

La conjecture de corrélation gaussienne

Houcine Ben Dali, Séverin Benzoni

Encadrant : Joseph Lehec

18 juin 2018

Ce mémoire est principalement basé sur le résumé de la preuve de la corrélation gaussienne par Franck Barthe dans le séminaire Bourbaki [1].

1 Lois gaussiennes et vecteurs gaussiens

Sur \mathbb{R} , la loi gaussienne $\mathcal{N}(m, \sigma^2)$ de moyenne m et de variance σ^2 est la loi à densité :

$$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}} dx.$$

Pour les dimensions supérieures, on a la définition suivante :

Définition 1.1. Soit γ une mesure sur \mathbb{R}^n . On dit que γ est une mesure gaussienne si son image par toute forme linéaire est une mesure gaussienne sur \mathbb{R} .

On a une définition équivalente pour les vecteurs aléatoires :

Définition 1.2. Soit X un vecteur aléatoire. On dit que $X = (X_1, \dots, X_n)$ est un vecteur gaussien sur \mathbb{R}^n si la loi de toute combinaison linéaire des X_i est gaussienne sur \mathbb{R} . On note alors : $X \sim \mathcal{N}(m, C)$ où m est la valeur moyenne de X et C est sa matrice de covariance, i.e. $\forall i, j$ $C_{i,j} = Cov(X_i, X_j)$.

Remarque. Vérifions qu'une loi gaussienne est caractérisée par sa moyenne et sa matrice de covariance. Pour cela, on calcule la fonction caractéristique d'un vecteur gaussien X , en remarquant que : $\forall \xi \in \mathbb{R}^n$ $\phi_X(\xi) = \phi_{\langle \xi, X \rangle}(1)$. Or, $\langle \xi, X \rangle$ est de moyenne $\langle \xi, m \rangle$ et sa variance se déduit de C :

$$\sigma^2 = Var(\langle \xi, X \rangle) = \sum_{i,j} \xi_i \xi_j Cov(X_i, X_j) = \sum_{i,j} \xi_i \xi_j C_{i,j} = \langle \xi, C \xi \rangle.$$

Par la transformée de Fourier d'une gaussienne en dimension 1, on en déduit que :

$$\phi_X(\xi) = e^{i\langle \xi, m \rangle - \frac{\langle \xi, C \xi \rangle}{2}}.$$

La loi de X est donc déterminée par m et C .

De plus, par le calcul de la variance ci-dessus, on a la proposition :

Proposition 1.1. Soit X un vecteur aléatoire et C sa matrice de covariance. Alors $\langle \xi, C\xi \rangle = \text{Var}(\langle \xi, X \rangle) \geq 0$, et donc C est symétrique positive.

Exemple 1.1. Si X_1, \dots, X_n sont indépendants tels que $X_i \sim \mathcal{N}(m_i, \sigma_i^2)$, alors $(X_1, \dots, X_n) \sim \mathcal{N}(m, \text{diag}(\sigma_1^2, \dots, \sigma_n^2))$

Proposition 1.2. Soient $G_i \sim \mathcal{N}(0, 1)$, $1 \leq i \leq n$ indépendantes. Alors $G = (G_1, \dots, G_n) \sim \mathcal{N}(0, I_n)$ et pour tout m et $C \geq 0$, $m + \sqrt{C}G \sim \mathcal{N}(m, C)$. Si de plus, C est inversible, la loi de X est à densité :

$$\mathcal{N}(m, C)(dx) = \frac{1}{\sqrt{2\pi^d \det(C)}} e^{-\frac{\langle x-m, C^{-1}(x-m) \rangle}{2}} dx$$

Démonstration. Tout d'abord, par l'exemple 1.1, $G \sim \mathcal{N}(0, I_n)$. Soit $m \in \mathbb{R}^n$ et $C \geq 0$ et $X = m + \sqrt{C}G$. On remarque que les composantes de X sont des combinaison affines des composantes de G , donc X est bien un vecteur gaussien. Par linéarité, $\mathbb{E}(X) = m$ et pour $1 \leq i, j \leq n$:

$$\begin{aligned} \text{Cov}(X_i, X_j) &= \text{Cov}(X_i - m_i, X_j - m_j) \\ &= \text{Cov}\left(\sum_k (\sqrt{C})_{i,k} G_k, \sum_l (\sqrt{C})_{j,l} G_l\right) \\ &= \sum_{k,l} (\sqrt{C})_{i,k} (\sqrt{C})_{j,l} \text{Cov}(G_k, G_l) \\ &= \sum_k (\sqrt{C})_{i,k} (\sqrt{C})_{k,j} = C_{i,j} \end{aligned}$$

Donc $X \sim \mathcal{N}(m, C)$. Calculons la loi de X dans le cas où C est inversible : soit $f \geq 0$ mesurable :

$$\begin{aligned} \mathbb{E}(f(X)) &= \mathbb{E}(f(m + \sqrt{C}G)) \\ &= \int f(m + \sqrt{C}g) d\mathbb{P}_{G_1}(g_1) \dots d\mathbb{P}_{G_n}(g_n) \\ &= \int f(m + \sqrt{C}g) e^{-\frac{g_1^2}{2}} \dots e^{-\frac{g_n^2}{2}} \frac{dg_1 \dots dg_n}{\sqrt{2\pi^d}} \\ &= \int f(m + \sqrt{C}g) e^{-\frac{\|g\|^2}{2}} \frac{dg}{\sqrt{2\pi^d}} \end{aligned}$$

On applique ensuite la formule du changement de variable en posant $x = m + \sqrt{C}g$ (qui est valable car C est inversible) :

$$\begin{aligned} \mathbb{E}(f(X)) &= \int f(x) e^{-\frac{\|\sqrt{C}^{-1}(x-m)\|^2}{2}} \det(\sqrt{C}^{-1}) \frac{dx}{\sqrt{2\pi^d}} \\ &= \int f(x) e^{-\frac{\langle x-m, C^{-1}(x-m) \rangle}{2}} \frac{dx}{\sqrt{2\pi^d} \sqrt{\det(C)}} \end{aligned}$$

□

Maintenant que les définitions sont posées, il est important pour la suite de remarquer que, pour les vecteurs gaussiens, il y a une équivalence entre indépendance et décorrélation, au sens précis suivant :

Proposition 1.3. *Soit $\{Y^{(j)} \sim \mathcal{N}(m_j, C_j), 1 \leq j \leq p\}$ des vecteurs aléatoires et $X = (Y^{(1)}, \dots, Y^{(p)})$ la concaténation des $Y^{(j)}$. Supposons que X est encore un vecteur gaussien. Alors les $Y^{(j)}$ sont indépendants si et seulement si $\forall i \neq j, \forall k, l : \text{Cov}(Y_k^{(i)}, Y_l^{(j)}) = 0$.*

Démonstration. Il suffit de montrer la réciproque : il suffit de calculer la fonction caractéristique de X . Par hypothèse, la matrice C de covariance de X est diagonale par blocs, ce qui donne :

$$\begin{aligned} \phi_X(\xi_1, \dots, \xi_p) &= e^{i \sum_{j=1}^p \langle \xi_j, m_j \rangle} e^{-\sum_{j=1}^p \frac{\langle \xi_j, C_j \xi_j \rangle}{2}} \\ &= \prod_{j=1}^p \phi_{Y^{(j)}}(\xi_j) \end{aligned}$$

Donc les $Y^{(j)}$ sont indépendants. □

2 Conjecture de corrélation gaussienne

2.1 Enoncé de la conjecture

Théorème 2.1. *Soit A et B deux convexes symétriques de \mathbb{R}^n , et γ une mesure de probabilité gaussienne centrée sur \mathbb{R}^n , on a alors*

$$\gamma(A \cap B) \geq \gamma(A)\gamma(B)$$

Remarque. On peut toujours supposer que $C > 0$. En effet, on considère une suite de vecteurs gaussiens $(X_k)_{k \geq 0}$ de loi $\mathcal{N}_n(0, C + \frac{1}{k}I_n)$. Pour k assez grand, $C + \frac{1}{k}I_n$ est inversible, de plus la suite $(X_k)_k$ converge étroitement vers X . En effet $\Phi_{X_k}(\xi) = e^{-\frac{\langle \xi, (C + \frac{1}{k}I_n) \xi \rangle}{2}} \rightarrow \Phi_X(\xi)$, on conclut par le théorème de Levy. Dans la suite, on suppose que $C > 0$.

Remarque. D'après la proposition 1.2, si γ a pour matrice de covariance C , alors γ est la mesure image de γ_n , la mesure gaussienne standard sur \mathbb{R}^n , par l'application linéaire

$$X \rightarrow \sqrt{C}X$$

ainsi

$$\gamma(A \cap B) \geq \gamma(A)\gamma(B) \iff \gamma_n((\sqrt{C}^{-1}A) \cap (\sqrt{C}^{-1}B)) \geq \gamma_n(\sqrt{C}^{-1}A)\gamma_n(\sqrt{C}^{-1}B)$$

Donc il suffit de montrer la conjecture pour la mesure gaussienne standard.

Définition 2.1. Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$. On dit que f est quasi-concave si les ensembles de niveaux $A_t := \{y \in \mathbb{R}^n; f(y) \geq t\}$ sont convexes, $\forall t \in \mathbb{R}$.

Exemple 2.1. Toute fonction concave sur \mathbb{R} est quasi-concave, mais la réciproque est fautive :

$$x \mapsto e^{-|x|}$$

est quasi-concave mais n'est pas concave.

On peut reformuler le théorème 2.1 avec les fonctions quasi-concaves ;

Théorème 2.2. Soient f et g deux fonctions paires et quasi-concaves sur \mathbb{R}^n . Alors

$$\int fgd\gamma_n \geq \int fd\gamma_n \int gd\gamma_n$$

En effet, il est facile de voir que le théorème 2.2 implique le théorème 2.1 (il suffit de prendre les fonctions indicatrices des convexes), montrons la réciproque. Soient f et g deux fonctions paires quasi-concaves. On pose $A_t := \{y \in \mathbb{R}^n; f(y) \geq t\}$ et $B_t := \{y \in \mathbb{R}^n; g(y) \geq t\}$.

$$\begin{aligned} \int_{\mathbb{R}^n} fgd\gamma_n &= \int_{\mathbb{R}^n} \int_{\mathbb{R}} \mathbb{1}_{A_t}(x) dt \int_{\mathbb{R}} \mathbb{1}_{B_s}(x) ds d\gamma_n(x) \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}^n} \mathbb{1}_{A_t \cap B_s}(x) d\gamma_n(x) ds dt \quad \text{car } f(x) = \int_{t \leq f(x)} dt = \int_{\mathbb{R}} \mathbb{1}_{A_t}(x) dt \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} \gamma_n(A_t \cap B_s) ds dt \\ &\geq \int_{\mathbb{R}} \int_{\mathbb{R}} \gamma_n(A_t) \gamma_n(B_s) ds dt \\ &= \int fd\gamma_n \int gd\gamma_n \end{aligned}$$

2.2 Approximation par des polytopes

Définition 2.2. Un polytope A dans \mathbb{R}^n est un corps convexe fermé qui s'écrit sous la forme d'une intersection finie de demi espaces, $A = \bigcap_{i=1}^n \{x \in \mathbb{R}^n; \langle x, u_i \rangle \leq a_i\}$, pour des vecteurs $u_i \in \mathbb{R}^n$ et des réels $(a_i)_{1 \leq i \leq n}$.

Le but de cette partie va être de montrer que l'on peut se contenter d'étudier la corrélation gaussienne sur les polytopes.

Définition 2.3. Soit $K \subset \mathbb{R}^n$. Le polaire de K est l'ensemble :

$$K^\circ = \{y \in \mathbb{R}^n; \langle x, y \rangle \leq 1, \forall x \in K\}$$

Théorème 2.3. Soit K un convexe fermé symétrique non vide. Alors $K = K^{\circ\circ}$.

Démonstration. La première inclusion $K \subset K^{\circ\circ}$ est facile à voir. Montrons l'autre inclusion. Soit $x_0 \notin K$. Comme K est fermé, par le théorème de Hahn-Banach, il existe $y \in \mathbb{R}^n$ et $\alpha \in \mathbb{R}$ tel que $\forall x \in K, \langle x, y \rangle < \alpha < \langle x_0, y \rangle$. Comme $0 \in K, \alpha \neq 0$, quitte alors à diviser par α , on peut supposer que $\alpha = 1$. Ainsi $y \in K^\circ$, et $\langle y, x_0 \rangle > 1$, donc $x_0 \notin K^{\circ\circ}$. \square

Théorème 2.4. *Soit K un corps convexe symétrique. Il existe une suite décroissante de polytope symétrique K_n tel que $K = \bigcap_{n \in \mathbb{N}} K_n$.*

Remarque. On peut toujours supposer que les convexes étudiés sont fermés, puisque $\gamma(K) = \gamma(\overline{K})$.

Démonstration. Soit $(x_n)_{n \geq 0}$ une suite de points dense dans K° . On pose :

$$K_n := \bigcap_{0 \leq i \leq n} \{x \in \mathbb{R}^n; |\langle x, x_i \rangle| \leq 1\}.$$

Il est clair que la suite des polytopes $(K_n)_{n \geq 0}$ est décroissante, et que $K^{\circ\circ} \subset K_n$. Or par le théorème précédent $K = K^{\circ\circ}$ donc $K \subset \bigcap_{n \in \mathbb{N}} K_n$. Réciproquement, si $x \in \bigcap_{n \in \mathbb{N}} K_n$ par densité de $(x_n)_{n \geq 0}$ on a que $\forall y \in K^\circ, \langle x, y \rangle \leq 1$ donc $x \in K^{\circ\circ} = K$. \square

Il suffit alors de prouver la corrélation gaussienne pour les polytopes. En effet, si A et B sont deux corps convexes symétriques, on les approche par deux suites décroissantes de polytopes symétriques $(A_n)_{n \geq 0}$ et $(B_n)_{n \geq 0}$. On a alors :

$$\gamma(A \cap B) = \lim_{n \rightarrow +\infty} \gamma(A_n \cap B_n) \geq \lim_{n \rightarrow +\infty} \gamma(A_n) \gamma(B_n) = \gamma(A) \gamma(B)$$

Remarque. Soient A et B sont deux polytopes symétriques de \mathbb{R}^d , tels que $A = \bigcap_{i=0}^{n_1} \{x \in \mathbb{R}^n; |\langle x, u_i \rangle| \leq 1\}$ et $B = \bigcap_{i=n_1+1}^n \{x \in \mathbb{R}^n; |\langle x, u_i \rangle| \leq 1\}$ où $u_i \in \mathbb{R}^d$. Soit Y un vecteur aléatoire sur \mathbb{R}^d de loi γ . On a :

$$\begin{aligned} \gamma(A \cap B) \geq \gamma(A) \gamma(B) &\iff \mathbb{P}(\forall i, |\langle Y, u_i \rangle| \leq 1) \\ &\geq \mathbb{P}(\forall i \leq n_1, |\langle Y, u_i \rangle| \leq 1) \mathbb{P}(\forall i > n_1, |\langle Y, u_i \rangle| \leq 1) \end{aligned}$$

En posant $X_i = \langle Y, u_i \rangle$, (qui est gaussien car ses composantes sont des combinaisons linéaires de celles de Y) on en déduit que la corrélation gaussienne est équivalente au théorème suivant :

Théorème 2.5. *Soient $1 \leq n_1 < n$ deux entiers et $X = (X_1, X_2, \dots, X_n)$ un vecteur gaussien centré à valeurs dans \mathbb{R}^n . Alors*

$$\mathbb{P}(\max_{1 \leq i \leq n} |X_i| \leq 1) \geq \mathbb{P}(\max_{1 \leq i \leq n_1} |X_i| \leq 1) \mathbb{P}(\max_{n_1 < i \leq n} |X_i| \leq 1)$$

3 Transformée de Laplace

3.1 Définition et propriétés

Définition 3.1. Soit γ une mesure finie sur \mathbb{R}_+^n . La *transformée de Laplace* de γ est définie par :

$$\begin{aligned} \mathcal{L}_\gamma : \mathbb{R}_+^n &\longrightarrow \mathbb{R} \\ \lambda &\longmapsto \int_{\mathbb{R}_+^n} \exp(-\langle x, \lambda \rangle) d\gamma(x) \end{aligned}$$

Remarque. Si X est un vecteur aléatoire, sa transformée de Laplace est celle de sa loi :

$$\mathcal{L}_X(\lambda) = \int_{\mathbb{R}_+^n} \exp(-\langle x, \lambda \rangle) d\mathbb{P}_X(x) = \mathbb{E}(\exp(-\langle x, \lambda \rangle))$$

Théorème 3.1. *L'application : $\mathcal{L} : \gamma \mapsto \mathcal{L}_\gamma$ est injective.*

Démonstration. On traite le cas $n = 1$ pour simplifier les écritures, mais le cas général se traite de la même manière.

Soit γ une mesure signée finie sur \mathbb{R}^+ telles que $\mathcal{L}\gamma = 0$, i.e. $\forall \lambda \geq 0, \int_{\mathbb{R}^+} e^{-\lambda x} d\gamma(x) = 0$.

On pose $f : x \mapsto e^{-x}$, puis on applique la formule de transfert :

$$\begin{aligned} \int_{\mathbb{R}^+} e^{-\lambda x} d\gamma(x) &= \int_{\mathbb{R}^+} f(x)^\lambda d\gamma(x) \\ &= \int_{]0,1]} t^\lambda d\nu(t) \end{aligned}$$

où ν est le mesure image de γ par f qui est donc une mesure signée finie sur $]0, 1]$. Par linéarité, en évaluant la transformée de Laplace sur les entiers, on voit que $\forall P \in \mathbb{R}[X]$,

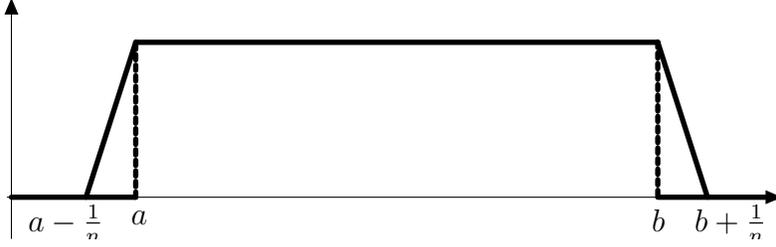
$$\int_{]0,1]} P d\nu = 0.$$

On va en déduire que ν est la mesure nulle. Soit g une fonction continue sur le segment $[0, 1]$. Par le théorème de Weierstrass, il existe une suite $(P_n)_{n \geq 0}$ de polynômes qui tend uniformément vers g sur le segment $[0, 1]$ (dans le cas général, il faut appliquer le théorème de Stone-Weierstrass pour avoir la densité des polynômes à plusieurs variables). Comme ν est une mesure finie, la convergence uniforme sur $[0, 1]$ donne la convergence des intégrales :

$$\int_{]0,1]} P_n d\nu \longrightarrow \int_{]0,1]} g d\nu.$$

Donc $\int_{]0,1]} g d\nu = 0$. On va maintenant passer aux indicatrices.

Soit $[a, b] \subset]0, 1]$ ($a > 0$). Choisissons g_n continue sur $[0, 1]$ comme ci-dessous :



On a convergence ponctuelle : $g_n(x) \rightarrow \mathbb{1}_{[a,b]}(x)$. De plus, $\forall n \in \mathbb{N}, \forall x, |g_n(x)| \leq 1$ qui est intégrable car ν est finie. Donc on peut appliquer le théorème de convergence dominée :

$$\int g_n d\nu \rightarrow \int \mathbb{1}_{[a,b]}(x) d\nu = \nu([a, b]).$$

Donc $\nu([a, b]) = 0$. Comme les segments forment un π -système qui engendre la tribu borélienne, $\nu = 0$. Donc γ est la mesure nulle, donc \mathcal{L} est injective. \square

3.2 Application aux lois gaussiennes

Pour prouver l'inégalité de corrélation, il va être intéressant d'étudier les carrés de vecteurs gaussiens. On va commencer cette étude en regardant leur transformée de Laplace.

Lemme 3.2. Soit $X \sim \mathcal{N}(0, C)$ un vecteur gaussien. On pose $Z = (\frac{X_1^2}{2}, \dots, \frac{X_n^2}{2})$. On a alors :

$$\forall \lambda \in \mathbb{R}_n^+, \mathcal{L}_Z(\lambda) = \det(I_n + \Lambda C)^{-1/2},$$

où $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$. Si $C > 0$, la loi de Z est alors à densité par rapport à la mesure de Lebesgue.

Démonstration. Par la proposition 1.2, le vecteur $\sqrt{C}G$, où $G \sim \mathcal{N}(0, I_n)$, suit la même loi que X . On a donc :

$$\begin{aligned} \mathcal{L}_Z(\lambda) &= \mathbb{E} \left[e^{-\sum_{i=1}^n \lambda_i \frac{X_i^2}{2}} \right] = \mathbb{E} \left[e^{-\frac{1}{2} \langle X, \Lambda X \rangle} \right] = \mathbb{E} \left[e^{-\frac{1}{2} \langle \sqrt{C}G, \Lambda \sqrt{C}G \rangle} \right] \\ &= \mathbb{E} \left[e^{-\frac{1}{2} \langle G, \sqrt{C} \Lambda \sqrt{C} G \rangle} \right] = \int_{\mathbb{R}^n} e^{-\frac{1}{2} \langle y, \sqrt{C} \Lambda \sqrt{C} y \rangle} e^{-\frac{\langle y, y \rangle}{2}} \frac{dy}{(2\pi)^{n/2}} \\ &= \int_{\mathbb{R}^n} e^{-\frac{1}{2} \langle y, M y \rangle} \frac{dy}{(2\pi)^{n/2}}, \text{ où } M = I_n + \sqrt{C} \Lambda \sqrt{C} > 0 \\ &= \int_{\mathbb{R}^n} e^{-\frac{1}{2} \langle x, x \rangle} \det(\sqrt{M})^{-1} \frac{dx}{(2\pi)^{n/2}} \text{ en posant } x = \sqrt{M} y \\ &= \det(I_n + \sqrt{C} \Lambda \sqrt{C})^{-1/2} = \det(I_n + \Lambda C)^{-1/2}, \end{aligned}$$

la dernière égalité découlant de la propriété du polynôme caractéristique : $\chi_{AB} = \chi_{BA}$.

Regardons maintenant le cas où $C > 0$. On prend un rectangle de \mathbb{R}_n^+ : $R = \prod_{i=1}^n [a_i, b_i] \subset$

\mathbb{R}_n^+ , et on regarde :

$$\begin{aligned} \mathbb{P}(Z \in R) &= \mathbb{P}\left(\frac{X_1^2}{2} \in [a_1, b_1], \dots, \frac{X_n^2}{2} \in [a_n, b_n]\right) = \sum_{\epsilon \in \{\pm 1\}^n} \mathbb{P}\left(X_i \in \epsilon_i [\sqrt{2a_i}, \sqrt{2b_i}]\right) \\ &= \sum_{\epsilon \in \{\pm 1\}^n} \int_{\prod_{i=1}^n \epsilon_i [\sqrt{2a_i}, \sqrt{2b_i}]} e^{-\frac{1}{2}\langle C^{-1}x, x \rangle} \frac{dx}{\sqrt{(2\pi)^n \det(C)}} \\ &= \sum_{\epsilon \in \{\pm 1\}^n} \int_{\prod_{i=1}^n [a_i, b_i]} \prod_{i=1}^n \frac{1}{\sqrt{2z_i}} e^{-\langle C^{-1}\epsilon \cdot \sqrt{z}, \epsilon \cdot \sqrt{z} \rangle} \frac{dz}{\sqrt{(2\pi)^n \det(C)}} \text{ en posant } x_i = \epsilon_i \sqrt{2z_i} \end{aligned}$$

où $\epsilon \cdot \sqrt{z} = (\epsilon_i \sqrt{z_i})_{1 \leq i \leq n}$. Donc, on a :

$$d\mathbb{P}_Z(z) = \mathbb{1}_{]0, +\infty[}(z) \frac{\det(C)^{-1/2} (4\pi)^{-n/2}}{\prod_{i=1}^n \sqrt{z_i}} \sum_{\epsilon \in \{\pm 1\}^n} e^{-\langle C^{-1}\epsilon \cdot \sqrt{z}, \epsilon \cdot \sqrt{z} \rangle} dz = f(z) dz$$

□

Pour la suite, on aura besoin de connaître une classe un peu plus générale de transformées de Laplace.

Lemme 3.3. *Fixons $k \in \mathbb{N}^*$ et $C \geq 0$. Alors il existe une mesure de probabilité γ dont la transformée de Laplace est :*

$$\forall \lambda \in \mathbb{R}_n^+, \mathcal{L}_\gamma(\lambda) = \det(I_n + \Lambda C)^{-k/2}.$$

Si $C > 0$, γ est une mesure à densité par rapport à la mesure de Lebesgue : $d\gamma(z) = h_{k/2}(z) dz$.

Démonstration. Dans le cas $k = 1$, on retrouve la variable aléatoire construite au lemme précédent. Si k est quelconque, on prend k copies indépendantes de $Z : Z_1, \dots, Z_k$. On a alors :

$$\mathcal{L}_{Z_1 + \dots + Z_k}(\lambda) = \prod_{i=1}^k \mathcal{L}_{Z_i}(\lambda) = \det(I_n + \Lambda C)^{-k/2}.$$

De plus, si $C > 0$, la loi Z_i est à densité, et alors la loi de $Z_1 + \dots + Z_k$ aussi car : $d\mathbb{P}_{Z_1 + \dots + Z_k}(z) = f * \dots * f(z) dz$. Ainsi, $h_{k/2} = f^{*k}$. □

Les lois construites ci-dessus s'appellent des *lois gamma*. Malheureusement, cette construction ne permet pas d'obtenir des propriétés précises sur $h_{k/2}$. On va donc développer une théorie plus générale sur ces loi.

4 Lois gamma

4.1 Lois gamma en dimension 1

Définition 4.1. Soit $\alpha, \beta > 0$. On définit la loi gamma en dimension 1 par :

$$d\gamma(\alpha, \beta)(x) = \beta^\alpha e^{-\beta x} x^{\alpha-1} \mathbb{1}_{x>0} \frac{dx}{\Gamma(\alpha)}.$$

Cependant, pour la suite on prendra toujours $\beta = 1$, et on notera $\gamma(\alpha) = \gamma(\alpha, 1)$.

Remarque. Pour correspondre aux lois obtenues au lemme 3.3, une loi gamma en dimension 1 doit avoir pour transformée de Laplace : $\mathcal{L}_\gamma(\lambda) = (1 + \lambda)^{-\alpha}$. Vérifions que la définition ci-dessus convient :

$$\begin{aligned} \mathcal{L}_{\gamma(\alpha)} &= \int_{\mathbb{R}^+} e^{-(\lambda+1)x} x^{\alpha-1} \frac{dx}{\Gamma(\alpha)} \\ &= \int_{\mathbb{R}^+} e^{-y} \frac{y^{\alpha-1}}{(1+\lambda)^\alpha \Gamma(\alpha)} dy \quad \text{en posant } y = (1+\lambda)x \\ &= (1+\lambda)^{-\alpha} \end{aligned}$$

En particulier, cela implique : $\gamma(\alpha) * \gamma(\beta) = \gamma(\alpha + \beta)$.

Les lois gamma ainsi définies peuvent s'obtenir comme carrés de loi gaussiennes centrées : en effet, si $G \sim \mathcal{N}(0, 1)$, on a vu que $\frac{G^2}{2} \sim \gamma(\frac{1}{2})$, et en sommant k copies indépendantes de $\frac{G^2}{2}$, on obtient une variable aléatoire de loi $\gamma(k/2)$. Pour cette raison, ces lois sont aussi appelées lois gamma *centrées*. Que se passe-t-il si on part d'une gaussienne non-centrée ? Regardons la transformée de Laplace que l'on obtient : soit $G \sim \mathcal{N}(m, 1)$. On a, par la formule de transfert :

$$\begin{aligned} \mathcal{L}_{\frac{G^2}{2}}(\lambda) &= \int_{\mathbb{R}} e^{-\frac{\lambda x^2}{2}} e^{-\frac{(x-m)^2}{2}} \frac{dx}{\sqrt{2\pi}} \\ &= \int_{\mathbb{R}} e^{-\frac{m^2}{2} \frac{\lambda}{(1+\lambda)}} e^{-\frac{\lambda+1}{2} (x - \frac{m}{\lambda+1})^2} \frac{dx}{\sqrt{2\pi}} \\ &= e^{-\frac{m^2}{2} \frac{\lambda}{(1+\lambda)}} \int_{\mathbb{R}} e^{-\frac{\lambda+1}{2} x^2} \frac{dx}{\sqrt{2\pi}} \\ &= (1+\lambda)^{-1/2} e^{-\frac{m^2}{2} \frac{\lambda}{(1+\lambda)}} \end{aligned}$$

Donc en sommant k copies indépendantes de $\frac{G^2}{2}$, on obtient la transformée de Laplace :

$$\mathcal{L}_Z(\lambda) = (1+\lambda)^{-k/2} e^{-\frac{km^2}{2} \frac{\lambda}{(1+\lambda)}}.$$

On pose donc la définition :

Définition 4.2. Une mesure de probabilité $\gamma(\alpha, y)$ sur \mathbb{R}^+ est une loi gamma de paramètres α et y si :

$$\mathcal{L}_{\gamma(\alpha, y)} = (1 + \lambda)^{-\alpha} e^{-y \frac{\lambda}{1+\lambda}}.$$

Si $y \neq 0$, $\gamma(\alpha, y)$ est une loi gamma *non centrée*.

On va maintenant construire des variables aléatoires suivant ces lois gamma non centrées. Pour cela, on a la proposition :

Proposition 4.1. Soit $\alpha > 0$ et $y > 0$. Il existe une variable aléatoire Z telle que $Z \sim \gamma(\alpha, y)$. De plus, on a :

$$d\gamma(\alpha, y)(x) = \sum_{n=0}^{\infty} e^{-x} \frac{x^{\alpha+n-1}}{\Gamma(\alpha+n)} e^{-y} \frac{y^n}{n!} dx = g_{\alpha, y}(x) dx$$

Remarque. Si $\alpha > 1$, on a les propriétés :

$$\forall x > 0, g_{\alpha, y}(x) \leq 2(1+x)x^{\alpha-1} \quad \text{et} \quad \partial_x g_{\alpha, y} = g_{\alpha-1, y} - g_{\alpha, y} \quad (1)$$

Démontrons la majoration :

On a $\Gamma(x) \geq \frac{1}{2}$, $\forall x \in \mathbb{R}_+^*$. En effet Γ est convexe sur \mathbb{R}_+^* . Or $\Gamma(1) = \Gamma(2) = 1$ donc elle présente un seul minimum entre 1 et 2. Donc il suffit de la minorer entre 1 et 2. Soit $x \in [1, 2]$, on a

$$\Gamma(x) = \frac{\Gamma(x+1)}{x} \geq \frac{\Gamma(2)}{2} \geq \frac{1}{2}$$

car elle est croissante sur $[2, +\infty[$.

Donc si $\alpha > 0$ et $n \in \mathbb{N}^*$ on a

$$\begin{aligned} \Gamma(\alpha+n) &\geq (n-1)! \Gamma(\alpha) \\ &\geq \frac{1}{2} (n-1)! \end{aligned}$$

Ainsi $\frac{x^n}{\Gamma(\alpha+n)} \leq 2xe^x$ si $n \geq 1$, et par suite $\frac{x^n}{\Gamma(\alpha+n)} \leq 2(1+x)e^x, \forall n \in \mathbb{N}$. Finalement,

$$\begin{aligned} g_{\alpha, y}(x) &= x^{\alpha-1} e^{-x} e^{-y} \sum_{n=0}^{\infty} \frac{x^n}{\Gamma(\alpha+n)} \frac{y^n}{n!} \\ &\leq 2(1+x)x^{\alpha-1} e^{-y} \sum_{n=0}^{\infty} \frac{y^n}{n!} \\ &= 2(1+x)x^{\alpha-1} \end{aligned}$$

Démonstration. Soit $(X_n)_{n \geq 1}$ une suite de variables positives i.i.d et N une variable indépendante des X_n qui suit le loi de Poisson $\mathcal{P}(y)$. On va alors regarder la variable

$S = \sum_{1 \leq i \leq N} X_i$, dont la transformée de Laplace est :

$$\begin{aligned}
\mathcal{L}_S(\lambda) &= \int_{\mathbb{R}^+} e^{-\lambda x} d\mathbb{P}_S(x) \\
&= \sum_{n=0}^{\infty} \mathbb{P}(N = n) \int_{\mathbb{R}^+} e^{-\lambda x} d\mathbb{P}_S(x|N = n) \\
&= \sum_{n=0}^{\infty} \mathbb{P}(N = n) \int_{\mathbb{R}^+} e^{-\lambda x} d\mathbb{P}_{\sum_{1 \leq i \leq n} X_i}(x) \text{ car } N \text{ est indépendante des } X_n \\
&= \sum_{n=0}^{\infty} \mathbb{P}(N = n) \mathbb{E}[e^{-\lambda \sum_{1 \leq i \leq n} X_i}] = \sum_{n=0}^{\infty} e^{-y} \frac{y^n}{n!} \mathbb{E}[e^{-\lambda X_1}]^n = e^{y(\mathcal{L}_{X_1}(\lambda) - 1)}
\end{aligned}$$

On voudrait alors que $\mathcal{L}_{X_1}(\lambda) = \frac{1}{1+\lambda}$, donc on prend les X_n tels que $X_n \sim \gamma(1)$. Pour trouver la transformée de Laplace recherchée, on prend une variable nouvelle X_0 indépendante des autres variables et telle que $X_0 \sim \gamma(\alpha)$ et on regarde $Z = X_0 + S$. Cela donne alors :

$$\mathcal{L}_Z(\lambda) = \mathcal{L}_{X_0}(\lambda) \mathcal{L}_S(\lambda) = (1 + \lambda)^{-\alpha} e^{y(\frac{1}{1+\lambda} - 1)} = (1 + \lambda)^{-\alpha} e^{-y \frac{\lambda}{1+\lambda}}.$$

De plus, si on prend A un borélien :

$$\begin{aligned}
\mathbb{P}(Z \in A) &= \sum_{n=0}^{\infty} \mathbb{P}(N = n) \mathbb{P}(Z \in A|N = n) \\
&= \sum_{n=0}^{\infty} \mathbb{P}(N = n) \int_A d\mathbb{P}_{\sum_{0 \leq i \leq n} X_i} \\
&= \sum_{n=0}^{\infty} \mathbb{P}(N = n) \int_A \gamma(\alpha) * \gamma(1)^{*n}(x) dx \\
&= \int_A \sum_{n=0}^{\infty} \mathbb{P}(N = n) \gamma(\alpha + n)(x) dx \text{ (par convergence monotone)} \\
&= \int_A \sum_{n=0}^{\infty} e^{-x} \frac{x^{\alpha+n-1}}{\Gamma(\alpha + n)} e^{-y} \frac{y^n}{n!} dx
\end{aligned}$$

□

4.2 Lois conditionnelles

Avant de pouvoir développer la théorie sur les lois gamma en dimension supérieure, il faut d'abord s'intéresser à la notion de loi conditionnelle. Le principal problème dans cette construction est celui de conditionner par des événements de probabilité nulle. On commence cette approche en introduisant la notion d'espérance conditionnelle.

4.2.1 Espérance conditionnelle

L'idée que l'on cherche à formaliser est de calculer l'espérance d'une variable aléatoire en fonction d'une information additionnelle.

Par exemple, si on partitionne Ω avec des événements de probabilité non-nulle : $\Omega = \bigsqcup_{i \in \mathbb{N}} B_i$. On peut alors regarder la variable : $Y = \sum_{i \in \mathbb{N}} \frac{1}{\mathbb{P}(B_i)} \left(\int_{B_i} X d\mathbb{P} \right) \mathbb{1}_{B_i}$. Par cette définition, on voit que si on a $\omega \in B_i$, Y donne la moyenne de X sur B_i . Pour développer cette notion, posons $\mathcal{B} = \sigma(B_i, i \in \mathbb{N})$, et regardons les propriétés de Y . Tout d'abord, Y est \mathcal{B} -mesurable. De plus, $\forall B \in \mathcal{B}$:

$$\int_B Y d\mathbb{P} = \sum_{i \in \mathbb{N}} \int_{B_i \cap B} X d\mathbb{P} = \int_B X d\mathbb{P}$$

car $\mathcal{B} = \{\bigcup_{i \in I} B_i; I \subset \mathbb{N}\}$. On va alors exploiter le théorème suivant :

Théorème 4.2. *On se place sur un espace de probabilités $(\Omega, \mathcal{A}, \mathbb{P})$. Soit \mathcal{B} une sous-tribu de \mathcal{A} et X \mathcal{A} -mesurable positive (ou L^1). Il existe une unique variable Y , \mathcal{B} -mesurable telle que :*

$$\forall B \in \mathcal{B}, \int_B Y d\mathbb{P} = \int_B X d\mathbb{P} \quad (2)$$

Définition 4.3. On définit l'espérance conditionnelle de X sachant \mathcal{B} , $\mathbb{E}(X|\mathcal{B})$, comme la variable aléatoire introduite dans le théorème précédent. On étend ensuite cette notion pour les vecteurs aléatoires : $\mathbb{E}((X_1, \dots, X_n)|\mathcal{B}) = (\mathbb{E}(X_1|\mathcal{B}), \dots, \mathbb{E}(X_n|\mathcal{B}))$.

Remarque. Toutes ces définitions sont posées à un ensemble de mesure nulle près.

Démonstration. Existence : On définit la mesure signée sur (Ω, \mathcal{B}) : $\nu_X(B) = \int_B X d\mathbb{P}$. Si $B \in \mathcal{B}$ et $\mathbb{P}_{\mathcal{B}}(B) = 0$, alors $\mathbb{P}(B) = 0$, donc $\nu_X(B) = 0$. Ainsi, ν_X est uniformément continue par rapport à $\mathbb{P}_{\mathcal{B}}$. Alors, par le théorème de Radon-Nikodym, il existe Y \mathcal{B} -mesurable telle que :

$$\forall B \in \mathcal{B}, \nu_X(B) = \int_B Y d\mathbb{P}_{\mathcal{B}} = \int_B Y d\mathbb{P}.$$

Unicité : Soient Y et Y' deux variables \mathcal{B} -mesurable telle que : $\forall B \in \mathcal{B}, \int_B Y d\mathbb{P} = \int_B X d\mathbb{P}$. Montrons que $Y = Y'$ presque sûrement. Supposons $\mathbb{P}(Y > Y') > 0$: il existe $\delta > 0$ tel que $\mathbb{P}(Y \geq Y' + \delta) > 0$. Alors $\int_{\{Y \geq Y' + \delta\}} Y - Y' d\mathbb{P} \geq \delta \mathbb{P}(Y \geq Y' + \delta) > 0$, ce qui contredit l'hypothèse. Donc $Y \leq Y'$ presque sûrement. En échangeant Y et Y' , on a $Y = Y'$ presque sûrement. \square

Remarque. La condition (2) équivaut à : $\forall Z$ \mathcal{B} -mesurable positive (ou L^1),

$$\mathbb{E}(ZX) = \mathbb{E}(ZY).$$

La principale formule que l'on retiendra est la suivante :

Proposition 4.3. Soit X une variable aléatoire positive ou L^1 :

$$\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|\mathcal{B}))$$

Démonstration. Il suffit d'appliquer (2) avec $B = \Omega$. □

On peut désormais définir le conditionnement par une variable aléatoire :

Définition 4.4. Soit X et Y deux variables aléatoires. On définit l'espérance de X sachant Y : $\mathbb{E}(X|Y) = \mathbb{E}(X|\sigma(Y))$.

Remarque. Heuristiquement, on calcule l'espérance de X connaissant la valeur de Y .

On va donc préciser cette idée à l'aide de la proposition suivante :

Proposition 4.4. Soit $Y : (\Omega, \mathcal{A}) \rightarrow (E, \mathcal{E})$ mesurable. Une variable Z est $\sigma(Y)$ -mesurable si et seulement si il existe une unique fonction f \mathcal{E} -mesurable définie sur $Y(\Omega)$ telle que $Z = f \circ Y$.

Démonstration. Posons $\mathcal{F} = \{f \circ Y ; f : E \mapsto \mathbb{R} \text{ mesurable}\}$ et montrons que \mathcal{F} contient toutes les fonctions $\sigma(Y)$ -mesurables.

Si $A = Y^{-1}(H) \in \sigma(Y)$, alors $\mathbb{1}_A = \mathbb{1}_H \circ Y \in \mathcal{F}$. De plus, \mathcal{F} est stable par combinaison linéaire, donc il contient toutes les fonction étagées $\sigma(Y)$ -mesurables.

Prenons maintenant $Z \geq 0$ $\sigma(Y)$ -mesurable. Il existe une suite de fonctions étagées $(Z_n)_{n \geq 0}$ telle que $Z_n \nearrow Z$. Or, on a vu que, pour chaque $n \in \mathbb{N}$, $Z_n \in \mathcal{F}$, i.e. il existe f_n telle que $Z_n = f_n \circ Y$. Mais alors, la suite f_n est croissante et on peut donc définir $f(y) = \lim_{n \rightarrow \infty} f_n(y)$, qui est mesurable en tant que limite simple de fonctions mesurables et $Z = f \circ Y$. Donc $Z \in \mathcal{F}$. On obtient ensuite toutes les fonction $\sigma(Y)$ -mesurables en passant par la partie positive et la partie négative. □

On applique ensuite cette proposition à $Z = \mathbb{E}(X|Y)$ pour poser la définition suivante :

Définition 4.5. On définit la variable aléatoire $\mathbb{E}(X|Y = \cdot) : y \mapsto \mathbb{E}(X|Y = y)$ comme l'unique fonction telle que $\mathbb{E}(X|Y) = \mathbb{E}(X|Y = \cdot) \circ Y$.

4.2.2 Définition des lois conditionnelles

On va définir la loi conditionnelle de X sachant Y à l'aide de l'espérance conditionnelle. Dans le cas classique, on détermine la loi d'une variable aléatoire en calculant $\mathbb{E}(f(X))$ pour $f \geq 0$. On pose donc la définition suivante :

Définition 4.6. Soit X une variable aléatoire et $Y : (\Omega, \mathcal{A}) \mapsto (E, \mathcal{E})$ mesurable. Le noyau $(\mu_y, y \in E)$ de mesures de probabilités est la loi de X sachant Y si pour tout borélien A , $y \mapsto \mu_y(A)$ est mesurable et si pour toute fonction $f \geq 0$: $\mathbb{E}(f(X)|Y = y) = \int f(x) d\mu_y(x)$, \mathbb{P}_Y -presque partout.

Remarque. L'existence d'un tel noyau de probabilité n'est pas évidente, mais dans les cas considérés par la suite on l'admettra.

On retiendra seulement cette conséquence de la proposition 4.3 :

Proposition 4.5. Soient X, Y deux variables aléatoires et f une fonction mesurable positive ou $L^1(\mathbb{P})$. Si $(\mu_y, y \in E)$ est la loi de X sachant Y , on peut calculer l'espérance de $f(X)$ par la formule :

$$\mathbb{E}(f(X)) = \mathbb{E} \left[\int f(x) d\mu_Y(x) \right]$$

4.3 Lois gamma en dimension n

Définition 4.7. Soit C une matrice symétrique positive et $\alpha > 0$. Un vecteur aléatoire X suit la loi gamma $\Gamma(\alpha, C)$ si sa transformée de Laplace est :

$$\mathcal{L}_X(\lambda) = \det(I_n + \Lambda C)^{-\alpha},$$

avec $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$.

Remarque. On peut construire des variables aléatoires de loi $\Gamma(\alpha, C)$ pour $\alpha \in \frac{\mathbb{N}^*}{2} \cup \frac{n-2}{2}, \infty[$. Cependant, on se contentera des α demi-entiers construites précédemment à partir de carrés de vecteurs gaussiens. Malheureusement, cette construction ne permet pas d'établir des propriétés précises sur les lois gamma, et on va donc tout de même s'intéresser à un passage de la construction générale des lois gamma dans le cas où α n'est pas demi-entier.

Proposition 4.6. Soit $\mu > 0$ et $Y \sim \Gamma(\alpha, C)$. Soit X une variable dont la loi conditionnelle sachant Y est : $\gamma(\alpha, Y_1/\mu) \otimes \dots \otimes \gamma(\alpha, Y_n/\mu)$. Alors $\mu X \sim \Gamma(\alpha, \mu I_n + C)$.

Démonstration. On calcule la transformée de Laplace de X en utilisant la proposition 4.5 :

$$\begin{aligned} \mathcal{L}_{\mu X}(\lambda) &= \mathbb{E}(e^{-\langle \lambda, \mu X \rangle}) \\ &= \mathbb{E} \left[\int e^{-\langle \lambda, \mu x \rangle} d\gamma(\alpha, Y_1/\mu) \otimes \dots \otimes d\gamma(\alpha, Y_n/\mu)(x) \right] \\ &= \mathbb{E} \left[\prod_{i=1}^n \int e^{-\lambda_i \mu x_i} d\gamma(\alpha, Y_i/\mu)(x_i) \right] \\ &= \mathbb{E} \left[\prod_{i=1}^n (1 + \mu \lambda_i)^{-\alpha} e^{-Y_i \frac{\lambda_i}{1 + \mu \lambda_i}} \right] \\ &= \det(I_n + \mu \Lambda)^{-\alpha} \mathbb{E} \left[\prod_{i=1}^n e^{-Y_i \frac{\lambda_i}{1 + \mu \lambda_i}} \right] \\ &= \det(I_n + \mu \Lambda)^{-\alpha} \mathcal{L}_Y \left(\frac{\lambda_1}{1 + \mu \lambda_1}, \dots, \frac{\lambda_n}{1 + \mu \lambda_n} \right) \\ &= \det(I_n + \mu \Lambda)^{-\alpha} \det(I_n + (I_n + \mu \Lambda)^{-1} \Lambda C)^{-\alpha} \\ &= \det(I_n + \Lambda(\mu I_n + C))^{-\alpha} \end{aligned}$$

□

On va faire apparaître cette construction de manière artificielle : on part d'une matrice $C > 0$, et alors il existe $\epsilon > 0$ tel que $C' = C - \epsilon I_n > 0$. On prend un vecteur $Y \sim \Gamma(k/2, C')$, et alors, par la proposition précédente, si X est un vecteur dont la loi conditionnelle sachant Y est $\gamma(k/2, Y_1/\epsilon) \otimes \dots \otimes \gamma(k/2, Y_n/\epsilon)$, alors ϵX suit la loi $\Gamma(k/2, C)$. L'intérêt de cette manipulation provient de la proposition :

Proposition 4.7. *Soit $Z : \Omega \rightarrow \mathbb{R}_+^n$ un vecteur aléatoire et $a \in]0, \infty[^n$. Posons X un vecteur dont la loi conditionnelle sachant Z est : $\gamma(a_1, Z_1) \otimes \dots \otimes \gamma(a_n, Z_n)$. Alors la densité de X est :*

$$f_a(x) = \mathbb{E} \left[\prod_{i=1}^n g_{a_i}(x_i, Z_i) \right],$$

où g_{a_i} est la densité des lois gamma en dimension 1 vue à la proposition 4.1.

Démonstration. On va de nouveau appliquer la formule de la proposition 4.5. Soit $\lambda \in \mathbb{R}_+^n$:

$$\begin{aligned} \mathbb{E}(e^{-\langle \lambda, X \rangle}) &= \mathbb{E} \left[\int e^{-\langle \lambda, x \rangle} d\gamma(a_1, Z_1) \otimes \dots \otimes d\gamma(a_n, Z_n)(x) \right] \\ &= \mathbb{E} \left[\int e^{-\langle \lambda, x \rangle} \prod_{i=1}^n g_{a_i}(x_i, Z_i) dx \right] \\ &= \int e^{-\langle \lambda, x \rangle} \mathbb{E} \left[\prod_{i=1}^n g_{a_i}(x_i, Z_i) \right] dx \text{ par le théorème de Fubini} \\ &= \int e^{-\langle \lambda, x \rangle} f_a(x) dx. \end{aligned}$$

□

Mais alors, par (1), on obtient les propriétés :

$$f_a(x) = \mathbb{E} \left[\prod_{i=1}^n g_{a_i}(x_i, Z_i) \right] \leq 2^n \prod_{i=1}^n x_i^{a_i-1} (1 + x_i),$$

et

$$\partial_i f_a = f_{a-e_i} - f_a.$$

On va maintenant pouvoir obtenir des informations précises sur la densité des lois gamma :

Proposition 4.8. *Soit $X \sim \Gamma(\alpha, C)$ avec $C > 0$ et $\alpha > 1$. Notons h_α la densité de X . On a alors, pour $S \subset [1, n]$:*

1. La dérivée $\frac{\partial^{|S|}}{\partial x_S} h_\alpha$ existe et appartient à $L^1(\mathbb{R}^n)$.
2. Ces dérivées vont être nulles en zéro : $\forall i \notin S, \frac{\partial^{|S|}}{\partial x_S} h_\alpha(x) \xrightarrow{x_i \rightarrow 0^+} 0$.
3. On peut alors calculer la transformée de Laplace suivante :

$$\mathcal{L}_{\frac{\partial^{|S|}}{\partial x_S} h_\alpha}(\lambda) = \det(I_n + \Lambda C)^{-\alpha} \prod_{i \in S} \lambda_i.$$

Démonstration. On reprend la construction de X sous la forme de loi conditionnelle vue précédemment. Alors, ϵX a pour densité f_a avec $a = (\alpha, \dots, \alpha)$. En itérant la propriété de dérivation de f_a , on obtient :

$$\frac{\partial^{|S|}}{\partial x_S} f_a = \sum_{A \subset S} (-1)^{|S|-|A|} f_{a-\sum_{j \in A} e_j}.$$

Donc $\frac{\partial^{|S|}}{\partial x_S} f_a$ est combinaison linéaire de densités de probabilités, donc est bien intégrable. Ensuite, comme $\alpha > 1$, la majoration obtenue sur f impose : $\forall i \notin S, \frac{\partial^{|S|}}{\partial x_S} f_a(x) \xrightarrow{x_i \rightarrow 0^+} 0$. Or, vu que ϵX a f_a comme densité, on vérifie que $h_\alpha(x) = f_a(\epsilon x) \epsilon^n$. Ainsi, h_α vérifie les points 1 et 2. Enfin, on détermine la transformée de Laplace de $\frac{\partial^{|S|}}{\partial x_S} h_\alpha$:

$$\begin{aligned} \mathcal{L}_{\frac{\partial^{|S|}}{\partial x_S} h_\alpha}(\lambda) &= \int_{\mathbb{R}_+^n} e^{-\langle \lambda, x \rangle} \frac{\partial^{|S|}}{\partial x_S} h_\alpha dx = \int_{\mathbb{R}_+^n} \prod_{i \in S} \lambda_i e^{-\langle \lambda, x \rangle} h_\alpha dx \\ &= \prod_{i \in S} \lambda_i \mathcal{L}_{h_\alpha}(\lambda) = \det(I_n + \Lambda C)^{-\alpha} \prod_{i \in S} \lambda_i \end{aligned}$$

La deuxième égalité s'obtient par intégrations par parties scécives, où les termes de bord sont nuls par le point 2. \square

5 Théorème d'interpolation

Pour toute matrice $M = (m_{i,j})_{i,j \in [1,n]}$ de $\mathcal{M}_n(\mathbb{R})$, et toute parties $S, S' \in [1, n]$, on note M_S la matrice carrée $M_S = (m_{i,j})_{i,j \in S}$ et $M_{S_1, S_2} = (m_{i,j})_{i \in S_1, j \in S_2}$

Théorème 5.1. *Pout $t \in [0, 1]$ on considère le vecteur gaussien $X(t) = (X_1(t), \dots, X_n(t))$ de loi $\mathcal{N}_n(0, C(t))$, où $C : [0, 1] \rightarrow M_n(\mathbb{R})$ est une application de classe \mathcal{C}^1 , tel que $\forall t, C(t) > 0$. On suppose que pour toute partie S de $[1, n]$ non vide l'application $t \mapsto \det(C(t)_S)$ est décroissante. Alors*

$$\mathbb{P}(\max_{1 \leq i \leq n} |X_i(1)| \leq 1) \geq \mathbb{P}(\max_{1 \leq i \leq n} |X_i(0)| \leq 1)$$

Démonstration. On considère le vecteur $Z(t) = (X_i(t)^2/2)_{1 \leq i \leq n}$. La densité de $Z(t)$, i.e. la densité associée à la loi gamma $\Gamma(1/2, C(t))$, est :

$$f(t, x) = \mathbb{1}_{]0, +\infty[^n}(x) \frac{\det(C(t))^{-1/2} (4\pi)^{-n/2}}{\prod_{1 \leq i \leq n} \sqrt{x_i}} \sum_{\varepsilon \in \{-1, 1\}^n} \exp(-\langle C(t)^{-1}(\varepsilon \cdot \sqrt{x}), (\varepsilon \cdot \sqrt{x}) \rangle)$$

où $\varepsilon \cdot \sqrt{x} = (\varepsilon_i \sqrt{x_i})$. On pose

$$\varphi(t) := \mathbb{P}(\max_{1 \leq i \leq n} |X_i(t)| \leq 1) = \mathbb{P}(Z(t) \in Q) = \int_Q f(t, x) dx$$

où $Q = [0, 1/2]^n$. Montrons que φ' est positive. On pose

$$\begin{aligned}\psi &: S^{n-1} \times [0, 1] \longrightarrow \mathbb{R} \\ (x, t) &\longmapsto \langle x, C(t)x \rangle\end{aligned}$$

Soient M et m , respectivement le maximum et le minimum de ψ sur $S^{n-1} \times [0, 1]$, ils sont bien définis et non nuls par continuité de ψ . On pose $\eta := \min(m, 1/M) > 0$.

On a alors $\eta I_n \leq C(t) \leq \eta^{-1} I_n$. D'autre part, l'application

$$t \longmapsto \det(C(t))^{-1/2}$$

est C^∞ , on note M_1 son maximum et M_2 le maximum de sa dérivée sur $[0, 1]$. Pour $\varepsilon \in \{-1, 1\}^n$ on pose M_ε le maximum de la dérivée de l'application

$$(x, t) \longmapsto \langle C(t)^{-1}(\varepsilon \cdot \sqrt{x}), (\varepsilon \cdot \sqrt{x}) \rangle$$

sur $S^{n-1} \times [0, 1]$.

On pose $g(x) := \sup_{t \in [0, 1]} \left| \frac{\partial}{\partial t} f(t, x) \right|$.

On a :

$$\begin{aligned}g(x) &\leq \mathbb{1}_{]0, +\infty[^n}(x) \frac{(4\pi)^{-n/2}}{\prod_{1 \leq i \leq n} \sqrt{x_i}} \left[M_2 \sum_{\varepsilon \in \{-1, 1\}^n} \exp(-\eta \langle \varepsilon \cdot \sqrt{x}, \varepsilon \cdot \sqrt{x} \rangle) \right. \\ &\quad \left. + M_1 \sum_{\varepsilon \in \{-1, 1\}^n} M_\varepsilon \|x\|^2 \exp(-\eta \langle \varepsilon \cdot \sqrt{x}, \varepsilon \cdot \sqrt{x} \rangle) \right] \\ &\leq \mathbb{1}_{]0, +\infty[^n}(x) \frac{(4\pi)^{-n/2}}{\prod_{1 \leq i \leq n} \sqrt{x_i}} \left[M_2 \sum_{\varepsilon \in \{-1, 1\}^n} \exp(-\eta \|x\|^2) \right. \\ &\quad \left. + M_1 \sum_{\varepsilon \in \{-1, 1\}^n} M_\varepsilon \|x\|^2 \exp(-\eta \|x\|^2) \right]\end{aligned}$$

Ainsi g est intégrable, on peut donc dériver sous le signe intégral, on a alors :

$$\varphi'(t) = \int_Q \frac{\partial}{\partial t} f(t, x) dx$$

On pose :

$$b : x \longmapsto \frac{\partial}{\partial t} f(t, x)$$

Le point crucial de la preuve va être de calculer b en regardant sa transformée de Laplace.

On a $\forall \lambda \in \mathbb{R}^n$:

$$\begin{aligned}
\mathcal{L}_b(\lambda) &= \int_{\mathbb{R}^n} e^{-\langle \lambda, x \rangle} \frac{\partial}{\partial t} f(t, x) dx \\
&= \frac{\partial}{\partial t} \int_{\mathbb{R}^n} e^{-\langle \lambda, x \rangle} f(t, x) dx \\
&= \frac{\partial}{\partial t} \mathbb{E}[e^{-\langle \lambda, Z(t) \rangle}] \\
&= \frac{\partial}{\partial t} (\det(I_n + \Lambda C(t))^{-1/2})
\end{aligned}$$

avec $\Lambda = \text{diag}(\lambda)$, or

$$\det(I_n + \Lambda C(t)) = \sum_{S \subset [1, n]} \det((\Lambda C(t))_S) = \sum_{S \subset [1, n]} \det((C(t))_S) \prod_{i \in S} \lambda_i.$$

Donc

$$\mathcal{L}_b(\lambda) = -\frac{1}{2} \det(I_n + \Lambda C(t))^{-3/2} \sum_{\emptyset \neq S \subset [1, n]} \frac{\partial}{\partial t} \det((C(t))_S) \prod_{i \in S} \lambda_i.$$

Or $\lambda \mapsto \det(I_n + \Lambda C(t))^{-3/2}$ est la transformation de Laplace d'une densité de probabilité h_t sur \mathbb{R}_+^n (h_t est la densité associée à $\Gamma(3/2, C(t))$), ainsi, par la proposition 4.8, \mathcal{L}_b est la transformée de Laplace de :

$$k_t := -\frac{1}{2} \sum_{\emptyset \neq S \subset [1, n]} \frac{d}{dt} \det(C(t)_S) \frac{\partial^S}{\partial x_S} h_t.$$

Par injectivité de la transformée de Laplace on obtient que $b = k_t$, par conséquent :

$$\varphi'(t) = \int_Q \frac{\partial}{\partial t} f(t, x) dx = -\frac{1}{2} \sum_{\emptyset \neq S \subset [1, n]} \frac{\partial}{\partial t} \det(C(t)_S) \int_Q \frac{\partial^S}{\partial x_S} h_t$$

Pour conclure il suffit de montrer que $\int_Q \frac{\partial^S}{\partial x_S} h_t \geq 0$, $\forall S \in [1, n]$. Or :

$$\int_Q \frac{\partial^S}{\partial x_S} h_t = \int_{[0, 1/2]^{S^c}} h_t^S$$

où on a posé h_t^S la restriction de h_t sur le sous espace affine $\{y \in \mathbb{R}^n; y_i = 1/2, \forall i \in S\}$. En effet, on commence par intégrer par rapport aux variables x_i pour $i \in S$, par exemple si $1 \in S$:

$$\begin{aligned}
\int_{[0, 1/2]} \frac{\partial^{|S|}}{\partial x_S} h_t &= \frac{\partial^{|S \setminus \{1\}|}}{\partial x_{S \setminus \{1\}}} h_t(1/2, x_2, \dots, x_n) - \lim_{x_1 \rightarrow 0^+} \frac{\partial^{|S \setminus \{1\}|}}{\partial x_{S \setminus \{1\}}} h_t(x_1, x_2, \dots, x_n) \\
&= \frac{\partial^{|S \setminus \{1\}|}}{\partial x_{S \setminus \{1\}}} h_t(1/2, x_2, \dots, x_n)
\end{aligned}$$

puisque $\lim_{x_1 \rightarrow 0^+} \frac{\partial^{|S \setminus \{1\}|}}{\partial x_{S \setminus \{1\}}} h_t(x_1, x_2, \dots, x_n) = 0$ par la proposition 4.8

□

6 Preuve de la conjecture de corrélation gaussienne

Lemme 6.1. Soit $n = n_1 + n_2$ Soit $A = \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix}$, tel que $A > 0$, et $A_{i,j} \in \mathcal{M}_{n_i, n_j}$. Alors $\det(A) = \det(A_{1,1}) \det(A_{2,2}) \det(I_n - A_{1,1}^{-1/2} A_{1,2} A_{2,2}^{-1} A_{2,1} A_{1,1}^{-1/2})$. On a de plus que $0 \leq A_{1,1}^{-1/2} A_{1,2} A_{2,2}^{-1} A_{2,1} A_{1,1}^{-1/2} \leq I_{n_1}$.

Démonstration. Il suffit de remarquer que

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix} = \begin{pmatrix} A_{1,1}^{1/2} & 0 \\ 0 & A_{2,2}^{1/2} \end{pmatrix} \begin{pmatrix} I_{n_1} & A_{1,1}^{-1/2} A_{1,2} A_{2,2}^{-1/2} \\ A_{2,2}^{-1/2} A_{2,1} A_{1,1}^{-1/2} & I_{n_2} \end{pmatrix} \begin{pmatrix} A_{1,1}^{1/2} & 0 \\ 0 & A_{2,2}^{1/2} \end{pmatrix}$$

et que si on pose $B = A_{1,1}^{-1/2} A_{1,2} A_{2,2}^{-1/2}$ alors

$$\begin{pmatrix} I_{n_1} & -B \\ 0 & I_{n_2} \end{pmatrix} \begin{pmatrix} I_{n_1} & B \\ B^T & I_{n_2} \end{pmatrix} \begin{pmatrix} I_{n_1} & 0 \\ -B^T & I_{n_2} \end{pmatrix} = \begin{pmatrix} I_{n_1} & -BB^T & 0 \\ 0 & I_{n_2} & \end{pmatrix}$$

La deuxième relation montre que $I_{n_1} - BB^T = I_{n_1} - A_{1,1}^{-1/2} A_{1,2} A_{2,2}^{-1} A_{2,1} A_{1,1}^{-1/2} \geq 0$. □

Soit X un vecteur gaussien centré sur \mathbb{R}^n , de loi $\mathcal{N}_n(0, C)$, $C > 0$ (d'après la remarque de 2.1). On note $N_1 = \{1, \dots, n_1\}$ et $N_2 = \{1 + n_1, \dots, n\}$. Soit $X(t)$ un vecteur gaussien de loi $\mathcal{N}_n(0, C(t))$, où

$$C(t) := \begin{pmatrix} C_{N_1} & tC_{N_1, N_2} \\ tC_{N_2, N_1} & C_{N_2} \end{pmatrix}$$

On commence par remarquer que $C(t) > 0$, $\forall t \in [0, 1]$. En effet, il est clair que $C(t)$ est symétrique, que $C(1) = C$ et $C(0)$ sont symétriques définies positives (pour $C(0)$ il suffit de voir que $C_{N_1}, C_{N_2} > 0$), de plus si on fixe $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ dans $\mathcal{M}_{n,1}(\mathbb{R})$, on a que

$$X^T C(t) X = X_1^T C_{N_1} X_1 + X_2^T C_{N_2} X_2 + t(X_1^T C_{N_1, N_2} X_2 + X_2^T C_{N_2, N_1} X_1)$$

donc l'application $t \mapsto X^T C(t) X$ est affine, comme $C(0)$ et $C(1)$ sont symétriques positifs, alors $C(t) > 0$, $\forall t \in [0, 1]$. Pour appliquer le théorème 5.1, on cherche à montrer que l'application $t \mapsto \det(C(t)_S)$ est décroissante sur $[0, 1]$ pour toute partie $S \subset [1, n]$. Soit $S \subset [1, n]$, on note $S_i = S \cap N_i$. Comme $C(t)_S > 0$, par le lemme précédent on a

$$\det(C(t)_S) = \det(C_{S_1}) \det(C_{S_2}) \det(I_{|S_1|} - t^2 C_{S_1}^{-1/2} C_{S_1, S_2} C_{S_2}^{-1} C_{S_2, S_1} C_{S_1}^{-1/2})$$

et $0 \leq C_{S_1}^{-1/2} C_{S_1, S_2} C_{S_2}^{-1} C_{S_2, S_1} C_{S_1}^{-1/2} \leq I_{|S_1|}$, ainsi les valeurs propres de cette matrice, qu'on note $(\lambda_i)_{1 \leq i \leq |S_1|}$ sont tous dans $[0, 1]$. Ainsi les valeurs propres de $I_{|S_1|} - t^2 C_{S_1}^{-1/2} C_{S_1, S_2} C_{S_2}^{-1} C_{S_2, S_1} C_{S_1}^{-1/2}$ sont $(1 - t^2 \lambda_i)_{1 \leq i \leq |S_1|}$. On a alors

$$\det(C(t)_S) = \det(C_{S_1}) \det(C_{S_2}) \prod_{i=1}^{|S_1|} (1 - t^2 \lambda_i)$$

Donc, l'application $t \mapsto \det(C(t)_S)$ est bien décroissante sur $[0, 1]$. Par le théorème 5.1 on déduit que

$$\mathbb{P}(\max_{1 \leq i \leq n} |X_i(1)| \leq 1) \geq \mathbb{P}(\max_{1 \leq i \leq n} |X_i(0)| \leq 1)$$

Or

$$\mathbb{P}(\max_{1 \leq i \leq n} |X_i(0)| \leq 1) = \mathbb{P}(\max_{1 \leq i \leq n_1} |X_i| \leq 1) \mathbb{P}(\max_{n_1 < i \leq n} |X_i| \leq 1)$$

puisque les deux blocs $1 \leq i \leq n_1$ et $n_1 + 1 \leq i \leq n$ sont décorrélés, et alors, par la proposition 1.3, ils sont indépendants. On déduit alors que

$$\mathbb{P}(\max_{1 \leq i \leq n} |X_i| \leq 1) \geq \mathbb{P}(\max_{1 \leq i \leq n_1} |X_i| \leq 1) \mathbb{P}(\max_{n_1 < i \leq n} |X_i| \leq 1)$$

Références

- [1] Franck Barthe. L'inégalité de corrélation gaussienne d'après Thomas Royen. *Séminaire Bourbaki*, 2016-2017(1124) :01–15, 2017.