Chapitre1
L'Analyse en composantes principales: ACP.

1 Introduction

L'ACP est une méthode descriptive permettant de traiter des tableaux de données quantitatives $X_{n,p}$(de grandes dimension) où n représente le nombre d'individus et p le nombre de variables quantitaves. Le but de l'ACP est de résumer la grande quantité d'informations contenues dans X, et cela dans un tableau de plus petite dimension $Y_{n,q}(q < p)$. Et ainsi fournir une représentation visuelle tels que:

♦ Y^j est une combinaison linéaire des p variables quantitatives, $X^j, j = 1, ..., p$.
♦ Les variables $(Y^j)_{j=1,...q}$ sont non correlées entre elles.
♦ Le tableau X peut être reconstitué à partir du nouveau tableau Y.
♦ Y contient le maximum d'informations sur X.

Exemple 1 - *Nous avons les notes de* 200 *tudiant en Informatiques dans* 8 *modules. Le tableau rsumant ces donnes est de dimension* $200 * 8$. *Pour pouvoir synthtiser l'information contenue dans ce grand tableau et visualiser les donnes, nous effectuons une ACP.*
- Relevés des dépenses de ménages en 10 postes.
- Teneur en mineraux de certaines eaux, ect.

1.1 Tableau de données

	X^1	$\cdots$	X^j	$\cdots$	X^p
1	$\cdots$	$\cdots$	x$_1^j$	$\cdots$	x$_1^p$
$\vdots$					
i	x_1^1		x$_i^j$		x$_i^p$
$\vdots$					
n	x$_n^1$		x$_n^j$		x$_n^p$

- x_1^j représente la mesure de la variable X^j sur l'individu "i".

A chaque individu "i" on associe le vecteur $X_i = \begin{pmatrix} x_i^1 \\ \vdots \\ x_i^p \end{pmatrix}$ et un poids p_i, tel que $0 \leq p_i \leq 1$. Ainsi on définit le nuage de n individus appartenant à $\mathbb{R}^p$:

$$\aleph(I) = \{X_i \in \mathbb{R}^p, i = 1, ..., n\} .$$

L'espace $\mathbb{R}^p$ est muni d'une métrique qu'on notera M. Cette métrique peut être euclidienne c'est à dire que:

$$M = \begin{pmatrix} 1 & 0 & 0 \\ & \ddots & \vdots \\ & & 1 \end{pmatrix}$$

ou

$$M = \begin{pmatrix} \frac{1}{\sigma_1} & 0 & 0 \\ & \ddots & \vdots \\ & & \frac{1}{\sigma p} \end{pmatrix}$$

où σ_j représente l'écart type de la variable X^j.

Remarque 1 Notons que le choix de l'une ou de l'autre des métriques se fera selon des cas qu'on citera ci après.

A chaque variable est associé le vecteur $X^j = \begin{pmatrix} x_1^j \\ \vdots \\ x_n^j \end{pmatrix}$ de $\mathbb{R}^n$ et on définit le nuage de variables par:

$$\aleph(J) = \left\{ X^j \in \mathbb{R}^n, j = 1, ..., p \right\}$$

$\mathbb{R}^n$ est muni de la métrique des poids $D_p = \begin{pmatrix} p_1 & & 0 \\ & \ddots & \\ 0 & & p_n \end{pmatrix}$. Lorsque les individus sont pris aléatoirement équiprobablement alors; $p_i = \frac{1}{n}, \forall i = 1, \ldots, n$.

1.2 Le centre de gravité du nuage N(I)

Il est défini par

$$g = \frac{1}{n} \sum_{i=1}^{n} x_i = \begin{pmatrix} \frac{1}{n} \sum\limits_{i=1}^{n} x_i^1 \\ \vdots \\ \frac{1}{n} \sum\limits_{i=1}^{n} x_i^p \end{pmatrix} = \begin{pmatrix} \overline{x}^1 \\ \vdots \\ \overline{x}^p \end{pmatrix}$$

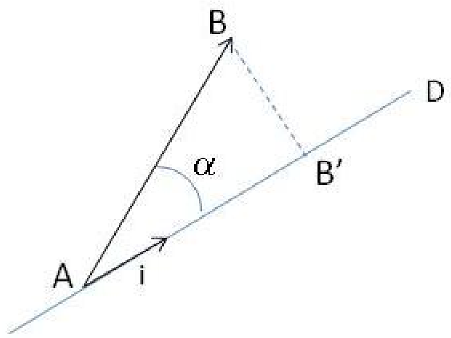
où $\overline{x}^j$ représente la moyenne arithmétique de la $j^{ième}$ variable.

1.3 L'inertie du nuage $N(I)$

L'inertie est une mesure de dispersion multidimentionnelle, elle est définie par:

$$I_g = \sum_{i=1}^{n} p_i \|x_i - g\|_M^2.$$

Remarque 2 La mesure de dispersion dans le cas unidimentionnel n'est rien d'autre que l' écart type.



1.4 Formulation du problème d'ACP

Le principe est d'obtenir une reprsentation approchée du nuage $N(I)$ $(N(J))$ dans un sous espace de plus faible dimension par projection. Ainsi, formellement:

1- On commence par rechercher un sous espace vectoriel de dimension 1, $E_1 = \Delta u_1$ engendré par un vecteur unitaire u_1, qui ajuste au mieux le $N(I)$ de $\mathbb{R}^n$

2- Ensuite rechercher un sous espace vectoriel de dimension 2, E_2 en déterminant Δu_2 orthogonal à Δu_1 qui ajuste au mieux le $N(I)$ de $\mathbb{R}^n$

3- En général rechercher un sous espace vectoriel E_k de dimension k en déterminant Δu_k orthogonal à Δu_{k-1} qui ajuste au mieux le $N(I)$ de $\mathbb{R}^n$ avec

$$E_k = E_{k-1} \oplus \Delta u_k$$

Dans ce qui suit, on expose comment déterminer ces sous espaces qui ajustent au mieux les nuages de points. Les (Δu_k) seront appelés les axes factoriels.

1.5 Détermination des axes factoriels

A partir de maintenant, on suppose que le tableau X est centré.

1.5.1 Ajustement sur $(\mathbb{R}^P, M)$

Dans ce cas le nuage $N(I)$ est ajusté.

On recherche le sous espace vectoriel de dim 1, Δu_1 passant par l'origine et engendré par le vecteur unitaire u_1 qui ajuste au mieux le nuage $N(I)$. Cela se fait, en déterminant u_1 qui maximise l'inertie du nuage $N(I)$, défini précedemment.

INCLURE un petit graphe

On commence par écrire l'inertie du nuage en fonction du vecteur u. Notons par α_i la valeur de projection du vecteur individu X_i du nuage $N(I)$ sur l'axe Δu_k engendré par le vecteur unitaire u_k, α_i est donnée par:

$$\alpha_i = \langle X_i, u_1 \rangle_M = X_i^t M u_1,$$

Le vecteur de projection de tout les individus est donc donné par:

$$Y = \begin{pmatrix} \alpha_i \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} X_i^t M u_1 \\ \vdots \\ X_i^t M u_1 \end{pmatrix} = XMu_1$$

Y est appelée composante principale.

Ainsi l'inertie du nuage $N(I)$ défini plus haut s'écrit:

$$I = \phi(u) = \parallel Y \parallel_{D_p}^2 = u_1^t M X^t D_p X M u_1.$$

Posons $V = X^t D_p X$. V représente la matrice de variance covariance des p variables.

Dans la suite, nous déterminons u_1 unitaire qui maximise $\phi(u)$. Commençons par écrire la fonction de Lagrange correspondant à notre problème d'optimisation:

$$L(u) = \phi(u) - \lambda(u^t M u - 1)$$

tel que λ est le multiplicateur de Lagrange. u_1 est solution du système suivant:

$$\begin{cases} \frac{dL(u)}{du}(u_1) = 0 \\ u_1^t M u_1 = 1 \end{cases}$$

Aprés résolution, nous obtenons l'équation suivante:

$$VMu_1 = \lambda u_1$$

. Ce qui indique que u_1 est vecteur propre de la matrice VM associé à la valeur propre λ. Laquelle des valeurs propres de VM?

En utilisant le contrainte sur le vecteur propre et en multipliant chaque coté de l'équation précédente par $u_1^T M$, nous obtenons:

$$u_1^t MVMu_1 = \lambda u_1^t M u_1 = \lambda.$$

Ainsi

$$\lambda = \max \phi(u).$$

d'où, λ est la plus grande valeur propre de VM.

Pour retrouver le sous espace vectoriel de dimension 2 qui ajuste au mieux le nuage de points $N(I)$, il suffit de trouver u_2 vecteur propre unitaire orthogonale à u_1 qui maximise $\phi(u)$.

Dans ce cas la fonction de Lagrange sous deux contraintes s'écrit:

$$L(u, v) = L(u) = \phi(u) - \lambda(u^t M u - 1) - \alpha u^t M v.$$

u_2 est solution du systhème

$$\begin{cases} \frac{dL(u,v)}{du}(u_2, u_1) = 0 \\ \frac{dL(u,v)}{dv}(u_2, u_1) = 0 \\ u_2^t M u_2 = 1, u_2^t M u_1 = 0 \end{cases}$$

Aprés résolution du système et en prenant en considération les contraintes, on déduit que u_2 est vecteur propre de VM associ é à la deuxième plus grande valeur propre.

En général, le sous espace vectoriel de dimension k qui ajuste au mieux le nuage de points $N(I)$ est engendré par les vecteurs propres $u_1, ..., u_k$ de VM unitaires et deux à deux orthogonaux associés aux valeurs propres $\lambda_1, ..., \lambda_k$, ordonnées de manière décroissantes, c'est à dire que $\lambda_1 \leq ... \leq \lambda_k$.

2-

1.5.2 Ajustement sur $(\mathbb{R}^n, D_p)$

Dans ce cas le nuage $N(J)$ des variables est ajusté.

On recherche le sous espace vectoriel de dim 1, Δv_1 passant par l'origine et engendré par le vecteur unitaire v_1 qui ajuste au mieux le nuage $N(J)$ et ceci en déterminant v_1 qui maximise l'inertie du nuage $N(J)$, définie dans ce qui suit.

On commence par écrire l'inertie du nuage en fontion du vecteur v.

Notons par α_j la valeur de projection du vecteur variable X^j sur Δv_k engendré par le vecteur unitaire v_k, α_j est donnée par :

$$\alpha_j = \left\langle X^j, v_1 \right\rangle_{D_p} = X^{j^t} D_p v_1,$$

Le vecteur de projection de toutes les variables est donc donné par:

$$W = \begin{pmatrix} \alpha_i \\ \vdots \\ \alpha_p \end{pmatrix} = \begin{pmatrix} X^t D_p v_1 \\ \vdots \\ X^t D_p v_1 \end{pmatrix} = X^t D_p v_1.$$

Ainsi l'inertie du nuage $N(j)$ défini plus haut s'écrit:

$$I = \psi(v) = \| W \|_M^2 = v_1^t D_p X M X^t D_p v_1.$$

Posons $T = XMX^t$ et déterminons maintenant v_1 unitaire qui maximise $\psi(v)$.

Commençons par écrire la fonction de Lagrange:

$$L(v) = \psi(v) - \lambda(v^t D_p v - 1).$$

v_1 est solution du système suivant:

$$\begin{cases} \frac{dL(v)}{dv}(v_1) = 0 \\ v_1^t D_p v_1 = 1 \end{cases}$$

De la mme manire que pour l'analyse du nuage $N(I)$, aprés résolution, on déduit que v_1 est le vecteur propre de la matrice TD_p associé à la plus grande valeur propre λ.

Pour retrouver le sous espace vectoriel de dimension 2 qui ajuste au mieux le nuage de points $N(J)$, il suffit de trouver v_2 vecteur propre unitaire orthogonale à v_1 qui maximise $\psi(v)$. Dans ce cas la fonction de Lagrange sous deux contraintes s'écrit:

$$L(u,v) = L(u) = \phi(v) - \lambda(v^t D_p v - 1) - \alpha v^t M u.$$

v_2 est solution du systhème

$$\begin{cases} \frac{dL(u,v)}{dv}(v_2, v_1) = 0 \\ \frac{dL(u,v)}{dv}(v_2, v_1) = 0 \\ v_2^t M v_2 = 1, v_2^t M v_1 = 0 \end{cases}$$

Après résolution du système et en prenant en considération les contraintes, on déduit que v_2 est vecteur propre de TD_p associé à la deuxième plus grande valeur propre.

En général, le sous espace vectoriel de dimension k qui ajuste au mieux le nuage de points $N(J)$ est engendrée par les vecteurs propres $v_1, ..., v_k$ de TD_p unitaires et deux à deux orthogonaux associés aux valeurs propres $\lambda_1, ..., \lambda_k$, ordonnées de manière décroissantes.

Remarque 3 Pour éviter la différence dans l'echelle de mesure de variables et pour faire jouer à chaque variable un rôle identique dans la définition des proximités entre individus, on passe à l'ACP normé qui consiste à réduire les variables, c'est à dire:

$$X_i^j \to \frac{X_i^j}{\sigma_j,}$$

ou bien utiliser la métrique $M = \begin{pmatrix} \frac{1}{\sigma_1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sigma_j} \end{pmatrix}$ *Ainsi deux cas se présente pour réduire les variables, ou bien les unités de mesure des variables sont différentes ou bien les écarts types sont très dispersés en valeurs (de petites et de très grandes valeurs). On dit que les donn´es sont normalisées.*

1.5.3 Propriétés des composantes principales

Nous rappellons que $Y_\alpha(i)$ qui représente le vecteur de projection des individus sur l'axe factoriel δ_α est appelé composante principale ou nouvelle variable, ses propriétés sont données dans la proposition suivante:

Proposition 2 *Les composantes principales vérifient ce qui suit* • $\overline{y}_\alpha = 0$, $\forall \alpha = 1, ..., p$,

- $\|y_\alpha\|^2 = var y_\alpha = \lambda_\alpha$,
- $cov(y_\alpha, y_{\alpha'}) = 0$.

Preuve. Nous vérifions dans ce qui suit la dernière propriété, les deux premières se démontrent de la même faon.

$$cov(y_\alpha, y_{\alpha'}) = \langle y_\alpha, y_{\alpha'} \rangle_{D_p}$$

or $Y_\alpha = Xu_\alpha$, donc

$$cov(y_\alpha, y_{\alpha'}) = \frac{1}{n} u_\alpha^t X^t X u_\alpha^{'} = u_\alpha^t V u_{\alpha'}$$

Rappelons que $u_{\alpha'}$ est vecteur propre de VM associé à la valeur propre $\lambda_{\alpha'}$, ainsi

$$cov(y_\alpha, y_{\alpha'}) = \lambda u_\alpha^t u_{\alpha'} = 0$$

car deux vecteurs propres associés à deux valeurs propres distinctes sont othogonaux.

■

1.5.4 Représentation d'un individu supplémentaire

Soit x_i un individu supplémentaire, sa projection sur l'axe factoriel est donnée par:

$$\alpha_{x_i} = \widetilde{x_i}^t u, \text{ tel que } \widetilde{x_i} = x_i - g^t.$$

1.5.5 Représentation d'une variable supplémentaire

Soit x^j une variable supplémentaire, sa projection sur l'axe factoriel est donnée par:

$\alpha_{x_i} = X^{jt} D_p v$, tel que $\widetilde{x^j} = x^j - \overline{x}^j$ est la variable centrée.

Remarque 4 Si l'ACP est normée en plus d'être centrés les vecteurs sont réduits.

1.5.6 Formules de transitions

Ces dernières permettent de passer de l'analyse d'un nuage à un autre.

Proposition 3 *Les matrices XX^tD_p et X^tD_pX ont les mêmes valeurs propres.*

Preuve. Soit u un vecteur propre de VM associé à la plus grande valeur propre λ, c'est à dire que

$$Vu = \lambda u \Longleftrightarrow X^t D_p X u = \lambda u,$$

on multipie les deux cotés de l'égalité par X, on obtient

$$XX^t D_p X u = \lambda X u,$$

En posant $w = Xu$, on déduit que w est vecteur propre de XX^tD_passocié à la même valeur propre λ. ■

Et on écrit que:

$$v = \frac{w}{\|w\|},$$

or
$$\|w\|_{D_p}^2 = w^t D_p w = u^t X^t D_p X u = \lambda \|u\|^2 = \lambda$$
D'où la première formule de transition:
$$v = \frac{1}{\sqrt{\lambda}} X u$$
En procédant de la même manière mais dans le sens inverse dans la preuve, on obtient la deuxième formule donnée par
$$u = \frac{1}{\sqrt{\lambda}} X^t D_p v$$

1.5.7 Les aides à l'interprétation

Qualité globale de représentation d' un axe factoriel: Elle est mesurée par le pourcentage d'inertie et elle est donnée par

$$I = \frac{\lambda_\alpha}{\sum_{r=1}^{p} \lambda_r} \times 100$$

a- **Individu**:
$$C_{re}^{\alpha}(i) = \cos_i^2(\theta) = \frac{(y_\alpha(i))^2}{\|x_i\|^2}$$
b-**Variable**:
$$C_{re}^{\alpha}(j) = \cos_j^2(\theta) = \frac{(V_\alpha(j))^2}{\|x^j\|_{D_p}^2}$$
Dés que $\cos_i^2(\theta) \simeq 1$, on dira que l'individu ou la variable sont trés bien représenté par le α^{ieme} axe factoriel.

1.5.8 Cercle de corrélation

Il y a une relation très étroite entre le coéffcient de corrélation entre l'ancienne et la nouvelle variable et la projection de cette dernière sur l'axe factoriel, en effet:
$$r(X^j, Y_\alpha) = \frac{V_\alpha(j)}{\sigma_j}$$
Lorsque l'ACP est normée, ce coéfficient de corrélation qui varie entre -1 et 1 correspond à la projection de la j^{ime} variable sur l'axe Δ_α. Ainsi les projections des variables varient à l'intérieur d'un cercle unité dont le centre est l'origine des axes. Ce cercle est appelé cercle de corrélation.

- Si la projection de la variable se retrouve proche du cercle alors, la variable est bien représentée par le plan factoriel.

- Les variables qui se retrouvent à l'intérieur du cercle, sont donc mal représentées.

1.5.9 Reconstitution du tableau de données

Le tableau de données est complètement reconstitué à partir de la formule suivante:

$$X = \sum_{\alpha=1}^{p} \sqrt{\lambda_\alpha} v_\alpha u_\alpha^t$$

En effet à partir des formules de transition, on a

$$v = \frac{1}{\sqrt{\lambda}} Xu \Leftrightarrow \sqrt{\lambda} v = Xu$$

On multiplie les deux cotés de l'égalité par u^t on aura

$$\sqrt{\lambda} v u^t = X \, \mathrm{u} u^t \Leftrightarrow X \sum_\alpha \mathrm{u}_\alpha u_\alpha^t = \sum_\alpha \sqrt{\lambda_\alpha} v_\alpha u_\alpha^t$$

Montrons que $\sum_\alpha \mathrm{u}_\alpha u_\alpha^t = I$,

Notons par $S = [u_1, ..., u_p]$ la matrice contenant en colonnes les vecteurs propres orthonormés de V. S est une matrice orthogonale, en effet,

$$S^t S = \begin{pmatrix} u_1^t \\ \vdots \\ u_p^t \end{pmatrix} \begin{pmatrix} u_1 & \cdots & u_p \end{pmatrix} = \begin{pmatrix} u_1^t u_1^t & \cdots & u_1^t u_p^t \\ \vdots & \ddots & \vdots \\ u_p^t u_1^t & \cdots & u_p^t u_p^t \end{pmatrix} = I = \sum_\alpha \mathrm{u}_\alpha u_\alpha^t.$$

Ainsi,

$$X = \sum_\alpha \sqrt{\lambda_\alpha} v_\alpha u_\alpha^t.$$

Exemple 4 *Reconstitution à partir de la plus grande valeur propre:*

$$X_1 = \sqrt{\lambda_1} v_1 u_1^t$$

Le taux de reconstitution de ce tableau est de I_1 *pour cent.*

1.6 Récapitulation

Algorithme ACP:

1. Calculer les moyennes des variables $\overline{X}^j, j = 1, ..., p$.
2. Centrer le tableau X (réduire si les données sont hétérogènes).
3. Calculer la matrice de variance covariance $V = X^t D_p X = \frac{1}{n} X^t X$.
4. Calculer les valeurs propres et les vecteurs propres de V.
5. Calculer les projections des individus et des variables sur les axes factoriels: $Y_\alpha = Xu, V_\alpha = \sqrt{\lambda_\alpha} u$.

6. Représenter graphiquement les individus et les variables.

7. Interpréter les résultats de l'analyse.

Exemple 5 *Exemple d'application sur le logiciel R: Le tableau suivant représente la structure du bilan dun groupe pétrolier de 1969 1984. telles que les variables représentent:*

NET : Situation nette ; représente l'ensemble des capitaux propres de l'entreprise.

INT : Intérêts ; représente l'ensemble des frais financiers supportés par l'entreprise.

SUB : Subventions ; représente le montant total des subventions accordées par l'Etat.

LMT : Dettes à long et moyen terme.

DCT : Dettes à court terme.

IMM : Immobilisations ; représente lensemble des terrains et du matériel de l'entreprise.

EXP : Valeurs dexploitation.

VRD : Valeurs réalisables et disponibles ; ensemble des créances court terme de l'entreprise.

```
#############Expliquer en quoi consiste le tableau#############
> read.table("C:/Users/Admin/Documents/tableau petrole.csv",sep=";",header=TRUE)
   Anne   NET  INT  SUB   LMT   DCT   IMM   EXP   VRD
1   1969 17.93 3.96 0.88  7.38 19.86 25.45  5.34 19.21
2   1970 16.21 3.93 0.94  9.82 19.11 26.58  5.01 18.40
3   1971 19.01 3.56 1.91  9.43 17.87 25.94  5.40 16.88
4   1972 18.05 3.33 1.73  9.72 18.83 26.05  5.08 17.21
5   1973 16.56 3.10 2.14  9.39 20.36 23.95  6.19 18.31
6   1974 13.09 2.64 2.44  8.10 25.05 19.48 11.61 17.59
7   1975 13.43 2.42 2.45 10.83 22.07 22.13 11.17 15.49
8   1976  9.83 2.46 1.79 11.81 24.10 22.39 11.31 16.30
9   1977  9.46 2.33 2.30 11.46 24.45 23.07 11.16 15.77
10  1978 10.93 2.95 2.25 10.72 23.16 24.17  9.64 16.20
11  1979 13.02 3.74 2.21  7.99 23.04 19.53 12.60 17.87
12  1980 13.43 3.60 2.29  7.09 23.59 17.61 16.67 15.72
13  1981 13.37 3.35 2.58  6.76 23.94 18.04 15.42 16.54
14  1982 11.75 2.74 3.11  7.37 25.04 18.11 14.71 17.18
15  1983 12.59 3.05 3.85  7.12 23.40 19.17 11.86 18.97
16  1984 13.00 3.00 4.00  7.00 24.00 20.00 12.00 17.00
###############commande pour effectuer l'ACP##########
> pca<-PCA(Tab)
###########"Calcul des valeurs propres###########"
> pca$eig
        eigenvalue        percentage of variance                cum perc of vari
comp 1 5.259233e+00         5.843592e+01                          58.43592
comp 2 2.184428e+00         2.427142e+01                          82.70734
comp 3 6.871060e-01         7.634511e+00                          90.34185
```

```
comp 4 5.225378e-01          5.805976e+00                              96.14783
comp 5 2.590453e-01          2.878281e+00                              99.02611
comp 6 6.432384e-02          7.147093e-01                              99.74082
comp 7 1.672358e-02          1.858175e-01                              99.92663
comp 8 6.602393e-03          7.335993e-02                              99.99999
comp 9 6.148855e-07          6.832061e-06                             100.00000
##coordonnes des variables
> pca$var
$coord
          Dim.1       Dim.2       Dim.3       Dim.4       Dim.5
Anne  0.9109872  0.23602348  0.08511972  0.15382329  0.27769809
NET  -0.8198942  0.41802161 -0.04250705  0.32241103 -0.21234399
INT  -0.5663198  0.67761605 -0.36411816 -0.04846668  0.28593427
SUB   0.7895101  0.22889450  0.42669755  0.36829627  0.04193376
LMT  -0.2772180 -0.93670765  0.06887425 -0.03582595  0.13810083
DCT   0.9374734 -0.09025863  0.01364877 -0.30107880 -0.07250136
IMM  -0.8914564 -0.36218983  0.15509629  0.06053271  0.16424251
EXP   0.9325853  0.12256690 -0.32142955 -0.04610497 -0.04337964
VRD  -0.4392585  0.64082024  0.48072776 -0.39901674  0.01168683

$cor
          Dim.1       Dim.2       Dim.3       Dim.4       Dim.5
Anne  0.9109872  0.23602348  0.08511972  0.15382329  0.27769809
NET  -0.8198942  0.41802161 -0.04250705  0.32241103 -0.21234399
INT  -0.5663198  0.67761605 -0.36411816 -0.04846668  0.28593427
SUB   0.7895101  0.22889450  0.42669755  0.36829627  0.04193376
LMT  -0.2772180 -0.93670765  0.06887425 -0.03582595  0.13810083
DCT   0.9374734 -0.09025863  0.01364877 -0.30107880 -0.07250136
IMM  -0.8914564 -0.36218983  0.15509629  0.06053271  0.16424251
EXP   0.9325853  0.12256690 -0.32142955 -0.04610497 -0.04337964
VRD  -0.4392585  0.64082024  0.48072776 -0.39901674  0.01168683
#Contribution relative
$cos2
          Dim.1      Dim.2       Dim.3       Dim.4        Dim.5
Anne 0.82989774 0.05570708 0.007245366 0.023661604 0.0771162271
NET  0.67222658 0.17474207 0.001806849 0.103948875 0.0450899691
INT  0.32071806 0.45916351 0.132582038 0.002349019 0.0817584046
SUB  0.62332620 0.05239269 0.182070799 0.135642146 0.0017584406
LMT  0.07684982 0.87742122 0.004743663 0.001283499 0.0190718398
DCT  0.87885635 0.00814662 0.000186289 0.090648446 0.0052564475
IMM  0.79469453 0.13118147 0.024054858 0.003664209 0.0269756028
EXP  0.86971527 0.01502264 0.103316954 0.002125668 0.0018817936
VRD  0.19294800 0.41065059 0.231099181 0.159214361 0.0001365821
########Contribution absolue###########""
$contrib
         Dim.1      Dim.2       Dim.3      Dim.4       Dim.5
```

```
Anne 15.779826  2.5501910  1.05447580  4.5282088 29.76939745
NET   12.781838  7.9994432  0.26296517 19.8930813 17.40620957
INT    6.098191 21.0198519 19.29571831  0.4495406 31.56143050
SUB   11.852037  2.3984628 26.49821122 25.9583400  0.67881586
LMT    1.461236 40.1670947  0.69038293  0.2456280  7.36235681
DCT   16.710734  0.3729407  0.02711212 17.3477290  2.02916146
IMM   15.110466  6.0053011  3.50089484  0.7012332 10.41346902
EXP   16.536924  0.6877153 15.03653798  0.4067970  0.72643415
VRD    3.668748 18.7989994 33.63370163 30.4694421  0.05272517
############Coordonnes individus
> pca$ind
$coord
        Dim.1        Dim.2       Dim.3       Dim.4       Dim.5
1  -3.8351154  1.652459236 -0.14154403 -1.02431449 -0.29938900
2  -3.8413051  0.182500857 -0.22565075 -0.70913558  0.59998318
3  -3.3337769  0.003773478 -0.22527265  1.39710806 -0.09762941
4  -3.0462168 -0.320436443  0.13020511  0.87136251 -0.09285819
5  -1.9912301  0.141236546  0.95538144  0.07832976 -0.33630677
6   1.0961077  0.037057135  0.38051562 -0.85238086 -1.28622815
7   0.4953239 -2.079776592 -0.17704893  0.86050204 -0.58304750
8   0.8768239 -2.625233250 -0.05351164 -0.87394426  0.12069605
9   1.3626321 -2.787505314  0.17471845 -0.32207266  0.18936592
10  0.3871842 -1.702731342  0.01220138 -0.05467552  0.88135554
11  0.6146480  1.361573682 -0.56505888 -0.70498334  0.58911643
12  1.9695831  1.064309914 -2.07820803  0.34097186 -0.01838120
13  2.0598931  1.296561148 -1.09078708  0.16509957 -0.11918392
14  2.8238016  0.691720319  0.26379560 -0.15616781 -0.30555878
15  1.9388097  1.911620313  1.73814421 -0.05755114  0.37636420
16  2.4228371  1.172870315  0.90212017  1.04185187  0.38170160
######## Contibution relative Individus############
$cos2
        Dim.1        Dim.2        Dim.3        Dim.4        Dim.5
1  0.78345657 1.454520e-01 1.067189e-03 0.0558887874 4.774520e-03
2  0.93875801 2.118975e-03 3.239438e-03 0.0319929730 2.290204e-02
3  0.84521754 1.082877e-06 3.859334e-03 0.1484415302 7.248644e-04
4  0.90894028 1.005767e-02 1.660618e-03 0.0743723472 8.446060e-04
5  0.77662231 3.907160e-03 1.787807e-01 0.0012017677 2.215329e-02
6  0.31316126 3.579359e-04 3.774041e-02 0.1893777556 4.312187e-01
7  0.04249293 7.491547e-01 5.429061e-03 0.1282454411 5.887704e-02
8  0.09033061 8.097407e-01 3.364392e-04 0.0897382620 1.711576e-03
9  0.18849642 7.888189e-01 3.099013e-03 0.0105306128 3.640404e-03
10 0.03794082 7.337764e-01 3.767816e-05 0.0007565842 1.965956e-01
11 0.10618746 5.210778e-01 8.974447e-02 0.1396940819 9.754894e-02
12 0.41020981 1.197826e-01 4.567047e-01 0.0122940366 3.572772e-05
13 0.59067077 2.340142e-01 1.656289e-01 0.0037944425 1.977386e-03
14 0.91657341 5.499966e-02 7.998969e-03 0.0028033799 1.073219e-02
```

```
15 0.35082655 3.410557e-01 2.819640e-01 0.0003091221 1.322022e-02
16 0.61732731 1.446662e-01 8.558464e-02 0.1141508112 1.532198e-02
#########"Contribution absolue individus##############"
$contrib
        Dim.1        Dim.2        Dim.3       Dim.4        Dim.5
1  17.4789169 7.812748e+00  0.182238189 12.54957206  2.162598784
2  17.5353830 9.529544e-02  0.463158729  6.01479511  8.685252191
3  13.2078071 4.074047e-05  0.461607887 23.34652653  0.229967057
4  11.0275554 2.937826e-01  0.154209917  9.08155094  0.208038969
5   4.7119485 5.707376e-02  8.302519007  0.07338644  2.728823867
6   1.4277892 3.929036e-03  1.317046944  8.69019772 39.915383625
7   0.2915655 1.237587e+01  0.285129982  8.85658087  8.201856513
8   0.9136554 1.971869e+01  0.026046701  9.13544585  0.351471721
9   2.2065554 2.223175e+01  0.277673108  1.24070922  0.865182906
10  0.1781529 8.295347e+00  0.001354174  0.03575593 18.741595796
11  0.4489630 5.304258e+00  2.904314738  5.94456369  8.373490951
12  4.6100567 3.240996e+00 39.285683478  1.39059084  0.008151776
13  5.0425127 4.809814e+00 10.822715654  0.32602743  0.342720113
14  9.4760205 1.369000e+00  0.632982031  0.29170598  2.252650480
15  4.4671239 1.045552e+01 27.480706124  0.03961596  3.417597390
16  6.9759938 3.935884e+00  7.402613337 12.98297541  3.515217862

$dist
       1        2        3        4        5        6        7        8
4.332823 3.964624 3.626204 3.195164 2.259522 1.958706 2.402874 2.917393
       9       10       11       12       13       14       15       16
3.138535 1.987761 1.886209 3.075185 2.680229 2.949515 3.273324 3.083660

> pca$svd
$vs
[1] 2.2933016700 1.4779810197 0.8289185716 0.7228677803 0.5089649372
[6] 0.2536214448 0.1293196685 0.0812551130 0.0007841464

$U
            [,1]        [,2]        [,3]        [,4]        [,5]
 [1,] -1.6723118  1.11805173 -0.17075746 -1.41701501 -0.58823108
 [2,] -1.6750108  0.12347984 -0.27222306 -0.98100317  1.17883008
 [3,] -1.4537019  0.00255313 -0.27176693  1.93272974 -0.19181952
 [4,] -1.3283105 -0.21680687  0.15707828  1.20542447 -0.18244516
 [5,] -0.8682809  0.09556046  1.15256368  0.10835973 -0.66076608
 [6,]  0.4779605  0.02507281  0.45905066 -1.17916565 -2.52714491
 [7,]  0.2159872 -1.40717409 -0.21359026  1.19040033 -1.14555534
 [8,]  0.3823413 -1.77622934 -0.06455596 -1.20899600  0.23714020
 [9,]  0.5941792 -1.88602240  0.21077879 -0.44554851  0.37206083
[10,]  0.1688327 -1.15206577  0.01471964 -0.07563696  1.73166259
[11,]  0.2680188  0.92123895 -0.68168201 -0.97525904  1.15747940
```

```
[12,]  0.8588417  0.72011068 -2.50713170  0.47169326 -0.03611487
[13,]  0.8982216  0.87725156 -1.31591584  0.22839525 -0.23416921
[14,]  1.2313258  0.46801705  0.31824067 -0.21603925 -0.60035329
[15,]  0.8454229  1.29339977  2.09688173 -0.07961503  0.73946980
[16,]  1.0564843  0.79356250  1.08830976  1.44127585  0.74995657

$V
           [,1]        [,2]        [,3]        [,4]        [,5]
 [1,]  0.3972383  0.15969318  0.10268767  0.21279588  0.54561339
 [2,] -0.3575170  0.28283287 -0.05128013  0.44601661 -0.41720750
 [3,] -0.2469452  0.45847412 -0.43926892 -0.06704779  0.56179561
 [4,]  0.3442679  0.15486971  0.51476413  0.50949328  0.08239028
 [5,] -0.1208816 -0.63377515  0.08308928 -0.04956087  0.27133663
 [6,]  0.4087876 -0.06106887  0.01646576 -0.41650605 -0.14244864
 [7,] -0.3887218 -0.24505716  0.18710678  0.08373967  0.32269907
 [8,]  0.4066562  0.08292860 -0.38776975 -0.06378064 -0.08523111
 [9,] -0.1915398  0.43357813  0.57994570 -0.55199132  0.02296196

cor(Tab)#########"calcul la matrice de corrlation#########"""""
#Tracer le cercle de corrlation.
s.corcircle(T$co, xax=1, yax=2)
####reprsentation simultane
 scatter(acp)
#########tracer le nuage des individus
 s.label(acp$li, xax=1,yax=2)
```

Références:

1- L. Lebart, M. Piron, A. Morineau, *Statistique Exploratoire Multidimensionnelle, [Visualisation et Inférence en Fouille de Données].*, Dunod, 2006

2- Michel Volle, *Analyse des données,* Economica 1997

3- Brigitte Escoufier, Jérôme Pagès, *Analyses factorielles simples et multiples: Objectifs, méthodes et interpré tation.*

4- Philippe Besse, *Analyse en composantes principales.*www.math.univ-toulouse.fr/~besse/enseignement.htm

individus.jpeg

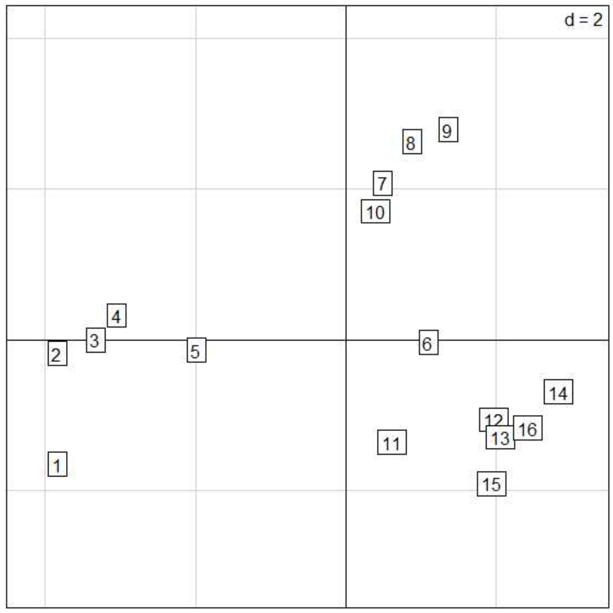
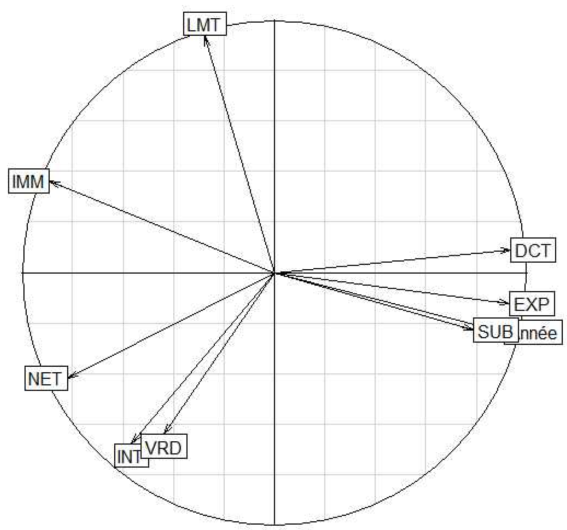


Figure 1: Représentation des individus.



de corre.JPEG

Figure 2: Représentation des variables avec cercle de corrélation.

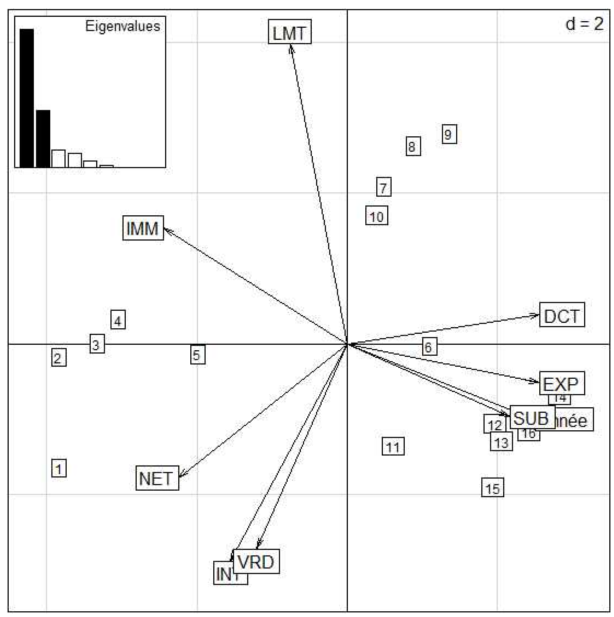


Figure 3: Représentation simultanée des individus et des variables.