

3. Vecteurs Gaussiens

3.1. Généralités :

Définition 3.1-1 On dit qu'un vecteur aléatoire $X \in \mathbb{R}^d$ est gaussien si pour tout $a \in \mathbb{R}^d$, la variable aléatoire $\langle X, a \rangle$ est gaussienne.

Théorème 3.1-1 Soit vecteur gaussien X d'espérance m_X et de matrice de covariance C_X alors L'image Y de X par une application affine $x \mapsto Ax + b$ est un vecteur gaussien d'espérance $m_Y = Am_X + b$ et de matrice de covariance $C_Y = AC_X A^T$.

Preuve : On a $Y = AX + b$. Soit $a \in \mathbb{R}^d$. Donc $\langle Y, a \rangle = \langle AX, a \rangle + \langle b, a \rangle = \langle X, A^T a \rangle + \langle b, a \rangle$. Puisque X est un vecteur gaussien, $\langle X, A^T a \rangle$ est une variable aléatoire gaussienne. Le deuxième terme est constant donc la somme de ces deux terme reste une variable aléatoire gaussienne. Ainsi, pour tout a , $\langle Y, a \rangle$ est une variable aléatoire gaussienne, donc Y est un vecteur gaussien.

Espérance :

Calculons l'espérance de Y :

$$\mathbb{E}[Y] = \mathbb{E}[AX + b] = A\mathbb{E}[X] + b = Am_X + b.$$

Matrice de covariance :

Calculons la matrice de covariance de Y :

$$\text{Cov}(Y) = \mathbb{E}[(Y - \mathbb{E}[Y])(Y - \mathbb{E}[Y])^T].$$

Comme $Y = AX + b$, nous avons :

$$Y - \mathbb{E}[Y] = AX + b - (Am_X + b) = A(X - m_X).$$

Par conséquent,

$$\text{Cov}(Y) = \mathbb{E}[A(X - m_X)(X - m_X)^\top A^\top].$$

En utilisant la linéarité de l'espérance et la définition de la covariance de X , nous obtenons :

$$\text{Cov}(Y) = A\mathbb{E}[(X - m_X)(X - m_X)^\top]A^\top = A\text{Cov}(X)A^\top = AC_XA^\top.$$

Ainsi, Y est un vecteur gaussien d'espérance $m_Y = Am_X + b$ et de matrice de covariance $C_Y = AC_XA^\top$.

Corollaire 3.1-1 Si $X = (X_1, \dots, X_d)$ est un vecteur gaussien, alors pour tout $I \subset \{1, \dots, d\}$, le vecteur $(X_i)_{i \in I}$ est gaussien.

Preuve : $(X_i)_{i \in I}$ est l'image de $X = (X_1, \dots, X_d)$ par l'application linéaire $\mathbb{R}^d \rightarrow \mathbb{R}^I$. $(x_i)_{i \in I} \mapsto (x_i)_{i \in I}$.

Théorème 3.1-2 Soient X_1, \dots, X_d des variables aléatoires gaussiennes indépendantes. Alors $X = (X_1, \dots, X_d)$ est un vecteur gaussien.

Preuve : Soit $a \in \mathbb{R}^d$. Pour $k \in \{1, \dots, d\}$, on pose

$$S_k = \sum_{i=1}^k a_i X_i.$$

On montre par récurrence sur k que S_k est une variable aléatoire gaussienne.

Initialisation : Pour $k = 1$,

$$S_1 = a_1 X_1.$$

Quand on multiplie une variable gaussienne par une constante, on obtient toujours une variable aléatoire gaussienne.

Hérédité : Supposons S_k gaussienne. On a

$$S_{k+1} = S_k + a_{k+1}X_{k+1}.$$

Or, $a_{k+1}X_{k+1}$ est une variable gaussienne indépendante de S_k , car S_k est $\mathcal{F}(X_1, \dots, X_k)$ -mesurable. La somme de deux variables aléatoires gaussiennes indépendantes étant une variable aléatoire gaussienne, il s'ensuit que S_{k+1} est une variable aléatoire gaussienne.

Comme $\langle X, a \rangle = S_d$ quel que soit a , on en déduit que X est un vecteur gaussien.

Théorème 3.1-3 *Soit C une matrice symétrique positive de \mathbb{R}^d et $m \in \mathbb{R}^d$. Alors, on peut construire un vecteur gaussien de moyenne m et de covariance C comme une application affine d'un vecteur gaussien de moyenne nulle et de covariance identité.*

Preuve : C une matrice symétrique positive. Soit O une matrice orthogonale et $\lambda_1, \dots, \lambda_d$ les valeurs propres de C , telles que $C = O\Lambda O^T$, où Λ est la matrice diagonale des valeurs propres. On pose alors $A = O\Lambda^{1/2}$ et $X \sim \mathcal{N}(0, I_d)$. Le vecteur $Y = AX + m$ est alors gaussien de moyenne m et de covariance

$$ACA^T = O\Lambda^{1/2}(O\Lambda^{1/2})^T = O\Lambda O^T = C$$

.

Lemme 3.1-1 *Soient X et Y deux vecteurs aléatoires sur \mathbb{R}^d tels que pour tout $a \in \mathbb{R}^d$, $\langle X, a \rangle$ et $\langle Y, a \rangle$ ont même loi. Alors X et Y ont même loi.*

Preuve : On va montrer que X et Y ont même fonction caractéristique, ce qui assurera qu'ils ont même loi. Soit $a \in \mathbb{R}^d$. On pose $Z = \langle X, a \rangle$ et $T = \langle Y, a \rangle$. Comme Z et T ont même loi, on a $\mathbb{E}(e^{iZ}) = \mathbb{E}(e^{iT})$. Mais on a aussi par définition de Z et T :

$$\mathbb{E}(e^{iZ}) = \mathbb{E}(e^{i\langle X, a \rangle}) = \varphi_X(a) \quad \text{et} \quad \mathbb{E}(e^{iT}) = \mathbb{E}(e^{i\langle Y, a \rangle}) = \varphi_Y(a),$$

donc $\varphi_X(a) = \varphi_Y(a)$. Ainsi $\varphi_X = \varphi_Y$, donc X et Y ont même loi.

Théorème 3.1-4 Soient X et Y deux vecteurs gaussiens ayant même espérance et même matrice de covariance. Alors X et Y ont même loi.

Preuve : Soit $a \in \mathbb{R}^d$. On pose $V = \langle X, a \rangle$ et $W = \langle Y, a \rangle$. L'espérance de V est $\langle m_X, a \rangle$ et la covariance de V est $\text{Cov}(\langle X, a \rangle, \langle X, a \rangle) = \langle C_X a, a \rangle$. Comme X est gaussien, V est gaussien, donc $V \sim \mathcal{N}(\langle m_X, a \rangle, \langle C_X a, a \rangle)$. De même, $W \sim \mathcal{N}(\langle m_Y, a \rangle, \langle C_Y a, a \rangle)$. Comme $m_X = m_Y$ et $C_X = C_Y$, on en déduit que $V \sim W$. Comme c'est vrai quel que soit a , on en déduit du lemme précédent que X et Y ont même loi.

3.2. Indépendance de vecteurs gaussiens

Théorème 3.2-1 Soient d_1, \dots, d_n des entiers positifs dont la somme est égale à d . Soient C_1, \dots, C_n des matrices symétriques positives, et m_1, \dots, m_n des vecteurs de tailles respectives d_1, \dots, d_n . Alors, on a :

$$\mathcal{N}(m_1, C_1) \otimes \mathcal{N}(m_2, C_2) \otimes \dots \otimes \mathcal{N}(m_n, C_n) = \mathcal{N}(m, C),$$

où

$$m = \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix} \quad \text{et} \quad C = \begin{pmatrix} C_1 & 0 & \dots & 0 \\ 0 & C_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & C_n \end{pmatrix}.$$

Preuve : Soient X_1, \dots, X_n des vecteurs aléatoires gaussiens indépendants tels que $X_k \sim \mathcal{N}(m_k, C_k)$ pour tout k entre 1 et n . On pose $X = (X_1, \dots, X_n)$. Soit $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ avec a_k de taille d_k . On a donc :

$$\langle X, a \rangle = \sum_{k=1}^n \langle X_k, a_k \rangle.$$

Puisque X_k est gaussien, donc $\langle X_k, a_k \rangle$ est également gaussienne. Puisque les X_k sont indépendants, les variables aléatoires $\langle X_k, a_k \rangle$ le sont aussi. Or, la somme de variables aléatoires gaussiennes indépendantes est une variable gaussienne, donc $\langle X, a \rangle$ est une

variable gaussienne. Don on déduit que X est un vecteur gaussien. Calculons $\mathbb{E}(X)$:

$$\mathbb{E}(X) = \mathbb{E} \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} \mathbb{E}(X_1) \\ \vdots \\ \mathbb{E}(X_n) \end{pmatrix} = \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix} = m.$$

Ensuite, calculons $\text{Cov}(X)$:

$$\text{Cov}(X) = \text{Cov} \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} \text{Cov}(X_1, X_1) & \cdots & \text{Cov}(X_1, X_n) \\ \vdots & \ddots & \vdots \\ \text{Cov}(X_n, X_1) & \cdots & \text{Cov}(X_n, X_n) \end{pmatrix}.$$

Puisque X_1, \dots, X_n sont indépendants, $\text{Cov}(X_i, X_j) = 0$ pour $i \neq j$. Donc,

$$\text{Cov}(X) = \begin{pmatrix} \text{Cov}(X_1) & 0 & \cdots & 0 \\ 0 & \text{Cov}(X_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \text{Cov}(X_n) \end{pmatrix} = \begin{pmatrix} C_1 & 0 & \cdots & 0 \\ 0 & C_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & C_n \end{pmatrix} = C.$$

Ainsi, nous avons montré que $\mathbb{E}(X) = m$ et $\text{Cov}(X) = C$.

Théorème 3.2-2 Soient d_1, \dots, d_n des entiers positifs dont la somme est égale à d . On suppose que $X = (X_1, \dots, X_n)$ est un vecteur gaussien dont la matrice de covariance est diagonale par blocs :

$$C_X = \begin{pmatrix} C_1 & 0 & \cdots & 0 \\ 0 & C_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & C_n \end{pmatrix}.$$

Soient $Z_1 = (X_1, \dots, X_{d_1})$, $Z_2 = (X_{d_1+1}, \dots, X_{d_1+d_2})$, et $Z_n = (X_{d_1+\dots+d_{n-1}+1}, \dots, X_d)$, alors les vecteurs Z_1, \dots, Z_n sont des vecteurs gaussiens indépendants.

Preuve : X a la même espérance et la même matrice de covariance que le vecteur aléatoire considéré dans le théorème précédent. Puisque X et ce vecteur aléatoire

sont tous deux gaussiens, ils ont la même loi. Par conséquent, la loi de X est

$$\mathcal{N}(m_1, C_1) \otimes \mathcal{N}(m_2, C_2) \otimes \cdots \otimes \mathcal{N}(m_n, C_n),$$

Et par conséquent les Z_i sont des vecteurs gaussiens indépendants.

Corollaire 3.2-1 *Si le vecteur gaussien $X = (X_1, \dots, X_d)$ a une matrice de covariance dont tous les termes non-diagonaux sont nuls, alors X_1, \dots, X_d sont des variables aléatoires indépendantes.*

Lemme 3.2-1 *Soit M une matrice inversible et $b \in \mathbb{R}^d$. Définissons $T(x) = Mx + b$. Soit μ_1 une mesure positive sur \mathbb{R}^d ayant pour densité f_1 par rapport à la mesure de Lebesgue sur \mathbb{R}^d . Alors, l'image de μ_1 par T a pour densité par rapport à la mesure de Lebesgue sur \mathbb{R}^d la fonction f_2 définie par*

$$f_2(y) = \frac{1}{|\det M|} f_1(T^{-1}(y)).$$

Preuve : Soit g une fonction mesurable positive sur \mathbb{R}^d . Désignons par μ_2 la mesure image de μ_1 par T . Selon le théorème de transfert,

$$\int_{\mathbb{R}^d} g d\mu_2 = \int_{\mathbb{R}^d} (g \circ T) d\mu_1 = \int_{\mathbb{R}^d} (g \circ T) f_1 d\lambda^d.$$

En utilisant le changement de variable $y = T(x)$, nous obtenons

$$\int_{\mathbb{R}^d} g d\mu_2 = \int_{\mathbb{R}^d} g(y) (f_1 \circ T^{-1}(y)) |\det(T^{-1})'(y)| d\lambda^d(y).$$

Or, $(T^{-1})'(y) = M^{-1}$, donc $|\det(T^{-1})'(y)| = \frac{1}{|\det M|}$. Ainsi,

$$\int_{\mathbb{R}^d} g d\mu_2 = \int_{\mathbb{R}^d} g(y) (f_1 \circ T^{-1}(y)) \frac{1}{|\det M|} d\lambda^d(y),$$

ce qui prouve le résultat souhaité.

Corollaire 3.2-2 *Soient $A \in GL_d(\mathbb{R})$ et $b \in \mathbb{R}^d$. Supposons que le vecteur aléatoire X a pour densité f par rapport à la mesure de Lebesgue sur \mathbb{R}^d . Alors,*

le vecteur aléatoire $Y = AX + b$ a pour densité

$$g(y) = \frac{1}{|\det A|} f(A^{-1}(y - b)).$$

Théorème 3.2-3 Soit C une matrice symétrique définie positive et $m \in \mathbb{R}^d$. Si Y est un vecteur aléatoire suivant la loi $\mathcal{N}(m, C)$, alors sa densité par rapport à la mesure de Lebesgue est donnée par la fonction

$$f_Y(y) = \frac{1}{(2\pi)^{d/2} \sqrt{\det C}} \exp \left(-\frac{1}{2} \langle C^{-1}(y - m), y - m \rangle \right).$$

Preuve : Soit $X \sim \mathcal{N}(0, I_d)$ tel que $Y = AX + m$. Puisque C est définie positive, les valeurs propres λ_i sont strictement positives, donc A est inversible. Puisque X est constitué de n variables aléatoires indépendantes et de densité, la densité de X est le produit des densités, soit

$$f_X(x) = \prod_{k=1}^n \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{x_k^2}{2} \right) = \frac{1}{(2\pi)^{d/2}} \exp \left(-\frac{1}{2} \|x\|^2 \right).$$

D'après le corollaire précédent, $Y = AX + m$ a pour densité

$$\frac{1}{\det A} f_X(A^{-1}(y - m)) = \frac{1}{\det A (2\pi)^{d/2}} \exp \left(-\frac{1}{2} \|A^{-1}(y - m)\|^2 \right).$$

Cependant, nous avons

$$\|A^{-1}(y - m)\|^2 = \langle (A^{-1})^T(y - m), y - m \rangle = \langle (A^{-1})(y - m), (A^{-1})(y - m) \rangle = \langle C^{-1}(y - m), y - m \rangle.$$

De plus,

$$\det C = \det(AA^T) = (\det A)^2, \quad \text{donc} \quad \det A = \sqrt{\det C}.$$

Il en découle que la densité de Y (c'est-à-dire la densité de la loi de Y) est

$$f_{m,C}(y) = \frac{1}{(2\pi)^{d/2} \sqrt{\det C}} \exp \left(-\frac{1}{2} \langle C^{-1}(y - m), y - m \rangle \right).$$

3.3. Caractérisation des vecteurs gaussiens

Théorème 3.3-1 *En dimension quelconque, la fonction caractéristique de la loi normale $\mathcal{N}(m, C)$ est*

$$t \mapsto \exp(i\langle t, m \rangle) \exp\left(-\frac{1}{2}\langle Ct, t \rangle\right).$$

Preuve : On a

$$\mathbb{E}(\exp(i\langle X, t \rangle)) = \mathbb{E}(\exp(i\langle Y, t \rangle)) = \varphi_Y(1),$$

où $Y = \langle X, t \rangle$. Comme X est gaussien de covariance C et d'espérance m , Y est gaussien de variance $\langle Ct, t \rangle$ et d'espérance $\langle t, m \rangle$. Donc

$$\mathbb{E}(\exp(i\langle X, t \rangle)) = \exp(im_Y) \exp\left(-\frac{1}{2}\sigma_Y^2\right) = \exp(i\langle t, m \rangle) \exp\left(-\frac{1}{2}\langle Ct, t \rangle\right).$$

Théorème 3.3-2 Théorème central limite en dimension d Soit $(X_n)_{n \geq 1}$ une suite de vecteurs aléatoires à valeurs dans \mathbb{R}^d , indépendants et identiquement distribués. On suppose que $\mathbb{E}[\|X_1\|^2] < +\infty$. On note m l'espérance et C la matrice de covariance. Alors

$$\frac{(X_1 + \cdots + X_n) - nm}{\sqrt{n}} \xrightarrow{L} \mathcal{N}(0, C).$$

Lemme 3.3-1 Soit $(X_n)_{n \geq 1}$ une suite de vecteurs aléatoires à valeurs dans \mathbb{R}^d et X un vecteur aléatoire à valeurs dans \mathbb{R}^d . Si, pour tout $a \in \mathbb{R}^d$, $\langle X_n, a \rangle$ converge en loi vers $\langle X, a \rangle$, alors X_n converge en loi vers X .

Preuve : Soit $a \in \mathbb{R}^d$. On pose $Y_n = \langle X_n, a \rangle$. Nous avons

$$\varphi_{X_n}(a) = \mathbb{E}[\exp(i\langle X_n, a \rangle)] = \mathbb{E}[\exp(iY_n)] = \varphi_{Y_n}(1).$$

Par hypothèse, Y_n converge en loi vers $\langle X, a \rangle$, donc $\varphi_{Y_n}(1)$ converge vers $\varphi_Y(1) = \varphi_X(a)$. Donc $\varphi_{X_n}(a)$ converge vers $\varphi_X(a)$. Comme c'est vrai pour tout $a \in \mathbb{R}^d$, le théorème de Lévy implique que $(X_n)_{n \geq 1}$ converge en loi vers X .

Preuve du théorème : Soit X un vecteur aléatoire suivant la loi $\mathcal{N}(0, C)$. On

pose

$$S_n = (X_1 + \cdots + X_n) - nm = \sum_{k=1}^n (X_k - m).$$

D'après le lemme précédent, il suffit de montrer que pour tout $a \in \mathbb{R}^d$, $\frac{\langle S_n, a \rangle}{\sqrt{n}}$ converge en loi vers $\langle X, a \rangle$. Fixons $a \in \mathbb{R}^d$ et posons $Y_n = \langle X_n - m, a \rangle = a^T(X_n - m)$. Les Y_n sont des variables aléatoires indépendantes ayant toutes la même loi, la loi image de P_{X_1} par $x \mapsto \langle x - m, a \rangle = a^T(x - m)$. Leur espérance est 0 et leur variance est $a^T C a = \langle C a, a \rangle$. Par conséquent, la suite $\frac{1}{\sqrt{n}} \sum_{k=1}^n Y_k$ converge en loi vers $\mathcal{N}(0, \langle C a, a \rangle)$. Mais nous avons

$$\frac{1}{\sqrt{n}} \sum_{k=1}^n Y_k = \frac{\langle S_n, a \rangle}{\sqrt{n}}.$$

et, d'après le théorème 6.35, la loi de $\langle X, a \rangle$ est précisément la loi $\mathcal{N}(0, \langle C a, a \rangle)$. Donc $\frac{\langle S_n, a \rangle}{\sqrt{n}}$ converge en loi vers $\langle X, a \rangle$, ce qui conclut la démonstration. \square

Définition 3.3-1 Un n -échantillon X_1, \dots, X_n de loi μ est un ensemble de n vecteurs aléatoires indépendants et identiquement distribués, suivant la loi μ .

Proposition 3.3-1 Soit X un vecteur gaussien dans \mathbb{R}^n de loi $\mathcal{N}(0, I_n)$. Alors la loi de $\|X\|^2$ est une loi à densité. On l'appelle loi du chi-deux à n degrés de liberté et on la note $\chi^2(n)$.

Proposition 3.3-2 On a $\chi^2(n) = \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$.

Théorème 3.3-3 (Théorème de Cochran) Soit $X = (X_1, \dots, X_n)$ un vecteur gaussien centré réduit. Pour F un sous-espace vectoriel de \mathbb{R}^n de dimension p , on note p_F (respectivement p_{F^\perp}) la projection orthogonale sur F (respectivement F^\perp). Alors les vecteurs aléatoires $p_F X$ et $p_{F^\perp} X$ sont gaussiens indépendants, $p_F X \sim \mathcal{N}(0, p_F)$ et $p_{F^\perp} X \sim \mathcal{N}(0, p_{F^\perp})$. De plus, les variables aléatoires $\|p_F X\|^2 = \sum_{i=1}^p Y_i^2$ et $\|p_{F^\perp} X\|^2 = \sum_{i=p+1}^n Y_i^2$ sont indépendantes, $\|p_F X\|^2$ suit la loi $\chi^2(p)$ et $\|p_{F^\perp} X\|^2$ suit la loi $\chi^2(n - p)$.

Preuve : Soient f_1, \dots, f_p une base orthonormée de F et f_{p+1}, \dots, f_n une base orthonormée de F^\perp . Pour tout i entre 1 et n , posons $Y_i = \langle X, f_i \rangle$. Le vecteur Y est

gaussien, en tant qu'image d'un vecteur gaussien par une transformation linéaire. On peut noter que Y_1, \dots, Y_p sont les coordonnées de $p_F X$ dans la base f_1, \dots, f_p , tandis que Y_{p+1}, \dots, Y_n sont les coordonnées de $p_{F^\perp} X$ dans la base f_{p+1}, \dots, f_n . L'application de \mathbb{R}^n dans lui-même qui à x associe $(\langle x, f_1 \rangle, \dots, \langle x, f_n \rangle)$ est une isométrie. Sa matrice O est orthogonale et la matrice de covariance de Y est $O \text{Id}_n O^T = \text{Id}_n$. Y_1, \dots, Y_n sont des variables indépendantes suivant la loi $\mathcal{N}(0, 1)$. Si l'on note encore p_F la matrice de p_F dans la base canonique, la matrice de covariance de $p_F X$ est $p_F \text{Id}_n p_F^T = p_F$ car p_F est un projecteur orthogonal. Ainsi $p_F X \sim \mathcal{N}(0, p_F)$. De même $p_{F^\perp} X \sim \mathcal{N}(0, p_{F^\perp})$. L'indépendance de $p_F X$ et de $p_{F^\perp} X$ découle des identités

$$p_F X = \sum_{i=1}^p Y_i f_i \quad \text{et} \quad p_{F^\perp} X = \sum_{i=p+1}^n Y_i f_i.$$

Comme les f_i sont orthonormés, on a aussi

$$\|p_F X\|^2 = \sum_{i=1}^p Y_i^2 \quad \text{et} \quad \|p_{F^\perp} X\|^2 = \sum_{i=p+1}^n Y_i^2,$$

ce qui donne les lois du chi-deux annoncées.