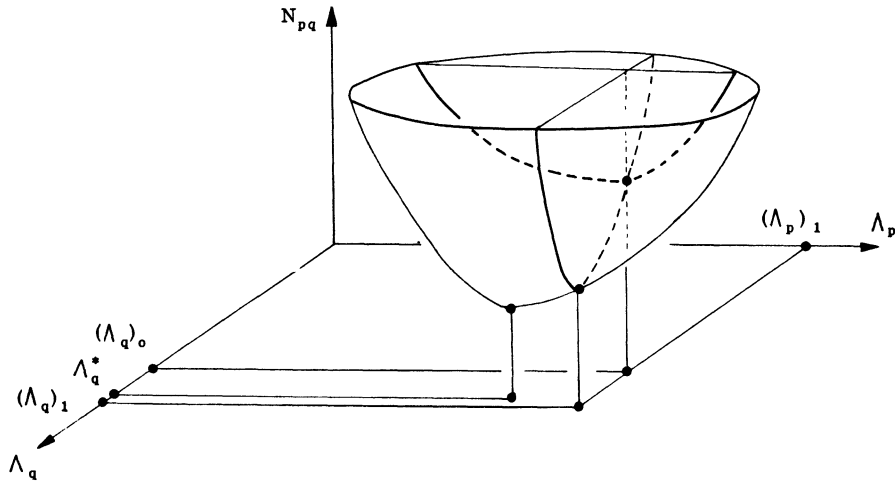


Figure 4

### Remarque

L'algorithme décrit ci-dessus est très efficace dans le cas où l'on a seulement deux facteurs  $p$  et  $q$  et que l'on opère la transformation  $(F^p, F^q) \xrightarrow{\Lambda} (\Phi^p, \Phi^q)$  dans le plan  $(p, q)$  ; on observe en effet une convergence très forte, et il faut prendre le test d'arrêt  $\varepsilon$  de l'ordre de  $10^{-2}$  ou  $10^{-3}$  si on veut espérer faire plus de 3 itérations.

Malheureusement il ne nous a pas été possible d'utiliser cet algorithme tel quel dans le cas où l'on a plus de deux facteurs ; en effet, quel que soit le vecteur initial  $(\Lambda_q)_0$ , on obtiendra toujours les 2 *mèmes* vecteurs  $\Lambda_p$  et  $\Lambda_q$ , et on voit ainsi que le procédé de calcul décrit ci-dessus ne nous permet de trouver que deux colonnes de la matrice  $\Lambda$ . Afin d'éviter cet ennui, nous avons été conduit à chercher un nouvel algorithme dérivé du précédent et permettant d'arriver à la solution de notre problème :

### ALGORITHME PROPOSE

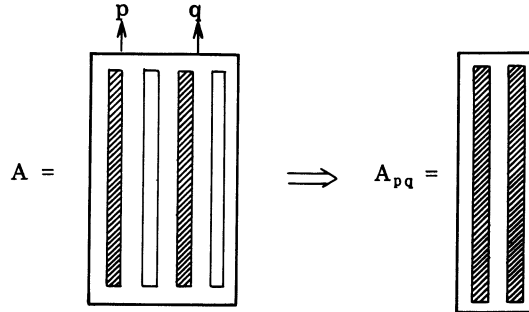
#### Introduction

L'idée directrice de la méthode proposée est assez proche de la méthode Jacobi servant à chercher les valeurs propres d'une matrice carrée symétrique  $M$ . Dans cette dernière, on fait tourner deux par deux dans leur plan les directions orthogonales de départ (qui restent orthogonales) de façon à les amener par itérations successives à être confondues avec les directions propres de la matrice  $M$  ; dans notre méthode, on va opérer de même : on fait tourner deux par deux les directions  $(F^p, F^q)$  orthogonales de départ (qui ne restent plus orthogonales) de façon à se rapprocher de la solution cherchée par itérations successives.

## Rotation dans le plan des facteurs $F^p$ et $F^q$

### Introduction

On commence par construire la matrice  $A_{pq}$  obtenue en sélectionnant les colonnes  $p$  et  $q$  de la matrice  $A$  :



on cherche ensuite, grâce à l'algorithme décrit par l'organigramme de la figure 3, la matrice carrée  $\Lambda_{pq}$  non orthogonale qui transforme  $A_{pq}$  en  $V_{pq}$ ...

$$\underbrace{\begin{array}{|c|c|} \hline \text{shaded bar} & \text{shaded bar} \\ \hline \end{array}}_{V_{pq}} = \underbrace{\begin{array}{|c|c|} \hline \text{shaded bar} & \text{shaded bar} \\ \hline \end{array}}_{A_{pq}} \cdot \underbrace{\begin{array}{|c|c|} \hline \lambda_1^1 & \lambda_1^2 \\ \hline \lambda_2^1 & \lambda_2^2 \\ \hline \end{array}}_{\Lambda_{pq}}$$

... de façon que si  $v^{js}$  est un élément de  $V_{pq}$ , la quantité...

$$N_{pq} = \sum_{j=1}^n (v^{jp})^2 \cdot (v^{jq})^2$$

... soit minimum.

### Remarque :

L'algorithme de la figure 3 ne suffit en fait pas à construire  $\Lambda_{pq}$ ; en effet, il nous fournit deux vecteurs  $\Lambda_1^*$  et  $\Lambda_2^*$ , ce qui donne 4 possibilités pour  $\Lambda_{pq}$  :

$$\begin{aligned} \Lambda_{pq} &= \begin{bmatrix} \boxed{\Lambda_1^*} & \boxed{\Lambda_2^*} \end{bmatrix} & \Lambda_{pq} &= \begin{bmatrix} \boxed{\Lambda_2^*} & \boxed{\Lambda_1^*} \end{bmatrix} \\ \Lambda_{pq} &= \begin{bmatrix} \boxed{-\Lambda_1^*} & \boxed{-\Lambda_2^*} \end{bmatrix} & \Lambda_{pq} &= \begin{bmatrix} \boxed{-\Lambda_2^*} & \boxed{-\Lambda_1^*} \end{bmatrix} \end{aligned}$$

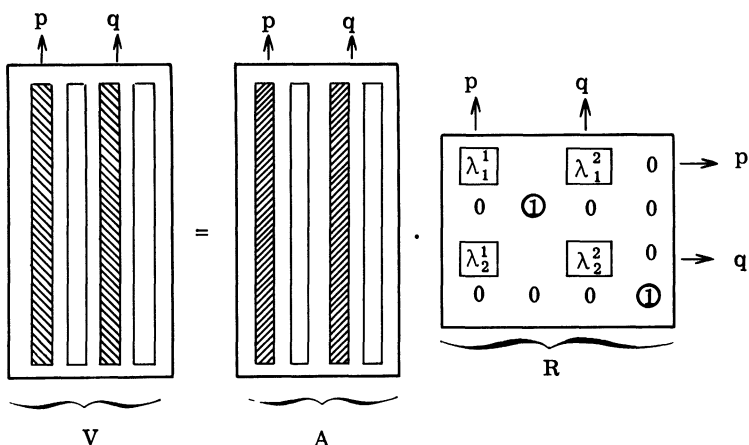
Afin d'éviter les permutations de colonnes et des changements de signe quand on passe de  $A_{pq}$  à  $V_{pq}$ , on choisira la combinaison  $\Lambda_{pq}$  telle que  $|\lambda_1^1| > |\lambda_1^2|$  et telle que l'on ait :  $\lambda_1^1 > 0$  et  $\lambda_1^2 > 0$

Matrice de "rotation oblique" R

Nous appellerons matrice de "rotation oblique" la matrice de transformation  $R = \{r_j^i\}$  dans l'espace des  $r$  facteurs  $F^p$  obtenue en remplaçant dans la matrice unité de taille  $(r \times r)$  les termes suivants par ceux de  $\Lambda_{pq}$  :

$$\left\{ \begin{array}{l} r_p^p = \lambda_1^1 \\ r_p^q = \lambda_1^2 \\ r_q^p = \lambda_2^1 \\ r_q^q = \lambda_2^2 \end{array} \right.$$

⇒ on constate que cette matrice ne transforme que les colonnes p et q de la matrice A et que ces deux colonnes transformées se trouvent être égales aux deux colonnes de  $V_{pq}$  :



Remarque

La matrice de "rotation oblique" R joue ici le même rôle que la matrice de rotation orthogonale  $\Omega$  utilisée dans la méthode de Jacobi pour faire tourner les axes p et q dans leur plan d'un angle  $\theta$  (cf. fig. 5) :

En effet, la matrice  $\Omega = \{\omega_j^i\}$  est alors obtenue en remplaçant dans la matrice unité de taille  $(r \times r)$  les termes suivants par...

$$\omega_p^p = \cos \theta$$

$$\omega_p^q = -\sin \theta$$

$$\omega_q^p = \sin \theta$$

$$\omega_q^q = \cos \theta$$

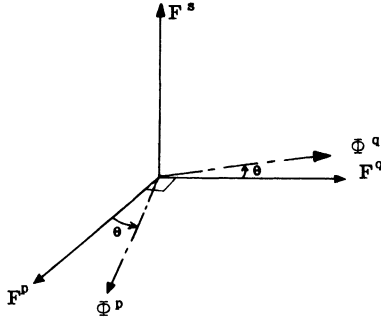


Figure 5

... c'est-à-dire par les éléments de la matrice  $M_{pq}$  de rotation dans le plan  $(F^p, F^q)$  :

$$M_{pq} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

On voit alors l'analogie étroite qu'il y a entre les matrices  $\Lambda_{pq}$  et  $M_{pq}$  d'une part et les matrices  $R$  et  $\Omega$  d'autre part.

#### Etude des tenseurs métriques fondamentaux associés au repère factoriel $\{F^p\}$

##### Introduction

Supposons pour fixer les idées que  $r = 3$ .

soient  $\vec{F}_1, \vec{F}_2, \vec{F}_3$  les vecteurs unitaires covariants de la base du repère factoriel  $F^1, F^2, F^3$ . A cette base de vecteurs  $\vec{F}_p$  covariants, on associe une autre base formée de vecteurs  $\vec{F}^p$  contravariants et tels que ...

$$\left[ \vec{F}_p \cdot \vec{F}^q = \delta_p^q = \begin{cases} 1 & \text{si } p = q \\ 0 & \text{si } p \neq q \end{cases} \right.$$

... c'est-à-dire que l'on se trouve dans le cas de la figure 6.

On sait que sous l'effet d'une matrice de transformation  $R = \{r_j^i\}$ , les vecteurs contravariants  $\vec{F}^p$  se transforment en vecteurs contravariants  $\vec{\Phi}^p$  suivant la relation suivante :

$$[\vec{\Phi}^1 \vec{\Phi}^2 \vec{\Phi}^3] = [\vec{F}^1 \vec{F}^2 \vec{F}^3] \cdot R$$

Evidemment, à ces vecteurs de base contravariants  $\vec{\Phi}^p$ , on peut associer les vecteurs covariants  $\vec{\Phi}_p$  transformés des vecteurs  $\vec{F}_p$  dans la transformation  $R$  :

$$\left[ \vec{\Phi}_p \cdot \vec{\Phi}^q = \delta_p^q = \begin{cases} 1 & \text{si } p = q \\ 0 & \text{si } p \neq q \end{cases} \right.$$

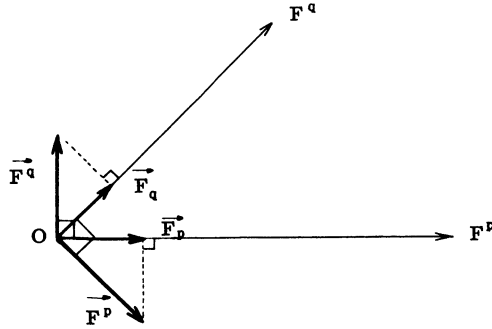


Figure 6

**Définition 1**

On appellera tenseurs métriques fondamentaux covariants associés aux repères  $\{\vec{F}_p\}$  et  $\{\vec{\Phi}_p\}$  les tableaux  $G^F$  et  $G^\Phi$  tels que ;

$$G^F = \begin{bmatrix} \vec{F}_1 \\ \vec{F}_2 \\ \vec{F}_3 \end{bmatrix} \cdot [\vec{F}_1 \ \vec{F}_2 \ \vec{F}_3] ; \quad G^\Phi = \begin{bmatrix} \vec{\Phi}_1 \\ \vec{\Phi}_2 \\ \vec{\Phi}_3 \end{bmatrix} \cdot [\vec{\Phi}_1 \ \vec{\Phi}_2 \ \vec{\Phi}_3]$$

on désignera par  $g_{pq}^F = \vec{F}_p \cdot \vec{F}_q$  et  $g_{pq}^\Phi = \vec{\Phi}_p \cdot \vec{\Phi}_q$  les éléments des tableaux  $G^F$  et  $G^\Phi$  :

$$\boxed{\begin{aligned} G^F &= \{g_{pq}^F\} = \{\vec{F}_p \cdot \vec{F}_q\} \\ G^\Phi &= \{g_{pq}^\Phi\} = \{\vec{\Phi}_p \cdot \vec{\Phi}_q\} \end{aligned}}$$

**Définition 2**

De même, on appellera tenseurs métriques fondamentaux contravariants associés aux repères  $\vec{F}^p$  et  $\vec{\Phi}^p$  les tableaux  $G_F$  et  $G_\Phi$  tels que :

$$G_F = \begin{bmatrix} \vec{F}^1 \\ \vec{F}^2 \\ \vec{F}^3 \end{bmatrix} \cdot [\vec{F}^1 \ \vec{F}^2 \ \vec{F}^3] ; \quad G_\Phi = \begin{bmatrix} \vec{\Phi}^1 \\ \vec{\Phi}^2 \\ \vec{\Phi}^3 \end{bmatrix} \cdot [\vec{\Phi}^1 \ \vec{\Phi}^2 \ \vec{\Phi}^3]$$

on désignera par  $g^{pq}_F = \vec{F}^p \cdot \vec{F}^q$  et  $g^{pq}_\Phi = \vec{\Phi}^p \cdot \vec{\Phi}^q$  les éléments des tableaux  $G_F$  et  $G_\Phi$  :

$$\boxed{\begin{aligned} G_F &= \{g^{pq}_F\} = \{\vec{F}^p \cdot \vec{F}^q\} \\ G_\Phi &= \{g^{pq}_\Phi\} = \{\vec{\Phi}^p \cdot \vec{\Phi}^q\} \end{aligned}}$$

### Remarque

Les relations...

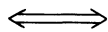
$$\left\{ \begin{array}{l} \vec{F}_p \cdot \vec{F}^q = \delta_p^q \\ \vec{\Phi}_p \cdot \vec{\Phi}^q = \delta_p^q \end{array} \right.$$

... nous indiquent que l'on a les relations suivantes entre les divers tenseurs métriques fondamentaux :

$$\boxed{\begin{array}{l} G^F = G_F^{-1} \\ G^\Phi = G_\Phi^{-1} \end{array}}$$

### Relation entre $G_F$ et $G_\Phi$

$$\begin{aligned} G_\Phi &= \begin{bmatrix} \vec{\Phi}^1 \\ \vec{\Phi}^2 \\ \vec{\Phi}^3 \end{bmatrix} \cdot [\vec{\Phi}^1 \ \vec{\Phi}^2 \ \vec{\Phi}^3] \\ &= \begin{bmatrix} \vec{\Phi}^1 \\ \vec{\Phi}^2 \\ \vec{\Phi}^3 \end{bmatrix} \cdot [\vec{F}^1 \ \vec{F}^2 \ \vec{F}^3] \cdot R \\ &= R^T \cdot \begin{bmatrix} \vec{F}^1 \\ \vec{F}^2 \\ \vec{F}^3 \end{bmatrix} \cdot [\vec{F}^1 \ \vec{F}^2 \ \vec{F}^3] \cdot R \end{aligned}$$



$$\boxed{G_\Phi = R^T \cdot G_F \cdot R} \quad \text{(III)}$$

### Normalisation des vecteurs $\vec{\Phi}_p$

#### Introduction

Si les vecteurs  $\vec{F}_p$  sont normés, il n'en est pas forcément de même des vecteurs  $\vec{\Phi}_p$  car la matrice de transformation  $R$  n'est en général pas orthogonale. Le problème est donc maintenant de trouver des vecteurs unitaires  $\vec{\Phi}_p^*$  colinéaires aux vecteurs  $\vec{\Phi}_p$ , ainsi que les matrices  $(G^\Phi)^*$ ,  $(G_\Phi)^*$ , et  $V^*$  associées.

Normalisation du tenseur  $G^\Phi$

Soient  $g_{pq}^\Phi = \vec{\Phi}_p \cdot \vec{\Phi}_q$  les éléments deux fois covariants du tenseur  $G^\Phi$ . Nous nous proposons de construire un nouveau tenseur  $(G^\Phi)^*$  dont les éléments  $(g_{pq}^\Phi)^* = \vec{\Phi}_p^* \cdot \vec{\Phi}_q^*$  sont obtenus à partir des vecteurs unitaires  $\vec{\Phi}_p^*$  et  $\vec{\Phi}_q^*$  colinéaires aux vecteurs  $\vec{\Phi}_p$  et  $\vec{\Phi}_q$ . Dans ce but construisons la matrice diagonale de transformation  $\Omega$  dont les éléments diagonaux  $\omega_p^p$  sont tels que  $\omega_p^p = \sqrt{g_{pp}^\Phi}$ .

$$\Omega = \begin{bmatrix} \sqrt{g_{11}^\Phi} & & & \\ & \sqrt{g_{22}^\Phi} & & \\ & & \ddots & \\ & & & \sqrt{g_{rr}^\Phi} \end{bmatrix}$$

Nous voyons que remplacer le vecteur  $\vec{\Phi}_p$  par le vecteur unitaire  $\vec{\Phi}_p^*$  équivaut à diviser la ligne et la colonne  $p$  de  $G^\Phi$  par  $1/\sqrt{g_{pp}^\Phi}$ .

Si on effectue cette opération pour tous les vecteurs  $\vec{\Phi}_p$  ( $p=1,2,\dots,r$ ) cela revient à effectuer la transformation :

$$(G^\Phi)^* = (\Omega^{-1})^T \cdot G^\Phi \cdot (\Omega^{-1})$$

Normalisation du tenseur  $G_\Phi$

Soit  $(G_\Phi)^*$  le tenseur "normalisé" associé au tenseur  $G_\Phi$  :

$$G_\Phi = R^T \cdot G_F \cdot R$$

$$\iff G_\Phi^{-1} = R^{-1} \cdot G_F^{-1} \cdot (R^T)^{-1}$$

$$\iff G^\Phi = R^{-1} \cdot G^F \cdot (R^T)^{-1}$$

$$\iff (\Omega^{-1})^T \cdot G^\Phi \cdot \Omega^{-1} = (\Omega^{-1})^T \cdot R^{-1} \cdot G^F \cdot (R^T)^{-1} \cdot \Omega^{-1}$$

mais  $\left\{ \begin{array}{l} (\Omega^{-1})^T = \Omega^{-1} \\ (R^T)^{-1} = (R^{-1})^T \\ (\Omega^{-1})^T \cdot G^\Phi \cdot \Omega^{-1} = (G^\Phi)^* \end{array} \right.$

$$\iff (G^\Phi)^* = (\Omega^{-1} \cdot R^{-1}) \cdot G^F \cdot (\Omega^{-1} \cdot R^{-1})^T$$

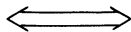
mais  $(G_\Phi)^* = (G^\Phi)^*^{-1}$

$$\iff (G_\Phi)^* = (R \cdot \Omega)^T \cdot (G^F)^{-1} \cdot (R \cdot \Omega)$$

$$\iff (G_\Phi)^* = (R \cdot \Omega)^T \cdot G_F \cdot (R \cdot \Omega) \tag{IV}$$

$$\iff (G_\Phi)^* = \Omega^T \cdot R^T \cdot G_F \cdot R \cdot \Omega$$





$$(G_\Phi)^* = \Omega^T \cdot G_\Phi \cdot \Omega$$

### Normalisation de la matrice V

En comparant les relations III et IV déjà obtenues...

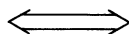
$$\begin{cases} \overline{G_\Phi} = R^T \cdot G_F \cdot R & \text{III} \\ (G_\Phi)^* = (R \cdot \Omega)^T \cdot G_F \cdot (R \cdot \Omega) & \text{IV} \end{cases}$$

... nous voyons que normaliser les vecteurs de base  $\vec{\Phi}_p$  revient à effectuer à la place de la transformation R, une transformation  $R^*$  telle que :

$$\boxed{R^* = R \cdot \Omega}$$

Il s'en suit que de même que nous avons la relation  $V = A \cdot R$ , nous aurons maintenant  $V^* = A \cdot R^*$ , c'est-à-dire :

$$V^* = A \cdot R \cdot \Omega$$



$$\boxed{V^* = V \cdot \Omega}$$

### Algorithme proposé

#### Introduction

L'algorithme proposé consiste à effectuer des "Rotations obliques" de facteurs  $(\vec{F}_p, \vec{F}_q)$  deux à deux dans leur plan ; après chaque transformation on fait :

$$\begin{cases} \vec{F}_p := \vec{\Phi}_p^* \\ \vec{F}_q := \vec{\Phi}_q^* \\ A := V^* \\ G_F := (G_\Phi)^* \end{cases}$$

on arrête les itérations lorsque l'hypervolume  $DG = \det(G_F)$  engendré par les vecteurs contravariants  $\vec{F}_p$  se stabilise, c'est-à-dire quand les vecteurs  $\vec{F}_p$  ne changent plus de direction au cours des itérations.

Pour obtenir  $G^F$ , il suffit alors de faire

$$\boxed{G^F = G_F^{-1}}$$

#### Organigramme

L'algorithme que nous avons utilisé pour résoudre notre problème est celui représenté par l'organigramme de la figure 7 ; dans cet organigramme,

$f_1$  et  $f_2$  représentent les procédés de calculs déjà décrits, les matrices  $\Lambda_1$ ,  $\Lambda_2$  des tableaux de la forme...

$$\Lambda_1 = \begin{bmatrix} \lambda_1^1 \\ \lambda_1^2 \end{bmatrix} ; \quad \Lambda_2 = \begin{bmatrix} \lambda_2^1 \\ \lambda_2^2 \end{bmatrix}$$

... et  $I$  la matrice unité de taille  $(r \times r)$ . La matrice  $G = G_F$  est initialisée en faisant  $G = I$  car initialement le repère  $\{\vec{F}^p\}$  est supposé orthonomé ; il s'en suit que l'on doit prendre  $DG1 = \det(G) = 1$  car l'hypervolume engendré par les  $\{\vec{F}_p\}$  est alors un hypercube.

Dans cet organigramme, nous avons introduit les paramètres  $\varepsilon_1$ , NIT1,  $\varepsilon_2$ , NIT2 qui servent à contrôler le nombre d'itérations du cycle majeur et des cycles mineurs

{	$\varepsilon_1$ = seuil d'arrêt des itérations pour trouver la matrice $\Lambda_{pq}$ dans un cycle mineur
	NIT1 = nombre maximum d'itérations tolérés dans la recherche de $\Lambda_{pq}$
{	$\varepsilon_2$ = seuil d'arrêt des itérations dans le cycle majeur
	NIT2 = nombre maximum d'itérations tolérés dans le cycle majeur .

#### Remarque

L'algorithme proposé s'est avéré très fortement convergent dans les applications que nous avons faites, et il nous est apparu qu'il suffisait en général de trois itérations majeures pour arriver à un résultat stable. Pour cette raison, et dans le but d'alléger la programmation on pourra, au lieu de calculer le déterminant de  $G$ , se contenter de fixer un nombre maximum d'itérations majeures NIT2 par exemple égal à 5.

### RESULTAT OBTENUS

#### Modèle test

Pour tester notre méthode, nous nous sommes donnés dans un repère initial  $F^1, F^2, F^3$  orthonormé, trois points  $X^1, X^8, X^{15}$  formant avec l'origine  $O$  un tétraèdre régulier d'arête égale à 0.5 et tel que les côtés  $OX^1$  et  $OX^8$  soient situés dans le plan  $(F^1, F^2)$  comme l'indique la figure 8.

Afin de construire trois nuages de points centrés aux points  $X^1, X^8$  et  $X^{15}$  nous avons entouré chacun de ces trois points de 6 points disposés parallèlement aux axes  $F^1, F^2, F^3$  à une distance  $D = 0.025$  du point central comme l'indique la figure 9 .

Les résultats obtenus sont consignés dans le tableau de la figure 10. Le tenseur métrique covariant associé au repère oblique  $\mathcal{Q}^p$  est alors égal à ...

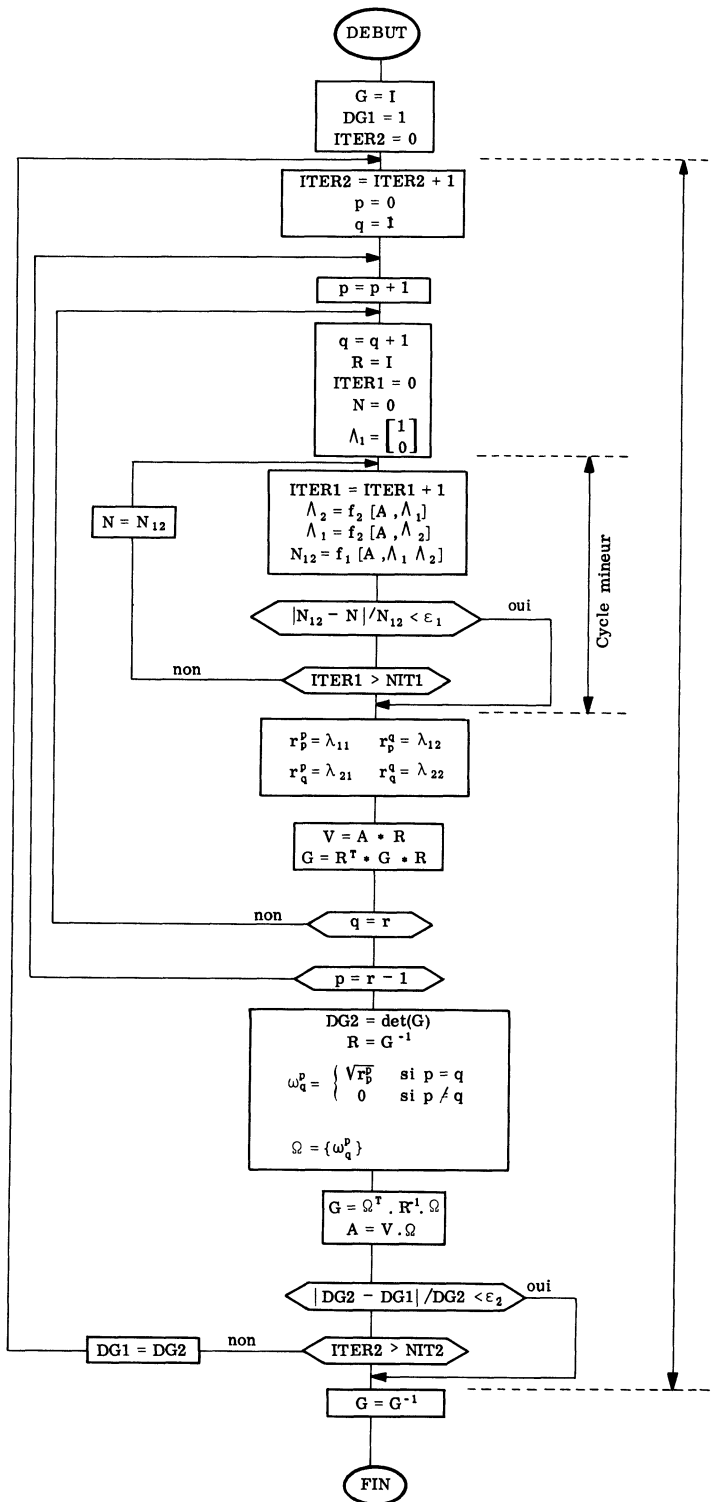


Figure 7

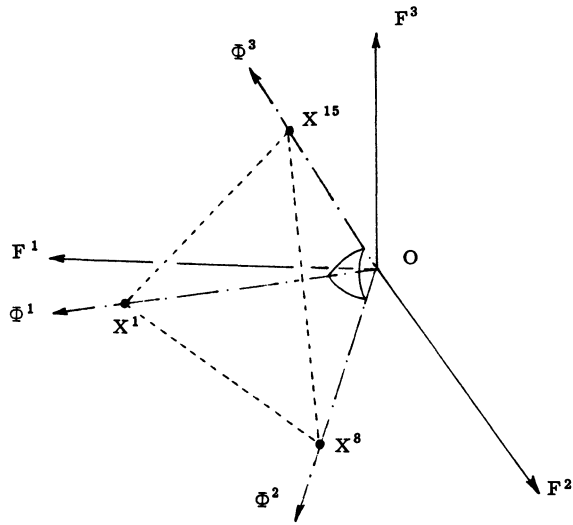


Figure 8

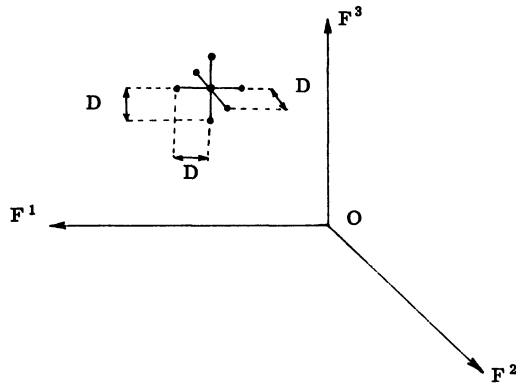


Figure 9

1.000	0.497	0.498
0.497	1.000	0.496
0.498	0.496	1.000

... ce qui montre que les vecteurs unitaires  $\vec{\Phi}_p$  du repère  $\vec{\Phi}$  sont, comme on l'espérait, sensiblement disposés à  $60^\circ$  les uns des autres ; en effet,  $\cos(60^\circ) = 0,500$  ce qui est très près des valeurs obtenues qui sont respectivement égales à 0.497, 0.498, 0.496.

### Conclusion

D'autres essais sur des cas réels nous ont montré l'efficacité de cette méthode qui nous semble non biaisée et qui de ce fait est certainement appelée à remplacer les méthodes classiques citées dans les ouvrages spécialisés (méthodes quartimin, covarimin, oblamin) et dont nous avons montré brièvement les principaux défauts.

Résultats obtenus en 5 itérations majeures

Variables	F <sup>1</sup>	F <sup>2</sup>	F <sup>3</sup>	Φ <sup>1</sup>	Φ <sup>2</sup>	Φ <sup>3</sup>
X <sup>1</sup>	0.483	0.129	0.000	0.499	0.000	0.000
X <sup>2</sup>	0.458	0.129	0.000	0.471	0.008	0.000
X <sup>3</sup>	0.508	0.129	0.000	0.527	-0.007	0.000
X <sup>4</sup>	0.483	0.104	0.000	0.507	-0.027	0.000
X <sup>5</sup>	0.483	0.151	0.000	0.492	0.028	0.000
X <sup>6</sup>	0.483	0.129	-0.025	0.509	0.011	0.031
X <sup>7</sup>	0.483	0.129	0.025	0.489	-0.010	0.030
X <sup>8</sup>	0.129	0.483	0.000	0.000	0.499	0.001
X <sup>9</sup>	0.104	0.483	0.000	-0.027	0.506	0.001
X <sup>10</sup>	0.154	0.483	0.000	0.028	0.491	0.001
X <sup>11</sup>	0.129	0.458	0.000	0.008	0.471	0.001
X <sup>12</sup>	0.129	0.508	0.000	-0.007	0.526	0.001
X <sup>13</sup>	0.129	0.483	-0.025	0.010	0.509	-0.029
X <sup>14</sup>	0.129	0.483	0.025	-0.010	0.488	0.032
X <sup>15</sup>	0.204	0.204	0.408	0.000	0.000	0.499
X <sup>16</sup>	0.179	0.204	0.408	-0.027	0.008	0.499
X <sup>17</sup>	0.229	0.204	0.408	0.028	-0.007	0.499
X <sup>18</sup>	0.204	0.179	0.408	0.008	-0.027	0.499
X <sup>19</sup>	0.204	0.229	0.408	-0.007	0.028	0.499
X <sup>20</sup>	0.204	0.204	0.383	0.010	0.011	0.468
X <sup>21</sup>	0.204	0.204	0.433	-0.010	-0.010	0.530
	matrice factorielle A avant "rotation oblique"			matrice factorielle A après "rotation oblique"		

BIBLIOGRAPHIE

- HARMAN - Modern Factor analysis (Chicago Press University)
- DEREC and MAXWELL - Factor analysis (Woods Hole, Massachusetts)
- LAWLEY and MAXWELL - Factor analysis as a statistical method (Butterworths, London)
- J. B. CAROLL - Biquartimin Criterion for rotation to oblique simple structure in factor analysis (Science vol. 126, Pg 1114).