

# Fiche TP (R)

## 1 simulation :

1.1 générer un tableau  $X$  de  $\frac{n=100}{\text{individus}} \times \frac{p=4}{\text{variables}}$

provenant d'une distribution normale multivariée

$N(\mu, \Sigma)$  avec

$$\mu = (1, 2, 4, 3) \text{ et}$$

$$\Sigma = \begin{pmatrix} 0.7 & 0 & 1.3 & 0.5 \\ 0 & 1.2 & -0.3 & -0.1 \\ 1.3 & -0.3 & 3.1 & 1.3 \\ 0.5 & -0.1 & 1.3 & 0.6 \end{pmatrix}.$$

Nb : utiliser le package "mvtnorm"

## 1.2 Réaliser l'ACP normée de $X$

- utiliser le package "FactoMineR"

- regarder les composantes principales et les facteurs propres.

1.3 programmer NIPALS. Comparer les résultats fournis par NIPALS avec ceux donnés par FactoMineR.

1.4. générer aléatoirement des valeurs manquantes dans le tableau  $X$ . On va supposer que  $P(X_{ij} \text{ soit manquante}) = 0.05$  (Données MCAR donc !).

①.5 On considère maintenant

$\tilde{X}$  le tableau avec données manquantes obtenu en 1.4.

Applique NIPALS pour estimer les données manquantes. Comparer avec les vraies valeurs

②. Application réelle.

Imputer par NIPALS les données manquantes présentes dans la base "airquality" (base de données existante en R)

> help(airquality)

— enjoy! —

Fin imputation simple. Passons à l'imputation multiple.