Probabilités
Chapitre 4 : Vecteurs gaussiens

Lucie Le Briquer

23 novembre 2017

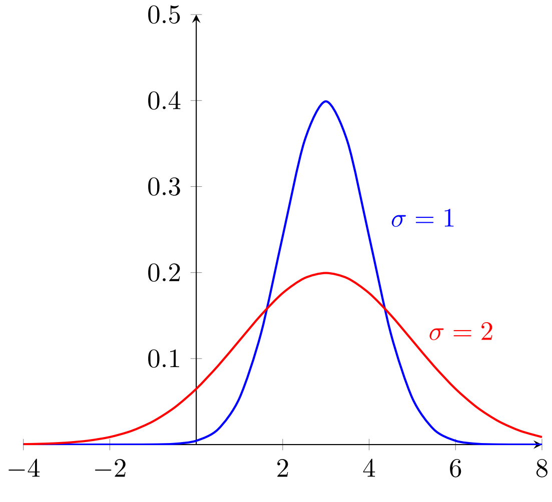
1 Lois gaussiennes réelles

$d = 1$, loi de probabilité sur $\mathbb{R}$.

- $\mathcal{N}(0,1)$: loi de $\frac{e^{\frac{-x^2}{2}}}{\sqrt{2\pi}}$
- $\mathcal{N}(m,\sigma^2)$: loi de $m + \sigma Z$ où $Z \sim \mathcal{N}(0,1)$

(par convention $\mathcal{N}(m,0) = \delta_m$)
On a l'équivalence :

$$\begin{aligned} X \sim \mathcal{N}(m,\sigma^2) &\Leftrightarrow X \underset{\text{loi}}{=} m + \sigma Z \text{ avec } Z \sim \mathcal{N}(0,1) \\ &\Leftrightarrow \phi_X(\lambda) = e^{i\lambda m} e^{-\frac{\sigma^2\lambda^2}{2}} \\ &\Leftrightarrow X \sim \frac{e^{-(x-m)^2/2\sigma^2}}{\sigma\sqrt{2\pi}} dx \qquad \text{si } \sigma \neq 0 \end{aligned}$$



On retrouve des courbes centrées sur m d'étalement σ. On a $\mathbb{E}[X] = m$ et $\text{Var}(X) = \sigma^2$.

2 Gaussiennes en dimension $d \geqslant 2$

Définition 1 (vecteur gaussien)

Soit $X = (X_1, ..., X_d)$ v.a. dans $\mathbb{R}^d$. On dit que c'est un vecteur gaussien si toute combinaison linéaire (ou affine) des X_i a une loi gaussienne sur $\mathbb{R}$.
C'est-à-dire $\forall \lambda \in \mathbb{R}^d, \langle \lambda | X \rangle = \sum_{i=1}^d \lambda_i X_i \sim \mathcal{N}(m_\lambda, \sigma_\lambda^2)$ pour des certains m_λ et σ_λ.

Remarque. C'est équivalent de demander linéaire ou affine.

Définition 2

$\mathcal{N}_{\mathbb{R}^d}(0_{\mathbb{R}^d}, I_d)$ la loi de $Z = (Z_1, ..., Z_d)$ où $Z_1, ..., Z_d$ v.a. indépendantes de loi $\mathcal{N}(0,1)$. On a :

$$Z \sim \frac{e^{-\frac{x_1^2}{2}}}{\sqrt{2\pi}} ... \frac{e^{-\frac{x_d^2}{2}}}{\sqrt{2\pi}} dx_1 ... dx_d = \frac{e^{-\|x\|_2^2}}{\sqrt{2\pi}^n} \underbrace{dx}_{\text{Lebesgue sur } \mathbb{R}^d}$$

Propriété 1

Soit $X = (X_1, ..., X_d)$ à valeurs dans $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, $m \in \mathbb{R}^d$, $A \in M_{d\times d}(\mathbb{R})$, $\Sigma = AA^T$. On a l'équivalence :

$$\begin{aligned}
& X \underset{\text{loi}}{=} m + AZ \quad \text{avec } Z \sim \mathcal{N}(0,1) \\
\Leftrightarrow\ & \phi_X(\underset{\in \mathbb{R}^d}{\lambda}) = e^{i\langle \lambda | m \rangle} e^{-(\lambda^T)\Sigma \frac{\lambda}{2}} \\
\Leftrightarrow\ & X \text{ vecteur gaussien et } \mathbb{E}[X] = m \text{ et } \Gamma_X = (\mathrm{Cov}(X_i, X_j))_{1 \leqslant i,j \leqslant d} = \Sigma \\
\Leftrightarrow\ & X \sim \frac{e^{-((x.m)^T)\Sigma^{-1} \frac{x.m}{2}}}{(2\pi)^{\frac{d}{2}} \sqrt{\det \Sigma}} dx \qquad (\text{si } \Sigma \text{ inversible})
\end{aligned}$$

Définition 3 (vecteur gaussien)

Si X vérifie une de ces propriétés, on dit que X est un vecteur gaussien de loi $\mathcal{N}(m, \Sigma)$. Où m est son espérance et Σ sa matrice de covariance.

Remarque. Σ joue le rôle de σ^2.

Preuve. (de la propriété)

1. Première équivalence :

   En général. $\phi_{m+AY}(\lambda) = e^{i\langle \lambda | m \rangle} \phi(A^T \lambda)$
   Raison.
   $$\mathbb{E}\left[e^{i<\lambda|m+AY>}\right] = e^{i\langle \lambda | m \rangle} \mathbb{E}[e^{i\lambda AY}] = e^{i<\lambda|m>} \underbrace{\mathbb{E}[e^{i(A^T\lambda)^T Y}]}_{\phi_Y(A^T\lambda)}$$

Si $Z = (Z_1, ..., Z_d) \sim \mathcal{N}_d(0, I_d)$,

$$\begin{aligned}\phi_Z(\lambda) &\underset{\text{indép}}{=} \phi_{Z_1}(\lambda_1)...\phi_{Z_d}(\lambda_d)\\ &\underset{Z_i\sim\mathcal{N}(0,1)}{=} e^{-\frac{\lambda_1^2}{2}}...e^{-\frac{\lambda_d^2}{2}}\\ &= e^{-\frac{\lambda^T\lambda}{2}}\end{aligned}$$

Donc,

$$\phi_{m+AZ}(\lambda) = e^{i\langle\lambda|m\rangle}e^{-(A^T\lambda)^T A^T\lambda/2} = e^{i\langle\lambda|m\rangle}e^{-\lambda^T\Sigma\frac{\lambda}{2}}$$

D'où :

$$\phi_X(\lambda) = e^{i\langle\lambda|m\rangle}e^{-\lambda^T\Sigma\frac{\lambda}{2}} \Leftrightarrow \phi_X = \phi_{m+AZ} \underset{\phi\text{ injectif}}{\Leftrightarrow} X \underset{\text{loi}}{=} m + AZ$$

2. ? ? : Il suffit de vérifier que $m+AZ$ a une loi de densité qui suit cette formule si A inversible.

$$\mathbb{E}[\phi(m+AZ)] \underset{\text{transfert}}{=} \int_{\mathbb{R}^d}\phi\left(m + A\begin{pmatrix}x_1\\ \vdots\\ x_d\end{pmatrix}\right)\underbrace{d\mu_Z(x_1...x_d)}_{\frac{e^{-\frac{\|x\|_2^2}{2}}}{(2\pi)^{\frac{d}{2}}}dx_1...dx_d}$$

Effectuons le changement de variable affine (donc $\mathcal{C}_1$, inversible car A inversible) :

$$\begin{pmatrix}y_1\\ \vdots\\ y_d\end{pmatrix} = m + A\begin{pmatrix}x_1\\ \vdots\\ x_d\end{pmatrix}$$

On obtient :

$$\mathbb{E}[\phi(m+AZ)] = \int_{\mathbb{R}^d}\varphi(y)e^{-(y-m)^T\Sigma^{-1}\frac{y-m}{2}}\frac{dy}{(2\pi)^{\frac{d}{2}}}\underbrace{\det A^{-1}}_{\frac{1}{\det A}=\frac{1}{\sqrt{\det\Sigma}}}$$

$x = A^{-1}(y-m)$, $\|x\|_2^2 = x^Tx = (y-m)^T(A^{-1})^TA^{-1}(y-m) = (y-m)\Sigma^{-1}(y-m)$. Ok.

3. $(i \Rightarrow iii)$

Fait général.

- $\mathbb{E}[m+AY] = m + A\mathbb{E}[Y]$
- $\Gamma_{m+AY} = A\Gamma_Y(A^T)$

En effet,

- $(\mathbb{E}[m+AY])_i = \mathbb{E}[m_i + \sum_{j=1}^d A_{ij}Y_j] = m_i + \sum_{j=1}^d A_{ij}\mathbb{E}[Y_j] = (m + A\mathbb{E}[Y])_i$
-

$$\begin{aligned}(\Gamma_{m+AY})_{ij} &= \text{Cov}((m+AY)_i, (m+AY)_)\\ &= \sum_{k=1}^d\sum_{l=1}^d A_{ik}A_{jl}\text{Cov}(Y_k, Y_l)\\ &= \sum_{k=1}^d\sum_{l=1}^d A_{ik}(\Gamma_Y)_{kl}(A^T)_{lj}\\ &= (A\Gamma_YA^T)_{ij}\end{aligned}$$

Si $Z \sim \mathcal{N}_{\mathbb{R}^d}(0, I_d)$, $\mathbb{E}[Z] = 0$, $\Gamma_Z = I_d$. Si $X \underset{\text{loi}}{=} m + AZ$, alors $\forall \lambda \in \mathbb{R}^d$, $\langle \lambda | X \rangle \underset{\text{loi}}{=} \langle \lambda | m + AZ \rangle$ qui est une combinaison affine de Z_i, gaussienne car c'est une famille stable par application affine et convolution. Donc X vecteur gaussien et $\mathbb{E}[X] = m + A\mathbb{E}[Z] = m$, $\Gamma_X = A\Gamma_Z(A^T) = AA^T = \Sigma$.

4. *iii* $\Rightarrow$ *ii* : On veut la fonction caractéristique de X. Soit $\lambda \in \mathbb{R}^d$, on sait que $\langle \lambda | X \rangle \sim \mathcal{N}(m_\lambda, \sigma_\lambda^2)$ et :

$$m_\lambda = \mathbb{E}[\langle \lambda | X \rangle] = \mathbb{E}[\sum_{i=1}^d \lambda_i X_i] = \sum_{i=1}^d \lambda_i m_i = \langle \lambda | m \rangle$$

$$\sigma_2^2 = \mathrm{Var}[\sum_{i=1}^d \lambda_i X_i] = \sum_{i=1}^d \sum_{j=1}^d \lambda_i \lambda_j \underbrace{\mathrm{Cov}(X_i, X_j)}_{\Sigma_{ij}} = \lambda^T \Sigma \lambda$$

$(X \sim \mathcal{N}(m, \mathcal{E}) \Rightarrow \langle \lambda | X \rangle \sim \mathcal{N}(\langle m | \lambda \rangle, \lambda^T \Sigma \lambda))$.

Alors :

$$\phi_{X(\in \mathbb{R}^d)}(\lambda) = \mathbb{E}[e^{i\langle \lambda | X \rangle}] = \phi_{\langle \lambda | X \rangle (\in \mathbb{R})}(1) = e^{i.1.m_\lambda} e^{-1\sigma_\lambda^2 \frac{1}{2}} = e^{i\langle \lambda | m \rangle} e^{-\lambda^T \Sigma \frac{\lambda}{2}}$$

$\square$

Proposition 1 (conséquences)

Soit $X = \begin{pmatrix} X_1 \in \mathbb{R}^p \\ X_2 \in \mathbb{R}^q \end{pmatrix}$ avec :

$$X \sim \mathcal{N}\left(\begin{pmatrix} m_1 \in \mathbb{R}^p \\ m_2 \in \mathbb{R}^q \end{pmatrix} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \right)$$

avec $\Sigma_{11} \in M_{p \times p}$ et $\Sigma_{22} \in M_{q \times q}$, etc. Alors :

1. $X_1 \sim \mathcal{N}(m_1, \Sigma_{11})$
2. X_1 indépendante de X_2 $\Leftrightarrow$ $\Sigma_{12} = 0$

Preuve.

1. trivial, X_1 vecteur gaussien et $\mathbb{E}[X_1] = m_1$, $\Gamma_{X_1} = \Sigma_{11}$
2. - $\Rightarrow$: s'ils sont indépendants alors les covariances entre coordonnées de X_1 et X_2 sont nulles
   -

$$\begin{aligned}
\phi_{(X_1, X_2)}(\underbrace{\lambda_1}_{\in \mathbb{R}^p}, \underbrace{\lambda_2}_{\in \mathbb{R}^q}) &= e^{i(\langle \lambda_1, m_1 \rangle + \langle \lambda_2, m_2 \rangle)} \exp\left(-(\lambda_1 | \lambda_2) \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} \times \frac{1}{2} \right) \\
&= ... \\
&= \phi_{X_1}(\lambda_1) \phi_{X_2}(\lambda_2)
\end{aligned}$$

$\square$

Contre-exemple. Si $X \sim \mathcal{N}(0,1)$ indépendante de $Z \sim \mathcal{U}\{-1,+1\}$, posons $Y = ZX$.

- X et Y ne sont pas indépendants (car $|Y| = |X|$ mais $\mathbb{P}(|X| < 1 \cap |Y| > 1) = 0$ et $\mathbb{P}(|X| > 1)\mathbb{P}(|Y| > 1) > 0$ car gaussienne a une densité > 0).
- Mais $X \sim \mathcal{N}(0,1)$,

$$\begin{aligned}\mathbb{P}(Y \in A) &= \mathbb{P}(X \in A \cap (Z = 1)) + \mathbb{P}(X \in -A \cap Z = -1) \\ &= \frac{1}{2}\mathbb{P}(X \in A) + \frac{1}{2}\mathbb{P}(X \in A) \\ &= \mathbb{P}(X \in A)\end{aligned}$$

  Donc $Y \sim \mathcal{N}(0,1)$. Or :

$$\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \underbrace{\mathbb{E}[X]\mathbb{E}[Y]}_{0} = \mathbb{E}[ZX^2] = \underbrace{\mathbb{E}[Z]}_{0}\mathbb{E}[X^2] = 0$$

Cela montre que $\mathrm{Cov}(X,Y) = 0$ n'implique pas X indépendant de Y, et montre que les coordonnées sont gaussiennes n'implique pas que c'est un vecteur gaussien.