

VECTEURS GAUSSIENS

Préparation à l'Agrégation Bordeaux 1

Année 2012 - 2013

Jean-Jacques Ruch

Table des Matières

Chapitre I. Vecteurs aléatoires gaussiens	5
1. Vecteurs aléatoires	5
2. Vecteurs aléatoires gaussiens	7
3. Théorème de Cochran	10

CHAPITRE I

Vecteurs aléatoires gaussiens

Les vecteurs gaussiens sont associés aux lois gaussiennes multivariées, et de ce fait jouent un rôle important en probabilités et en statistique. Ils apparaissent naturellement comme des objets limites et serviront en particulier dans le prochain chapitre :

soit (Y_n) une suite de vecteurs aléatoires à valeurs dans \mathbb{R}^k , dont les composantes sont de carré intégrable, indépendants et de même loi avec $\mathbb{E}(Y_1) = M$ et de matrice de covariance Γ ; alors on a

$$\sqrt{n} \left(\frac{Y_1 + \dots + Y_n}{n} - M \right) \xrightarrow{\mathcal{L}} \mathcal{N}_k(0, \Gamma).$$

1. Vecteurs aléatoires

Soit $X = \begin{pmatrix} X_1 \\ \vdots \\ X_d \end{pmatrix} \in \mathbb{R}^d$ un vecteur aléatoire tel que $\mathbb{E}(X_j^2) < +\infty$. Pour tout $1 \leq i, j \leq d$, la *covariance* entre X_i et X_j est donnée par

$$\text{Cov}(X_i, X_j) = \mathbb{E}[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])] = \mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\mathbb{E}[X_j].$$

Avec l'hypothèse faite ci-dessus cette covariance est toujours bien définie car, par l'inégalité de Cauchy-Schwarz,

$$|\text{Cov}(X_i, X_j)|^2 \leq \text{Var}(X_i)\text{Var}(X_j).$$

Il est facile de voir que l'on a $\text{Cov}(X_i, X_i) = \text{Var}(X_i)$. De plus, si les variables X_i et X_j sont indépendantes on a $\text{Cov}(X_i, X_j) = 0$. En général, la réciproque est fautive, sauf pour des vecteurs gaussiens, comme nous le verrons plus loin.

Définition 1. Soit X un vecteur (colonne) aléatoire de \mathbb{R}^d dont les composantes sont de carré intégrable. Le vecteur moyenne de X est défini par :

$$\mathbb{E}(X) = \begin{pmatrix} \mathbb{E}(X_1) \\ \vdots \\ \mathbb{E}(X_d) \end{pmatrix}$$

et sa matrice de covariance par :

$$\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])(X - \mathbb{E}[X])^t] = \begin{pmatrix} \text{Var}(X_1) & \text{Cov}(X_1, X_2) & \cdots & \text{Cov}(X_1, X_d) \\ \text{Cov}(X_2, X_1) & \text{Var}(X_2) & \cdots & \text{Cov}(X_2, X_d) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_d, X_1) & \text{Cov}(X_d, X_2) & \cdots & \text{Var}(X_d) \end{pmatrix}$$

On dira que X est centré si $\mathbb{E}[X] = 0$. Si les composantes de X sont indépendantes la matrice de covariance de X est diagonale. Comme avant la réciproque est fautive en général sauf pour les vecteurs gaussiens.

Avant d'énoncer le résultat suivant rappelons quelques propriétés des matrices symétriques semi-définies positives.

Une matrice symétrique réelle M est semi-définie positive si et seulement si elle vérifie l'une des deux propriétés équivalentes suivantes :

1. Pour tout vecteur colonne V on a $V^t M V \geq 0$
2. Toutes les valeurs propres de M sont positives ou nulles i.e. $\text{Sp}(M) \subset [0, +\infty[$

D'autre part, si M_1 et M_2 sont deux matrices symétriques semi-définies positives de même dimension, alors pour tout $a \geq 0$ et $b \geq 0$, $aM_1 + bM_2$ est encore une matrice symétrique semi-définie positive. Le théorème spectral entraîne que si M est symétrique semi-définie positive, alors il existe une matrice orthogonale P et une matrice diagonale $D = \text{Diag}(\lambda_1, \dots, \lambda_d)$ (positive) telle que $M = PDP^t$. On en déduit qu'il existe une matrice A (pas unique en général), telle que $M = AA^t$; par exemple $A = P\text{Diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_d})$.

Théorème 2.

|| La matrice de covariance est une matrice symétrique semi-définie positive.

DÉMONSTRATION. Par construction la matrice de covariance est symétrique. Elle s'écrit comme le produit d'un vecteur et de sa transposée (l'espérance s'applique ensuite à chaque composante et ne change donc pas la symétrie). Pour le deuxième point il suffit de remarquer que si M est la matrice de covariance du vecteur aléatoire X et si V est un vecteur constant de \mathbb{R}^d , alors

$$V^t M V = \text{Var}(v_1 X_1 + v_2 X_2 + \dots + v_d X_d) \geq 0.$$

□

Théorème 3.

|| Si X est un vecteur (colonne) aléatoire de \mathbb{R}^p de vecteur moyenne m et de matrice de covariance Γ . Alors si A est une matrice réelle $q \times p$, le vecteur aléatoire AX de \mathbb{R}^q a pour vecteur moyenne Am et pour matrice de covariance $A\Gamma A^t$.

DÉMONSTRATION. C'est une simple conséquence de la linéarité de l'espérance. Pour la moyenne on a :

$$\mathbb{E}[AX] = A\mathbb{E}[X] = Am$$

et pour la matrice de covariance :

$$\mathbb{E}[(AX - Am)(AX - Am)^t] = \mathbb{E}[A(X - m)(X - m)^t A^t] = A\mathbb{E}[(X - m)(X - m)^t] A^t = A\Gamma A^t.$$

□

Théorème 4.

|| Toute matrice symétrique semi-définie positive Γ de dimension $d \times d$ est la matrice de covariance d'un vecteur aléatoire de \mathbb{R}^d .

DÉMONSTRATION. Soit A une racine carrée matricielle de Γ . On note X un vecteur aléatoire de \mathbb{R}^d dont les composantes sont indépendantes, centrées et de variance 1. L'espérance de X est donc le vecteur nul, et la matrice de covariance est égale à Id_d . Le vecteur aléatoire AX est alors centré de matrice de covariance $A\text{Id}_d A^t = AA^t = \Gamma$. □

Définition 5. Soit X un vecteur aléatoire de \mathbb{R}^d centré et de matrice de covariance Γ . On considère (X_1, \dots, X_n) un n -échantillon issu de X . La matrice de covariance empirique Γ_n est définie par

$$\Gamma_n = \frac{1}{n} (X_1 X_1^t + \dots + X_n X_n^t)$$

Proposition 6. La matrice de covariance empirique associée à vecteur aléatoire X de \mathbb{R}^d est une matrice symétrique semi-définie positive. C'est un estimateur sans biais et fortement consistant de Γ .

DÉMONSTRATION. La forme de la matrice Γ_n montre que c'est une matrice symétrique semi-définie positive. De plus comme $\mathbb{E}(X_i X_i^t) = \Gamma$ on obtient que $\mathbb{E}(\Gamma_n) = \Gamma$. Enfin la loi forte des grands nombres entraîne que Γ_n converge presque sûrement vers Γ . \square

Définition 7. Soit X un vecteur aléatoire de \mathbb{R}^d . On définit sa fonction caractéristique $\Phi_X : \mathbb{R}^d \rightarrow \mathbb{C}$ par :

$$\forall u \in \mathbb{R}^d \quad \Phi_X(u) = \mathbb{E}[e^{iu^t X}] = \mathbb{E}[e^{i(u_1 X_1 + \dots + u_d X_d)}]$$

2. Vecteurs aléatoires gaussiens

Définition 8. Un vecteur aléatoire de \mathbb{R}^d est un vecteur aléatoire gaussien si et seulement si toute combinaison linéaire de ses composantes est une variable aléatoire réelle gaussienne, i.e. :

$$\forall a \in \mathbb{R}^d, \quad a^t X \stackrel{\mathcal{L}}{\sim} \mathcal{N}(m, \sigma^2)$$

En particulier tout vecteur gaussien de dimension 1 est une variable aléatoire gaussienne réelle, éventuellement dégénérée. D'autre part, toutes les composantes de X sont aussi des variables aléatoires gaussiennes réelles.

Exemple : soit X un vecteur aléatoire de \mathbb{R}^d dont les composantes X_i sont indépendantes et de loi $\mathcal{N}(m_i, \sigma_i^2)$. Le vecteur X est alors un vecteur gaussien. En effet toute combinaison linéaire de ses composantes s'écrit $a_1 X_1 + \dots + a_d X_d$, dont la loi est par indépendance des variables

$$\mathcal{N}(a_1 m_1 + \dots + a_d m_d, a_1^2 \sigma_1^2 + \dots + a_d^2 \sigma_d^2).$$

Théorème 9.

Soit X un vecteur aléatoire de \mathbb{R}^d de moyenne m et de matrice de covariance Γ . Les assertions suivantes sont équivalentes :

1. Le vecteur X est un vecteur gaussien.
2. La fonction caractéristique de X est donnée par : pour tout $u \in \mathbb{R}^d$

$$\Phi_X(u) = \mathbb{E}[\exp(iu^t X)] = \exp\left(iu^t m - \frac{1}{2}u^t \Gamma u\right)$$

3. Le vecteur X a même loi que le vecteur $m + AZ$, où Z est un vecteur aléatoire de \mathbb{R}^d à composantes indépendantes et de loi $\mathcal{N}(0, 1)$ et A est une matrice racine carrée de Γ .

DÉMONSTRATION. La première équivalence provient de la définition d'un vecteur aléatoire gaussien et de celle de la fonction caractéristique d'une variable aléatoire gaussienne réelle. Pour l'équivalence entre 1 et 3 cela provient simplement du théorème 3 et des propriétés des matrices symétriques semi-définies positives. \square

La loi d'un vecteur gaussien est caractérisée par son vecteur moyenne m et sa matrice de covariance Γ . Elle est appelée loi gaussienne sur \mathbb{R}^d et est notée $\mathcal{N}(m, \Gamma)$. La loi $\mathcal{N}(0, I_d)$ est appelée loi gaussienne standard sur \mathbb{R}^d . Un vecteur gaussien ayant cette loi est appelé vecteur aléatoire gaussien standard.

Proposition 10. *Toute matrice symétrique semi-définie positive Γ de dimension $d \times d$ est la matrice de covariance d'un vecteur aléatoire gaussien de \mathbb{R}^d .*

DÉMONSTRATION. Soit Z le vecteur gaussien de \mathbb{R}^d dont toutes les composantes sont indépendantes et de loi $\mathcal{N}(0, 1)$. C'est un vecteur aléatoire gaussien standard. D'après le théorème 3 et le 3 du théorème 9, AZ est un vecteur gaussien de \mathbb{R}^d de loi $\mathcal{N}(0, \Gamma)$. \square

Attention, les composantes d'un vecteur gaussien sont gaussiennes mais la réciproque est fautive. En effet, soit $X = (Y, \varepsilon Y)$ un vecteur aléatoire de \mathbb{R}^2 tel que Y et ε sont deux variables aléatoires réelles indépendantes avec $Y \sim \mathcal{N}(0, 1)$ et ε suit une loi de Rademacher c'est-à-dire $\mathbb{P}(\varepsilon = 1) = \mathbb{P}(\varepsilon = -1) = 1/2$. Il est facile de voir que Y et εY sont deux variables aléatoires gaussiennes, mais la combinaison linéaire $Y + \varepsilon Y$ ne l'est pas car $\mathbb{P}(Y + \varepsilon Y = 0) = \mathbb{P}(\varepsilon = -1) = 1/2$. De plus, $\text{Cov}(Y, \varepsilon Y) = \mathbb{E}[Y^2]\mathbb{E}[\varepsilon] = 0$, mais Y et εY ne sont pas indépendantes.

Le théorème suivant caractérise les lois gaussiennes de matrice de covariance inversible.

Théorème 11.

La loi gaussienne $\mathcal{N}(m, \Gamma)$ sur \mathbb{R}^d admet une densité de probabilité par rapport à la mesure de Lebesgue de \mathbb{R}^d si et seulement si Γ est inversible. Dans ce cas la densité de f est donnée par : pour tout $x \in \mathbb{R}^d$

$$f(x) = \frac{1}{\sqrt{(2\pi)^d \det(\Gamma)}} \exp\left(-\frac{1}{2}(x - m)^t \Gamma^{-1}(x - m)\right)$$

DÉMONSTRATION. La loi gaussienne $\mathcal{N}(m, \Gamma)$ peut être vue comme la loi du vecteur aléatoire gaussien $m + AZ$, où Z est un vecteur gaussien de \mathbb{R}^d dont toutes les composantes sont indépendantes et de loi $\mathcal{N}(0, 1)$. Par conséquent, Z a pour densité

$$f_Z(z) = \prod_{i=1}^d \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z_i^2\right) = \frac{1}{\sqrt{(2\pi)^d}} \exp\left(-\frac{1}{2}\|Z\|^2\right).$$

Si $X \sim \mathcal{N}(m, \Gamma)$ alors pour toute fonction continue bornée $h : \mathbb{R}^d \rightarrow \mathbb{R}$,

$$\mathbb{E}[h(X)] = \mathbb{E}[h(AZ + m)] = \int_{\mathbb{R}^d} h(Az + m) f_Z(z) dz.$$

La décomposition $\Gamma = AA^t$ entraîne que $|\det(A)| = \sqrt{\det(\Gamma)}$. De plus Γ est inversible si et seulement si A l'est et alors $\Gamma^{-1} = (A^{-1})^t A^{-1}$.

D'autre part, le changement de variable affine $x = Az + m$ est un difféomorphisme de \mathbb{R}^d dans lui-même si et seulement si A est inversible. Son jacobien est alors égal à $\det(A^{-1})$. On en déduit que

$$\mathbb{E}[h(X)] = \frac{1}{\sqrt{(2\pi)^d \det(\Gamma)}} \int_{\mathbb{R}^d} h(x) \exp\left(-\frac{1}{2}(x - m)^t \Gamma^{-1}(x - m)\right) dx$$

d'où la formule annoncée pour la densité de f . \square

Le théorème suivant caractérise les vecteurs gaussiens à matrice de covariance diagonale.

Théorème 12.

- Pour tout vecteur gaussien X de \mathbb{R}^d , les trois propriétés suivantes sont équivalentes :
1. Les composantes X_1, \dots, X_d sont mutuellement indépendantes.
 2. Les composantes X_1, \dots, X_d sont deux à deux indépendantes.
 3. La matrice de covariance Γ de X est diagonale.

En particulier, un vecteur gaussien est à composantes indépendantes si et seulement si pour tout $i \neq j$, $\text{Cov}(X_i, X_j) = 0$. Par exemple, un vecteur gaussien standard a toujours des composantes indépendantes et de loi $\mathcal{N}(0, 1)$ sur \mathbb{R} .

DÉMONSTRATION. Les implications 1. \Rightarrow 2. \Rightarrow 3. sont évidentes. Montrons que 3. \Rightarrow 1.. Si on a $\Gamma = \text{Diag}(\sigma_1^2, \dots, \sigma_n^2)$, alors pour tout $u \in \mathbb{R}^d$

$$\begin{aligned} \Phi_X(u) &= \exp\left(iu^t m - \frac{1}{2}u^t \Gamma u\right) = \exp\left(i \sum_{j=1}^d u_j m_j - \frac{1}{2} \sum_{j=1}^d \sigma_j^2 u_j^2\right) \\ &= \prod_{j=1}^d \exp\left(iu_j m_j - \frac{1}{2} \sigma_j^2 u_j^2\right) = \prod_{j=1}^d \Phi_{X_j}(u_j) \end{aligned}$$

□

Le théorème suivant montre que, pour un vecteur gaussien, la loi conditionnelle d'un bloc de composantes par rapport à un autre bloc de composantes disjoint s'obtient par un calcul d'algèbre matricielle.

Théorème 13.

Soit $Z = (X, Y) : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow \mathbb{R}^d$ un vecteur gaussien de $\mathbb{R}^d = \mathbb{R}^p \times \mathbb{R}^q$. On peut décomposer la matrice de covariance Γ de Z en quatre blocs :

$$\Gamma = \begin{pmatrix} \text{Cov}(X, X) & \text{Cov}(X, Y) \\ \text{Cov}(Y, X) & \text{Cov}(Y, Y) \end{pmatrix} = \begin{pmatrix} K & C \\ C^t & L \end{pmatrix}$$

Si Γ est inversible, alors K et L le sont. De plus, on a les propriétés suivantes :

1. Les vecteurs aléatoires X et Y sont des vecteurs gaussiens de \mathbb{R}^p et \mathbb{R}^q , de matrices de covariance respectives K et L . Ils sont indépendants si et seulement si $C = 0$.
2. Si K est inversible, alors

$$\mathbb{E}[Y | X] = \mathbb{E}[Y] + C^t K^{-1}(X - \mathbb{E}[X]).$$

De plus, $Y - \mathbb{E}[Y | X]$ est un vecteur gaussien de \mathbb{R}^q , centré, indépendant de X , et de matrice de covariance $L - C^t K^{-1} C$.

DÉMONSTRATION. La propriété 1. s'établit comme le théorème 12. Pour le deuxième point on pose

$$V = Y - \mathbb{E}[Y] - C^t K^{-1}(X - \mathbb{E}[X]).$$

On peut remarquer que

$$\begin{pmatrix} X \\ V \end{pmatrix} = \begin{pmatrix} \mathbb{E}[X] \\ 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ -C^t K^{-1} & 1 \end{pmatrix} \begin{pmatrix} X - \mathbb{E}[X] \\ Y - \mathbb{E}[Y] \end{pmatrix}$$

Comme $\begin{pmatrix} X - \mathbb{E}[X] \\ Y - \mathbb{E}[Y] \end{pmatrix}$ est un vecteur gaussien centré, $\begin{pmatrix} X \\ V \end{pmatrix}$ est un vecteur gaussien de moyenne $\begin{pmatrix} \mathbb{E}[X] \\ 0 \end{pmatrix}$ et de matrice de covariance

$$\begin{pmatrix} 1 & 0 \\ -C^t K^{-1} & 1 \end{pmatrix} \begin{pmatrix} K & C \\ C^t & L \end{pmatrix} \begin{pmatrix} 1 & -K^{-1}C \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} K & 0 \\ 0 & L - C^t K^{-1}C \end{pmatrix}$$

Donc V est un vecteur gaussien de \mathbb{R}^q , centré, indépendant de X , et de matrice de covariance $L - C^t K^{-1}C$. De plus on a

$$\begin{aligned} \mathbb{E}[Y|X] &= \mathbb{E}[V + \mathbb{E}[Y] + C^t K^{-1}(X - \mathbb{E}[X])|X] = \mathbb{E}[V|X] + \mathbb{E}[Y] + C^t K^{-1}(X - \mathbb{E}[X]) \\ &= \mathbb{E}[V] + \mathbb{E}[Y] + C^t K^{-1}(X - \mathbb{E}[X]) = \mathbb{E}[Y] + C^t K^{-1}(X - \mathbb{E}[X]) \end{aligned}$$

□

3. Théorème de Cochran

C'est un analogue du théorème de Pythagore pour les vecteurs gaussiens.

Théorème 14.

Théorème de Cochran

Soit X un vecteur aléatoire de \mathbb{R}^n de loi $\mathcal{N}(m, \sigma^2 I_n)$ avec $\sigma^2 > 0$. Soit $\mathbb{R}^n = E_1 \oplus E_2 \oplus \dots \oplus E_p$ une décomposition de \mathbb{R}^n en somme directe de p sous-espaces vectoriels orthogonaux de dimensions respectives d_1, \dots, d_p , avec $d_1 + \dots + d_p = n$. Soit P_k la matrice du projecteur orthogonal sur E_k et $Y_k = P_k X$ la projection orthogonale de X sur E_k .

1. Les projections Y_1, \dots, Y_p sont des vecteurs gaussiens indépendants et $Y_k \sim \mathcal{N}(P_k m, \sigma^2 P_k)$.
2. Les variables aléatoires $\|Y_1 - P_1 m\|^2, \dots, \|Y_p - P_p m\|^2$ sont indépendantes et

$$\|Y_k - P_k m\|^2 / \sigma^2 \sim \chi^2(d_k).$$

DÉMONSTRATION. Par translation on se ramène au cas où $m = 0$. On a

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_p \end{pmatrix} = \begin{pmatrix} P_1 \\ \vdots \\ P_p \end{pmatrix} X = AX$$

Par conséquent la loi de Y est $\mathcal{N}(0, \sigma^2 AA^t)$. Or pour tout $1 \leq i \leq p$, on a $P_i = P_i^t = P_i^2$. De plus, $P_i P_j = 0$ si $1 \leq i \neq j \leq p$ car $E_i \perp E_j$. Par conséquent, $AA^t = \text{Diag}(P_1, \dots, P_p)$ est diagonale par blocs. On en déduit que Y_1, \dots, Y_p sont des vecteurs gaussiens indépendants avec $Y_k \sim \mathcal{N}(0, \sigma^2 P_k)$ pour tout $1 \leq k \leq p$. En particulier les variables aléatoires $\|Y_1 - P_1 m\|^2, \dots, \|Y_p - P_p m\|^2$ sont indépendantes. Il reste à déterminer leur loi. Cela se fait en prenant une base orthonormée $\{e_{k,1}, \dots, e_{k,d_k}\}$ de chaque E_k . Alors on a $Y_k = y_{k,1}e_{k,1} + \dots + y_{k,d_k}e_{k,d_k}$ où les $y_{k,i} = \langle X, e_{k,i} \rangle$ sont des variables aléatoires indépendantes de loi $\mathcal{N}(0, \sigma^2)$. ON obtient alors le résultat. □

Ce théorème a de nombreuses des applications, dont en particulier le corollaire suivant.

Corollaire 15. Soit (X_1, \dots, X_n) un échantillon de loi $\mathcal{N}(m, \sigma^2)$ avec σ^2 . On définit :

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{et} \quad \bar{S}_n = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2.$$

Alors, les variables aléatoires sont indépendantes de loi $\mathcal{N}(m, \sigma^2/n)$ et $\chi^2(n-1)$.