IUP SID
Statistique Multidimensionnelle
2010-2011

# Examen du 9 Décembre 2011 de 7h45 à 9h45

*Les notes de cours sont autorisées. Les calculatrices sont autorisées.*

## 1 Algèbre linéaire

Soit $n$ un entier naturel non nul et soit $u_1$ un vecteur colonne, donné, de $\mathbb{R}^n$ de norme 1. La matrice identité de taille $n$ sera notée $I_n$. Pour $\lambda \in \mathbb{R}$, on considère la matrice $n \times n$

$$A := I_n + \lambda u_1 u_1^T$$

1. Dire pourquoi $A$ est diagonalisable ?
2. Montrer que $1 + \lambda$ est une valeur propre de $A$ associé au vecteur propre $u_1$.
3. Soit $u_2, \cdots, u_n$ des vecteurs colonnes tous de norme 1 tels que, pour $i, j \in \{1, 2, \cdots, n\}$, $i \neq j$ on a $u_i^T u_j = 0$.
4. Montrer que $u_2, \cdots, u_n$ sont des vecteurs propres de $A$.
5. Déterminer une matrice inversible $M$ telle que la matrice
$$D := M^{-1} A M$$
soit diagonale.
6. Que vaut le déterminant de $A$ ?
7. Pour quelles valeurs de $\lambda$ la matrice $A$ est inversible ? positive ? définie positive ?

## 2 Dentiste

On considère les 7 points de $\mathbb{R}^2$ suivants :

$$A := \begin{pmatrix} 0 \\ 0 \end{pmatrix},\ B := \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ C := \begin{pmatrix} \sqrt{2^{-1}} \\ \sqrt{2^{-1}} \end{pmatrix},\ D := \begin{pmatrix} 0 \\ 1 \end{pmatrix},\ E := \begin{pmatrix} -1 \\ 0 \end{pmatrix},\ F := \begin{pmatrix} -\sqrt{2^{-1}} \\ -\sqrt{2^{-1}} \end{pmatrix},\ G := \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

Soit les points de $\mathbb{R}^2$, $u := \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ et $v := \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$. On considère les quatre distances suivantes :

$$\begin{aligned} d_0(u,v) &:= \sharp\{i \in \{1,2\} : u_i \neq v_i\},\ d_1(u,v) := |u_1 - v_1| + |u_2 - v_2|, \\ d_2(u,v) &:= \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2},\ d_\infty(u,v) := \max(|u_1 - v_1|, |u_2 - v_2|). \end{aligned}$$

1. Représenter graphiquement les points $A, B, C, D, E, F, G$.
2. Pour chacune des distances $d_0, d_1, d_2, d_\infty$, construire le dendogramme obtenu lorsque que l'on effectue la classification hiérarchique ascendante des points $A, B, C, D, E, F, G$ (la méthode d'agrégation est bien sûr le saut minimum).
3. Bâtir la matrice de variance covariance empirique des observations $A, B, C, D, E, F, G$.
4. Effectuer l'ACP pour les observations $A, B, C, D, E, F, G$. Quel pourcentage de la variance est expliquée par le premier axe principale ?
5. On reprend les deux questions précédentes pour les observations $A, B, D, E, G$. Comment interprétez-vous ce résultat ?