CHAPITRE 6 : VECTEURS GAUSSIENS

I Variables Gaussiennes réelles et loi du χ^2

Une variable aléatoire réelle X suit la loi Gaussienne d'espérance m et de variance σ^2, notée $\mathcal{N}(m, \sigma^2)$, si elle admet la densité

$$f_X : x \longmapsto \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right).$$

Nous utiliserons la convention que $\sigma = 0$ correspond à δ_m la masse de Dirac en m (c'est-à-dire la loi d'une variable aléatoire égale à m presque sûrement).

Proposition 1. *(i) Si $Y \sim \mathcal{N}(0,1)$ alors $m + \sigma Y \sim \mathcal{N}(m, \sigma^2)$,*

$$\forall k \in \mathbb{N} \qquad \mathbb{E}(Y^{2k+1}) = 0 \qquad \textit{et} \qquad \mathbb{E}(Y^{2k}) = 2^{-k}(2k)!/k!$$

(ii) Si $X_1 \sim \mathcal{N}(m_1, \sigma_1^2)$ et $X_2 \sim \mathcal{N}(m_2, \sigma_2^2)$ sont indépendantes, alors $X_1 + X_2 \sim \mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$.

Démonstration :

Rappelons aussi que, pour toute constante $c \in \mathbb{R}$, $cX \sim \mathcal{N}(cm, c^2\sigma^2)$, de telle sorte qu'on obtient :

Corollaire 2. *Toute combinaison linéaire (et même affine) de Gaussiennes indépendantes est une Gaussienne.*

Introduisons maintenant la loi du χ^2 (prononcé *"Khi-2"*) :

Définition 3. *La loi du χ^2 à $k \in \mathbb{N}^*$ degrés de liberté est la loi d'une variable aléatoire Y qui s'écrit $Y = X_1^2 + \cdots + X_k^2$, où $X_1, \ldots, X_k$ sont des variables aléatoires indépendantes et de même loi $\mathcal{N}(0,1)$. Nous la notons $\chi^2(k)$. Sa densité est*

$$f_k : x \longmapsto \frac{1}{2^{k/2}\,\Gamma(k/2)} x^{-1+k/2} e^{-x/2} \mathbb{1}_{x>0}.$$

Proposition 4. *Si Y suit la loi $\chi^2(k)$ alors $\mathbb{E}(Y) = k$ et $\mathbb{V}\mathrm{ar}(Y) = 2k$.*

II Vecteurs Gaussiens

Comme nous allons le constater, les vecteurs Gaussiens constituent la généralisation naturelle en n dimensions des variables Gaussiennes uni-dimensionnelles. La convention prise pour $\sigma = 0$ va s'avérer bien pratique, permettant d'éviter d'avoir à évoquer les cas particuliers. Nous commençons par donner une définition assez formelle, mais nous verrons ensuite des manières plus naturelles de les voir.

Avant cela, rappelons quelques définitions et notations :

⋆ *Si A est une matrice de taille $n \times p$ alors sa transposée ${}^tA$ est la matrice de taille $p \times n$ telle que $({}^tA)_{i,j} = A_{j,i}$ pour tous $i \in \{1, \ldots, p\}$ et $j \in \{1, \ldots, n\}$.* ***Dans ce chapitre, nous noterons en colonne les vecteurs de $\mathbb{R}^n$.***

⋆ *Pour $n \geq 1$, nous notons $\langle \cdot, \cdot \rangle$ le produit scalaire usuel sur $\mathbb{R}^n$ défini par*

$$\forall x = {}^t(x_1, \ldots, x_n) \in \mathbb{R}^n \quad \forall y = {}^t(y_1, \ldots, y_n) \in \mathbb{R}^n \qquad \langle x, y \rangle = {}^t y x = \sum_{i=1}^{n} x_i y_i$$

La norme euclidienne de $x \in \mathbb{R}^n$ est définie par $\|x\| = \sqrt{\langle x, x \rangle}$.

⋆ *Une matrice carrée M de taille n est dite symétrique si ${}^tM = M$. Elle est dite positive si $\langle Mx, x \rangle \geq 0$ pour tout $x \in \mathbb{R}^n$. Elle est dite définie positive si $\langle Mx, x \rangle > 0$ pour tout $x \in \mathbb{R}^n \backslash \{0\}$.*

1) Définitions et propriétés

Définition 5. *Une variable aléatoire $X = {}^t(X_1, \ldots, X_n)$ dans $\mathbb{R}^n$ est un vecteur Gaussien, si pour tout $a = {}^t(a_1, \ldots, a_n) \in \mathbb{R}^n$, la variable aléatoire réelle $\langle a, X \rangle = a_1 X_1 + \ldots a_n X_n$ est Gaussienne.*

Le corollaire précédent montre que si l'on considère des variables aléatoires $X_i \sim \mathcal{N}(m_i, \sigma_i^2)$ indépendantes, alors $X = {}^t(X_1, \ldots, X_n)$ est un vecteur Gaussien.

Remarque :

Définition 6. *Soit $X = {}^t(X_1, \ldots, X_n)$ un vecteur Gaussien. Son espérance est le vecteur*

$$m = \mathbb{E}(X) = {}^t(\mathbb{E}(X_1), \ldots, \mathbb{E}(X_n)) \in \mathbb{R}^n$$

et sa matrice de covariance est $\Gamma = (\mathrm{Cov}(X_i, X_j))_{1 \leq i,j \leq n}$ définie par

$$\forall (i,j) \in \{1, \ldots, n\}^2 \qquad \mathrm{Cov}(X_i, X_j) = \mathbb{E}([X_i - \mathbb{E}(X_i)][X_j - \mathbb{E}(X_j)]) = \mathbb{E}(X_i X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j)$$

Proposition 7. *Pour tout $u \in \mathbb{R}^n$, $\langle \Gamma u, u \rangle = \mathbb{V}\mathrm{ar}(\langle u, X \rangle)$. Par conséquent la matrice Γ est symétrique positive.*

Démonstration :

Le fait que Γ soit symétrique positive permet de lui associer une unique racine carrée, c'est-à-dire une matrice A symétrique positive telle que $A^2 = \Gamma$. C'est un résultat classique qui découle de la diagonalisation des matrices symétriques.

Rappelons que la transformée de Fourier d'un vecteur aléatoire X dans $\mathbb{R}^n$ caractérise la loi de X. Il s'agit de la fonction

$$\phi_X : u \longmapsto \mathbb{E}(e^{i<u,X>}) = \mathbb{E}\left(e^{i(u_1 X_1 + \ldots + u_n X_n)}\right)$$

Proposition 8. *Si X est un vecteur Gaussien d'espérance m et de matrice de covariance Γ, sa transformée de Fourier est*

$$\phi_X : u \longmapsto \exp\left(i\langle u, m \rangle - \frac{1}{2}\langle \Gamma u, u \rangle\right)$$

Démonstration :

Corollaire 9. *Un vecteur Gaussien X est caractérisé par son espérance m et sa matrice de covariance Γ symétrique positive.*

Définition 10. *Nous notons $\mathcal{N}_n(m, \Gamma)$ la loi d'un vecteur Gaussien $X = {}^t(X_1, \ldots, X_n)$ d'espérance m et sa matrice de covariance Γ. Si $m = 0$ et $\Gamma = I_n$, on dit que X est un vecteur Gaussien centré réduit.*

2) Caractérisation de l'indépendance

Si $X_1, \ldots, X_n$ sont des variables aléatoires réelles Gaussiennes indépendantes alors X est Gaussien et

$$\forall i \neq j \qquad \operatorname{Cov}(X_i, X_j) = 0.$$

Ainsi Γ est une matrice diagonale. La réciproque est vraie :

Proposition 11. *Soient $X_1, \ldots, X_n$ des variables aléatoires réelles Gaussiennes. Alors les X_i sont indépendantes si et seulement si le vecteur $X = {}^t(X_1, \ldots, X_n)$ est Gaussien de matrice de covariance diagonale.*

Démonstration : utilise la transformée de Fourier.

3) Existence de vecteurs Gaussiens

Notons $\mathcal{M}_{p,n}(\mathbb{R})$ l'ensemble des matrices de taille $p \times n$ à coefficients réels.

Proposition 12. *Soient $X \sim \mathcal{N}_n(m, \Gamma)$ et $Y = AX + b$, avec $A \in \mathcal{M}_{p,n}(\mathbb{R})$ et $b \in \mathbb{R}^p$. Alors $Y \sim \mathcal{N}_p(Am + b, A\Gamma\, {}^tA)$.*

Démonstration :

Théorème 13. *Si $m \in \mathbb{R}^n$ et Γ est une matrice symétrique positive alors il existe un vecteur Gaussien X d'espérance m et de matrice de covariance Γ.*

Démonstration :

4) Densité d'un vecteur Gaussien

Proposition 14. *La loi $\mathcal{N}_n(m, \Gamma)$ admet une densité si et seulement si Γ est inversible (c'est-à-dire symétrique définie positive). Dans ce cas, sa densité (par rapport à $dx_1 \ldots dx_n$) est*

$$(x_1, \ldots, x_n) \longmapsto \frac{1}{\sqrt{(2\pi)^n \det\Gamma}} \exp\left(-\frac{1}{2}\langle \Gamma^{-1}(x - m), x - m\rangle\right)$$

Démonstration : Admis (utilise le changement de variable en dimension supérieure).

5) Théorème Central Limite Vectoriel

Théorème 15. *Soient $X_1, \dots, X_n$ des vecteurs aléatoires indépendants et de même loi dans $\mathbb{R}^d$ tels que $\mathbb{E}(X_1^2(j)) < +\infty$ pour tout $j \in \{1, \dots, n\}$. Notons m l'espérance de X_1 et Γ sa matrice de covariance. Alors*

$$\sqrt{n}\left(\frac{X_1 + \ldots + X_n}{n} - m\right) \xrightarrow[n \to +\infty]{\mathcal{L}} \mathcal{N}_d(0, \Gamma)$$

Démonstration : voir TD.

III Théorème de Cochran et modèles Gaussiens

1) Rappel : projection orthogonale sur un sous-espace vectoriel

Soit E un $\mathbb{R}$-espace vectoriel de dimension finie n, muni d'un produit scalaire (on parle d'espace Euclidien). On peut par exemple prendre $E = \mathbb{R}^n$ et le produit scalaire défini au début du paragraphe II.

Définition 16. ⋆ *Deux vecteurs x et y de E sont dits orthogonaux si $\langle x, y \rangle = 0$.*
⋆ *Deux parties A et B de E sont dites orthogonales si $\langle x, y \rangle = 0$ pour tout $(x, y) \in A \times B$.*
⋆ *Si A est une partie de E alors son orthogonal est le sous-espace vectoriel*

$$A^\perp = \{x \in E : \forall y \in A \ \langle x, y \rangle = 0\}$$

⋆ *Une base orthonormée (BON) de E est une base $(e_i)_{1 \leq i \leq n}$ telle que $\|e_i\| = 1$ pour tout i, et $\langle e_i, e_j \rangle = 0$ pour $i \neq j$*

Notons qu'une telle base existe toujours dans un espace Euclidien.

Définition 17. *Si il existe $p \in \{2, \dots, n\}$ et $E_1, \dots E_p$ des sous-espaces vectoriels de E deux à deux orthogonaux tels que pour tout $x \in E$ s'écrit de manière unique*

$$x = x_1 + \cdots + x_p \qquad \text{avec } (x_1, \dots, x_p) \in E_1 \times \cdots \times E_p$$

alors on dit que E est somme directe orthogonale de $E_1, \dots E_p$ et on note $E = E_1 \overset{\perp}{\oplus} \cdots \overset{\perp}{\oplus} E_p$. De plus, pour tout $i \in \{1, \dots, p\}$, l'application $\Pi_{E_i} : x \in E \longmapsto x_i$ est appelée projection orthogonale sur E_i.

Notons que, dans ce cas, on peut former une BON de E en réunissant des BON des E_i et

$$\forall i \in \{1, \dots, p\} \qquad E_i^\perp = E_1 \overset{\perp}{\oplus} \cdots E_{i-1} \overset{\perp}{\oplus} E_{i+1} \overset{\perp}{\oplus} \cdots \overset{\perp}{\oplus} E_p$$

Proposition 18. *Si F est un sous-espace vectoriel de E, l'orthogonal de F est un sous-espace vectoriel de dimension $n - dimF$ et $E = F \oplus F^\perp$. Mentionnons aussi que $(F^\perp)^\perp = F$.*

2) Théorème de Cochran

Théorème 19 (Cochran)**.** *Soit $X \sim \mathcal{N}_n(0, \sigma^2 I_n)$ (c'est-à-dire $X_1, \dots, X_n$ est un n-échantillon de loi $\mathcal{N}(0, \sigma^2)$). Supposons que $\mathbb{R}^n$ est muni de sa base canonique et que*

$$\mathbb{R}^n = E_1 \overset{\perp}{\oplus} \cdots \overset{\perp}{\oplus} E_p$$

Alors les projections $\Pi_{E_1} X, \dots, \Pi_{E_p} X$ de X sont des vecteurs Gaussiens indépendants de lois respectives $\mathcal{N}_n(0, \sigma^2 \Pi_{E_1}), \dots, \mathcal{N}_n(0, \sigma^2 \Pi_{E_p})$. En particulier

$$\forall i \in \{1, \dots, p\} \qquad \frac{1}{\sigma^2} \|\Pi_{E_i} X\|^2 \sim \chi^2(\text{dim} E_i)$$

Démonstration :

3) Intervalles de confiance et tests pour les paramètres d'une loi $\mathcal{N}(m,\sigma^2)$

Introduisons la loi de Student (dont nous avons à disposition la table de loi) :

Définition 20. *La loi de Student à n degrés de liberté est la loi d'une variable aléatoire réelle qui s'écrit $\sqrt{n}X/\sqrt{Y}$ avec $X \sim \mathcal{N}(0,1)$ et $Y \sim \chi^2(n)$ indépendants. On la note $\mathcal{T}(n)$. Sa densité est*

$$x \in \mathbb{R} \longmapsto c_n \left(1 + \frac{x^2}{n}\right)^{-(n+1)/2}$$

où c_n est une constante de renormalisation.

Les vecteurs Gaussiens sont souvent utilisés dans les modèles statistiques multi-dimensionnels car ils s'avèrent assez faciles à manipuler, notamment grâce au théorème de Cochran.

Considérons par exemple un n-échantillon $X_1, \dots, X_n$ de loi $\mathcal{N}(m,\sigma^2)$. L'estimateur du maximum de vraisemblance de (m,σ^2) est $(\overline{X}_n, \overline{V}_n)$ avec

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i \qquad \text{et} \qquad \overline{V}_n = \frac{1}{n}\sum_{i=1}^n (X_i - \overline{X}_n)^2$$

Considérons plutôt l'estimateur sans biais

$$\widehat{\sigma}_n^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2$$

Rappelons que le lemme de Slutsky entraîne que

$$\sqrt{n}\,\frac{\overline{X}_n - m}{\sqrt{\widehat{\sigma}_n^2}} \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1)$$

Cela nous a permis de construire des intervalles de confiance asymptotiques. La proposition suivante va nous permettre de construire, dans le cas Gaussien, des intervalles de confiance (non asymptotiques) et des tests à partir de la loi de Student.

Proposition 21. *Dans le cas Gaussien que nous venons de décrire, $\overline{X}_n$ et $\widehat{\sigma}_n^2$ sont indépendantes et*

$$\overline{X}_n \sim \mathcal{N}\Big(m, \frac{\sigma^2}{n}\Big) \qquad \textit{et} \qquad \frac{n-1}{\sigma^2}\widehat{\sigma}_n^2 \sim \chi^2(n-1)$$

Démonstration :