Faculté d'Electronique er d'Informatique
Département d'Informatique
MASTER 1 : MIV Série N°=2 USTHB : 03/18

EXERCICE 1: On considère la matrice de données suivante:

$$X = {}^{t}\begin{pmatrix} 1 & -1 & 0 & 1 & 0 & 1 & -1 & -1 & 0 & 0 \\ 0 & 1 & 1 & -1 & -1 & 1 & -1 & 0 & 1 & -1 \\ 1 & -1 & 0 & 1 & 0 & 1 & -1 & -1 & 0 & 0 \end{pmatrix}$$

1) Donner le nuage des points $N(I)$.
2) Calculer le centre de gravité de ce nuage. Que peut-on déduire ?.
3) Déterminer la matrice des variances-covariances V.
4) Posons $S = {}^{t}X.X$. Montrer que S possède une valeur propre nulle (sans faire de calculs).
5) Calculer les valeurs propres de S et en déduire celles de V.
6) Déterminer le meilleur plan qui ajuste $N(I)$.(ACP non normée).

EXERCICE 2: Soit X la matrice de données suivante :

$$X = \begin{pmatrix} 8 & 30 & 55 \\ 2 & 6 & 40 \\ 5 & 15 & 30 \\ 7 & 22 & 40 \end{pmatrix}$$

1) Donner le centre de gravité du nuage de points.
2) Déterminer la matrice des données centrées-réduites.
3) Donner l'expression et le type de la matrice des coefficients de corrélation.
4) Après calculs, nous trouvons les vecteurs propres u_i associés aux valeurs propres λ_i de R : ${}^{t}u_1 = (-0.5988, -0.6277, -0.4974)$, ${}^{t}u_2 = (-0.6515, 0.7430, -0.1534)$, ${}^{t}u_3 = (-0.4658, -0.2322, 0.8539)$.

   et $\lambda_1 = 2.4598$, $\lambda_2 = 0.0034$, $\lambda_3 = 0.5368$. Donner la valeur de la variance expliquée par les axes factoriels choisis pour le plan principal.
5) Déterminer les axes factoriels choisis.
6) Calculer l'inertie expliquée par les axes choisis. Commenter le résultat obtenu.
7) Déterminer les composantes principales.

EXERCICE 3 : Soient X une matrice de données de type $(4,3)$: $X = \begin{pmatrix} 0 & 2 & 3 \\ 0 & 0 & 3 \\ 4 & 2 & 1 \\ 4 & 0 & 1 \end{pmatrix}$

1) Déterminer la matrice des variances-covariances et celle des corrélations.
2) Calculer les variances des différentes composantes principales.
3) En calculant le taux d'inertie expliquée par chaque axe, donner le nombre d'axes à retenir.
4) Déterminer les coordonnées des individus le long des axes choisis.
5) Etudier la corrélation entre les variables initiales et les composantes principales.

EXERCICE 4 : Soit X une matrice de données de type (96, 8). La matrice de corrélation entre les variables est donnée par la matrice suivante :

$$R = \begin{pmatrix} 1 & 0.56 & 0.28 & 0.20 & 0.43 & 0.29 & 0.31 & 0.43 \\ 0.56 & 1 & 0.34 & 0.32 & 0.48 & 0.32 & 0.47 & 0.49 \\ 0.28 & 0.34 & 1 & 0.64 & 0.72 & 0.52 & 0.70 & 0.45 \\ 0.20 & 0.32 & 0.64 & 1 & 0.67 & 0.46 & 0.58 & 0.44 \\ 0.43 & 0.48 & 0.72 & 0.67 & 1 & 0.56 & 0.68 & 0.53 \\ 0.29 & 0.32 & 0.52 & 0.46 & 0.56 & 1 & 0.64 & 0.42 \\ 0.31 & 0.47 & 0.70 & 0.58 & 0.68 & 0.64 & 1 & 0.51 \\ 0.43 & 0.49 & 0.45 & 0.44 & 0.53 & 0.42 & 0.51 & 1 \end{pmatrix}$$

1) Commenter la matrice de corrélation.
2) La diagonalisation de R donne les valeurs propres suivantes : 4.4242, 1.1366, 0.583, 0.5302, 0.4580, 0.3758, 0.2582, 0.2287. Calculer le pourcentage d'inertie expliquée par chaque axe.
3) Déterminer le meilleur sous espace principal de dim 2 ainsi que le taux expliqué de l'inertie initiale.
4) Donner les expressions des coordonnées des individus dans le plan principal ainsi que celles des variables.
5) Après calculs, on trouve les coordonnées des variables :

Variables	1^{er} axe	$2^{ème}$ axe
X^1	0.55	0.68
X^2	0.65	0.55
X^3	0.81	-0.32
X^4	0.75	-0.35
X^5	0.87	-0.1
X^6	0.72	-0.2
X^7	0.84	-0.18
X^8	0.71	0.25

Représenter graphiquement les variables sur le plan factoriel (cercle de corrélation). Que constater-vous ?.

Exo 1 :

Solution :

$$\begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & -1 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & -1 & 0 \\ 1 & 1 & 1 \\ -1 & -1 & -1 \\ -1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \end{pmatrix}$$

1- $N(I) = \{(x_i, w_i)\} = \{(X_1, X_2, X_3, \frac{1}{10})\}$

2- Centre de gravité $= (0,0,0)$

3- Matrice de variances-Covariance $= \frac{1}{10}\, {}^t X_0 . X_0$

$$\frac{1}{10}\begin{pmatrix} 6 & 0 & 6 \\ 0 & 8 & 0 \\ 6 & 0 & 6 \end{pmatrix}$$

4- On remarque que $L1 = L3$ dans la matrice S et donc $\det S = 0 \Rightarrow \exists$ 0 valeur propre de S.

5- $\det [S - (X * I)]$

$$\det\begin{pmatrix} 6-X & 0 & 6 \\ 0 & 8-X & 0 \\ 6 & 0 & 6-X \end{pmatrix} = (8-X)((6-X)^2 - 36)$$
$$(8-X)(X^2 - 12X)$$
$$X(8-X)(X-12)$$

$\Rightarrow$ $X = 0$, $X = 12$, $X = 8$ les valeurs de S. Donc les valeurs de V $\Rightarrow$ $X = 0$, $X = 12/10 = 6/5$, $X = 8/10 = 4/5$

1

Scanned by CamScanner

6- $$\begin{pmatrix} 6/10 & 0 & 6/10 \\ 0 & 8/10 & 0 \\ 6/10 & 6/10 & 6/10 \end{pmatrix}\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \frac{12}{10}\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$$

$$\begin{cases} 6x + 6z = 12x \\ 8y = 12y \\ 6x + 6z = 12z \end{cases} \Rightarrow \begin{matrix} x = z \\ y = 0 \end{matrix} \Rightarrow X\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

$$V_1 = \left(\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2}\right)$$

$$\begin{cases} 6x + 6z = 8x \\ 8y = 8y \\ 6x + 6z = 8z \end{cases} \Rightarrow \begin{cases} -2x + 6z = 0 \\ 6x - 2z = 0 \\ y = y \end{cases} \Rightarrow \begin{cases} 16z = 0 \\ z = 0 \\ x = 0 \end{cases}$$

$$\Rightarrow \begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix} \Rightarrow V_2\,(0, 1, 0)$$

Exo 2 :

Exercice 2 TD 2 (2016)

$$X = \begin{pmatrix} 8 & 30 & 55 \\ 2 & 6 & 40 \\ 5 & 15 & 30 \\ 7 & 22 & 40 \end{pmatrix}$$

1) Centre de gravité du nuage de points.

$$\bar{X}_1 = \frac{8+2+5+7}{4} = \frac{22}{4} = \frac{11}{2} = 5,5$$

$$\bar{X}_2 = \frac{30+6+15+22}{4} = \frac{73}{4} = 18,25$$

$$\bar{X}_3 = \frac{55+40+30+40}{4} = \frac{165}{4} = 41,25$$

$$G = (5,5 \; ; \; 18,25 \; ; \; 41,25)$$

2)

* matrice centrée :

$$X_0 = \begin{pmatrix} 8-5,5 & 30-18,25 & 55-41,25 \\ 2-5,5 & 6-18,25 & 40-41,25 \\ 5-5,5 & 15-18,25 & 30-41,25 \\ 7-5,5 & 22-18,25 & 40-41,25 \end{pmatrix}$$

$$X_0 = \begin{pmatrix} 2,5 & 11,75 & 13,75 \\ -3,5 & -12,75 & -1,75 \\ -0,5 & -2,75 & -11,25 \\ 1,5 & 3,75 & -1,25 \end{pmatrix}$$

Matrice centrée réduite Z.

la variance

$Var_1 = \frac{1}{4}\left[(2,5)^2 + (-3,5)^2 + (-0,5)^2 + (1,5)^2\right] = 5,25$

l'écart-type $= \sigma_1 = \sqrt{Var_1} = 2,29$.

$V_2 = 80,56$ $\quad V_3 = 79,68$

$\sigma_2 = 8,97$ $\quad \sigma_3 = 8,92$

$$Z = X_0 / \sigma_n = \begin{pmatrix} 2,5/\sigma_1 = 1,09 & 1,31 & 1,54 \\ -3,5/\sigma_1 = -1,53 & -1,42 & -0,14 \\ -0,22 & -0,3 & -1,26 \\ 0,65 & 0,41 & -0,14 \end{pmatrix}$$

3) $R = \frac{1}{m} Z^T Z$ $\quad$ le type = nombres de variables

$\Rightarrow$ type $(3,3)$

4) Z matrice centrée réduite.

les variances sont les valeurs propres de la matrice

$v_1 = 2,4598$, $v_3 = 0,5368$

5) calculer les vecteurs propres

6) a) Inertie = 0,5368 + 2,4598) = 2,9966

b) ~~Tout~~ Somme des deux grandes valeurs propres / Σ Toutes les valeurs propres

$(0,5368 + 2,4598) / (0,5368 + 2,4598 + 0,0034) = 0,9988$

7/ $C_k = Z \cdot U_k$

$C_1 = Z \cdot U_1$

$$\begin{pmatrix} 1,09 & 1,31 & 1,54 \\ -1,53 & -1,42 & -0,14 \\ -0,22 & -0,3 & -1,26 \\ 0,65 & 0,41 & -0,11 \end{pmatrix} \cdot (-0,5988 ; -0,6277 ; -0,4974)$$

$C_1 = (-2,24092 ; 1,7378 ; -0.3066 ; -0,7162)$

$C_2 = Z \cdot U_2$

$C_3 = Z \cdot U_3$

8/ $r(x^j, C^k) = \varphi^k_j$

$\varphi^k = \sqrt{\lambda_k}\, u_k$

C^1, C^2

$u_1 = (-0{,}5988, -0{,}6277, -0{,}4974)$

$r(x^1, C^1) = \varphi^1_1$

$\varphi^1 = \sqrt{\lambda_1}\, u_1$

$$= \sqrt{2{,}4598} \times \begin{pmatrix} -0{,}5988 \\ -0{,}6277 \\ -0{,}4974 \end{pmatrix} = \begin{pmatrix} -0{,}9391 \\ -0{,}9845 \\ -0{,}7801 \end{pmatrix}$$

$r(x^1, C^1) = -0{,}9391$

$r(x^2, C^1) = -0{,}9845$

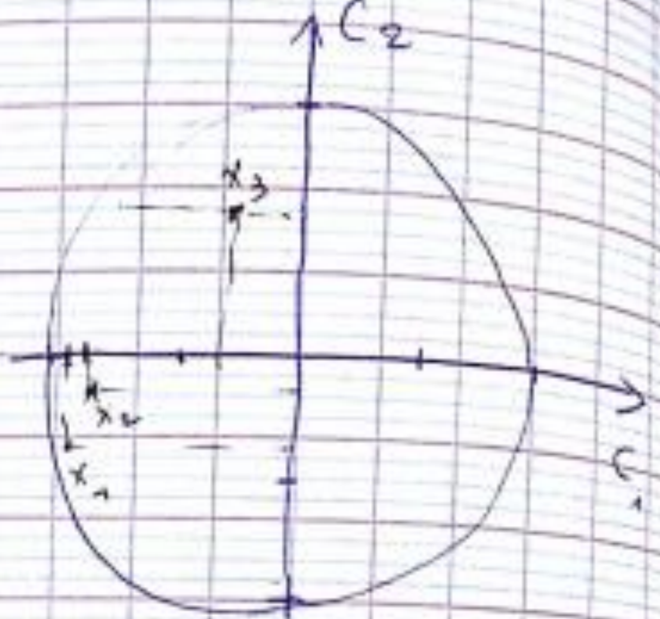
$r(x^3, C^1) = -0{,}7801$

$$r(x^1, C^2) = \sqrt{\lambda_3}\, u_3 = \sqrt{0{,}5368} \begin{pmatrix} -0{,}4658 \\ -0{,}2322 \\ 0{,}8539 \end{pmatrix} = \begin{pmatrix} -0{,}3413 \\ -0{,}1701 \\ 0{,}6256 \end{pmatrix}$$

Scanned by CamScanner

Les variables dans ce plan $C(C^1, C^2)$

	C^1	C^2
x_1	-0,9391	-0,3413
x_2	-0,9845	-0,1701
x_3	-0,1801	0,6256



Exo 3 :

Exercice 3 :

$$X = \begin{pmatrix} 0 & 2 & 3 \\ 0 & 0 & 3 \\ 4 & 2 & 1 \\ 4 & 0 & 1 \end{pmatrix}$$

$g(2, 1, 2)$

$$Y \begin{pmatrix} -2 & 1 & 1 \\ -2 & -1 & 1 \\ 2 & 1 & -1 \\ 2 & -1 & -1 \end{pmatrix}$$

$$V = \frac{1}{m} \cdot {}^tY \cdot Y$$

$$V = \frac{1}{4} \begin{pmatrix} -2 & -2 & 2 & 2 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} -2 & 1 & 1 \\ -2 & -1 & 1 \\ 2 & 1 & -1 \\ 2 & -1 & -1 \end{pmatrix}$$

$$V = \frac{1}{4} \begin{pmatrix} 16 & 0 & -8 \\ 0 & 4 & 0 \\ -8 & 0 & 4 \end{pmatrix} = \begin{pmatrix} 4 & 0 & -2 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}$$

$$Z = \begin{pmatrix} -1 & 1 & 1 \\ -1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \end{pmatrix}$$

$$R = \frac{1}{m}\,{}^{t}Z \cdot Z$$

$$R = \frac{1}{4}\begin{pmatrix} -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{pmatrix}\begin{pmatrix} -1 & 1 & 1 \\ -1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 4 & 0 & -4 \\ 0 & 4 & 0 \\ -4 & 0 & 4 \end{pmatrix}$$

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

$$R \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

$$\det(R - x\,Id) = \det\begin{pmatrix} 1-x & 0 & -1 \\ 0 & 1-x & 0 \\ -1 & 0 & 1-x \end{pmatrix}$$

$= (1-x)\left((1-x)^2 - 1\right)$

$= (1-x)(-x)(2-x)$

$$\begin{cases} x_1 = 1 \\ x_2 = 0 \\ x_3 = 2 \end{cases}$$

$Var(C_1) = 2$

$Var(C_2) = 1$

$Var(C_3) = 0$

$\Sigma\lambda_i = 0+1+2 = 3$

Taux
$$\boxed{\begin{aligned} T_1 &= \frac{\lambda_1}{\Sigma\lambda_i} = \frac{2}{3} \\ T_2 &= \frac{1}{3} \end{aligned}}$$

$T_3 = 0$

$C_1 = 2 . U_1$

V_1 vecteur propre normalisée associé à la plus grande valeur propre $\lambda_1 = 2$

$$RV = V\lambda_1 \Rightarrow \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 2\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$\begin{cases} x - z = 2x \\ y = 2y \Rightarrow y = 0 \qquad x = -z \quad x\begin{pmatrix}1\\0\\-1\end{pmatrix} \Rightarrow V_1 \begin{pmatrix}\sqrt{2}/2\\0\\-\sqrt{2}/2\end{pmatrix} \\ -x + z = 2z \end{cases}$$

$$\boxed{C^1 = ZV_1} = \begin{pmatrix} -1 & 1 & 1 \\ -1 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \end{pmatrix} \cdot \begin{pmatrix}\sqrt{2}/2\\0\\\sqrt{2}/2\end{pmatrix} = \begin{pmatrix}-\sqrt{2}\\-\sqrt{2}\\+\sqrt{2}\\\sqrt{2}\end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix} = 1\begin{pmatrix}x\\y\\z\end{pmatrix}$$

$$\begin{cases} x - z = x \Rightarrow z = 0 \\ y = y \qquad\qquad y = \begin{pmatrix}0\\1\\0\end{pmatrix} \\ -x + z = z \Rightarrow x = 0 \end{cases}$$

$$C^2 = ZV_2 = \begin{pmatrix} -1 & 1 & 1 \\ -1 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \end{pmatrix} \begin{pmatrix}0\\1\\0\end{pmatrix} = \begin{pmatrix}1\\-1\\1\\-1\end{pmatrix}$$

5. Etudier la corrélation entre les variables et les composantes principales

Réponse

$$f^k = \sqrt{\lambda_k} \cdot u_k$$

Pour C^1

$$f^1 = \sqrt{\lambda_1} \cdot {}^t\left(\frac{\sqrt{2}}{2}, 0, -\frac{\sqrt{2}}{2}\right)$$

$$= \sqrt{2} \cdot {}^t\left(\frac{\sqrt{2}}{2}, 0, -\frac{\sqrt{2}}{2}\right) = {}^t(1, 0, -1)$$

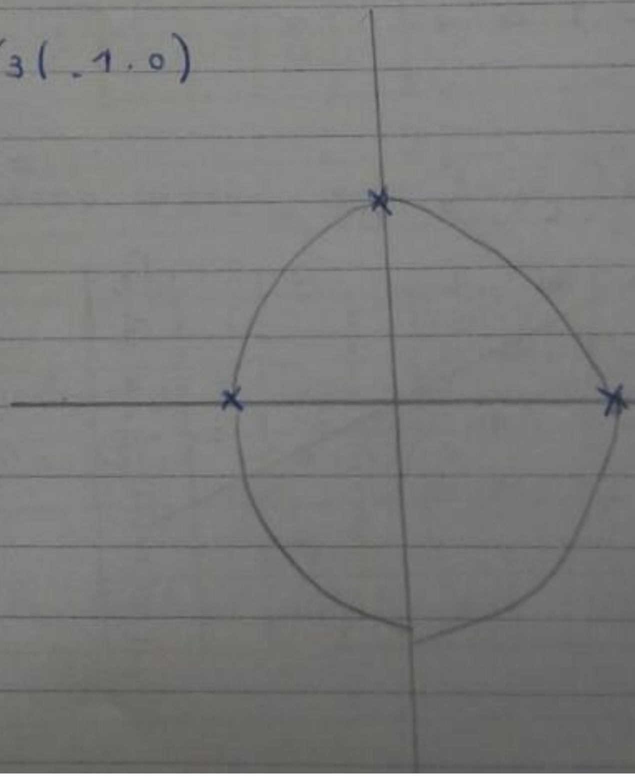
$r(X^1, C^1)$ $r(X^2, C^1)$ $r(X^3,$

Pour C^2

$$f^2 = \sqrt{\lambda_2} \cdot u_2 = \sqrt{1} \cdot (0, 1, 0)$$

	C_1	C_2
X_1	1	0
X_2	0	1
X_3	-1	0

$X_1(1, 0)$ $X_2(0, 1)$ $X_3(-1, 0)$



toutes les variables sont dans la surface du cercle

$\Longrightarrow$ Les données sont très bien représentée

Exo 4 :

Exercice N°: 04

Soit La Matrice de données de type (8.8). La Matrice de corrélation entre Les variables est données par:

$$R = \begin{array}{cccccccc} 1 & 0,56 & 0,28 & 0,20 & 0,43 & 0,29 & 0,31 & 0,43 \\ 0,56 & 1 & 0,34 & 0,32 & 0,48 & 0,32 & 0,47 & 0,49 \\ 0,28 & 0,34 & 1 & 0,64 & 0,72 & 0,52 & 0,70 & 0,45 \\ 0,20 & 0,32 & 0,64 & 1 & 0,67 & 0,46 & 0,58 & 0,44 \\ 0,43 & 0,48 & 0,72 & 0,67 & 1 & 0,56 & 0,68 & 0,53 \\ 0,29 & 0,32 & 0,52 & 0,46 & 0,56 & 1 & 0,64 & 0,42 \\ 0,31 & 0,47 & 0,70 & 0,58 & 0,68 & 0,64 & 1 & 0,51 \\ 0,43 & 0,49 & 0,45 & 0,44 & 0,53 & 0,42 & . & 1 \\ x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 \end{array}$$

1- Commenter La Matrice de corrélation.

On commente comme suite: La variable x_1 est moyennement correlé a x_2 faiblement a x_3...

2- La diagonalisation de R donnes Les valeurs propres suivantes: 4,4242 . 1,1366 . 0,583 0,5302 0,4580 0,3758 0,2582 0,2287. Calculer Le pourcentage dinertie expliquée par chaque axe.

Réponse

$$\text{taux dinertie} = \frac{\lambda_i}{\sum \lambda_i} = \frac{\lambda_i}{n} \qquad n = 8$$

$$I_1 = \frac{4,4242}{8} = 0,553 \ (55\%) \qquad I_5 = \frac{0,4580}{8} = 0,05 \ (\simeq 5\%)$$

$$I_2 = \frac{1,1366}{8} = 0,1420 \ (\simeq 14\%) \qquad I_6 = \frac{0,3758}{8} = 0,0469 \ (\simeq 4\%)$$

$$I_3 = \frac{0,583}{8} = 0,07 \ (\simeq 7\%) \qquad I_7 = \frac{0,2582}{8} = 0,03 \ (3\%)$$

$$I_4 = \frac{0,5302}{8} = 0,06 \ (\simeq 6\%) \qquad I_8 = \frac{0,2287}{8} = 0,028 \ (\simeq 2\%)$$

3 - Determiner Le meilleur sous espace principal de dim 2 ainsi que Le taux expliqué de L'inertie initiale.

Réponse.

Le meilleur sous espace principal de dim 2 est engendré par Les deux plus grands valeurs propres autrement dit λ_1 et λ_2

$$\text{taux dinertie} = \frac{\lambda_1 + \lambda_2}{8} = 0,675 \approx 67.5\%$$

4 Donner Les expressions des coordonnées des individus dans Le plan principal ainsi que celle des variables.

Réponse

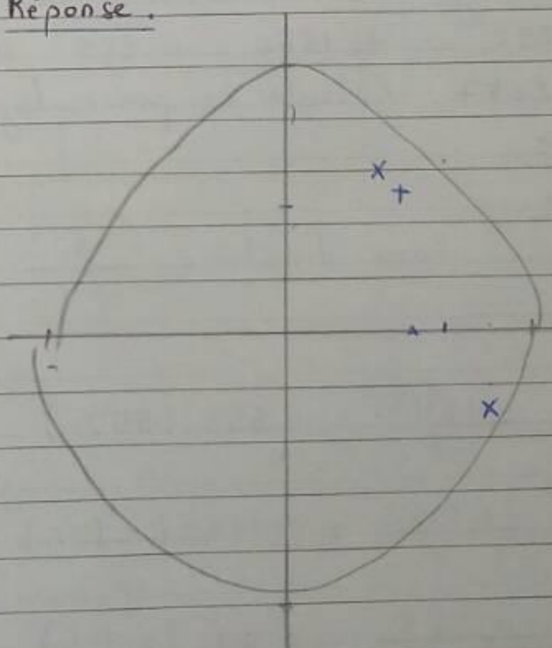
Les coordonnee des individus sont C^1. C^2 : Les composantes principales. $C^i = Z.U^i$ $i.1.2$

Les coordonnés des variables sont $X_j^k = \sqrt{\lambda_k}.U^k$

5 Aprés calcul, on trouve les coordonnées suivantes. Representez graphiquement sur un cercle de corrélation. Que constatez vous?

Varia	C^1	C^2
X^1	0,55	0,68
X^2	0,65	0,55
X^3	0,81	-0,32
X^4	0,75	-0,35
X^5	0,87	-0,1
X^6	0,72	-0,2
X^7	0,84	-0,18

Réponse.



Remarques

Bonne réduction. Néamoins, certaines variables sont mieux representé que d'autre.