

Examen du 19 décembre 2008-de 14h à 16h

Documents autorisés : notes de cours et de TP. Les calculatrices sont autorisées.

1 Problème 1

Calcul matriciel

On considère les matrices suivantes :

$$M := \begin{pmatrix} 91 & 69 & \frac{9}{\sqrt{2}} & 9\sqrt{\frac{3}{2}} \\ 69 & 91 & -\frac{9}{\sqrt{2}} & -9\sqrt{\frac{3}{2}} \\ \frac{9}{\sqrt{2}} & -\frac{9}{\sqrt{2}} & \frac{29}{5} & \frac{27\sqrt{3}}{5} \\ 9\sqrt{\frac{3}{2}} & -9\sqrt{\frac{3}{2}} & \frac{27\sqrt{3}}{5} & \frac{83}{5} \end{pmatrix}$$

$$O := \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2\sqrt{2}} & -\frac{\sqrt{3}}{2\sqrt{2}} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2\sqrt{2}} & \frac{\sqrt{3}}{2\sqrt{2}} \\ 0 & 0 & -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}$$

- 1) Expliquer pourquoi la matrice M ne possède que des valeurs propres réelles. Est-elle diagonalisable ?
- 2) Pour $\lambda \in \mathbb{R}$, Le polynôme caractéristique de M est donné par

$$P(\lambda) := \frac{1}{5} (5\lambda^4 - 1022\lambda^3 + 36408\lambda^2 - 142400\lambda + 51200).$$

Dire pourquoi le déterminant de M vaut $\frac{51200}{5} = 10240$. Que vaut le produit des valeurs propres ?

- 3) Montrer que les trois réels suivants sont des valeurs propres de M : 160, 40, 4. Montrer que la somme des valeurs propres vaut 204.4 et en déduire la dernière valeur propre.
- 4) La matrice O contient les coordonnées des 4 vecteurs propres. Ces vecteurs propres sont-ils écrits en ligne ou en colonne ?

Analyse en composantes principales

On a effectué une enquête auprès de $n = 1000$ individus. Sur chacun de ces n on a mesuré 4 variables quantitatives. On note $X_{ij}, j = 1, \dots, 4, i = 1, \dots, 1000$ les observations que l'on range dans la matrice $X = (X_{ij})$. La matrice de variance covariance obtenue à partir de ces données est M . On travaille avec des poids uniformes et la distance euclidienne habituelle et l'on s'intéresse à l'ACP associée à notre enquête.

- 1) Comment a-t-on calculé M à partir de X ?
- 2) Que vaut l'inertie totale du nuage de points ?

- 3) Quel pourcentage de cette inertie est expliqué par les deux premières composantes principales ?
- 4) Représenter les points qui suivent dans le premier plan principal :

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

2 Problème 2

Le tableau suivant fournit, pour 15 pays d'Afrique de l'ouest, 5 variables démographiques (espérance de vie masculine à la naissance, taux de natalité, de mortalité, de mortalité infantile, indice conjoncturel de fécondité), 1 variable sociale (PNB/habitant) et 1 variable sanitaire (VIH % : proportion des 15-54 ans atteints pas le virus du SIDA).

	ESPVIEH	PNB par hab	VIH %	NATAL	MORTAL	MORT INFANT	FECOND
Bénin	49	920	2.5	45	15	94	6.3
Burkina	47	960	6.4	47	17	105	6.8
Cote d'Ivoire	45	1540	10.8	36	16	112	5.2
Gambie	51	1550	2	43	14	82	5.9
Ghana	56	1850	3.6	32	10	56	4.3
Guinée	43	1870	1.5	41	19	98	5.5
Guinée- Bissau	44	630	2.5	42	20	131	5.8
Liberia	49	1500	2.8	49	17	139	6.6
Mali	45	740	2	50	20	123	7
Mauritanie	49	1550	1.5	43	15	106	6
Niger	41	740	1.4	53	24	123	7.5
Nigeria	52	770	5.1	41	14	75	5.8
Sénégal	51	1400	1.8	41	13	68	5.7
Sierra Leone	42	440	3	47	20	153	6.3
Togo	53	1380	6	40	11	80	5.8

Tableau 1 Variables pour 15 pays ouest africains (source ONU 2001)

	ESPVIEH	PNB par hab	VIH %	NATAL	MORTAL	MORT INFANT	FECOND
Moyenne	47.80	1189	3.527	43.33	16.33	103.00	6.033
Ecart type	4.39	466	2.578	5.42	3.792	27.73	0.775
Coef de variation	0.092	0.392	0.731	0.125	0.232	0.269	0.128

Tableau 2 paramètres résumant les 7 distributions

r	ESPVIEH	PNB par hab	VIH %	NATAL	MORTAL	MORT INFANT	FECOND
ESPVIEH	1.00	0.48	0.11	-0.58	-0.94	-0.80	-0.55
PNB par hab		1.00	0.08	-0.59	-0.57	-0.53	-0.62
VIH %			1.00	-0.42	-0.28	-0.08	-0.29
NATAL				1.00	0.73	0.65	0.98
MORTAL					1.00	0.80	0.69
MORT INFANT						1.00	0.59
FECOND							1.00

Tableau 3 matrice de corrélations entre les variables

r	F1	F2	F3
ESPVIEH	-0.84	-0.33	-0.37
PNB par hab	-0.71	-0.11	0.50
VIH %	-0.32	0.86	-0.33
NATAL	0.90	-0.29	-0.20
MORTAL	0.93	0.13	0.27
MORT INFANT	0.84	0.31	0.17
FECOND	0.87	-0.22	-0.32

Tableau 4 Corrélations variables-axes factoriels

Les pourcentages de variance sont : axe F1 64%, axe F2 16% et axe F3 10%.

- 1) Présenter les données et indiquer la problématique.
- 2) Présenter le type d'analyse effectué en la justifiant.
- 3) Quelles sont, dans les 15 pays ouest africains, les distributions les plus variables ?
- 4) Commenter les corrélations entre les 7 distributions.
- 5) Interpréter les 3 premiers axes factoriels.