TD (3) Algèbre Linéaire

matrice de Variance-Covariance, base des vecteurs propres

Lorsque l'on simule ou que l'on observe p **variables aléatoires** sur n **individus**,
on peut représenter le nuage des points observés dans l'espace $\mathbb{R}^p$. Lorsque $p = 2$ ou $p = 3$, ce nuage de points peut correspondre par exemple à « un objet » dans une image binaire.

Ces données sont rangées dans une matrice X, un tableau $n \times p$ tel que :
chaque colonne contient les observations d'une variable sur les n individus
chaque ligne contient les observations des p variables pour un individu.

$$X = \begin{pmatrix} x_1^1 & \dots & x_1^j & \dots & x_1^p \\ \vdots & & \vdots & & \vdots \\ x_i^1 & \dots & x_i^j & \dots & x_j^p \\ \vdots & & \vdots & & \vdots \\ x_n^1 & \dots & x_n^j & \dots & x_n^p \end{pmatrix} = \begin{pmatrix} | & & | & & | \\ X^1 & \dots & X^j & \dots & X^p \\ | & & | & & | \end{pmatrix}$$

Dans ce cours-TD, nous allons nous intéresser à la matrice de Variance-Covariance associée à ces p variables.

Exercice 1

Dans un premier temps nous allons supposer que l'on observe 2 variables X et Y pour simplifier les écritures (éviter l'indice correspondant au numéro de la variable).
Chaque individu est donc représenté par un vecteur $\begin{pmatrix} x_i \\ y_i \end{pmatrix}$

1. Ecrire la matrice de Variance-Covariance empirique associées à l'observation de ces n individus. On notera V cette matrice.

2. Si chacune des 2 variables X et Y est **centrée**, vérifier que l'on peut écrire V sous la forme :

$$V = \frac{1}{n} B^t \times B$$

   où B contient dans chacune de ses colonnes les observations d'une variable.
   La matrice V est une matrice réelle symétrique. En plus des propriétés que nous avons vues, on peut montrer que toutes ses valeurs propres sont toutes **positives**.

3. Nous allons dans ce qui suit effectuer un **changement de base** en exprimant les coordonnées des vecteurs représentant les individus dans **une base de vecteurs propres** de la matrice V.

   Nous choisirons une **base orthonormée** notée $\mathcal{B}' = (u_1, u_2)$
   Comment s'écrit la matrice de passage vers cette nouvelle base et quelles sont ses propriétés ?

   Exprimer les nouvelles coordonnées d'un individu dans cette base en fonction des anciennes :

$$\begin{pmatrix} x' \\ y' \end{pmatrix} \quad \text{en fonction de} \quad \begin{pmatrix} x \\ y \end{pmatrix}$$

$$\begin{pmatrix} x_1' & x_2' & .. & x_n' \\ y_1' & y_2' & .. & y_n' \end{pmatrix} \quad \text{en fonction de} \quad \begin{pmatrix} x_1 & x_2 & .. & x_n \\ y_1 & y_2 & .. & y_n \end{pmatrix}$$

4. Pour finir, exprimer la matrice de Variance-Covariance dans la base orthonormée des vecteurs propres $\mathcal{B}'$. Que pouvez vous dire de cette matrice notée V' ? (commenter)