

Cours 6 – Vecteurs Gaussiens.

*Eya ZOUGAR **

Institut National des Sciences appliquées-INSA

Génie mathématiques GM3
Thursday 2nd March, 2023



*Basé sur le cours de Bruno PORTIER

1. Introduction.

1. On présente dans ce cours quelques notions et résultats sur:
 - ☐ les vecteurs aléatoires de \mathbb{R}^n
 - ☐ les vecteurs gaussiens.
2. On s'intéresse à l'estimation de l'espérance et la variance d'un vecteur aléatoire quelconque et Gaussien par la méthode du maximum de vraisemblance pour le cas gaussien.

2. Quelques rappels sur les vecteurs aléatoires.

2.1. Vecteur aléatoire.

On se place dans la base canonique orthonormale de \mathbb{R}^n .

Vecteur aléatoire

Soient Z_1, Z_2, \dots, Z_n , des n variables aléatoires réelles. Alors, le vecteur Z défini par:

$$Z = \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_n \end{pmatrix} \quad \text{ou} \quad Z^\top = (Z_1, Z_2, \dots, Z_n)$$

est un vecteur aléatoire à valeurs sur \mathbb{R}^n .

2.2. Espérance d'un vecteur aléatoire.

Espérance d'un vecteur aléatoire

On peut définir l'espérance mathématique, notée $\mathbb{E}[Z]$, du vecteur aléatoire Z , comme étant le vecteur de \mathbb{R}^n dont les composantes sont les espérances mathématiques des composantes du vecteur Z , c'est à dire :

$$\mathbb{E}[Z] = \begin{pmatrix} \mathbb{E}(Z_1) \\ \mathbb{E}(Z_2) \\ \vdots \\ \mathbb{E}(Z_n) \end{pmatrix}$$

2.3. Matrice de variance-covariance.

Matrice de variance-covariance

La matrice de Variance-Covariance du vecteur aléatoire Z est définie de la manière suivante :

$$\begin{aligned}\text{Var}(Z) &= \mathbb{E} \left[(Z - \mathbb{E}(Z))(Z - \mathbb{E}(Z))^T \right] = \mathbb{E} [Z Z^T] - \mathbb{E}(Z)\mathbb{E}(Z)^T \\ &= \begin{pmatrix} \text{Var}(Z_1) & \text{Cov}(Z_1, Z_2) & \dots & \text{Cov}(Z_1, Z_n) \\ \text{Cov}(Z_2, Z_1) & \text{Var}(Z_2) & \dots & \text{Cov}(Z_2, Z_n) \\ \vdots & & \ddots & \vdots \\ \text{Cov}(Z_n, Z_1) & \text{Cov}(Z_n, Z_2) & \dots & \text{Var}(Z_n) \end{pmatrix}\end{aligned}$$

où la notation U^T désigne la transposée de U .

Remarque.

1. La matrice de variance-covariance est symétrique, semi définie positive.
2. Toutes ses valeurs propres sont donc réelles et positives ou nulles.

2.4. Propriétés.

- Soient Z_1 et Z_2 deux vecteurs aléatoires de \mathbb{R}^n , alors pour tous réels a_1, a_2 , on a :

$$\mathbb{E}[a_1 Z_1 + a_2 Z_2] = a_1 \mathbb{E}[Z_1] + a_2 \mathbb{E}[Z_2].$$

- Soient Z un vecteur aléatoire de \mathbb{R}^n et A et B des matrices à coefficients réels déterministes, de formats respectifs $p \times n$ et $1 \times q$. Alors, on a :

$$\begin{aligned} \mathbb{E}(A Z B) &= A \mathbb{E}(Z) B \\ \text{Var}(A Z) &= A \text{Var}(Z) A^T \end{aligned}$$

3. Vecteurs Gaussiens.

3.1. Définition.

Soient Z_1, Z_2, \dots, Z_n des variables aléatoires gaussiennes.

On considère le vecteur aléatoire $Z = (Z_1, Z_2, \dots, Z_n)^T$ de \mathbb{R}^n .

Vecteur Gaussien

On dit que Z est un vecteur gaussien si et seulement si toute combinaison linéaire de ses composantes est une gaussienne.

En particulier, si les variables aléatoires Z_1, Z_2, \dots, Z_n sont indépendantes, alors Z est un vecteur gaussien. On retiendra donc que toute combinaison linéaire de variables gaussiennes indépendantes est une gaussienne.

3.2. Densité d'un vecteur gaussien..

Soit Z un vecteur gaussien de \mathbb{R}^n ,

- d'espérance le vecteur μ de \mathbb{R}^n ,
- et de matrice de variance-covariance Γ .

On note alors : $Z \sim \mathcal{N}(\mu, \Gamma)$.

On notera que la matrice Γ est définie positive, donc inversible.

La densité de probabilité f de Z est une fonction de \mathbb{R}^n dans \mathbb{R}_+ .

Elle est définie pour tout $z \in \mathbb{R}^n$ par :

$$f(z) = \frac{1}{\sqrt{(2\pi)^n \det(\Gamma)}} \exp \left(-\frac{1}{2} (z - \mu)^T \Gamma^{-1} (z - \mu) \right)$$

3.3. Un résultat utile.

Lemme

Si A est une matrice déterministe de format $p \times n$ et de rang p , et si Z est un vecteur gaussien d'espérance μ et de matrice de variance-covariance Γ .

Alors AZ est un vecteur gaussien de \mathbb{R}^p et on a :

$$AZ \sim \mathcal{N}(A\mu, A\Gamma A^T)$$

4. Estimation dans le cas multivarié

4.1 Introduction.

On dispose de n données multidimensionnelles

$$(x_{j,1}, x_{j,2}, \dots, x_{j,p})_{j=1, \dots, n}.$$

Les données $(x_{j,1}, x_{j,2}, \dots, x_{j,p})_{j=1, \dots, n}$ sont les mesures de p variables quantitatives.

Pour tout j , on pose $x_j = (x_{j,1}, x_{j,2}, \dots, x_{j,p})$.

On suppose que les données x_1, x_2, \dots, x_n sont en fait les réalisations de n vecteurs aléatoires de \mathbb{R}^p , X_1, X_2, \dots, X_n indépendants et de même loi, d'espérance $\mu \in \mathbb{R}^p$ et de matrice de variance covariance Γ (matrice carrée d'ordre p).

On s'intéresse à l'estimation des paramètres $\mu = \mathbb{E}(X_1)$ et $\Gamma = \mathbb{E}[(X_1 - \mathbb{E}(X_1))(X_1 - \mathbb{E}(X_1))^T]$.

4.2 Estimation du vecteur d'espérance μ .

Pour estimer le vecteur μ , on utilise la moyenne empirique, c'est à dire qu'on estime μ par:

$$\bar{X}_n = \frac{1}{n} \sum_{j=1}^n X_j$$

- ❑ Le vecteur \bar{X}_n est un vecteur aléatoire de \mathbb{R}^p .
- ❑ C'est un estimateur sans biais et convergent du vecteur μ .
- ❑ Sa matrice de variance-covariance est égale à :

$$\text{Var}(\bar{X}_n) = \frac{1}{n} \Gamma$$

4.3 Estimation de la matrice de variance-covariance.

Pour estimer la matrice de variance-covariance

$$\Gamma = \mathbb{E} [(X_1 - \mathbb{E}(X_1))(X_1 - \mathbb{E}(X_1))^T] = \mathbb{E} [X_1 X_1^T] - \mu \mu^T$$

on utilise l'estimateur classique

$$\begin{aligned} G_n &= \frac{1}{n} \sum_{j=1}^n X_j X_j^T - \bar{X}_n \bar{X}_n^T \\ &= \frac{1}{n} \sum_{j=1}^n (X_j - \bar{X}_n)(X_j - \bar{X}_n)^T \end{aligned}$$

La matrice G_n est une matrice aléatoire carrée d'ordre p .

C'est un estimateur biaisé de la matrice Γ mais asymptotiquement sans biais.

On peut en effet montrer que :

$$\mathbb{E} [G_n] = \frac{n-1}{n} \Gamma$$

4.4. Éléments de preuve.

On établit facilement la décomposition :

$$\sum_{j=1}^n (X_j - \mu)(X_j - \mu)^T = \sum_{j=1}^n (X_j - \bar{X}_n)(X_j - \bar{X}_n)^T + n(\bar{X}_n - \mu)(\bar{X}_n - \mu)^T$$

et par suite

$$\sum_{j=1}^n (X_j - \bar{X}_n)(X_j - \bar{X}_n)^T = \sum_{j=1}^n (X_j - \mu)(X_j - \mu)^T - n(\bar{X}_n - \mu)(\bar{X}_n - \mu)^T$$

Finalement, puisque $\text{Var}(\bar{X}_n) = \mathbb{E}[(\bar{X}_n - \mu)(\bar{X}_n - \mu)^T] = \Gamma/n$, on déduit facilement que

$$\mathbb{E}[G_n] = \frac{n-1}{n} \Gamma$$

5. Estimation dans le cas multivarié

5.1 Introduction.

On dispose de n données multidimensionnelles

$(x_{j,1}, x_{j,2}, \dots, x_{j,p})_{j=1, \dots, n}$.

On notera $x_j = (x_{j,1}, \dots, x_{j,p})^T$ pour tout entier j .

Hypothèse: On suppose que ces données sont les réalisations de n vecteurs gaussiens X_1, X_2, \dots, X_n indépendants et de même loi $\mathcal{N}(\mu, \Gamma)$.

Le vecteur $\mu \in \mathbb{R}^p$ est le vecteur d'espérance et Γ est la matrice de variance-covariance, de format $p \times p$.

But: On s'intéresse à l'estimation des paramètres μ et Γ par la méthode du maximum de vraisemblance.

5.2 Densité gaussienne multivariée et Vraisemblance.

On rappelle que si X est un vecteur gaussien d'espérance μ et de variance Γ , alors, sa densité de probabilité f est définie pour tout $x \in \mathbb{R}^p$ par :

$$f(x) = \frac{1}{(2\pi)^{p/2} (\det(\Gamma))^{1/2}} \exp \left(-\frac{1}{2} (x - \mu)^T \Gamma^{-1} (x - \mu) \right)$$

La vraisemblance de l'échantillon s'écrit alors

$$\begin{aligned} & L(\mathbf{x}, \mu, \Gamma) \\ &= \prod_{j=1}^n \left(\frac{1}{(2\pi)^{p/2} (\det(\Gamma))^{1/2}} \exp \left(-\frac{1}{2} (x_j - \mu)^T \Gamma^{-1} (x_j - \mu) \right) \right) \\ &= \frac{1}{(2\pi)^{np/2} (\det(\Gamma))^{n/2}} \exp \left(-\frac{1}{2} \sum_{j=1}^n (x_j - \mu)^T \Gamma^{-1} (x_j - \mu) \right) \end{aligned}$$

5.3 Log-vraisemblance.

La log-vraisemblance s'écrit alors :

$$\begin{aligned} LL(\mathbf{x}, \mu, \Gamma) \\ &= -\frac{np}{2} \log(2\pi) - \frac{n}{2} \log(\det(\Gamma)) - \frac{1}{2} \sum_{j=1}^n (x_j - \mu)^T \Gamma^{-1} (x_j - \mu) \\ &= -\frac{1}{2} \sum_{j=1}^n (\mu^T \Gamma^{-1} \mu - 2\mu^T \Gamma^{-1} x_j) + C(\mathbf{x}, \Gamma) \end{aligned}$$

où la quantité $C(\mathbf{x}, \Gamma)$ ne dépend pas de μ .

5.4 Estimateur du vecteur d'espérance μ .

En utilisant la dérivée vectorielle et en partant de la dernière expression de $LL(\mathbf{x}, \mu, \Gamma)$, on montre facilement que le gradient de LL par rapport à μ s'écrit :

$$\nabla_{\mu} LL(\mathbf{x}, \mu, \Gamma) = \sum_{j=1}^n (-\Gamma^{-1} \mu + \Gamma^{-1} x_j)$$

La résolution de l'équation $\nabla_{\mu} LL(\mathbf{x}, \mu, \Gamma) = 0$ conduit à la solution

$$\mu = \frac{1}{n} \sum_{j=1}^n x_j = \bar{x}_n$$

5.5. Estimation de la matrice Γ .

Il s'agit donc ici de trouver la valeur de Γ qui maximise la quantité

$$-\frac{n}{2} \log(\det(\Gamma)) - \frac{1}{2} \sum_{j=1}^n (x_j - \mu)^T \Gamma^{-1} (x_j - \mu)$$

ce qui revient à minimiser, puisque $\log(\det(\Gamma)) = -\log(\det(\Gamma^{-1}))$,

$$\sum_{j=1}^n (x_j - \mu)^T \Gamma^{-1} (x_j - \mu) - n \log(\det(\Gamma^{-1}))$$

On pose $\Sigma = \Gamma^{-1}$. Nous allons dériver par rapport à Σ . En utilisant le fait que le gradient de $\log(\det(M))$ est égale à M^{-1} pour toute matrice M définie positive et que le gradient de $u^T M v$ est égal à $u v^T$ pour tout vecteur u, v , on déduit que la matrice Σ est solution de l'équation:

$$\sum_{j=1}^n (x_j - \mu)(x_j - \mu)^T - n \Sigma^{-1} = 0 \iff \Sigma^{-1} = \Gamma = \frac{1}{n} \sum_{j=1}^n (x_j - \mu)(x_j - \mu)^T$$

En remplaçant alors μ par son estimation \bar{x}_n , on en déduit l'estimation de Γ .

5.6. Conclusion.

Les estimateurs du maximum de vraisemblance des paramètres μ et Γ sont donc respectivement \bar{X}_n et $\hat{\Gamma}_n$ données par :

$$\bar{X}_n = \frac{1}{n} \sum_{j=1}^n X_j$$

$$\hat{\Gamma}_n = \frac{1}{n} \sum_{j=1}^n (X_j - \bar{X}_n)(X_j - \bar{X}_n)$$