

Vecteurs aléatoires Gaussiens

Préliminaires

Notation pour les vecteurs et les matrices. Tout vecteur $u \in \mathbb{R}^d$ est un vecteur colonne (autrement dit une matrice avec une colonne et d lignes, ou $d \times 1$). Si $A \in \mathbb{R}^{m \times n}$ est une matrice $m \times n$ (m lignes et n colonnes) d'élément courant A_{ij} pour $i = 1, \dots, m, j = 1, \dots, n$ alors A^T est la matrice *transposée* qui est une matrice $n \times m$ d'élément courant $(A^T)_{ij} = A_{ji}, i = 1, \dots, n, j = 1, \dots, m$.

Si A est une matrice $n \times m$ et B une matrice $m \times k$ alors le produit (lignes par colonnes) AB est défini étant la matrice $n \times k$ d'élément courant $(AB)_{ij} = \sum_{\ell=1}^m A_{i\ell} B_{\ell j}$ pour $i = 1, \dots, n$ et $j = 1, \dots, k$. On a que $(AB)^T = B^T A^T$. On note \mathbb{I}_d la matrice identité $d \times d$, i.e. $(\mathbb{I}_d)_{ij} = 1$ si $i = j = 1, \dots, d$ et $(\mathbb{I}_d)_{ij} = 0$ si $i \neq j, i, j = 1, \dots, d$.

La transposée u^T d'un vecteur $u \in \mathbb{R}^d$ est un vecteur ligne (ce qui revient au même à une matrice $1 \times d$); si $u \in \mathbb{R}^n$ et $v \in \mathbb{R}^m$ alors le produit matriciel uv^T est une matrice $n \times m$ d'élément courant $(uv^T)_{ij} = u_{i1}(v^T)_{1j} = u_{i1}v_{j1} = u_i v_j$ pour $i = 1, \dots, n$ et $j = 1, \dots, m$; si $u \in \mathbb{R}^n$ et $v \in \mathbb{R}^n$ le produit matriciel $u^T v$ est une matrice 1×1 d'élément $(u^T v)_{11} = \sum_{i=1}^n (u^T)_{1i} v_{i1} = \sum_{i=1}^n u_i v_i$ qui n'est rien d'autre que le produit scalaire des deux vecteurs u et v .

Si X est un vecteur aléatoire de dimension d alors $\mathbb{E}[X]$ est le vecteur de dimension d tel que $(\mathbb{E}[X])_i = \mathbb{E}[X_i]$. Si A est une matrice aléatoire alors $\mathbb{E}[A]$ est la matrice d'élément courant $(\mathbb{E}[A])_{ij} = \mathbb{E}[A_{ij}]$.

Matrice de covariance. Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire dans \mathbb{R}^d tel que $\mathbb{E}[X_j^2] < +\infty$ pour tout $j = 1, \dots, d$. On appelle matrice de covariance du vecteur X , et on la notera Σ , la matrice d'élément courant $\Sigma_{ij} = \text{Cov}(X_i, X_j), i, j = 1, \dots, d$.

Proposition 1.

1. $\Sigma = \mathbb{E}[(X - \mathbb{E}[X])(X - \mathbb{E}[X])^T]$
2. Les éléments diagonaux sont les variances des composantes de X : $\Sigma_{ii} = \text{Var}(X_i)$.
3. Pour tout $u \in \mathbb{R}^d$: $\text{Var}(u^T X) = u^T \Sigma u$.
4. Σ est symétrique et semi-définie positive: $\Sigma_{ij} = \Sigma_{ji}$ et $u^T \Sigma u \geq 0$ pour tout $u \in \mathbb{R}^d$.

Démonstration.

1. $\mathbb{E}[(X - \mathbb{E}[X])(X - \mathbb{E}[X])^T]$ est une matrice d'élément courant $\mathbb{E}[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])] = \text{Cov}(X_i, X_j) = \Sigma_{ij}$.
2. $\Sigma_{ii} = \text{Cov}(X_i, X_i) = \text{Var}(X_i)$
3. $\text{Var}(u^T X) = \text{Var}(u_1 X_1 + \dots + u_d X_d) = \text{Cov}(\sum_{i=1}^d u_i X_i, \sum_{j=1}^d u_j X_j) = \sum_{i,j=1}^d u_i u_j \text{Cov}(X_i, X_j) = \sum_{i,j=1}^d u_i u_j \Sigma_{ij} = u^T \Sigma u$. La covariance étant une fonction bilinéaire.
4. $\Sigma_{ij} = \text{Cov}(X_i, X_j) = \text{Cov}(X_j, X_i) = \Sigma_{ji}$. $u^T \Sigma u = \text{Var}(u^T X) \geq 0$. □

Remarque 2. Si X est tel que les composantes X_j sont indépendantes, alors la matrice de covariance Σ est diagonale car $\Sigma_{ij} = \text{Cov}(X_i, X_j) = 0$ si $i \neq j$. Et donc $\text{Var}(u^T X) = \sum_{i=1}^d u_i^2 \text{Var}(X_i)$ pour tout $u \in \mathbb{R}^d$.

Définition 3. La fonction caractéristique $\phi_X: \mathbb{R} \rightarrow \mathbb{C}$ d'une v.a. réelle X est la transformée de Fourier de sa loi de probabilité:

$$\phi_X(t) = \mathbb{E}[e^{itX}], \quad t \in \mathbb{R}$$

Si X admet une densité f_X alors $\phi_X(t) = \int_{\mathbb{R}} e^{itx} f_X(x) dx$.

La fonction caractéristique détermine de manière unique une loi de probabilité, d'où son nom.

Proposition 4. Deux v.a. réelles X et Y telles que $\phi_X(t) = \phi_Y(t)$ pour tout $t \in \mathbb{R}$ ont la même loi, c-à-d: $\mathbb{P}(X \in A) = \mathbb{P}(Y \in A)$ pour tout Borélien A de \mathbb{R} . Si X ou Y admet une densité alors l'autre v.a. admet une densité aussi et $f_X(z) = f_Y(z)$ pour tout $z \in \mathbb{R}$.

Quelques propriétés calculatoires...

Proposition 5.

1. Pour tout $\lambda \in \mathbb{R}$: $\phi_{\lambda X}(t) = \phi_X(\lambda t)$.
2. Pour tout $a \in \mathbb{R}$: $\phi_{X+a}(t) = e^{iat} \phi_X(t)$.
3. Soit $\mu = \mathbb{E}[X]$ et $\sigma^2 = \text{Var}(X)$ et $U = (X - \mu)/\sigma$ alors $\phi_X(t) = \phi_U(\sigma t) e^{it\mu}$ et $\phi_U(t) = \phi_X(t/\sigma) e^{-it\mu/\sigma}$.
4. Si X et Y sont indépendantes alors $\phi_{X+Y}(t) = \phi_X(t) \phi_Y(t)$.

Démonstration. $\phi_{\lambda X}(t) = \mathbb{E}[e^{it\lambda X}] = \phi_X(\lambda t)$. $\phi_{X+a}(t) = \mathbb{E}[e^{it(X+a)}] = e^{ita} \mathbb{E}[e^{itX}] = e^{ia} \phi_X(t)$.

$\phi_X(t) = \mathbb{E}[e^{itX}] = \mathbb{E}[e^{it(\mu + \sigma U)}] = e^{it\mu} \mathbb{E}[e^{i\sigma U}] = e^{it\mu} \phi_U(\sigma t)$. Pour X et Y indépendantes on a $\mathbb{E}[e^{i(X+Y)t}] = \mathbb{E}[e^{iXt} e^{iYt}] = \mathbb{E}[e^{iXt}] \mathbb{E}[e^{iYt}] = \phi_X(t) \phi_Y(t)$. \square

Exemple 6. Si $X \sim \mathcal{E}(\lambda)$ alors

$$\phi_X(t) = \mathbb{E}[e^{iXt}] = \lambda \int_0^{\infty} e^{ixt - \lambda x} dx = \frac{\lambda}{\lambda - it}$$

En effet si $F(x) = e^{(it-\lambda)x}/(it-\lambda)$ alors $F'(x) = e^{(it-\lambda)x}$ et

$$\lim_{x \rightarrow +\infty} F(x) = \lim_{x \rightarrow +\infty} \frac{\cos(t)e^{-\lambda x}}{it-\lambda} + i \lim_{x \rightarrow +\infty} \frac{\sin(t)e^{-\lambda x}}{it-\lambda} = 0$$

donc

$$\int_0^{\infty} e^{ixt - \lambda x} dx = F(+\infty) - F(0) = -F(0) = \frac{1}{\lambda - it}$$

Exemple 7. Si $Z \sim \mathcal{N}(0, 1)$ alors

$$\phi_Z(t) = \int_{\mathbb{R}} e^{itx - x^2/2} \frac{dx}{\sqrt{2\pi}} = e^{-t^2/2}, \quad \forall t \in \mathbb{R}$$

Si $X = \mu + \sigma Z$ alors $X \sim \mathcal{N}(\mu, \sigma^2)$ et donc

$$\phi_X(t) = e^{i\mu t} \phi_Z(\sigma t) = e^{i\mu t - \sigma^2 t^2/2} = \exp(it\mathbb{E}[X] - \text{Var}(X)t^2/2)$$

Exemple 8. Soit $Y \sim \mathcal{P}(\lambda)$ (Poisson) et $X|Y \sim \mathcal{Bin}(Y, p)$ c-à-d

$$\mathbb{P}(X = k|Y = n) = \binom{n}{k} p^k (1-p)^{n-k}$$

alors

$$\mathbb{E}[e^{iXt}|Y] = (1 + p(e^{it} - 1))^Y, \quad \mathbb{E}[e^{itY}] = e^{\lambda(e^{it} - 1)}$$

et

$$\phi_X(t) = \mathbb{E}[e^{iXt}] = \mathbb{E}[\mathbb{E}[e^{iXt}|Y]] = \mathbb{E}[(1 + p(e^{it} - 1))^Y] = e^{\lambda(1+p(e^{it}-1))-1} = e^{\lambda p(e^{it} - 1)}$$

et donc la loi marginale de X est la loi de Poisson de paramètre λp .

Définition 9. Soit X un vecteur aléatoire dans \mathbb{R}^d . La fonction caractéristique $\phi_X: \mathbb{R}^d \rightarrow \mathbb{C}$ de X est donnée par

$$\phi_X(t) = \mathbb{E}[e^{it^T X}] = \mathbb{E}[e^{i(t_1 X_1 + \dots + t_d X_d)}] \quad \forall t \in \mathbb{R}^d.$$

Proposition 10. Soit X un vecteur aléatoire dans \mathbb{R}^d . Les v.a. X_1, \dots, X_d sont indépendantes ssi

$$\phi_X(t) = \prod_{i=1}^d \phi_{X_i}(t_i) \quad \forall t \in \mathbb{R}^d.$$

Proposition 11. Deux vecteurs aléatoires X, Y ont la même loi ssi $\phi_X(t) = \phi_Y(t)$ pour tout $t \in \mathbb{R}^d$. En particulier si l'une des deux v.a. admet densité alors il en est de même pour l'autre v.a. et $f_X = f_Y$.

Vecteurs Gaussiens

Proposition 12. (Stabilité) Soient X_1, \dots, X_d des v.a. Gaussiennes indépendantes. Pour tout $u \in \mathbb{R}^d$ la combinaison linéaire $u^T X = u_1 X_1 + \dots + u_d X_d$ est une v.a. Gaussienne réelle.

Démonstration. Il suffi de montrer que la fonction caractéristique de $Y = u^T X$ est la f.c. d'une v.a. Gaussienne réelle, i.e.

$$\phi_Y(t) = \exp(it \mathbb{E}[Y] - \text{Var}(Y)t^2/2)$$

(voir l'Exemple 7). Or: $\mathbb{E}[Y] = \sum_{i=1}^d u_i \mathbb{E}[X_i]$ et $\text{Var}(Y) = \sum_{i=1}^d u_i^2 \text{Var}(X_i)$ par l'indépendance des X_i . Donc

$$\begin{aligned} \phi_Y(t) &= \mathbb{E}[e^{it(u_1 X_1 + \dots + u_d X_d)}] = \mathbb{E}\left[\prod_{j=1}^d e^{it u_j X_j}\right] = \prod_{j=1}^d \mathbb{E}[e^{it u_j X_j}] \\ &= e^{(t \sum_{j=1}^d \mathbb{E}[u_j X_j] - \sum_{j=1}^d \text{Var}(u_j X_j) t^2/2)} = \exp(it \mathbb{E}[Y] - \text{Var}(Y)t^2/2) \end{aligned}$$

On en déduit par l'unicité de la fonction caractéristique que Y est une v.a. Gaussienne. \square

Définition 13. On appelle vecteur Gaussien de dimension d un vecteur aléatoire $X = (X_1, \dots, X_d)$ tel que toute combinaison linéaire de ses composantes $(X_j)_{j=1, \dots, d}$ est une v.a. réelle Gaussienne, c-à-d $\forall u \in \mathbb{R}^d$ la v.a. $u^T X$ est Gaussienne. (En particulier, toutes ses composantes X_j sont Gaussiennes.)

Par la Proposition 12 de tels vecteurs existent bien. Si $g: \mathbb{R}^d \rightarrow \mathbb{R}^m$ est une application linéaire et X est un vecteur Gaussien de dimension d alors la composée $Y = g(X)$ est un vecteur Gaussien de dimension m . En effet si $g_k(x) = \sum_{j=1}^d g_{kj} x_j$ alors pour tout $u \in \mathbb{R}^m$, la v.a.

$$u^T Y = \sum_{k=1}^m u_k Y_k = \sum_{k=1}^m u_k g_k(X) = \sum_{k=1}^m u_k \sum_{j=1}^d g_{kj} X_j = \sum_{j=1}^d \left(\sum_{k=1}^m u_k g_{kj} \right) X_j = (g^T u)^T X$$

est Gaussienne, car combinaison linéaire des composantes de X .

Si X est un vecteur Gaussien (de dimension d) alors $Y = X - \mathbb{E}[X]$ est encore un vecteur Gaussien: $u^T Y = u^T X - u^T \mathbb{E}[X]$ est Gaussienne car somme d'une Gaussienne et une constante (à vérifier à l'aide de la fonction caractéristique).

Théorème 14. Soit X un vecteur Gaussien de dimension d . Soit $\mu = \mathbb{E}[X] \in \mathbb{R}^d$ l'espérance de X et Σ la matrice de covariance de X . La fonction caractéristique du vecteur X est donnée par

$$\phi_X(t) = \exp\left(i t^T \mu - \frac{1}{2} t^T \Sigma t \right) \quad \forall t \in \mathbb{R}^d.$$

Démonstration. $Y = t^T X$ est une v.a. Gaussienne. $\mathbb{E}[Y] = t^T \mu$. $\text{Var}(Y) = t^T \Sigma t$.

$$\phi_X(t) = \phi_Y(1) = e^{i \mathbb{E}[Y] - \text{Var}(Y)/2} = e^{i t^T \mu - t^T \Sigma t / 2}. \quad \square$$

Remarque: la loi du vecteur Gaussien ne dépend que de son espérance et de sa covariance. La quantité $t^T \Sigma t$ est toujours ≥ 0 .

Définition 15. Soit X un vecteur Gaussien d'espérance μ et matrice de covariance Σ . On note sa loi par $\mathcal{N}_d(\mu, \Sigma)$.

Lemme 16. Soit Σ une matrice $d \times d$, symétrique et semi-définie positive. Alors il existe une matrice carrée A de dimension $d \times d$ telle que $\Sigma = A A^T$. On dit que A est une racine carrée de Σ . De plus si Σ est inversible alors il en est de même de A .

Démonstration. Du fait que Σ est symétrique on déduit que elle est diagonalisable et donc qu'il existent une matrice orthogonale O (i.e. telle que $O^T = O^{-1}$) et une matrice diagonale Λ avec $\Lambda_{ii} = \lambda_i$ telles que $\Sigma = O^T \Lambda O$. Puisque Σ est semi-définie positive on a que les valeurs propres de Σ sont ≥ 0 et donc que $\lambda_i \geq 0$ pour tout $i = 1, \dots, d$. Soit $\Lambda^{1/2}$ la matrice diagonale telle que $(\Lambda^{1/2})_{ii} = \sqrt{\lambda_i}$, donc $\Lambda^{1/2} \Lambda^{1/2} = \Lambda$ et si on pose $A = O^T \Lambda^{1/2}$ on a que $A A^T = O^T \Lambda^{1/2} (O^T \Lambda^{1/2})^T = O^T \Lambda^{1/2} \Lambda^{1/2} O = O^T \Lambda O = \Sigma$. Si Σ est inversible alors $\lambda_i > 0$ pour tout $i = 1, \dots, d$ et donc $\Lambda^{1/2}$ est inversible et $(\Lambda^{1/2})^{-1}$ est la matrice diagonale avec éléments $1/\sqrt{\lambda_i}$ sur la diagonale. Donc $A^{-1} = (O^T \Lambda^{1/2})^{-1} = (\Lambda^{1/2})^{-1} (O^T)^{-1} = (\Lambda^{1/2})^{-1} O$ qui montre que A est inversible (étant le produit de deux matrices inversibles). \square

Exemple 17. Si $X \sim \mathcal{N}_d(\mu, \Sigma)$, A est une matrice $n \times d$ et $v \in \mathbb{R}^n$ alors

$$v + A X \sim \mathcal{N}_n(v + A \mu, A \Sigma A^T).$$

En effet on remarque que $t^T A X = (A^T t)^T X$ et donc

$$\begin{aligned} \phi_{v+AX}(t) &= e^{i t^T v} \phi_X(A^T t) = \exp\left(i [t^T v + (A^T t)^T \mu] - \frac{1}{2} (A^T t)^T \Sigma (A^T t) \right) \\ &= \exp\left(i t^T [v + A \mu] - \frac{1}{2} t^T A \Sigma A^T t \right). \end{aligned}$$

En particulier si on considère un vecteur aléatoire $Z = (Z_1, \dots, Z_d)$ dans \mathbb{R}^d tel que $Z_i \sim \mathcal{N}(0, 1)$ pour tout $i = 1, \dots, d$ et les v.a. $Z_i, i = 1, \dots, d$ sont indépendantes alors $Z \sim \mathcal{N}_d(0, \mathbb{I}_d)$. Si A est une matrice telle que $AA^T = \Sigma$ (une racine carrée de Σ donné par le Lemme 16) alors $X = \mu + AZ \sim \mathcal{N}_d(\mu, \Sigma)$. D'une famille de v.a. Gaussiennes indépendantes on peut donc construire n'importe quel vecteur Gaussien. Si Σ est inversible alors il en est de même de A et

$$Z = A^{-1}(X - \mu).$$

Théorème 18. Soit $\mu = (\mu_1, \dots, \mu_d) \in \mathbb{R}^d$ un vecteur et Σ une matrice $d \times d$, la fonction

$$f_{\mu, \Sigma}(t) = \exp(it^T \mu - t^T \Sigma t / 2)$$

est la fonction caractéristique d'un vecteur Gaussien d'espérance μ et matrice de covariance Σ ssi Σ est symétrique et semi-définie positive. Pour tout $\mu \in \mathbb{R}^d$ et Σ matrice $d \times d$ symétrique et semi-définie positive il existe un vecteur $X \sim \mathcal{N}_d(\mu, \Sigma)$.

Démonstration. On a déjà vu que si $X \sim \mathcal{N}_d(\mu, \Sigma)$ alors $\phi_X(t) = f_{\mu, \Sigma}(t)$. Il nous reste donc de montrer que si Σ est semi-définie positive et symétrique alors il existe bien un vecteur aléatoire Gaussien X de moyenne μ et matrice de covariance Σ (et donc tel que $\phi_X(t) = f_{\mu, \Sigma}(t)$). Mais par l'exemple 17 et le lemme 16 on a que il existe une matrice A telle que $\Sigma = AA^T$ et que si $Z \sim \mathcal{N}_d(0, \mathbb{I}_d)$ alors $X = \mu + AZ$ est un vecteur aléatoire gaussien de moyenne μ et matrice de covariance $AA^T = \Sigma$. \square

Proposition 19. Le vecteur aléatoire $X \sim \mathcal{N}_d(\mu, \Sigma)$ admet une densité si et seulement si Σ est inversible (i.e. définie positive) et alors

$$f_X(x) = (2\pi)^{-d/2} \det(\Sigma)^{-1/2} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right) \quad \forall x \in \mathbb{R}^d \quad (1)$$

Démonstration. On montre seulement que si Σ est inversible alors X admet la densité donnée en eq. (1). On considère le vecteur aléatoire $Z = (Z_1, \dots, Z_d)$ dans \mathbb{R}^d tel que $Z_i \sim \mathcal{N}(0, 1)$ pour tout $i = 1, \dots, d$ et les v.a. $Z_i, i = 1, \dots, d$ sont indépendantes. Donc la densité de Z est donnée par

$$f_Z(z) = f_{Z_1}(z_1) \cdots f_{Z_d}(z_d) = \frac{1}{(2\pi)^{d/2}} \exp\left(-\frac{1}{2}(z_1^2 + \cdots + z_d^2)\right) = \frac{1}{(2\pi)^{d/2}} \exp\left(-\frac{t^T t}{2}\right)$$

pour tout $z \in \mathbb{R}^d$. Par l'exemple 17 on a que la v.a. $X = \mu + AZ$ (ou $\Sigma = AA^T$) est bien un vecteur gaussien de moyenne μ et matrice de covariance Σ . Donc la densité de X est donnée par la formule de changement de variables à partir de la densité de Z . Si l'on pose $\Psi(z) = \mu + Az$ alors $z = \Psi^{-1}(x) = A^{-1}(x - \mu)$ et

$$f_X(x) = f_Z(\Psi^{-1}(x)) J \Psi^{-1}(x) = \frac{1}{(2\pi)^{d/2}} \det(A^{-1}) \exp\left(-\frac{(\Psi^{-1}(x))^T \Psi^{-1}(x)}{2}\right).$$

Mais

$$\begin{aligned} (\Psi^{-1}(x))^T \Psi^{-1}(x) &= [A^{-1}(x - \mu)]^T A^{-1}(x - \mu) = (x - \mu)^T (A^{-1})^T A^{-1}(x - \mu) \\ &= (x - \mu)^T (A^T)^{-1} A^{-1}(x - \mu) = (x - \mu)^T (AA^T)^{-1}(x - \mu) = (x - \mu)^T \Sigma^{-1}(x - \mu) \end{aligned}$$

et $\det(\Sigma) = \det(AA^T) = \det(A) \det(A^T) = [\det(A)]^2$ et $\det(A^{-1}) = 1/\det(A)$. Donc $\det(A^{-1}) = [\det \Sigma]^{-1/2}$ et on obtient la formule (1). \square

Lemme 20. Soit $X \sim \mathcal{N}_d(\mu, \Sigma)$. Les composantes $(X_j)_{j=1, \dots, d}$ de X sont affinement indépendantes (c-à-d il n'existe pas de vecteur $u \in \mathbb{R}^d$ et $c \in \mathbb{R}$ tels que $u \neq 0$ et $u^T X = c$) ssi Σ est inversible.

Démonstration. Montrons que si Σ est inversible les composantes de X sont affinement indépendantes: en effet si tel vecteur existait alors pour tout $j = 1, \dots, d$:

$$0 = \text{Cov}(c, X_j) = \text{Cov}(u^T X, X_j) = \sum_{k=1}^d u_k \text{Cov}(X_k, X_j) = \sum_{k=1}^d u_k \Sigma_{jk} = (\Sigma u)_j$$

et donc $\Sigma u = 0$ qui montre que Σ a une valeur propre nulle et donc ne peut pas être inversible. Réciproquement si Σ est singulière (c-à-d elle n'est pas inversible) alors il existe $u \in \mathbb{R}^d$, $u \neq 0$ tel que $\Sigma u = 0$ et donc $\text{Var}(u^T X) = u^T \Sigma u = 0$ ce qui implique que la v.a. $u^T X$ est constante et égale à $\mathbb{E}[u^T X] = u^T \mu$ donc $u^T X = c = u^T \mu$ qui montre que les composantes de X sont affinement dépendantes. \square

Exemple 21. Soit (X, Y) un couple aléatoire Gaussien de matrice de covariance égale à

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

1. Condition nécessaire et suffisante pour que Σ soit vraiment la matrice de covariance de (X, Y) est que Σ soit symétrique et semi-définie positive ($\det(\Sigma) \geq 0$ et $\text{Tr}(\Sigma) \geq 0$), i.e. $1 - \rho^2 \geq 0 \Leftrightarrow |\rho| \leq 1$.
2. Condition nécessaire et suffisante sur Σ pour qu'en plus le couple (X, Y) admette une densité est que Σ soit inversible (i.e. définie positive $\Leftrightarrow \det(\Sigma) > 0$ et $\text{Tr}(\Sigma) > 0$). Donc $1 - \rho^2 > 0 \Leftrightarrow |\rho| < 1$.
3. Si $|\rho| < 1$ et (X, Y) est supposé centré ($\mathbb{E}[X] = \mathbb{E}[Y] = 0$), alors (X, Y) admet pour densité

$$f_{(X, Y)}(x, y) = \frac{e^{-\frac{1}{2}(x, y) \Sigma^{-1} \begin{pmatrix} x \\ y \end{pmatrix}}}{2\pi \det(\Sigma)^{1/2}}$$

et on a

$$\Sigma^{-1} = \frac{1}{1 - \rho^2} \begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix}$$

$$(x, y) \Sigma^{-1} \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{1 - \rho^2} (x, y) \begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{1 - \rho^2} (x, y) \begin{pmatrix} x - \rho y \\ y - \rho x \end{pmatrix}$$

$$= \frac{1}{1 - \rho^2} (x^2 - 2\rho x y + y^2)$$

$$f_{(X, Y)}(x, y) = \frac{e^{-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho x y + y^2)}}{2\pi \sqrt{1 - \rho^2}}, \quad x, y \in \mathbb{R}$$

Proposition 22. Soit X un vecteur Gaussien de dimension d . Les v.a. Gaussiennes X_i et X_j sont indépendantes ssi $\text{Cov}(X_i, X_j) = \Sigma_{ij} = 0$. D'une manière générale X_{i_1}, \dots, X_{i_k} sont indépendantes ssi $\text{Cov}(X_{i_a}, X_{i_b}) = 0$ pour tout $a \neq b$ et $a, b = 1, \dots, k$.

Démonstration. Il est clair que si X_i et X_j sont indépendantes alors $\Sigma_{ij} = 0$. Montrons alors que $\Sigma_{ij} = 0$ implique l'indépendance de X_i et X_j . Soit $t = (t_1, t_2) \in \mathbb{R}^2$ et $\mu = \mathbb{E}[X]$. La v.a. $Y = t_1 X_i + t_2 X_j$ est une v.a. Gaussienne de moyenne $\mathbb{E}[Y] = t_1 \mu_i + t_2 \mu_j$ et variance $\text{Var}(Y) = t_1^2 \Sigma_{ii} + t_2^2 \Sigma_{jj}$ et donc

$$\phi_{(X_i, X_j)}(t) = \phi_Y(1) = \mathbb{E}[e^{i(t_1 X_i + t_2 X_j)}] = e^{i(t_1 \mu_i + t_2 \mu_j) - (t_1^2 \Sigma_{ii} + t_2^2 \Sigma_{jj})/2} = \phi_{X_i}(t_1) \phi_{X_j}(t_2)$$

ce qui implique l'indépendance. (En effet, par la propriété fondamentale des fonctions caractéristiques, on a montré que le couple (X_i, X_j) a une densité égale à la densité d'un couple de v.a. indépendantes.) \square

Autres distributions classiques

Loi Gamma

La v.a. X suit une loi gamma de paramètres $\alpha > 0$ et $\beta > 0$ ssi sa densité est donnée par

$$f_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \mathbb{I}_{x>0}$$

et on note $X \sim \mathcal{G}(\alpha, \beta)$. Le paramètre α est la forme de la loi Gamma et β son intensité. On rappelle que

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt, \quad \forall \alpha > 0$$

et que

1. $\Gamma(\alpha + 1) = \alpha \Gamma(\alpha)$. $\Gamma(1) = 1$, $\Gamma(n + 1) = n!$ si $n \in \mathbb{N}$.
2. $\mathbb{E}[X] = \alpha/\beta$, $\text{Var}(X) = \alpha/\beta^2$.

Loi Bêta

On dit que X suit une loi bêta de paramètres $a > 0$ et $b > 0$ ssi la densité de X est donnée par

$$f_X(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1} \mathbb{I}_{0 < x < 1}$$

et on note $X \sim \mathcal{B}(a, b)$.

Proposition 23. On a $\mathbb{E}[X] = a/(a+b)$, $\text{Var}(X) = ab/(a+b)^2(a+b+1)$. Si X et Y sont indépendantes et $X \sim \mathcal{G}(\alpha, \beta)$ et $Y \sim \mathcal{G}(\alpha', \beta)$ alors

$$S = X + Y \sim \mathcal{G}(\alpha + \alpha', \beta), \quad R = \frac{X}{X+Y} \sim \mathcal{B}(\alpha, \alpha')$$

et S et R sont indépendantes.

Démonstration. Si $X \sim \mathcal{B}(a, b)$

$$\begin{aligned}\mathbb{E}[X] &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 x x^{a-1} (1-x)^{b-1} dx = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 x^{(a+1)-1} (1-x)^{b-1} dx \\ &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(a+1)\Gamma(b)}{\Gamma(a+b+1)} = \frac{a}{a+b}\end{aligned}$$

$$\mathbb{E}[X^2] = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 x^{(a+2)-1} (1-x)^{b-1} dx = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(a+2)\Gamma(b)}{\Gamma(a+b+2)} = \frac{a(a+1)}{(a+b)(a+b+1)}.$$

Si $X \sim \mathcal{G}(\alpha, \beta)$ et $Y \sim \mathcal{G}(\alpha', \beta)$ alors on considère le changement de variables $\Psi(x, y) = (s, r)$ avec $s = x + y$ et $r = x/(x + y)$. $\Psi: \mathbb{R}_+^2 \rightarrow \mathbb{R}_+ \times (0, 1)$ et $\Psi^{-1}(s, r) = (x, y)$ avec $x = rs$ et $y = s(1-r)$. La matrice Jacobienne de Ψ^{-1} est donnée par

$$J\Psi^{-1} = \frac{D\Psi^{-1}(s, r)}{D(s, r)} = \begin{vmatrix} \frac{\partial x}{\partial s} & \frac{\partial x}{\partial r} \\ \frac{\partial y}{\partial s} & \frac{\partial y}{\partial r} \end{vmatrix} = \begin{vmatrix} r & s \\ (1-r) & -s \end{vmatrix} = -rs - s(1-r) = -s$$

et donc

$$\begin{aligned}f_{(S,R)}(s, r) &= f_{(X,Y)}(x, y) s = \frac{\beta^{\alpha+\alpha'}}{\Gamma(\alpha)\Gamma(\alpha')} s x^{\alpha-1} y^{\alpha'-1} e^{-\beta(x+y)} \mathbb{I}_{x>0, y>0} \\ &= \frac{\beta^{\alpha+\alpha'}}{\Gamma(\alpha)\Gamma(\alpha')} s^{\alpha+\alpha'-1} e^{-\beta s} r^{\alpha-1} (1-r)^{\alpha'-1} \mathbb{I}_{s>0} \mathbb{I}_{0<r<1} = f_R(r) f_S(s)\end{aligned}$$

avec

$$f_R(r) = \frac{\Gamma(\alpha+\alpha')}{\Gamma(\alpha)\Gamma(\alpha')} r^{\alpha-1} (1-r)^{\alpha'-1} \mathbb{I}_{0<r<1}, \quad f_S(s) = \frac{\beta^{\alpha+\alpha'}}{\Gamma(\alpha+\alpha')} s^{\alpha+\alpha'-1} e^{-\beta s} \mathbb{I}_{s>0}$$

et donc $R \sim \mathcal{B}(\alpha, \alpha')$ et $S \sim \mathcal{G}(\alpha + \alpha', \beta)$ et S et R sont indépendantes. \square

Loi du Khi-deux (χ^2)

Définition 24. Soit $X = (X_1, \dots, X_d)$ un vecteur aléatoire Gaussien centré de matrice de covariance identité (i.e. les composantes de X sont indépendantes et $X_i \sim \mathcal{N}(0, 1)$). On appelle la loi du Khi-deux de d degrés de liberté la loi de la v.a.

$$Y = X_1^2 + \dots + X_d^2$$

et on la note $Y \sim \chi_d^2$.

Proposition 25. Si $Y \sim \chi_d^2$ alors $Y \sim \mathcal{G}(d/2, 1/2)$, $\mathbb{E}[Y] = d$, $\text{Var}(Y) = 2d$.

Démonstration. La loi du carré de la Gaussienne standard $Q = X_1^2$ est $\mathcal{G}(1/2, 1/2)$, en effet la méthode de la fonction muette donne

$$\mathbb{E}[h(Q)] = \int_{\mathbb{R}} h(x^2) \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx = \frac{\sqrt{2}}{\sqrt{\pi}} \int h(q) e^{-q/2} q^{-1/2} dq.$$

Donc la loi de la somme de d carrés de Gaussiennes standards est $\mathcal{G}(d/2, 1/2)$ par les propriétés des lois Gamma (voir la Proposition 23).

On a que $\mathbb{E}[Q] = 1$ et $\text{Var}(Q) = 2$ par les propriétés des Gammas, et donc $\mathbb{E}[Y] = d\mathbb{E}[Q]$ et $\text{Var}(Y) = d\text{Var}(Q)$. \square

Loi de Student

Définition 26. On appelle loi de Student de paramètre d , notée \mathcal{T}_d , la loi de la v.a.

$$T = \frac{X}{\sqrt{Y/d}}$$

où $X \sim \mathcal{N}(0, 1)$, $Y \sim \chi_d^2$ et, X et Y sont indépendantes. T admet pour densité

$$f_T(t) = \frac{\Gamma((d+1)/2)}{\Gamma(d/2) \sqrt{\pi d}} \left(1 + \frac{t^2}{d}\right)^{-(d+1)/2}.$$

Remarque 27. Lorsque $d \rightarrow \infty$ on a que

$$\lim_{d \rightarrow \infty} \frac{f_T(t)}{f_T(0)} = \lim_{d \rightarrow \infty} \left(1 + \frac{t^2}{d}\right)^{-(d+1)/2} = e^{-t^2/2}$$

limite qui est proportionnel à la densité d'une Gaussienne standard (centrée réduite).