Université de Montpellier **HMMA307**
2020/2021 **Joseph Salmon**

TD N° 3 : Modèles à effets aléatoires

Exercice 1. Soit $\mathbb{1}_n \in \mathbb{R}^n$ le vecteur de taille n ne contenant que des 1. On définit alors la matrice J_n par $J_n = \mathbb{1}_n \mathbb{1}_n^\top \in \mathbb{R}^n$.

1) Montre que la matrice J_n est symétrique définie positive, et donner J_n^2. Qu'en est il pour $\bar{J_n}^2$ avec $\bar{J_n} = \frac{1}{n} J_n$?
2) Interpréter géométriquement $\bar{J_n}$ et $\overline{C}_n = \operatorname{Id}_n - \bar{J_n}$.
3) Donner $\operatorname{rang}(J_n)$ et $\operatorname{rang}(\bar{J_n}^2)$.
4) Vérifier que

$$J_n = H_n^\top \operatorname{diag}(n, 0, \dots, 0) H_n \ , \tag{1}$$
$$\bar{J_n} = H_n^\top \operatorname{diag}(1, 0, \dots, 0) H_n \ . \tag{2}$$

où $H_n \in \mathbb{R}^{n \times n}$ est la matrice de Helmert d'ordre n, c'est-à-dire la matrice orthogonale définie par :

$$H_n = \begin{bmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} \\ \frac{1}{\sqrt{n}} & \frac{-1}{\sqrt{n}} & 0 & 0 & \cdots & 0 & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & 0 & \cdots & 0 & 0 \\ \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} & \frac{-3}{\sqrt{12}} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \cdots & \frac{1}{\sqrt{n(n-1)}} & \frac{-(n-1)}{\sqrt{n(n-1)}} \end{bmatrix}$$

5) Montrer les propriétés suivantes pour tout $(a, b, a', b') \in \mathbb{R}^4$,

$$(a \operatorname{Id}_n + bJ_n)(a' \operatorname{Id}_n + b'J_n) = aa' \operatorname{Id}_n + (ab' + a'b + nbb')J_n, \tag{3}$$
$$(a \operatorname{Id}_n + bJ_n)^{-1} = \frac{1}{a}\left(\operatorname{Id}_n - \frac{b}{a + nb} J_n\right), \quad \text{pour } a \neq 0 \text{ et } a \neq -nb, \tag{4}$$
$$a \operatorname{Id}_n + bJ_n = H_n^\top \operatorname{diag}(a + nb, \underbrace{a, \dots, a}_{(n-1)\text{fois}}) H_n, \quad \text{(décomp. spectrale)} \tag{5}$$

Exercice 2. Prenons un modèle à effet aléatoire pour une catégorie à J modalités $C_1, \dots, C_J$ et n observations (avec $n = n_1 + \cdots + n_J$) :

$$y = \mu \mathbb{1}_n + \sum_{j=1}^{J} \mathbb{1}_{C_j} A_j + \varepsilon \ . \tag{6}$$

avec $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2 \operatorname{Id}_n)$ et $A \sim \mathcal{N}(0, \sigma_A^2 \operatorname{Id}_J)$.

1) Expliciter $\operatorname{Var}(y)$, la matrice de covariance de $y \in \mathbb{R}^n$.
2) En déduire grâce à l'exercice précédent l'expression de l'inverse de cette matrice de covariance.

Exercice 3. Montrer que le gradient (pour le produit scalaire $\langle A, B \rangle = \operatorname{tr}(A^\top B)$) de la fonction logdet définie par

$$\begin{aligned} \operatorname{logdet} : \mathcal{S}_n^{++} &\mapsto \mathbb{R}^{++} \\ X &\to \log(\det(X)) \ . \end{aligned}$$

est donné par

$$\nabla \operatorname{logdet}(X) = X^{-1} \quad .$$

Exercice 4. On suppose que l'on observe un n-échantillon $x_1, \ldots, x_n$ de la loi gaussienne $\mathcal{N}(\mu, \Lambda^{-1})$, avec $\mu \in \mathbb{R}^d$ et $\Lambda \in \mathcal{S}_d^{++}$. La matrice Λ s'appelle la matrice de précision et est l'inverse de la matrice de variance-covariance Σ. Donner alors l'estimateur $(\hat{\mu}, \hat{\Lambda})$ du maximum de vraisemblance du couple (μ, Λ).