Vecteurs gaussiens

24 novembre 2013 - v1 non relue

1 Variables aléatoires gaussiennes réelles

Définition. On appelle loi gaussienne centrée réduite la loi admettant la densité f définie par :

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

On note $\mathcal{N}(0,1)$ cette loi.

Proposition 1.1 *Soit G une v.a. suivant une loi gaussienne centrée réduite. Alors :*

1. $E(G) = 0$.
2. $\mathrm{var}(G) = 1$.
3. *Pour tout réel λ, on a $E(e^{i\lambda G}) = \exp(-\lambda^2/2)$.*

Définition. Soient $m \in \mathbb{R}$ et $\sigma \geq 0$. Soit G une v.a. suivant la loi $\mathcal{N}(0,1)$. On appelle loi gaussienne de paramètres m et σ^2 la loi de la v.a. $\sigma G + m$. On note $\mathcal{N}(m, \sigma^2)$ cette loi.

Remarque. Une v.a. suivant la loi $\mathcal{N}(m, 0)$ est constante égale à m.

Proposition 1.2 *Soit G une v.a. suivant une loi $\mathcal{N}(m, \sigma^2)$. Alors :*

1. $E(G) = m$.
2. $\mathrm{var}(G) = \sigma^2$.
3. *Pour tout réel λ, on a $E(e^{i\lambda G}) = \exp(im\lambda - \lambda^2\sigma^2/2)$.*

C'est une conséquence immédiate des résultats dans le cas centré réduit.

Remarque. Soient $m \in \mathbb{R}$ et $\sigma \geq 0$. La loi $\mathcal{N}(m, \sigma^2)$ admet une densité si et seulement si σ est non nul.

2 Intermède : variance d'un vecteur aléatoire

On munit $\mathbb{R}^d$ du produit scalaire canonique $\langle \cdot, \cdot \rangle$ et de la norme euclidienne correspondante $\| \cdot \|$. En identifiant un vecteur avec une matrice colonne, on peut écrire $\langle x, y \rangle = {}^t x y$. J'identifie de même application linéaire et matrice dans les bases canoniques. Par défaut, on considère dans cette section des vecteurs aléatoires de $\mathbb{R}^d$.

Définitions

- Soit $X = (X_1, \ldots, X_d)$ un vecteur aléatoire. On dit que X est de carré intégrable si $\|X\|^2$ est intégrable, i.e. si $X_1^2 + \cdots + X_d^2$ est intégrable, i.e. si chacun des X_i est de carré intégrable.
- Soit $X = (X_1, \ldots, X_d)$ un v.a. de carré intégrable.
  - On appelle matrice de (co-)variance de X et on note $\mathrm{var}(X)$ la matrice symétrique $d \times d$ dont les éléments sont les covariances $\mathrm{cov}(X_i, X_j), 1 \leq i, j \leq d$.
  - On appelle forme quadratique de variance la forme quadratique $q_X : \mathbb{R}^d \to \mathbb{R}$ définie par

$$q_X(\lambda) = {}^t\lambda \mathrm{var}(X)\lambda = \sum_{i,j} \lambda_i\lambda_j \mathrm{cov}(X_i, X_j) = \mathrm{var}(\sum_i \lambda_i X_i) = \mathrm{var}(\langle \lambda, X \rangle).$$

La proposition suivante est une conséquence simple de l'égalité $q_X(\lambda) = \mathrm{var}(\langle \lambda, X \rangle)$.

Proposition 2.1 *Soit X un vecteur aléatoire de carré intégrable de $\mathbb{R}^d$.*

1. *La matrice $\mathrm{var}(X)$ est positive (autrement dit, q_X est une application positive).*
2. *Soit A une matrice $p \times d$. Alors $Y = AX$ est un vecteur aléatoire de carré intégrable de $\mathbb{R}^p$ et on a :*
$$\mathrm{var}(Y) = A \, \mathrm{var}(X) \, {}^tA.$$

La proposition suivante est une formulation de la propriété classique suivante : si X et Y sont deux variables aléatoires indépendantes, alors elles sont décorrélées.

Proposition 2.2 *Soit X un vecteur aléatoire de carré intégrable. Si les coordonnées de X sont indépendantes, alors $\mathrm{var}(X)$ est diagonale.*

3 Vecteurs aléatoires gaussiens

Par défaut, on considère dans cette section des vecteurs aléatoires de $\mathbb{R}^d$.

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

Remarques.

- L'image d'un vecteur gaussien par une application affine est un vecteur gaussien.
- Chacune des marginales d'un vecteur gaussien est gaussienne. Par contre, le fait que les marginales d'un vecteur aléatoire X soient gaussiennes n'entraîne pas le caractère gaussien du vecteur aléatoire X.
- Chacune des marginales d'un vecteur gaussien est gaussienne et donc de carré intégrable. Par conséquent un vecteur aléatoire gaussien est de carré intégrable.

Proposition 3.1 *Soit q_X la forme quadratique de variance d'un vecteur gaussien X, m son vecteur espérance. La transformée de Fourier de X est donnée par :*

$$E(e^{i\langle \lambda, X \rangle}) = \exp(-\frac{1}{2}q_X(\lambda) + i\langle \lambda, m \rangle).$$

Cela se montre par exemple en notant que $E(e^{i\langle \lambda, X \rangle})$ est la transformée de Fourier en 1 de la variable aléatoire gaussienne $\langle \lambda, X \rangle$.

Remarque. En particulier, la donnée de la forme quadratique de variance (ou de la matrice de variance) et du vecteur espérance d'un vecteur gaussien suffit pour décrire complètement sa loi. Cela permet de poser la définition suivante.

Définition. On appelle loi gaussienne sur $\mathbb{R}^d$ de vecteur espérance m et de matrice de variance K la loi d'un vecteur gaussien de vecteur espérance m et de matrice de variance K. On note $\mathcal{N}(m, K)$ cette loi.

Définition. On appelle loi gaussienne centrée réduire la loi $\mathcal{N}(0, Id)$.

Une matrice de variance est nécessaire symétrique et positive. Le résultat suivant établit que c'est la seule restriction.

Théorème 3.1 *Soit K une matrice symétrique et positive de taille $d \times d$. Soit m un vecteur de $\mathbb{R}^d$. Alors il existe un vecteur aléatoire X gaussien de moyenne m et de matrice de variance K.*

Plan de la preuve. Soient $X_1, \ldots, X_d$ des v.a. réelles gaussiennes centrées réduites indépendantes. Le vecteur $X = (X_1, \ldots, X_d)$ est un vecteur gaussien. Cela se montre par exemple *via* les transformées de Fourier. Par ailleurs X est centré et de matrice de variance l'identité. Ainsi, X est un vecteur gaussien centré réduit.

Comme K est symétrique positive, il existe[1] une matrice A tel que $K = A\,{}^tA$. On pose alors :

$$G = AX + m.$$

Ce vecteur suit la loi $\mathcal{N}(m, K)$. $\square$

Remarque. La preuve montre que l'on peut obtenir un vecteur gaussien de loi donnée comme image d'un vecteur gaussien centré réduit par une application affine.

Remarque. Si K est inversible, alors la loi $\mathcal{N}(m, K)$ admet une densité de la forme $x \mapsto C \exp(-q(x))$ où C est une constante et où q est une forme quadratique définie positive. Ce résultat semble très peu utile.

Théorème 3.2
- *Pour que les marginales d'un vecteur gaussien soient indépendantes, il faut et il suffit que sa matrice de variance soit diagonale.*
- *Soit X un vecteur aléatoire de $\mathbb{R}^n$. Soit Y un vecteur aléatoire de $\mathbb{R}^p$. On suppose que (X, Y) est un vecteur gaussien de $\mathbb{R}^{n+p}$. Alors X et Y sont indépendants si et seulement si la matrice de variance de (X, Y) est une matrice composée de deux blocs diagonaux formés par les matrices de variances de X et Y.*

Le sens direct est valable sans l'hypothèse gaussienne (si deux variables aléatoires sont indépendantes alors elles sont décorrélées). Le sens réciproque peut par exemple s'établir *via* les transformées de Fourier.

Une conséquence de l'agréable résultat précédent est que, dans le cas des vecteurs gaussiens, le calcul des espérances conditionnelles est particulièrement simple.

1. Comme K est symétrique, on peut écrire $K = BD\,{}^tB$ où B est une matrice orthogonale et D est une matrice diagonale. Comme K est positive, tous les coefficients de D sont positifs. On peut donc écrire $D = C^2$ où C est une matrice diagonale. On a donc $K = BC^2\,{}^tB = BC\,{}^tC\,{}^tB = A\,{}^tA$ où $A = BC$.

Théorème 3.3 *Soit $(X_1, \dots, X_d, Y)$ un vecteur gaussien de $\mathbb{R}^{d+1}$. On pose $X = (X_1, \dots, X_d)$. Alors $E(Y|X)$ est la projection orthogonale*[2] *de Y sur l'espace vectoriel engendré par la constante 1 et les variables aléatoires $X_1, \dots, X_d$.*

Remarque. Sans l'hypothèse gaussienne, mais en supposant toujours les variables aléatoires de carré intégrable, on sait simplement que Y est la projection orthogonale sur l'espace vectoriel des variables aléatoires de carré intégrable et mesurables par rapport à la tribu engendrée par $X_1, \dots, X_d$.

Preuve. Notons Z la projection orthogonale de Y sur $\text{vect}(1, X_1, \dots, X_d)$. On a alors :

$$E(Y|X) = E(Z|X) + E(Y - Z|X).$$

Mais $E(Z|X) = Z$. Il reste à prouver $E(Y - Z|X) = 0$.

On a $E(Y - Z) = \langle Y - Z, 1\rangle = 0$. Par conséquent, pour tout indice k, on a $\text{cov}(Y - Z|X_k) = \langle Y - Z, X_k\rangle = 0$. Mais $(Y - Z, X_1, \dots, X_d)$ est un vecteur gaussien. Par conséquent, $Y - Z$ et X sont indépendants. On en déduit $E(Y - Z|X) = E(Y - Z) = 0$. □

4 Théorème de Cochran

Définition. La loi de Pearson à d degrés de liberté est la loi de la somme des carrés de d variables aléatoires gaussiennes réelles centrées réduites.

Théorème 4.1 *. Soit X un vecteur gaussien centré réduit de $\mathbb{R}^n$. Soit L un sous-espace de $\mathbb{R}^n$ de dimension d avec $1 \leq d \leq n - 1$. On désigne par Y et Z les projections de X respectivement sur L et $L^\perp$. Alors Y et Z sont indépendants. De plus, les variables aléatoires $\|Y\|^2$ et $\|Z\|^2$ suivent respectivement des lois de Pearson à d et $n - d$ degrés de liberté.*

Plan de la preuve. Si X est un vecteur gaussien centré réduit de $\mathbb{R}^n$ et si A est une matrice orthogonale de $\mathbb{R}^n$ alors AX est un vecteur gaussien centré réduit de $\mathbb{R}^n$. Cela permet de se ramener au cas où L est engendré par les d premiers vecteurs de la base canonique. L'espace vectoriel $L^\perp$ est alors engendré par les n_d derniers vecteurs de la base canonique. On a donc simplement $Y = (X_1, \dots, X_d, 0, \dots, 0)$ et $Z = (0, \dots, 0, X_{d+1}, \dots, X_n)$. Le résultat est alors immédiat. □

5 Théorème central limite

Théorème 5.1 *Soit $(X_n)_n$ une suite de vecteurs aléatoires i.i.d. de $\mathbb{R}^d$. On suppose que les X_n sont de carré intégrable. On note m le vecteur espérance et K la matrice de variance des X_n. On a alors la convergence en loi, lorsque n tend vers l'infini, de :*

$$\frac{X_1 + \cdots + X_n - nm}{\sqrt{n}}$$

vers la loi $\mathcal{N}(0, K)$.

On peut par exemple déduire ce résultat de sa version en dimension 1 *via* les transformées de Fourier.

2. On se place dans l'espace de Hilbert des variables aléatoires de carré intégrable. Le produit scalaire de deux variables aléatoires X, Y est $\langle X, Y\rangle = E(XY)$.