

## TP 9 : Réduction de la dimensionnalité de données

### Énoncé

1. Générer un ensemble  $X_1$  de  $N = 500$  vecteurs 2-dimensions à partir d'une distribution gaussienne avec la moyenne zéro et la matrice de covariance :

$$S_1 = \begin{bmatrix} 0.3 & 0.2 \\ 0.2 & 1.0 \end{bmatrix}$$

Effectuer une analyse ACP sur  $X_1$ ; C'est-à-dire, calculer les deux valeurs propres/vecteurs propres de l'estimateur  $\hat{S}_1$  de  $S_1$  obtenue en utilisant les vecteurs de  $X_1$ . Tenant compte du fait que la  $i^{\text{ème}}$  valeur propre "explique" la variance selon la direction du  $i^{\text{ème}}$  vecteur propre de  $S_1$ , calcule le pourcentage de la variance totale expliqué par chacune des deux composantes comme étant le rapport  $\frac{\lambda_i}{\lambda_0 + \lambda_1}$ ,  $i = 0, 1$ . Dessiner l'ensemble de données  $X_1$  ainsi que les vecteurs propres de  $S_1$ . Commenter les résultats.

2. De même, générer un ensemble de données  $X_2$ , maintenant avec la matrice de covariance

$$S_2 = \begin{bmatrix} 0.3 & 0.2 \\ 0.2 & 9.0 \end{bmatrix}$$

Répétez l'expérience précédente.