UNIVERSITÉ MOHAMMED SEDDIK BENYAHIA

JIJEL

FACULTÉ DE SCIENCES EXACTES ET D'INFORMATIQUE

MEMOIRE DE MASTER

Présentée pour l'obtention du diplôme de :

MASTER

En **Mathématiques**

Option : Probabilités et Statistique

Par :

BERKANE Rayene

Thème

Les lois normales multidimentionnelles

Soutenue le : 03/ 07/ 2022 , devant le jury composé de :

Dr.	Cheraitia Hassen	M.C.A	Président
Dr.	Roula Amel	M.A.B	Encadreur
Dr.	Djeridi Zohra	M.C.B	Co-encadreur
Dr.	Aroud Chems-Edinne	M.C.B	Examinateur

Remerciements

Après avoir rendu grâce à ALLAH le tout puissant et le Miséricordieux.

J'adresse mes plus vifs remerciements à ma promotrice, ***Dr. Roula Amel*** *pour avoir m'encadrer, pour sa patience, son aide, ses conseils et encouragements,pour le temps qu'elle m'a consacré pour réalisé ce modeste travail.*

Je remercie, chaleureusement, ***Dr. Djeridi Zohra****, de m'avoir donné la force, la patience et la volonté pour accomplir ce travail.*

J'exprime mes remerciements aux honorables membres du jury, ***Dr.Cheraitia Hassen*** *pour m'avoir fait l'honneur de présider le jury et* ***Dr. Aroud Chems-Edinne*** *d'avoir accordé le temps et la patience pour l'évaluer cette étude.*

J'exprime ma profonde gratitude à tous mes enseignant(e)s du département de mathématiques, Surtout ***Dr. Laoudj Farida*** *pour ses lucides conseils et sa précieuse assistance.*

Sans oublier de remercier touts ceux qui ont, de prés ou de loin, contribué à la réalisation de ce travail.

Dédicace

À ma très chère mère
Quoi que je fasse ou que je dise, je ne saurai point te remercier comme il se doit. Ton affection me couvre, ta bienveillance me guide et ta présence à mes cotés a toujours été ma source de force pour affronter les différents obstacles.

À mon très cher père
Tu as toujours été à mes cotés pour me soutenir et m'encourager. Que ce travail traduit ma gratitude et mon affection.

À mon très cher frère ***Aimen*** *et mes belles sœurs* ***Kawthar Hayam*** *et* ***Maria Nour Alhouda****.*

À mon grand-père et ma grand-mère, qu'ALLAH leur donne une longue et joyeuse vie.

À tous mes amis, et toute personne qui occupe une place dans mon cœur. Qu'ALLAH vous donne la santé, le bonheur, le courage et surtout la réussite.

À Berkane Rayene

Table des matières

Résumé

Les lois Normales réelles se généralisent à plusieurs dimensions. Cette démarche n'est pas gratuite, elle correspond à une réalité. En effet, sur chaque unité extraite d'une population, nous effectuons en général non pas une mesure, mais plusieurs mesures de différentes caractéristiques. Dans ce mémoire, nous présentons un aperçu sur les lois normales multidimensionnelles. Ces lois, jouent un rôle fondamental en probabilités et statistique multivariée et elles constituent des modèles fréquemment utilisés dans divers domaines.

Introduction Générale

Les lois normales multivariées jouent un rôle central dans la théorie de l'analyse statistique multivariée. Ces lois servent à modéliser et décrire des phénomènes économiques, physiques, biologiques, ...etc où "big data" apparaissent souvent. Nous trouvons que les observations multivariées sont, au moins, approximativement normalement distribuées, lors du travail avec plusieurs variables aléatoires corrélées.

Même si les données d'origine ne suivent pas une loi normale multivariée, la loi d'échantillonnage de certaines statistiques peut être rapprochée par une loi normale, en raison du théorème central limite. Il est donc nécessaire d'étudier ces modèles probabilistes.

La plupart des procédures d'inférence existantes, pour l'analyse des données à valeurs vectorielles, ont été développées sous l'hypothèse de normalité. Dans les problèmes de modèles linéaires, tels que l'analyse de la variance et l'analyse de régression, le vecteur d'erreur est souvent supposé être normalement distribué afin que l'analyse statistique puisse être effectuée à l'aide des lois dérivées de la loi normale. Mais le vecteur aléatoire dont certain états est généralisé au concept des matrices aléatoires grâce à l'augmentation de nombre d'échantillons.

Les matrices aléatoires sont apparues pour la première fois dans les statistiques multivariées, au court des années 1930, avec les travaux de Wishart [14], Hsu et d'autres. Elles ont connu un énorme boom dans les années 1950 et 1960 grâce aux importantes contributions de Dyson, Goldin, Mehta et Wigner [4]. Dans les années 1990 et au-delà, il y a eu une résurgence de la théorie des matrices aléatoires, en raison du développement rapide de la théorie des codes de faible dimension. À ce jour, les matrices aléatoires sont toujours étudiées et utilisées dans divers problèmes.

L'observation simultanément de plusieurs caractéristiques d'un même phénomène sont représentées sous la forme des vecteurs ou des matrices dont tous ses éléments sont des variables aléatoires réelles. Alors une question naturelle à se poser est : quelle est la meilleure généralisation de la loi normale multivariée aux matrices ? Pour traiter la loi matricielle, nous devons introduire deux opérateurs mathématiques utiles : le produit de Kronecker et l'opérateur Vec.

Notre mémoire comporte deux chapitres ;

Le premier chapitre est dédié à la présentation des notions et définitions portant sur les vecteurs aléatoires. Une partie de ce chapitre sera réservée à la définition de la loi normale vectorielle, ses lois marginales et ses différentes propriétés. Pour illustrer la structure de cette loi, un exemple de loi normale bivariée sera donné.

Dans le deuxième chapitre, nous donnerons les concepts de base des matrices aléatoires. Nous présenterons entre autres la relation entre la loi normale vectorielle et la loi normale matricielle et nous étudierons les différentes propriétés de la loi normale matricielle et ses lois marginales.

Chapitre 1

Loi normale vectorielle

Dans le premier chapitre, nous introduisons les notions fondamentales et nécessaires pour la bonne compréhension de ce manuscrit. Ce chapitre est constitué de trois parties : Dans la première partie, nous rappelons les principales définitions et propriétés des vecteurs aléatoires. Dans la deuxième partie, nous définissons les concepts liées aux vecteurs aléatoires gaussiens. La troisième partie est consacrée à la loi normale vectorielle et ses différentes propriétés.

1.1 Vecteurs aléatoires

Un vecteur aléatoire est une application d'un espace probabilisé dans un espace vectoriel réel, en général $\mathbb{R}^d$,muni de sa tribu borélienne.
En pratique $\mathbb{R}^d$ est muni de sa base canonique et son produit scalaire euclidien.On donne dans cette section les différentes caractéristiques statistiques des vecteurs aléatoires (Voir [17] et [5]).

Définition 1.1.1.
Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et $X_1, X_2, ..., X_d$, d-variables aléatoires définies sur Ω à valeurs dans $\mathbb{R}$. Soit l'application X de Ω dans $\mathbb{R}^d$ définie par :

$$\begin{array}{rcl} X:\Omega & \to & \mathbb{R}^d \\ \omega & \mapsto & X(\omega)=\begin{pmatrix} X_1(\omega) \\ \vdots \\ X_d(\omega) \end{pmatrix}. \end{array}$$

On dit que X est un vecteur aléatoire (une variable aléatoire multivariée où multidimensionnelle) dans $\mathbb{R}^d$, on le note simplement par :

$$X = \begin{pmatrix} X_1 \\ \vdots \\ X_d \end{pmatrix},$$

où chaque composante $X_i, (i = 1, ..., d)$ est une variable aléatoire réelle.

A tout vecteur aléatoire,on associe sa fonction de répartition et sa densité de probabilité :

1.1.1 Fonction de répartition conjointe

Définition 1.1.2. *[5]*
On appelle fonction de répartition du vecteur aléatoire $X = (X_1, ..., X_d)'$ l'application F_X définie sur $\mathbb{R}^d$ et à valeurs dans $[0, 1]$ par :

$$\begin{aligned} F_X(x_1, ..., x_d) &= P_X\left(X \in \prod_{1 \leq i \leq d}]-\infty, x_i]\right) \\ &= P_X\left(X_1 \leq x_1, \ldots, X_d \leq x_d\right) \end{aligned}$$

Proposition 1.1.1. *Pour $i = 1, ..., d$, on a*

$$\lim_{x_i \to -\infty} F_X(x_1, ..., x_d) = 0 \quad et \quad \lim_{x_i \to +\infty} F_X(x_1, ..., x_d) = 1.$$

1.1.2 Fonction de densité

Définition 1.1.3. *[5]*
On dit que le vecteur aléatoire $X = (X_1, ..., X_d)'$ est absolument continu si il existe une fonction mesurable

$$f_X : (\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d}) \to (\mathbb{R}^+, \mathcal{B}_{\mathbb{R}^+})$$

telle que, pour tout $x = (x_1, ..., x_d)'$ dans $\mathbb{R}^d$, on ait :

$$F_X(x_1, ..., x_d) = P_X\left(X_1 \leq x_1, \ldots, X_d \leq x_d\right) = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_d} f_X(u_1, ..., u_d) du_d \ldots du_1.$$

La fonction f_X est appelée densité de probabilité conjointe du vecteur X.

Proposition 1.1.2. *[5]*

Toute densité de probabilité conjointe f_X de $\mathbb{R}^d$ vérifie les trois assertions suivantes :

i) *f_X est positive,*

ii) *f_X est mesurable,*

iii) *f_X est intégrable et*

$$\int_{\mathbb{R}^d} f_X(x_1, ..., x_d)dx_1 \dots dx_d = 1.$$

Réciproquement toute fonction f_X dans $\mathbb{R}^d$ vérifiant les trois assertions est une densité de probabilité.

Proposition 1.1.3. *[17]*

Si la fonction de densité f_X existe, elle est définie par :

$$f_X(x_1, \dots, x_d) = \frac{\partial^d F_X(x_1, ..., x_d)}{\partial x_1 \partial x_2 \dots \partial x_d}.$$

Définition 1.1.4. *(Densité conditionnelle)*

Soit $Z = (X, Y)$ une variable aléatoire à valeurs dans $\mathbb{R}^d \times \mathbb{R}^m$ de loi absolument continue et de densité f_Z, les variables X et Y sont également absolument continues et possèdent donc des densités f_X et f_Y, ces dernières sont appelées densités marginales et elles sont définies par

$$f_X(x) = \int_{\mathbb{R}^m} f_Z(x, y)dy \quad et \quad f_Y(y) = \int_{\mathbb{R}^d} f_Z(x, y)dx.$$

La fonction

$$\begin{aligned} f_{Y|X=x} : \mathbb{R}^m &\rightarrow \mathbb{R}^+ \\ y &\mapsto \frac{f_Z(x, y)}{f_X(x)} \end{aligned}$$

est une densité de probabilité sur $\mathbb{R}^m$ et elle est appelée densité conditionnelle de Y sachant que $X = x$. On définit de manière tout à fait similaire la loi de X sachant que $Y = y$.

Il résulte de cette définition que, pour presque-tout x de $\mathbb{R}^d$, on a :

$$f_Z(x, y) = f_{Y|X=x}(y) f_X(x).$$

Ainsi, si on connaît la densité marginale f_X et la densité conditionnelle $f_{Y|X}$ sachant $X = x$ on a, immédiatement, l'expression de la densité conjointe.

1.1.3 Changement de variables

Soit X un vecteur aléatoire dans $\mathbb{R}^d$ et φ une application mesurable de $\mathbb{R}^d$ vers $\mathbb{R}^d$. On veut déterminer la loi du vecteur aléatoire $\varphi(X)$.
Rappelons en premier lieu que le jacobien d'une fonction

$$\begin{array}{rcl} H: \quad \mathbb{R}^d & \to & \mathbb{R}^d \\ x = (x_1, ..., x_d) & \mapsto & H(x) = (h(x_1), ..., h(x_d)) \end{array}$$

est le déterminant de la matrice des dérivées premières, i.e.

$$\det J = \begin{vmatrix} \frac{\partial h_1}{\partial x_1} & \cdots & \cdots & \frac{\partial h_1}{\partial x_d} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \frac{\partial h_d}{\partial x_1} & \cdots & \cdots & \frac{\partial h_d}{\partial x_d} \end{vmatrix}.$$

Théorème 1.1.1. *[5]*
Soit X un vecteur aléatoire à valeurs dans un ouvert S de $\mathbb{R}^d$, de loi absolument continue et de densité f_X. Soit φ un $\mathcal{C}^1$-difféomorphisme de S vers un ouvert T (i.e. une bijection continûment différentiable et de réciproque également continûment différentiable). Alors, le vecteur aléatoire $Y = \varphi(X)$ est absolument continu et de densité, pour $y = (y_1, ..., y_d) \in \mathbb{R}^d$,

$$f_Y(y) = f_X\left(\varphi^{-1}(y)\right) |\det J_{\varphi^{-1}}|, \tag{1.1}$$

où $J_{\varphi^{-1}}$ est le jacobien de la fonction φ^{-1}.
Notons que, parfois, pour des raisons de calcul, il est plus simple de déterminer J_φ et on utilise alors l'égalité $J_{\varphi^{-1}} = J_\varphi^{-1}$.

1.1.4 Fonction caractéristique

Soit a un vecteur non aléatoire de composantes $(a_1, ..., a_d)$.

Définition 1.1.5. *On appelle fonction caractéristique du vecteur aléatoire X, la fonction de l'argument vectoriel a définie par :*

$$\varphi_X(a) = E\left(\exp(ia'X)\right) = E\left(\exp(i(a_1X_1 + a_2X_2 + \cdots + a_dX_d))\right).$$

Théorème 1.1.2. *[17]*
La fonction caractéristique φ_X caractérise la loi de la variable aléatoire X : si deux variables aléatoires ont même fonction caractéristique, elles ont même loi.

Théorème 1.1.3. *Les composantes $X_1, ..., X_d$ de X sont indépendantes si et seulement si la fonction caractéristique de X est égale au produit des fonctions caractéristiques de ses composantes :*

$$\varphi_X(a) = \prod_{i=1}^{d} \varphi_{X_i}(a_i).$$

Avec a un vecteur non aléatoire .

Démonstration.

i) Condition nécessaire : Les composantes de X sont indépendantes, alors

$$\begin{aligned}\varphi_X(a) &= E\left(\exp(i(a_1X_1 + \ldots + a_dX_d))\right)\\ &= \prod_{i=1}^{d} E\left(\exp(i(a_iX_i))\right)\\ &= \prod_{i=1}^{d} \varphi_{X_i}(a_i).\end{aligned}$$

ii) Condition sufisante : On peut construire des variables aléatoires X_i' , $i \in 1, \ldots, d$, indépendantes, telles que pour tout i, X_i et X_i' ont même loi. On en déduit que $\varphi_{X_i} = \varphi_{X_i'}$. En utilisant la condition nécessaire, on en déduit que :

$$\varphi_{X'} = \prod_{i=1}^{d} \varphi_{X_i'} = \varphi_X.$$

Ainsi les vecteurs aléatoires X et X' ont même loi.

Pour tous boréliens A_i,

$$\begin{aligned}\mathbb{P}(\bigcap_i \{X_i \in A_i\}) &= \mathbb{P}(\bigcap_i \{X_i' \in A_i\})\\ &= \prod_{i=1}^{d} \mathbb{P}(\{X_i' \in A_i\})\\ &= \prod_{i=1}^{d} \mathbb{P}(\{X_i \in A_i\}).\end{aligned}$$

Nous avons donc montré l'indépendance cherchée.

□

Théorème 1.1.4. *(Cramer-Wold) [17]*

La loi de X est entièrement déterminée par celles de toutes les combinaisons linéaires de ses composantes.

En effet, posons $Y = a'X = \sum_{i=1}^{d} a_i X_i$ et cherchons la fonction caractéristique de Y.

$$\varphi_Y(t) = E\left(\exp(itY)\right) = E\left(\exp(ita'X)\right),$$

d'où $\varphi_Y(1) = \varphi_X(a)$. Si la loi de Y est connue, pour tout a, on connait donc la fonction caractéristique de X donc la loi de X.

1.1.5 Moments d'un vecteur aléatoire

Le moment du vecteur aléatoire X peut être utilisé en statistique pour estimer les paramètres de la loi. Par conséquent, les moments sont un facteur majeur pour déterminer ou connaître la loi des vecteurs normaux [17, 16] .

1.1.5.1 Moyenne d'un vecteur aléatoire

Soit $X = (X_1, \ldots, X_d)'$ un vecteur aléatoire de $\mathbb{R}^d$ dont les composantes sont intégrable. On définit classiquement son espérance (sa moyenne) par :

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

et on la note simplement par :

$$\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_d \end{pmatrix}.$$

Remarque 1.1.1. *On dit que X est centré si $E(X) = 0_{\mathbb{R}^d}$.*

1.1.5.2 Matrice de variance-covariance

Soit $X = (X_1, \ldots, X_d)'$ un vecteur aléatoire de $\mathbb{R}^d$ tel que $E(X_j^2) < +\infty$, pour tout $1 \leq i, j \leq d$, la covariance entre X_i et X_j est donnée par :

$$Cov(X_i, X_j) = E[(X_i - E(X_i))(X_j - E(X_j))] = E(X_i X_j) - E(X_i)E(X_j).$$

Définition 1.1.6.

Soit X et Y deux vecteurs aléatoires de $\mathbb{R}^d$ dont les composantes sont de carré intégrable. La matrice de covariance(variance-covariance) de X et Y est définie par :

$$\begin{aligned}
\Sigma = Cov(X,Y) = Var(X) &= E[(X - E(X))(Y - E(Y))^{'}] \\
&= E(XY) - E(X)E(Y) \\
&= \begin{pmatrix} Cov(X_1, Y_1) & Cov(X_1, Y_2) & \cdots & Cov(X_1, Y_p) \\ Cov(X_2, Y_1) & Cov(X_2, Y_2) & \cdots & Cov(X_2, Y_p) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(X_d, Y_1) & Cov(X_d, Y_2) & \cdots & Cov(X_d, Y_p) \end{pmatrix}.
\end{aligned}$$

Si $X = Y$ alors :

$$\begin{aligned}
\Sigma = Cov(X,X) = Var(X) &= E[(X - E(X))(X - E(X))^{'}] \\
&= \begin{pmatrix} Var(X_1) & Cov(X_1, X_2) & \cdots & Cov(X_1, X_d) \\ Cov(X_2, X_1) & Var(X_2) & \cdots & Cov(X_2, X_d) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(X_d, X_1) & Cov(X_d, X_2) & \cdots & Var(X_d) \end{pmatrix}.
\end{aligned}$$

Théorème 1.1.5.

La matrice de covariance est une matrice symétrique d'ordre d et 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 : Soit X une vecteur aléatoire de $\mathbb{R}^d$, son vecteur moyen $\mu = E(X)$ et Σ sa matrice de covariance et soit v est un vecteur constant non nul de $\mathbb{R}^d$, alors

$$\begin{aligned}
v'\Sigma v &= v'E((X-\mu)(X-\mu)^{'})v \\
&= E(v'(X-\mu)(X-\mu)^{'}v) \\
&= E(v'(X-\mu)(v'(X-\mu))^{'}) \\
&= E((v'X - v'\mu)(v'X - v'\mu)^{'}) \\
&= Var(v'X) \geq 0.
\end{aligned}$$

□

Remarque 1.1.2. *[16]*
La covariance entre deux variables aléatoires indépendantes est nulle, mais deux variables aléatoires, dont la covariance est nulle, ne sont pas nécessairement indépendantes. En d'autres termes, si les composantes $X_1, ..., X_d$ d'un vecteur X sont indépendantes deux à deux, alors Σ est une matrice diagonale ayant pour coefficients diagonaux $Var(X_1), ..., Var(X_d)$.

Définition 1.1.7. *(Matrice de corrélation)*
La matrice de corrélation linéaire de X est la matrice de terme général :

$$\rho_{ij} = \frac{Cov(X_i, X_j)}{\sqrt{Var(X_i)Var(X_j)}},$$

où $-1 \leq \rho_{ij} \leq 1$, où $1 \leq i, j \leq d$.
Si les variables X_i sont réduites, Σ s'identifie avec la matrice de corrélation :

$$\begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1d} \\ \rho_{21} & 1 & \cdots & \rho_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{d1} & \rho_{d2} & \cdots & 1 \end{pmatrix}.$$

1.2 Vecteurs aléatoires gaussiens

Les vecteurs aléatoires gaussiens sont associés aux lois gaussiennes normales et de ce fait, ils jouent un rôle important en probabilités et statistique.

1.2.1 Rappel sur la loi normale unidimensionnelle

Les lois normales sont parmi les lois de probabilité les plus utilisées pour modéliser des phénomènes naturels issus de plusieurs événements aléatoires.
Une loi normale est une loi de probabilité absolument continue qui dépend de deux paramètres : son espérance, un nombre réel noté μ, et son écart type, un nombre réel positif noté σ. La densité de probabilité de la loi normale d'espérance μ, et d'écart type σ est donnée par :

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right).$$

Une telle variable aléatoire est alors dite variable gaussienne.
On note habituellement cela de la manière suivante :

$$X \sim \mathcal{N}(\mu, \sigma^2).$$

La loi normale de moyenne nulle et d'écart type unitaire est appelée loi normale centrée réduite ou loi normale standard. Sa fonction de densité est donnée par :

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

On écrit simplement

$$X \sim \mathcal{N}(0,1).$$

Proposition 1.2.1.
Soit $X \sim \mathcal{N}(0,1)$*, alors*

$$Y = \sigma X + \mu \sim \mathcal{N}(\mu, \sigma^2).$$

Définition 1.2.1. *[2]*
Soit X *une variable aléatoire réelle. On appelle fonction caractéristique de* X *la fonction* φ_X *définie sur* $\mathbb{R}$*, à valeurs complexes, par :*

$$\varphi_X(t) = E(e^{itX}) = \int_{-\infty}^{+\infty} e^{itx} f_X(x)dx, \qquad \forall t \in \mathbb{R}.$$

Proposition 1.2.2. *[2]*
Soit $X \sim \mathcal{N}(0,1)$*, alors :*

$$\varphi_X(t) = e^{-\frac{t^2}{2}}.$$

Démonstration. Pour tout $t \in \mathbb{R}$, on a :

$$\begin{aligned}
\varphi_X(t) &= E(e^{itX}) \\
&= \int_{-\infty}^{+\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\
&= \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(x-it)^2} dx \\
&= \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \sqrt{2\pi} \\
&= e^{-\frac{t^2}{2}}.
\end{aligned}$$

□

Proposition 1.2.3. *[2]*
Soit $X \sim \mathcal{N}(m, \sigma^2)$*, alors :*

$$\varphi_X(t) = e^{imt - \frac{t^2\sigma^2}{2}}.$$

Démonstration.
Soit $X = \sigma Z + m$, où $Z \sim \mathcal{N}(0,1)$, alors pour tout $t \in \mathbb{R}$ on a :

$$\varphi_X(t) = E(e^{itX}) = E(e^{it(\sigma Z+m)}) = E(e^{i(\sigma t)Z})e^{itm} = e^{imt-\frac{t^2\sigma^2}{2}}.$$

□

1.2.2 Propriétés des vecteurs gaussiens

Définition 1.2.2. *[16]*
Un vecteur aléatoire X à valeurs dans $\mathbb{R}^d$ est dit gaussien si toute combinaison linéaire de ses composantes est une variable aléatoire gaussienne réelle.
Autrement dit, pour tout vecteur constant $(a_1, ..., a_d)$ de $\mathbb{R}^d$, la variable aléatoire $\langle a, X\rangle = a_1X_1 + a_2X_2 + ... + a_dX_d$ est une variable aléatoire gaussienne réelle.
On écrit simplement :

$$\forall a \in \mathbb{R}^d, \qquad \langle a, X\rangle \sim \mathcal{N}(\mu, \sigma^2).$$

Remarque 1.2.1.
Les composantes d'un vecteur gaussien sont gaussiennes mais la réciproque est fausse.

Exemple 1.2.1. *(contre exemple) : il se peut que $X_1, .., X_n$ soient des v.a.réelles gaussiennes sans pour autant que le vecteur $(X_1, .., X_n)'$ soit gaussien : en effet, soient $X_1 \sim N_1(0,1)$ et $X_2 = \varepsilon X_1$, où $\varepsilon \amalg X_1$, et de loi définie par $P(\varepsilon = \pm 1) = 1/2$. Alors, $X_2 \sim N_1(0,1)$ mais la fonction caractéristique du vecteur $(X_1, X_2)'$ n'est pas une fonction caractéristique gaussienne.*

Théorème 1.2.1. *[16]*
Si X est un vecteur aléatoire de $\mathbb{R}^d$ de vecteur moyenne μ et de matrice de covariance Σ. Alors si A est une matrice réelle $d \times d$, le vecteur aléatoire AX de $\mathbb{R}^d$ a pour vecteur moyenne $A\mu$ et pour matrice de covariance $A\Sigma A'$.

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

$$E(AX) = AE(X) = A\mu.$$

et pour la matrice de covariance :

$$Var(AX) = E((AX-A\mu)(AX-A\mu)^{'}) = E(A(X-\mu)(X-\mu)^{'}A') = AE((X-\mu)(X-\mu)^{'})A' = A\Sigma A'.$$

□

Proposition 1.2.4. *[16]*

Si Σ est symétrique semi-définie positive de dimension $d \times d$, alors il existe une matrice A possédant d lignes (pas unique en général), telle que $\Sigma = AA'$.
Par exemple, il existe une matrice orthogonale P et une matrice diagonale $D = Diag(\lambda_1, ..., \lambda_d)$ (positive) telle que

$$\Sigma = PDP' = PDiag(\sqrt{\lambda_1}, ..., \sqrt{\lambda_d})Diag(\sqrt{\lambda_1}, ..., \sqrt{\lambda_d})P'.$$

Théorème 1.2.2. *[16]*

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

Démonstration.
Soit $\Sigma = AA'$. 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 la matrice identité I_d. Le vecteur aléatoire AX est alors centré de matrice de covariance $AI_dA' = AA' = \Sigma$. $\square$

1.3 Loi normale vectorielle

La loi d'un vecteur gaussien est caractérisée par son vecteur moyen μ et sa matrice de covariance Σ. Cette loi est appelée la loi normale (gaussienne) dans $\mathbb{R}^d$ qui est la généralisation multidimensionnelle de la loi normale réelle et elle est notée par $\mathcal{N}_d(\mu, \Sigma)$.
Pour trouver la forme générale de la fonction de densité d'un vecteur aléatoire gaussien, il convient de partir du cas le plus simple où les composantes du vecteur sont indépendantes et de même loi $\mathcal{N}(0, 1)$. La loi de ce vecteur est appelée la loi standard dans $\mathbb{R}^d$ et notée par $\mathcal{N}_d(0, I_d)$.
Dans la suite, nous donnons les définitions et les propriétés nécessaires de la loi normale vectorielle (voir

Théorème 1.3.1. *[12]*

Soit $X \sim \mathcal{N}_d(0, I_d)$, alors la fonction caractéristique de X est donnée par :

$$\varphi_X(a) = \exp\left(-\frac{1}{2}a'a\right).$$

Avec a un vecteur non aléatoire de composantes $(a_1, ..., a_d)$.

Démonstration.

$$\varphi_X(a) = E(\exp(ia'X)).$$

La variable X suit une loi normale standard où les composantes X_i sont des variables indépendantes de loi $\mathcal{N}(0,1)$. Alors par indépendance, on a

$$\varphi_X(a) = \prod_{i=1}^{d} E(\exp(ia_iX_i))$$

et par normalité, on trouve

$$\varphi_X(a) = \exp\left(-\frac{1}{2}a'a\right).$$

$\square$

Théorème 1.3.2. *[16]*
Soit X un vecteur aléatoire de $\mathbb{R}^d$ de moyenne $E(X) = \mu$ et de matrice de covariance $Var(X) = \Sigma$, telle que $\Sigma = AA'$ (Aest une matrice de dimension $d \times d$). Le vecteur X a même loi que le vecteur $AZ + \mu$, où Z est de loi $\mathcal{N}_d(0, I_d)$.
On dit que X suit une loi normale $\mathcal{N}_d(\mu, \Sigma)$ de paramètres μ et Σ.

Proposition 1.3.1. *(Propriété de linéarité)[10]*
Soit $X = (X_1, \dots, X_d)'$ un vecteur gaussien. On note $\mu = E(X)$ et $\Sigma = Var(X)$. Alors pour toute matrice A de dimension $d \times d$ et pour tout vecteur $b \in \mathbb{R}^d$, on a :

$$AX + b \sim \mathcal{N}_d(A\mu + b, A\Sigma A')$$

Théorème 1.3.3.
Soit $X \sim \mathcal{N}_d(\mu, \Sigma)$, alors la fonction caractéristique de X est donnée par :

$$\varphi_X(a) = \exp\left(ia'\mu - \frac{1}{2}a'\Sigma a\right), \qquad \forall a \in \mathbb{R}^d.$$

Démonstration.
Soient $X = AY + \mu$, $Y \sim \mathcal{N}_d(0, I_d)$ et $\Sigma = AA'$ (A est une matrice $d \times d$).
Par définition, on a

$$\begin{aligned}
\varphi_X(a) &= E(\exp(ia'X)) \\
&= E(\exp(ia'(AY + \mu))) \\
&= \exp(ia'\mu))E(\exp(ia'AY))) \\
&= \exp(ia'\mu))\varphi_Y(A'a).
\end{aligned}$$

En utilisant le Théorème (1.3.1), on a

$$\begin{aligned}\varphi_X(a) &= \exp(ia'\mu))\exp(ia'AA'a)) \\ &= \exp\left(ia'\mu - \frac{1}{2}a'\Sigma a\right).\end{aligned}$$

□

Théorème 1.3.4. *(Propriété d'indépendance)*
Les composantes d'un vecteur gaussien X sont indépendantes si et seulement si la matrice de covariance Σ est diagonale, c'est-à-dire si elles sont non corrélées.

Démonstration.
Il suffit de montrer la réciproque. Supposons donc que Σ soit diagonale, i.e.

$$\Sigma = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_d^2 \end{pmatrix}.$$

Comme X est un vecteur gaussien de loi $\mathcal{N}_d(\mu, \Sigma)$, chacune de ses composantes X_i, $i = 1, ..., d$, est de loi normale $\mathcal{N}(\mu_i, \sigma_i^2)$ et de fonction caractéristique :

$$\varphi_{X_i}(a_i) = \exp\left(ia_i\mu_i - \frac{1}{2}a_i^2\sigma_i^2\right),$$

pour tout a_i dans $\mathbb{R}$.
Par ailleurs, la fonction caractéristique du vecteur X est, pour tout a dans $\mathbb{R}$:

$$\begin{aligned}\varphi_X(a) &= \exp\left(ia'\mu - \frac{1}{2}a'\Sigma a\right) \\ &= \exp\left(i\sum_{i=1}^d a_i\mu_i - \frac{1}{2}\sum_{i=1}^d a_i^2\sigma_i^2\right) \\ &= \exp\left(\sum_{i=1}^d \left(ia_i\mu_i - \frac{1}{2}a_i^2\sigma_i^2\right)\right) \\ &= \prod_{i=1}^d \exp\left(ia_i\mu_i - \frac{1}{2}a_i^2\sigma_i^2\right) = \prod_{i=1}^d \varphi_{X_i}(a_i).\end{aligned}$$

Le Théorème (1.1.3) permet d'en déduire l'indépendance. □

Notons que, dans tout qui précède, la matrice Σ n'est pas supposée inversible. En revanche, la définition d'un vecteur gaussien par sa densité, par rapport à la mesure de Lebesgue dans $\mathbb{R}^d$, n'est possible que si cette matrice est inversible, comme l'affirme le théorème qui suit :

1.3.1 Fonction de densité

Pour une loi normale vectorielle quelconque, on donne la définition de la densité de probabilité :

Théorème 1.3.5.
Si $X \sim \mathcal{N}_d(\mu, \Sigma)$ et Σ est une matrice symétrique définie positive, X admet une fonction de densité par rapport à la mesure de Lebesgue sur $\mathbb{R}^d$, donnée par

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

où $x = (x_1, ..., x_d)^{'}$ et $\det \Sigma$ est le déterminant de la matrice de covariance Σ.

Démonstration.
D'après la Proposition (1.2.4), $\Sigma = AA'$, où A est une matrice $d \times d$ non singulière, alors $\Sigma^{-1} = \left(A^{-1}\right)^{'} A^{-1}$.
Les vecteurs X et $AY + \mu$ ont la même loi, où $Y \sim \mathcal{N}_d(0, I_d)$.
Par conséquent Y a pour densité

$$f_Y(y) = \frac{1}{(\sqrt{2\pi})^d} \exp\left(-\frac{1}{2}y'y\right).$$

La transformation inverse est $Y = B(X - \mu)$, où $B = A^{-1}$, et sa matrice jacobienne est donnée par

$$\begin{aligned}
\det \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \cdots & \frac{\partial y_1}{\partial x_d} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \frac{\partial y_d}{\partial x_1} & \cdots & \cdots & \frac{\partial y_d}{\partial x_d} \end{pmatrix} &= \det \begin{pmatrix} b_{11} & \cdots & \cdots & b_{1d} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ b_{d1} & \cdots & \cdots & b_{d1} \end{pmatrix} \\
&= \det B = \det A^{-1} = (\det A)^{-1} \\
&= (\det(AA'))^{-\frac{1}{2}} = (\det \Sigma)^{-\frac{1}{2}}.
\end{aligned}$$

En utilisant (1.1), on trouve la fonction de densité de la variable X

$$\begin{aligned}
f_X(x) &= \frac{(\det \Sigma)^{-\frac{1}{2}}}{(\sqrt{2\pi})^d} \exp\left(-\frac{1}{2}(x-\mu)^{'} B'B(x-\mu)\right) \\
&= \frac{1}{(\sqrt{2\pi})^d \sqrt{\det \Sigma}} \exp\left(-\frac{1}{2}(x-\mu)^{'}\Sigma^{-1}(x-\mu)\right).
\end{aligned}$$

□

Remarque 1.3.1.

Dans le cas où la matrice de covariance Σ *est singulière* $(\det \Sigma = 0)$, *on dit que la loi de* X *est singulière.*

Proposition 1.3.2.

Si $X \sim \mathcal{N}_d(\mu, \Sigma)$ *et ses composantes sont indépendantes, alors sa fonction de densité est donnée par :*

$$f_X(x_1, ..., x_d) = \frac{1}{(\sqrt{2\pi})^d \prod\limits_{i=1}^{d} \sigma_i} \exp\left(-\frac{1}{2} \sum_{i=1}^{d} \left(\frac{x_i - \mu_i}{\sigma_i}\right)^2\right).$$

Démonstration. Les composantes $X_1, ..., X_d$ du vecteur X sont indépendantes et de loi respectivement $\mathcal{N}(\mu_1, \sigma_1^2), ..., \mathcal{N}(\mu_1, \sigma_1^2)$, alors le vecteur moyen du vecteur X et sa matrice de covariance sont :

$$\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_d \end{pmatrix} \quad \text{et} \quad \Sigma = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_d^2 \end{pmatrix}.$$

Notons que la matrice Σ est diagonale en raison de l'indépendance des v.a.r. $(X_i)_{i=1,\dots,d}$. Comme toutes les variances σ_i^2 sont strictement positives, on obtient aisément la matrice inverse

$$\Sigma^{-1} = \begin{pmatrix} \frac{1}{\sigma_1^2} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sigma_d^2} \end{pmatrix}$$

et

$$\begin{aligned} (x-\mu)'\Sigma^{-1}(x-\mu) &= (x_1 - \mu_1, ..., x_d - \mu_d) \begin{pmatrix} \frac{1}{\sigma_1^2} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sigma_d^2} \end{pmatrix} \begin{pmatrix} x_1 - \mu_1 \\ \vdots \\ x_d - \mu_d \end{pmatrix} \\ &= \frac{(x_1 - \mu_1)^2}{\sigma_1^2} + ... + \frac{(x_d - \mu_d)^2}{\sigma_d^2} \\ &= \sum_{i=1}^{d} \left(\frac{x_i - \mu_i}{\sigma_i}\right)^2. \end{aligned}$$

D'où

$$f_X(x_1, ..., x_d) = \frac{1}{(\sqrt{2\pi})^d \prod\limits_{i=1}^{d} \sigma_i} \exp\left(-\frac{1}{2} \sum_{i=1}^{d} \left(\frac{x_i - \mu_i}{\sigma_i}\right)^2\right).$$

$\square$

1.3.1.1 Cas particuliers

i) Si $d = 1$, $X \sim \mathcal{N}(\mu, \sigma^2)$ et sa fonction de densité est donnée par

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right), \quad x \in \mathbb{R}.$$

ii) Si $d = 2$, $X = (X_1, X_2)' \sim \mathcal{N}_2(\mu, \Sigma)$, avec :

$$\mu = \begin{pmatrix} E(X_1) \\ E(X_2) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix},$$

$$\Sigma = \begin{pmatrix} Var(X_1) & Cov(X_1, X_2) \\ Cov(X_1, X_2) & Var(X_2) \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

et ρ_{12} est le coéfficient de corrélation entre X_1 et X_2 définit par :

$$\rho_{12} = \frac{Cov(X_1, X_2)}{\sqrt{Var(X_1)Var(X_2)}} = \frac{\sigma_{12}}{\sigma_1\sigma_2}.$$

On déduit que

$$\det\Sigma = \sigma_1^2\sigma_2^2 - (\rho_{12}\sigma_1\sigma_2)^2 = (\sigma_1\sigma_2)^2(1-\rho_{12}^2)$$

et

$$\Sigma^{-1} = \frac{1}{\det\Sigma}\begin{pmatrix} \sigma_2^2 & -\rho_{12}\sigma_1\sigma_2 \\ -\rho_{12}\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix} = \frac{1}{(\sigma_1\sigma_2)^2(1-\rho_{12}^2)}\begin{pmatrix} \sigma_2^2 & -\rho_{12}\sigma_1\sigma_2 \\ -\rho_{12}\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix}.$$

La fonction de densité du vecteur aléatoire $X = (X_1, X_2)'$ est définie par :

$$f_X(x) = \frac{1}{(2\pi)\sqrt{\det\Sigma}} \exp\left(-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\right).$$

On a

$$\begin{aligned}(x-\mu)'\Sigma^{-1}(x-\mu) &= \frac{1}{(\sigma_1\sigma_2)^2(1-\rho_{12}^2)}\begin{pmatrix} x_1-\mu_1 & x_2-\mu_2 \end{pmatrix}\begin{pmatrix} \sigma_2^2 & -\rho_{12}\sigma_1\sigma_2 \\ -\rho_{12}\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix}\begin{pmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{pmatrix} \\ &= \frac{1}{(1-\rho_{12}^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - \frac{2\rho}{\sigma_1\sigma_2}(x_1-\mu_1)(x_2-\mu_2) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right].\end{aligned}$$

D'où

$$f_X(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho_{12}^2}} \exp\left[-\frac{1}{2(1-\rho_{12}^2)}\left(\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - \frac{2\rho_{12}}{\sigma_1\sigma_2}(x_1-\mu_1)(x_2-\mu_2) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right)\right].$$

Exemple 1.3.1. *Soit* $X = (X_1, X_2)' \sim \mathcal{N}_2(\mu, \Sigma)$, *le vecteur moyen de* X *et sa matrice de covariance sont :*

$$\mu = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad et \quad \Sigma = \begin{pmatrix} 3 & 0.5 \\ 0.5 & 3 \end{pmatrix}.$$

La fonction de densité de la variable X *est :*

$$f_X(x_1, x_2) = \frac{1}{5.95\pi} \exp\left[-0.17\,(x_1 - 2)^2 + 0.2(x_1 - 2)(x_2 - 1) - 0.17\,(x_2 - 1)^2\right]$$

et sa représentation graphique se fait à l'aide du programme sous R donné dans l'annex 2 par la FIGURE1.1 suivante :

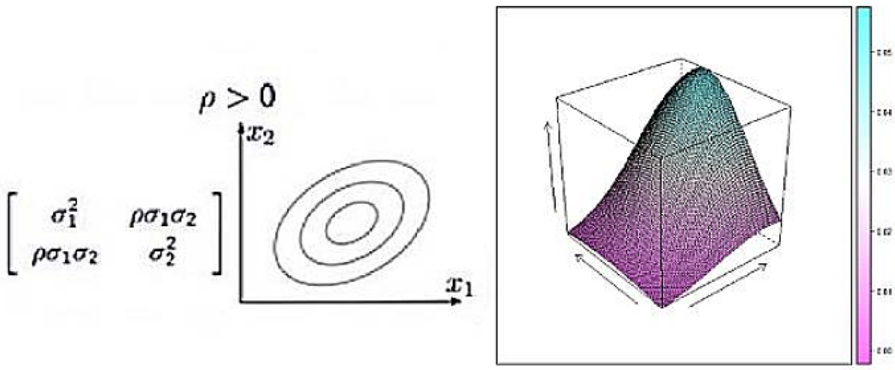

FIGURE 1.1 – Loi normale bivariée

iii) Si $d = 2$, $X = (X_1, X_2)' \sim \mathcal{N}_2(\mu, \Sigma)$ et les variables X_1 et X_2 sont indépendantes, le vecteur moyen et la matrice de covariance sont :

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad \text{et} \quad \Sigma = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}.$$

On déduit que

$$\det \Sigma = \sigma_1^2 \sigma_2^2$$

et

$$\Sigma^{-1} = \begin{pmatrix} \frac{1}{\sigma_1^2} & 0 \\ 0 & \frac{1}{\sigma_2^2} \end{pmatrix}.$$

$$\begin{aligned} (x - \mu)'\Sigma^{-1}(x - \mu) &= (x_1 - \mu_1, x_2 - \mu_2) \begin{pmatrix} \frac{1}{\sigma_1^2} & 0 \\ 0 & \frac{1}{\sigma_2^2} \end{pmatrix} \begin{pmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{pmatrix} \\ &= \frac{(x_1 - \mu_1)^2}{\sigma_1^2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2}. \end{aligned}$$

D'où

$$f_X(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left[-\frac{1}{2}\left(\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right)\right].$$

Exemple 1.3.2. *Soit $X = (X_1, X_2)' \sim \mathcal{N}_2(\mu, \Sigma)$, où X_1 et X_2 sont deux variables indépendantes. Le vecteur moyen de X et sa matrice de covariance sont :*

$$\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad et \quad \Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

La fonction de densité de la variable X est :

$$f_X(x_1, x_2) = \frac{1}{2\pi} \exp\left[-\frac{1}{2}\left(x_1^2 + x_2^2\right)\right]$$

et sa représentation graphique est donnée par la FIGURE1.2 suivante :

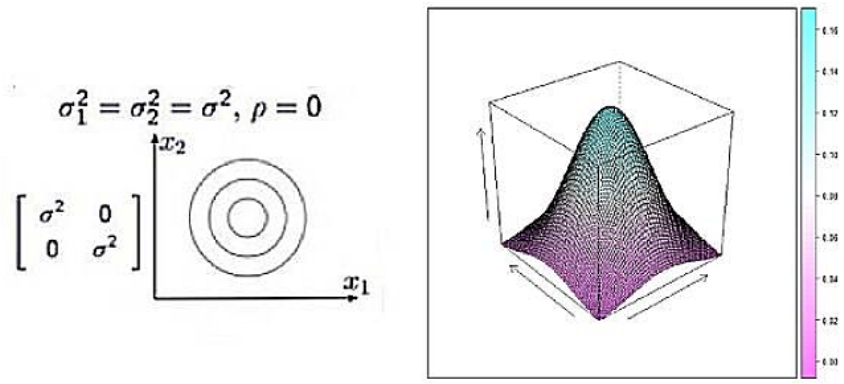

FIGURE 1.2 – Loi normale bivariée centrée et réduite

1.3.2 Lois marginales et lois conditionnelles

Soit $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ un vecteur aléatoire de $\mathbb{R}^d$ partitionné en deux sous vecteurs, $X_1 \in \mathbb{R}^r$ et $X_2 \in \mathbb{R}^{d-r}$, son vecteur moyen μ est :

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \text{où} \quad \mu_1 \in \mathbb{R}^r, \mu_2 \in \mathbb{R}^{d-r},$$

et sa matrice de covariance est donnée par :

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

où Σ_{11} est une matrice $r \times r$, Σ_{22} est une matrice $(d-r)\times(d-r)$, Σ_{12} est une matrice $r\times(d-r)$ et $\Sigma_{12} = \Sigma'_{21}$.

On écrit simplement :

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim \mathcal{N}_d\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right).$$

Théorème 1.3.6. *[12]*

Si $X \sim \mathcal{N}_d(\mu, \Sigma)$, la loi marginale de tout sous vecteur de k $(k < d)$ composantes est une loi normale à k variable.

Théorème 1.3.7. *[12]*

Soient $X \sim \mathcal{N}_d(\mu, \Sigma), \mu$ et Σ sont partitionnés comme

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad et \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

où $X_1 \in \mathbb{R}^r$ et $X_2 \in \mathbb{R}^{d-r}$, alors :

i) X_1 et X_2 sont deux vecteurs normales de lois

$$X_1 \sim \mathcal{N}_r(\mu_1, \Sigma_{11}) \quad et \quad X_2 \sim \mathcal{N}_{d-r}(\mu_2, \Sigma_{22}).$$

ii) Les vecteurs X_1 et X_2 sont indépendants, si et seulement si, $Cov(X_1, X_2) = \Sigma_{12} = 0$.

Théorème 1.3.8. *[10]*

Soit $X \sim \mathcal{N}_d(\mu, \Sigma), \mu$ et Σ sont partitionnés comme

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad et \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

où $X_1 \in \mathbb{R}^r$, $X_2 \in \mathbb{R}^{d-r}$ et $\det \Sigma \neq 0$. Alors :

i) $X_1 - \Sigma_{12}\Sigma_{22}^{-1}X_2 \sim \mathcal{N}_r\left(\mu_1 - \Sigma_{12}\Sigma_{22}^{-1}\mu_2, \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)$,

ii) $X_1 - \Sigma_{12}\Sigma_{22}^{-1}X_2$ et X_2 sont indépendants et $X_2 \sim \mathcal{N}_{d-r}(\mu_2, \Sigma_{22})$, *iii)* $X_1|X_2 = x_2 \sim \mathcal{N}_r\left(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)$.

Démonstration.
Soit A une matrice $d \times d$ avec :

$$A = \begin{pmatrix} I_r & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I_{d-r} \end{pmatrix},$$

où I_r est la matrice identité de dimension $r \times r$ et I_{d-r} est la matrice identité de dimension $(d-r) \times (d-r)$. Alors

$$\begin{aligned} AX &= \begin{pmatrix} I_r & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I_{d-r} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \\ &= \begin{pmatrix} X_1 - \Sigma_{12}\Sigma_{22}^{-1}X_2 \\ X_2 \end{pmatrix}. \end{aligned}$$

Le vecteur moyen du vecteur AX est donné par

$$E(AX) = AE(X) = \begin{pmatrix} \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}\mu_2 \\ \mu_2 \end{pmatrix}$$

et sa matrice de covariance est

$$\begin{aligned} Var(AX) = A\Sigma A' &= \begin{pmatrix} I_r & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I_{d-r} \end{pmatrix} \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \begin{pmatrix} I_r & 0 \\ -\Sigma_{12}\Sigma_{22}^{-1} & I_{d-r} \end{pmatrix} \\ &= \begin{pmatrix} \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} & 0 \\ 0 & \Sigma_{22} \end{pmatrix}. \end{aligned}$$

En utilisant le Théorème (1.3.6) et le fait que la covariance entre $X_1 - \Sigma_{12}\Sigma_{22}^{-1}X_2$ et X_2 est nulle, on trouve que les assertions $i)$ et $ii)$ sont vérifiée.
De plus, sachant que $X_2 = x_2$, la loi conditionnelle de $X_1 - \Sigma_{12}\Sigma_{22}^{-1}x_2$ est la même que la loi inconditionnelle de $X_1 - \Sigma_{12}\Sigma_{22}^{-1}X_2$.
Autrement dit, la variable $X_1|X_2 = x_2 \sim \mathcal{N}_r(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})$. □

Exemple 1.3.3.
Si $d = 2$, $X \sim \mathcal{N}_2(\mu, \Sigma)$, μ *et* Σ *sont partitionnés comme*

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad et \quad \Sigma = \begin{pmatrix} \sigma_1 & \sigma_{12} \\ \sigma_{21} & \sigma_2 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

où $X_1 \in \mathbb{R}$, $X_2 \in \mathbb{R}$ et $\det \Sigma \neq 0$. La fonction de densité de X :

$$f_X(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho_{12}^2}} \exp\left[-\frac{1}{2(1-\rho_{12}^2)}\left(\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - \frac{2\rho_{12}}{\sigma_1\sigma_2}(x_1-\mu_1)(x_2-\mu_2) + \left(\frac{x_2-\mu}{\sigma_2}\right)^2\right)\right]$$

En utilisant la formule

$$f_X(x_1,x_2) = f_{X_2}(x_2) f_{X_1|X_2=x_2}(x_1),$$

on obtient :

$$\begin{aligned}
f_X(x_1,x_2) &= \frac{1}{\sqrt{2\pi}\sigma_2}\exp\left(-\frac{(x_2-\mu_2)^2}{2\sigma_2^2}\right) \times \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho_{12}^2}}\exp\Bigg[-\frac{1}{2(1-\rho_{12}^2)}\Bigg(\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 \\
&\quad - 2\rho_{12}\,\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\Bigg) + \frac{1}{2}\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\Bigg] \\
&= \frac{1}{\sqrt{2\pi}\sigma_2}\exp\left(-\frac{(x_2-\mu_2)^2}{2\sigma_2^2}\right) \times \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho_{12}^2}}\exp\Bigg[-\frac{1}{2{\sigma_1}^2(1-\rho_{12}^2)}\Bigg((x_1-\mu_1)^2 \\
&\quad - 2\rho_{12}\frac{\sigma_1}{\sigma_2}\,(x_1-\mu_1)(x_2-\mu_2) + \frac{\sigma_1^2\rho_{12}^2}{\sigma_2^2}\,(x_2-\mu_2)^2\Bigg)\Bigg].
\end{aligned}$$

En remplaçant ρ_{12} par $\frac{\sigma_{12}}{\sigma_1\sigma_2}$ dans la formule précédente, on trouve

$$f_X(x_1,x_2) = \frac{1}{\sqrt{2\pi}\sigma_2}\exp\left(-\frac{(x_2-\mu_2)^2}{2\sigma_2^2}\right) \times \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\frac{\sigma_{12}}{\sigma_1\sigma_2}^2}}\exp\left(-\frac{\left(x_1-\mu_1-\frac{\sigma_{12}}{\sigma_2^2}(x_2-\mu_2)\right)^2}{2{\sigma_1}^2(1-\frac{\sigma_{12}}{\sigma_1\sigma_2}^2)}\right).$$

On en déduit que :

i) $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ *et sa fonction de densité est donnée par :*

$$f_{X_2}(x_2) = \frac{1}{\sqrt{2\pi}\sigma_2}\exp\left(-\frac{(x_2-\mu_2)^2}{2\sigma_2^2}\right).$$

ii) $X_1|X_2 = x_2 \sim \mathcal{N}\left(\mu_1 + \frac{\sigma_{12}}{\sigma_2^2}(x_2-\mu_2), \sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^2}\right)$ *et sa fonction de densité est :*

$$f_{X_1|X_2=x_2}(x_1) = \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{\left(1-\frac{\sigma_{12}^2}{\sigma_1^2\sigma_2^2}\right)}}\exp\left(-\frac{1}{2\left(\sigma_1^2-\frac{\sigma_{12}^2}{\sigma_2^2}\right)}\left(x_1-\mu_1-\frac{\sigma_{12}}{\sigma_2^2}(x_2-\mu_2)\right)^2\right).$$

Théorème 1.3.9. *[10]*
Soit $X \sim \mathcal{N}_d(\mu,\Sigma), \mu$ et Σ sont partitionnés comme

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad et \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

où $X_1 \in \mathbb{R}^r$, $X_2 \in \mathbb{R}^{d-r}$ et $\det \Sigma \neq 0$. Alors :

i) $X_2 - \Sigma_{21}\Sigma_{11}^{-1}X_1 \sim \mathcal{N}_{d-r}\left(\mu_2 - \Sigma_{21}\Sigma_{11}^{-1}\mu_1, \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right)$,

ii) $X_2 - \Sigma_{21}\Sigma_{11}^{-1}X_1$ *et* X_1 *sont indépendants et* $X_1 \sim \mathcal{N}_r(\mu_1, \Sigma_{11})$,

iii) $X_2|X_1 = x_1 \sim \mathcal{N}_{d-r}\left(\mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(x_1 - \mu_1), \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right)$.

Démonstration.
La démonstration se fait de la même façon que celle du théorème précédant. □

Chapitre 2

Loi normale matricielle

Un phénomène aléatoire matriciel est un phénomène observable qui peut être représenté sous forme de matrice et qui, sous des observations répétées, produit des résultats différents qui ne sont pas prévisibles de manière déterministe. Au lieu de cela, les résultats obéissent à certaines conditions de régularité statistique. Dans ce chapitre,nous présentons quelques définitions et propriétés du calcul matriciel, utilisées dans le calcul des matrices aléatoires, ainsi les concepts fondamentales liés à ces derniers, et nous présentons la loi normale matricielle et ces propriétés.

2.1 Calcul matriciel

La connaissance des différentes opérations avancées dans le calcul matriciel est nécessaires pour développé le travail avec les matrices aléatoires. Pour effectuer certaines opérations, il peut être utile de travailler sur le système des lignes ou colonnes d'une matrice.
Soit la matrice $A_{m,n}$ est un tableau d'éléments de $\mathbb{R}$, tel que

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \quad avec \quad 1 \leq i \leq n \; et \, 1 \leq j \leq m.$$

Alors la matrice A peut être écrite sous une des deux formes suivantes :
La première forme :

$$A = (a_1 \ \ a_2 \ \ldots \ a_n)', \quad \text{où} \ \ a_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \in \mathbb{R}^m, \ \text{pour tout} \ \ j = 1, ..., n.$$

La deuxième forme :

$$A = \begin{pmatrix} \alpha_1' \\ \vdots \\ \alpha_m' \end{pmatrix}, \quad \text{où} \ \ \alpha_i = \begin{pmatrix} a_{i1} \\ \vdots \\ a_{in} \end{pmatrix} \in \mathbb{R}^n, \ \text{pour tout} \ \ i = 1, ..., m.$$

L'ensemble des matrices à m lignes, n colonnes et à coefficients réel est noté $\mathcal{M}_{m,n}(\mathbb{R})$. Ce dernier est un espace vectoriel isomorphe à $\mathbb{R}^{m\times n}$.

La base canonique de cet espace est formée des matrices E_{ij} dont tous les coefficients sont nuls, sauf le ij-ème élément de l'intersection entre la i-ème ligne et la j-ème colonne qui vaut 1.

On peut l'écrit comme suit

$$E_{ij} = \begin{pmatrix} 0 & \ldots & 0 \\ \vdots & \vdots & \vdots \\ 0 & 1 & 0 \\ \vdots & \vdots & \vdots \\ 0 & \ldots & 0 \end{pmatrix}, \tag{2.1}$$

où $i = 1, ..., m$ et $j = 1, ..., n$.

2.1.1 Produit de Kronecker

De nombreux résultats liés aux matrices aléatoires sont exprimés en termes du produit de Kronecker. Ce produit doit être distingué de la multiplication matricielle habituelle, qui est une opération entièrement différente.

Le produit de Kronecker, noté par $\otimes$, est une opération portant sur les matrices. Il s'agit d'un cas particulier du produit tensoriel. Il est dénommé en hommage au mathématicien allemand Léopold Kronecker et il est aussi appelé produit direct ou produit de Zehfuss.

Définition 2.1.1. *[11]*

Soient A une matrice $(m \times n)$ et B une matrice $(p \times q)$. Le produit de Kronecker de A et B est *une matrice* $(mp \times nq)$, *noté* $A \otimes B$ *défini par :*

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & \vdots & \vdots & \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{pmatrix}.$$

Si A est la matrice identité $(n \times n)$, alors le produit de Kronecker A et B est une matrice $(np \times nq)$ et il est donné par :

$$I_n \otimes B = \begin{pmatrix} B & 0 & \dots & 0 \\ 0 & B & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & B \end{pmatrix},$$

où la matrice B apparait n fois sur la diagonale.

Exemple 2.1.1. *Soient* A *et* B *deux matrices carrées* (2×2), *pour*

i) $A = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}$ *et* $B = \begin{pmatrix} 0 & 7 \\ 6 & 4 \end{pmatrix}$. *Le produit de Kronecker* A *et* B *est donné par :*

$$\begin{aligned} A \otimes B &= \begin{pmatrix} 2\begin{pmatrix} 0 & 7 \\ 6 & 4 \end{pmatrix} & 1\begin{pmatrix} 0 & 7 \\ 6 & 4 \end{pmatrix} \\ \\ 5\begin{pmatrix} 0 & 7 \\ 6 & 4 \end{pmatrix} & 3\begin{pmatrix} 0 & 7 \\ 6 & 4 \end{pmatrix} \end{pmatrix} \\ &= \begin{pmatrix} 0 & 14 & 0 & 7 \\ 12 & 8 & 6 & 4 \\ 0 & 35 & 0 & 21 \\ 30 & 20 & 18 & 12 \end{pmatrix}. \end{aligned}$$

ii) $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ *et* $B = \begin{pmatrix} 5 & 2 \\ 1 & 4 \end{pmatrix}$. *Le produit de Kronecker des deux matrices est donné par :*

$$I_2 \otimes B = \begin{pmatrix} 1\begin{pmatrix} 5 & 2 \\ 1 & 4 \end{pmatrix} & 0\begin{pmatrix} 5 & 2 \\ 1 & 4 \end{pmatrix} \\ \\ 0\begin{pmatrix} 5 & 2 \\ 1 & 4 \end{pmatrix} & 1\begin{pmatrix} 5 & 2 \\ 1 & 4 \end{pmatrix} \end{pmatrix}$$
$$= \begin{pmatrix} 5 & 2 & 0 & 0 \\ 1 & 4 & 0 & 0 \\ 0 & 0 & 5 & 2 \\ 0 & 0 & 1 & 4 \end{pmatrix}.$$

Remarque 2.1.1. *Le produit de Kronecker est une opération non commutative. Pour deux matrices quelconques A et B, on a*

$$A \otimes B \neq B \otimes A.$$

Suit de l'exemple (2.1.5), on a

$$B \otimes A = \begin{pmatrix} 0\begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} & 7\begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} \\ \\ 6\begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} & 4\begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 0 & 14 & 7 \\ 0 & 0 & 35 & 21 \\ 12 & 6 & 8 & 4 \\ 30 & 18 & 20 & 12 \end{pmatrix}.$$

On remarque que $A \otimes B \neq B \otimes A$.

La proposition suivante donne les opérations élémentaires du produit de Kronecker.

Proposition 2.1.1. *[11]*

Soient A, B, C et D des matrices quelconques.

1. *Pour tout a et b dans $\mathbb{R}$, on a*

$$aA \otimes bB = abA \otimes B.$$

2. *Le produit de Kronecker est associatif :*

$$(A \otimes B) \otimes C = A \otimes (B \otimes C).$$

3. *Le produit de Kronecker est distributif :*

$$A \otimes (B + C) = (A \otimes B) + (A \otimes C),$$

$$A \otimes (BC) = (A \otimes B)(A \otimes C),$$

   avec B et C sont de mêmes dimensions.

4. *La propriété suivante mélange les aspects liés au produit matriciel usuel et au produit de Kronecker :*

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD),$$

   où A et C sont de mêmes dimensions et B et D sont aussi de mêmes dimensions.

Dans cette proposition, nous présentons les différentes propriétés du produit de Kronecker.

Proposition 2.1.2.

1. *La transposition est distributive sur le produit de Kronecker :*

$$(A \otimes B)' = A' \otimes B'.$$

2. *$A \otimes B$ est inversible si et seulement si A et B sont inversibles et on a :*

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}.$$

3. *Si A et B sont des matrices carrées (pas nécessairement du même ordre), alors :*

$$tr(A \otimes B) = tr(A)tr(B),$$

   avec $tr(A)$ est la trace de la matrice A.

4. *Soit A une matrice $n \times n$ et soit B une matrice $m \times m$, on a :*

$$\det(A \otimes B) = \det(A)^m \det(B)^n.$$

La propriété suivante est très importante aux calculs des matrices aléatoires.

Propriété 2.1.3.
Le produit de Kronecker de deux matrices symétriques définies positives est une matrice symétrique définie positive.

2.1.2 Opérateur Vec

La factorisation d'une matrice est une transformation linéaire qui convertit la matrice en un vecteur colonne. Plus précisément, L'opérateur Vec est un opérateur qui transforme une matrice en un vecteur colonne en empilant, verticalement, les colonnes de la matrice les unes au dessous des autres.

Définition 2.1.2. *[9]*
Soit $A = (a_1 \ \ a_2 \ \ \dots \ \ a_n)$, $a_j \in \mathbb{R}^m$, $\forall j = 1, ..., n$, *une matrice* $(m \times n)$, *alors l'opérateur* $Vec(A)$ *est un vecteur* $mn \times 1$, *tel que*

$$Vec(A) = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

Exemple 2.1.2. *Soit* A *une matrice* (2×2), *tel que*

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

alors

$$Vec(A) = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{12} \\ a_{22} \end{pmatrix}.$$

Remarque 2.1.2. *Soient* A *une matrice* $(m \times n)$ *et* B *une matrice* $(p \times q)$.

$$Vec(A) = Vec(B) \ \text{n'implique pas que} \ \ A = B.$$

Par exemple, soient $A = \begin{pmatrix} 1 & 3 \\ 2 & 0 \\ 0 & 4 \end{pmatrix}$ *et* $B = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & 4 \end{pmatrix}$. *Alors :*

$$Vec(A) = Vec(B) = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 3 \\ 0 \\ 4 \end{pmatrix} \quad ; \text{mais} \ \ A \neq B.$$

Si les deux matrices ont les mêmes dimensions, l'implication est vraie.

Proposition 2.1.4. *[11]*

Voici quelques propriétés de l'opérateur Vec :

i) Pour tout vecteur colonne a, on a

$$Vec(a') = Vec(a) = a.$$

ii) Pour a et b deux vecteurs colonnes, alors

$$Vec(ab') = b \otimes a.$$

iii) A et B deux matrices de mêmes dimensions, alors

$$(Vec(A))' Vec(B) = \text{ tr } (A'B).$$

Lemme 2.1.5. *[11]*

Si A une matrice $(m \times n)$, B une matrice $(n \times p)$ et C une matrice $(p \times q)$, alors

$$Vec(ABC) = (C' \otimes A)Vec(B).$$

Démonstration. Soient $(b_1, b_2, \ldots, b_p)$ les p colonnes de la matrice B et $(e_1, e_2, \ldots, e_p)$ les p colonnes de la matrice d'identité I_p, tel que $B = \sum\limits_{j=1}^{p} b_j\, e_j'$. Alors

$$Vec(ABC) = Vec\left(\sum_{j=1}^{p} Ab_j\, e_j' C\right) = \sum_{j=1}^{p} Vec\left((Ab_j)\,(C'e_j)'\right).$$

En utilisant la propriété $ii)$ de la Proposition (2.1.4), on trouve que :

$$\begin{aligned} Vec(ABC) &= \sum_{j=1}^{p} (C'e_j \otimes Ab_j) = (C' \otimes A) \sum_{j=1}^{p} (e_j \otimes b_j) \\ &= (C' \otimes A) Vec(\sum_{j=1}^{p} (b_j e_j')) = (C' \otimes A) Vec\,(B). \end{aligned}$$

□

Lemme 2.1.6. *[15]*

Soient $A \in \mathcal{M}_{p,m}(\mathbb{R})$, $B \in \mathcal{M}_{n,q}(\mathbb{R})$, $C \in \mathcal{M}_{q,m}(\mathbb{R})$, $D \in \mathcal{M}_{q,n}(\mathbb{R})$, $E \in \mathcal{M}_{m,m}(\mathbb{R})$ et $X \in \mathcal{M}_{m,n}(\mathbb{R})$, alors

i) $Vec(BC) = (I_m \otimes B)Vec(C) = (C' \otimes I_n)Vec(B) = (C' \otimes B)Vec(I_q)$,
où I_m est la matrice identité $(m \times m)$.

ii) tr $(CXB) = (Vec(C'))^{'} (I_q \otimes X)Vec(B)$.

iii) tr $(DX'EXB) = (Vec(X))^{'} (D'B' \otimes E)Vec(X)$.

Démonstration. $A(p \times m)$, $B(n \times q)$, $C(q \times m)$, $D(q \times n)$, $E(m \times m)$ et $X(m \times n)$, alors :

i) En utilisant le lemme précédent, on trouve que :
$Vec(BC) = Vec(BCI_m) = (I_m \otimes B)Vec(C)$.
$Vec(BC) = Vec(I_nBC) = (C' \otimes I_n)Vec(B)$.
$Vec(BC) = Vec(BI_qC) = (C' \otimes B)Vec(I_q)$.

ii) tr $(CXB) =$ tr $((C')^{'}XB)$.
En utilisant *iii*) de la Proposition (2.1.4), on a
tr $(CXB) =$ tr $((C')^{'}XB) = (Vec(c'))^{'} Vec(XB)$
et d'après la propriété *i*) de ce lemme, on trouve que :
tr $(CXB) = (Vec(C'))^{'} (I_q \otimes X)Vec(B)$.

iii) En utilisant la propriété *ii*) de ce lemme, on a
tr $(DX'EXB) =$ tr $((DX')E(XB)) = (Vec(XD'))^{'} (I_q \otimes E)Vec(XB)$.
D'après la propriété *i*) de ce lemme, on trouve que
tr $(DX'EXB) = ((D \otimes I_m)Vec(X))^{'} (I_q \otimes E)(B' \otimes I_m)Vec(X)$.
Appliquons les propriétés 4) de la Proposition 2.1.1 et 1) de la Proposition 2.1.2, on a :

$$\begin{aligned} \text{tr } (DX'EXB) &= (Vec(X))^{'} (D' \otimes I_m)(I_q \otimes E)(B' \otimes I_m)Vec(X) \\ &= (Vec(X))^{'} (D'B' \otimes E)Vec(X). \end{aligned}$$

□

2.1.3 Matrice de commutation

Le rôle principal de la matrice de commutation, aussi appelée matrice de permutation, est d'inverser l'ordre du produit de Kronecker. Cette matrice est carrée de dimension $(mn \times mn)$ et elle est notée K_{mn}. Tous ces coefficients sont égaux à 0, sauf un coefficient sur chaque ligne et sur chaque colonne est égal à 1.
La matrice de commutation peut être simplement définie comme suit :

$$K_{mn} = \sum_{i=1}^{m} \sum_{j=1}^{n} (E_{ij} \otimes E'_{ij}),$$

où E_{ij} est donnée dans (2.1).

Exemple 2.1.3. *Pour $m = 2$ et $n = 2$, on a :*

$$E_{11} \otimes E'_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$E_{12} \otimes E'_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$E_{21} \otimes E'_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$E_{22} \otimes E'_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Alors,

$$K_{22} = \sum_{i=1}^{2} \sum_{j=1}^{2} (E_{ij} \otimes E'_{ij}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Les vecteurs $Vec(A)$ et $Vec(A')$ contiennent évidemment les mêmes mn composantes, mais dans un ordre différent. La matrice de commutation est l'unique matrice qui transforme $Vec(A)$ en $Vec(A')$ comme le montre le lemme suivant :

Lemme 2.1.7. *[11]*
Soit A une matrice $(m \times n)$, alors

$$K_{mn} Vec(A) = Vec(A').$$

Exemple 2.1.4. *Soit* $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ *une matrice* (2×2), *telle que :*

$$Vec(A) = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{12} \\ a_{22} \end{pmatrix} \quad et \quad Vec(A') = \begin{pmatrix} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \end{pmatrix}.$$

Alors :

$$K_{22}Vec(A) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{21} \\ a_{12} \\ a_{22} \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \end{pmatrix} = Vec(A').$$

Lemme 2.1.8. *[11]*
Soient A *une matrice* $(m \times n)$ *et* B *une matrice* $(p \times q)$. *Alors*

$$K_{pm}(A \otimes B) = (B \otimes A)K_{qn}.$$

Démonstration. Soit X une matrice $(q \times n)$.
En utilisant les deux Lemme (2.1.5) et (2.1.7), on trouve que :

$$\begin{aligned} K_{pm}(A \otimes B)\,Vec\,(X) = K_{pm}\,Vec\,(BXA') &= Vec\,(AX'B') \\ &= (B \otimes A)\,Vec\,(X') \\ &= (B \otimes A)K_{qn}\,Vec\,(X). \end{aligned}$$

On déduit que :

$$K_{pm}(A \otimes B) = (B \otimes A)K_{qn}.$$

□

Lemme 2.1.9. *[11]*
Soient A *une matrice* $(m \times n)$ *et* B *une matrice* $(p \times q)$, *alors*

$$Vec(A \otimes B) = (I_n \otimes K_{qm} \otimes I_p)\,(Vec\,(A) \otimes Vec\,(B)).$$

Démonstration. Soient $(a_1, a_2, ..., a_n)$ les n colonnes de la matrice A, $(b_1, b_2, ..., b_q)$ les q colonnes de la matrice B, $(e_1, e_2, ..., e_n)$ les n colonnes de la matrice identité I_n, et $(u_1, u_2, ..., u_q)$ les q colonnes de la matrice identité I_q. Alors, nous pouvons écrire A et B comme :

$$A = \sum_{i=1}^{n} a_i\, e_i' \quad \text{et} \quad B = \sum_{j=1}^{q} b_j\, u_j'.$$

On obtient :

$$\begin{aligned}
Vec(A \otimes B) &= \sum_{i=1}^{n}\sum_{j=1}^{q} Vec(a_i e_i' \otimes b_j u_j') \\
&= \sum_{i=1}^{n}\sum_{j=1}^{q} Vec\left((a_i \otimes b_j)(e_i \otimes u_j)'\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{q} (e_i \otimes u_j \otimes a_i \otimes b_j) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{q} (I_n \otimes K_{qm} \otimes I_p)\,(e_i \otimes a_i \otimes u_j \otimes b_j) \\
&= (I_n \otimes K_{qm} \otimes I_p)\,\left(\sum_{i=1}^{n} Vec\,(a_i e_i') \otimes \sum_{j=1}^{q} Vec\,(b_j u_j')\right) \\
&= (I_n \otimes K_{qm} \otimes I_p)\,(Vec\,(A) \otimes Vec\,(B)).
\end{aligned}$$

□

2.2 Matrices aléatoires

L'ensemble des descriptions de tous les résultats possibles qui peuvent se produire lors de l'observation d'un phénomène aléatoire matriciel est l'espace d'échantillonnage $\mathcal{S}$. Un événement matriciel est un sous-ensemble de $\mathcal{S}$. Une mesure du degré de certitude avec laquelle un événement matriciel donné se produira lors de l'observation d'un phénomène aléatoire matriciel peut être trouvée en définissant une fonction de probabilité sur des sous-ensembles de $\mathcal{S}$, en attribuant une probabilité à chaque événement matriciel.

Dans cette section, nous définissons les concepts de base liés aux matrices aléatoires et nous présentons la fonction de densité d'une variable normale matricielle. Cette partie correspond au traitement standard du cas uni-varié et sa généralisation pas à pas.

Définition 2.2.1. *[8]*

Une matrice $X(m \times n)$ composée de mn éléments $X_{11}(.), X_{12}(.), \ldots, X_{mn}(.)$, qui sont des fonctions à valeurs réelles définies sur l'espace d'échantillon $\mathcal{S}$, est une matrice aléatoire réelle si l'intervalle $\mathbb{R}^{m\times n}$ de :

$$X = \begin{pmatrix} X_{11}(.) & \dots & X_{1n}(.) \\ \vdots & & \vdots \\ X_{m1}(.) & \dots & X_{mn}(.) \end{pmatrix}$$

est constitué d'ensembles boréliens d'espace réel à mn dimensions et si pour chaque ensemble borélien $\mathcal{B}$ de mn-uplets réels, arrangés dans une matrice

$$X = \begin{pmatrix} X_{11} & \dots & X_{1n} \\ \vdots & & \vdots \\ X_{m1} & \dots & X_{mn} \end{pmatrix}$$

dans $\mathbb{R}^{mn}$, l'ensemble

$$\left\{ s \in S : \begin{pmatrix} X_{11}(s_{11}) & \dots & X_{1n}(s_{1n}) \\ \vdots & & \vdots \\ X_{m1}(s_{m1}) & \dots & X_{mn}(s_{mn}) \end{pmatrix} \in \mathcal{B} \right\}$$

est un événement dans $\mathcal{S}$.

Proposition 2.2.1. *[6]*
Soit la variable $X \in \mathcal{M}_{m,n}(\mathbb{R})$, alors on peut écrire

$$X = \begin{pmatrix} X_1' \\ \vdots \\ X_m' \end{pmatrix} \in \mathcal{M}_{m,n}(\mathbb{R}) \Leftrightarrow X' = (X_1 \; ... \; X_m) \in \mathcal{M}_{n,m}(\mathbb{R}).$$

L'opérateur Vec permet de passer de la forme matricielle à la forme vectoriel. Il convient de l'appliquer non pas à X mais à sa transposé X', alors

$$Vec(X') = \begin{pmatrix} X_1 \\ \vdots \\ X_m \end{pmatrix} \in \mathbb{R}^{nm}.$$

Définition 2.2.2. *(Fonction marginale et la fonction conditionnelle)*
Soit $X \in \mathbb{R}^{m\times n}$ et $Y \in \mathbb{R}^{p\times q}$ deux matrices aléatoire, la fonction de densité de probabilité marginale de X et Y est définie par :

$$f_X(x) = \int_{\mathbb{R}^{p\times q}} f_{X,Y}(x,y)dy \quad et \quad f_Y(y) = \int_{\mathbb{R}^{m\times n}} f_{X,Y}(x,y)dx.$$

La fonction

$$f_{X|Y=y} = \frac{f_(X,Y)(x,y)}{f_Y(y)}$$

est la fonction de probabilité conditionnelle de $X|Y$.

2.2.1 Moments d'une matrice aléatoire

L'étude de tendance centrale et de la dispersion est une étape primordiale pour les matrices aléatoires, pour cela, on définit les moments et la fonction caractéristique pour les matrices aléatoires.

Moyenne

Soit X une matrice aléatoire de $\mathcal{M}_{m,n}(\mathbb{R})$ dont les éléments peuvent être intégrés, alors son espérance (sa moyenne) est définie par :

$$\begin{aligned}\mathbb{E}(X) &= \begin{pmatrix} E(X_{11}) & E(X_{12}) & \dots & E(X_{1n}) \\ E(X_{21}) & E(X_{22}) & \dots & E(X_{2n}) \\ \vdots & \vdots & & \vdots \\ E(X_{m1}) & E(X_{m2}) & \dots & E(X_{mn}) \end{pmatrix} \\ &= \begin{pmatrix} \mu_{11} & \mu_{12} & \dots & \mu_{1n} \\ \mu_{21} & \mu_{22} & \dots & \mu_{2n} \\ \vdots & \vdots & & \vdots \\ \mu_{m1} & \mu_{m2} & \dots & \mu_{mn} \end{pmatrix} = M\end{aligned}$$

On peut l'écrire autrement

$$\mathbb{E}(X) = \begin{pmatrix} \mu'_1 \\ \vdots \\ \mu'_m \end{pmatrix} = M,$$

avec $\mu'_i = (\mu_{i1}, \mu_{i2}, .., \mu_{in}) \in \mathbb{R}^n$.

Matrice de covariance

La matrice de covariance de la matrice $X(m \times n)$ est une matrice $mn \times mn$ définie par :

$$\begin{aligned}\Omega = Var(X) = Cov(X,X) &= Cov(Vec(X'), Vec(X')) \\ &= E((Vec(X') - E(Vec(X')))(Vec(X') - E(Vec(X')))') \\ &= E(Vec(X')(Vec(X'))') - E(Vec(X'))E((Vec(X'))').\end{aligned}$$

En effet, pour

$$X = \begin{pmatrix} X_1^{'} \\ \vdots \\ X_m^{'} \end{pmatrix} \quad avec \quad X_i \in \mathbb{R}^n, i = 1, \cdots, m,$$

on a

$$Vec(X^{'}) = \begin{pmatrix} X_1 \\ \vdots \\ X_m \end{pmatrix}.$$

Alors

$$\Omega = \begin{pmatrix} Cov(X_1, X_1) & Cov(X_1, X_2) & \dots & Cov(X_1, X_m) \\ \vdots & \vdots & & \vdots \\ Cov(X_m, X_1) & Cov(X_m, X_2) & \dots & Cov(X_m, X_m) \end{pmatrix}.$$

Exemple 2.2.1. *Pour $m = n = 2$, on a :*

$$X = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix} = \begin{pmatrix} X_1^{'} \\ X_2^{'} \end{pmatrix} \quad donc \quad Vec(X^{'}) = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

avec : $X_1^{'} = (X_{11} \quad X_{12})$ et $X_2^{'} = (X_{21} \quad X_{22})$.
La matrice de covariance de X est donnée par :

$$\Omega = Cov(X, X) = Var(X) = Var(Vec(X^{'})) = \begin{pmatrix} Cov(X_1, X_1) & Cov(X_1, X_2) \\ Cov(X_2, X_1) & Cov(X_2, X_2) \end{pmatrix},$$

où

$$Cov(X_1, X_1) = \begin{pmatrix} Cov(X_{11}, X_{11}) & Cov(X_{11}, X_{12}) \\ Cov(X_{12}, X_{11}) & Cov(X_{12}, X_{12}) \end{pmatrix}$$

$$Cov(X_1, X_2) = \begin{pmatrix} Cov(X_{11}, X_{21}) & Cov(X_{11}, X_{21}) \\ Cov(X_{12}, X_{21}) & Cov(X_{12}, X_{22}) \end{pmatrix}$$

$$Cov(X_2, X_1) = \begin{pmatrix} Cov(X_{21}, X_{11}) & Cov(X_{21}, X_{12}) \\ Cov(X_{22}, X_{11}) & Cov(X_{22}, X_{12}) \end{pmatrix}$$

$$Cov(X_2, X_2) = \begin{pmatrix} Cov(X_{21}, X_{21}) & Cov(X_{21}, X_{21}) \\ Cov(X_{22}, X_{21}) & Cov(X_{22}, X_{22}) \end{pmatrix}.$$

On écrit :

$$\Omega = \begin{pmatrix} Cov(X_{11}, X_{11}) & Cov(X_{11}, X_{12}) & Cov(X_{11}, X_{21}) & Cov(X_{11}, X_{22}) \\ Cov(X_{12}, X_{11}) & Cov(X_{12}, X_{12}) & Cov(X_{12}, X_{21}) & Cov(X_{12}, X_{22}) \\ Cov(X_{21}, X_{11}) & Cov(X_{21}, X_{12}) & Cov(X_{21}, X_{21}) & Cov(X_{21}, X_{22}) \\ Cov(X_{22}, X_{11}) & Cov(X_{22}, X_{12}) & Cov(X_{22}, X_{21}) & Cov(X_{22}, X_{22}) \end{pmatrix}.$$

2.2.2 Fonction caractéristique

Afin de déterminer les moments d'une matrice aléatoire ainsi que sa loi de probabilité, il convient de définir en premier lieu sa fonction caractéristique.

Définition 2.2.3. *[8]*
La fonction caractéristique d'une variable $X \in \mathcal{M}_{m,n}(\mathbb{R})$ est définie par

$$\varphi_X(Z) = E\left(\exp(\text{ tr } (iZ'X))\right),$$

où Z est une matrice aléatoire dans $\mathcal{M}_{m,n}(\mathbb{R})$.

2.3 Loi normale matricielle

La loi normale matricielle ou loi gaussienne matricielle est une loi de probabilité qui est une généralisation de la loi normale multivariée aux variables aléatoires à valeur matricielle. La loi normale d'une matrice aléatoire de $\mathcal{M}_{m,n}(\mathbb{R})$ est caractérisée par sa matrice moyenne $M(m \times n)$ et sa matrice de covariance $\Omega(mn \times mn)$ et elle est notée par $\mathcal{MN}_{m,n}(M, \Omega)$.

2.3.1 Fonction de densité

Soit $X_1, \ldots, X_m$ un échantillon de vecteurs dans $\mathbb{R}^n$, on organise l'ensemble de ces données sous la forme d'une matrice $(m \times n)$ donnée par

$$X = \begin{pmatrix} X'_1 \\ \vdots \\ X'_m \end{pmatrix} \in \mathcal{M}_{m,n}(\mathbb{R}),$$

avec : m est la taille de l'échantillon qui correspond au nombre des lignes de la matrice X et n est la dimension de l'espace des observations qui correspond au nombre de colonnes de la matrice X.

Définition 2.3.1.

On dit que la matrice aléatoire $X(m \times n)$ suit une loi normale matricielle de paramètres $M(m \times n)$ et $\Omega(mn \times mn)$, si et seulement si

$$Vec(X') \sim \mathcal{N}_{nm}(Vec(M'), \Omega).$$

On écrit simplement :

$$X \sim \mathcal{MN}_{m,n}(M, \Omega) \Leftrightarrow Vec(X') \sim \mathcal{N}_{nm}(Vec(M'), \Omega).$$

Exemple 2.3.1. *[15]*

Soit $X_1, \ldots, X_m$ un échantillon de vecteurs aléatoires indépendants de loi normale $\mathcal{N}_n(\mu, \Sigma)$. Alors, la moyenne de la matrice X est :

$$M = E(X) = \begin{pmatrix} E(X_1') \\ E(X_2') \\ \vdots \\ E(X_m') \end{pmatrix} = \begin{pmatrix} \mu' \\ \mu' \\ \vdots \\ \mu' \end{pmatrix} = 1_m \mu' = 1_m \otimes \mu',$$

où $1_m = (1, 1, ..., 1)' \in \mathbb{R}^m$.

La moyenne de $Vec(X')$ est donnée par :

$$E(Vec(X')) = \begin{pmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_m) \end{pmatrix} = \begin{pmatrix} \mu \\ \mu \\ \vdots \\ \mu \end{pmatrix} = 1_m \otimes \mu.$$

La covariance de $Vec(X')$ est donnée par :

,

$$\begin{aligned} Var(Vec(X')) &= \begin{pmatrix} Cov(X_1, X_1) & Cov(X_1, X_2) & \ldots & Cov(X_1, X_m) \\ Cov(X_2, X_1) & Cov(X_2, X_2) & \ldots & Cov(X_2, X_m) \\ \vdots & \vdots & & \vdots \\ Cov(X_m, X_1) & Cov(X_m, X_2) & \ldots & Cov(X_m, X_m) \end{pmatrix} \\ &= \begin{pmatrix} \Sigma & 0_n & \ldots & 0_n \\ 0_n & \Sigma & \ldots & 0_n \\ \vdots & \vdots & & \vdots \\ 0_n & 0_n & \ldots & \Sigma \end{pmatrix} \\ &= I_m \otimes \Sigma \end{aligned}$$

où 0_n est la matrice nulle $(n \times n)$ et I_m est la matrice identité $(m \times m)$.
On écrit :

$$X \sim \mathcal{MN}_{m,n}(1_m \otimes \mu', I_m \otimes \Sigma) \Leftrightarrow Vec(X') \sim \mathcal{N}_{nm}(1_m \otimes \mu, I_m \otimes \Sigma).$$

Remarque 2.3.1.
La matrice de covariance de X n'est que le produit de Kronecker de la matrice de covariance des lignes $I_m(m \times m)$ et la matrice de covariance des colonnes $\Sigma(n \times n)$.

Proposition 2.3.1. *[6]*
Soient $X \sim \mathcal{MN}_{m,n}(M, \Omega), B \in \mathcal{M}_{r,m}(\mathbb{R})$ et $C \in \mathcal{M}_{n,s}(\mathbb{R})$. Alors la transformation

$$Y = BXC \sim \mathcal{MN}_{r,s}(BMC, (B \otimes C')\Omega(B' \otimes C)).$$

Démonstration. D'après la Définition (2.3.1), il suffit de démontrer que

$$Vec(Y') \sim \mathcal{N}_{sr}(Vec((BMC)^{'}), (B \otimes C')\Omega(B' \otimes C)).$$

On a $Y = BXC$, alors

$$E(Y) = E(BXC) = BE(X)C = BMC.$$

En utilisant le Lemme (2.1.5), on trouve que

$$Vec(Y') = Vec((BXC)^{'}) = Vec(C'X'B') = (B \otimes C')Vec(X').$$

Alors

$$\begin{aligned} E(Vec(Y')) &= E((B \otimes C')Vec(X')) \\ &= (B \otimes C')E(Vec(X')) \\ &= (B \otimes C')Vec(M') \\ &= Vec((BMC)^{'}) \end{aligned}$$

et sa matrice de covariance est :

$$\begin{aligned} Var(Vec(Y')) &= Var((B \otimes C')Vec(X')) \\ &= (B \otimes C')Var(Vec(X'))(B \otimes C')^{'} \\ &= (B \otimes C')\Omega(B' \otimes C). \end{aligned}$$

On déduit que

$$Vec(Y') \sim \mathcal{N}_{sr}(Vec((BMC)^{'}), (B \otimes C')\Omega(B' \otimes C))$$

et

$$Y \sim \mathcal{MN}_{r,s}(BMC, (B \otimes C')\Omega(B' \otimes C)).$$

$\square$

Remarque 2.3.2. *[8]*
Les vecteurs lignes (colonnes) d'une matrice aléatoire sont indépendantes si et seulement si la matrice de covariance des lignes (colonnes) est diagonale.

Exemple 2.3.2. *[15]*
Soit $X_1, \ldots, X_m$ un échantillon de vecteurs aléatoires indépendants de loi normale $\mathcal{N}_n(\mu, \Sigma)$. Alors

$$X \sim \mathcal{MN}_{m,n}(1_m \otimes \mu', I_m \otimes \Sigma) \Leftrightarrow Vec(X') \sim \mathcal{N}_{nm}(1_m \otimes \mu, I_m \otimes \Sigma).$$

i) Pour $Y = BXC$, où $B \in \mathcal{M}_{r,m}(\mathbb{R})$ et $C \in \mathcal{M}_{n,s}(\mathbb{R})$. On a
$E(Vec(Y')) = (B \otimes C')E(Vec(X')) = (B \otimes C')(1_m \otimes \mu) = (B1_m) \otimes (C'\mu)$.
$Var(Vec(Y')) = (B \otimes C')Var(Vec(X'))(B' \otimes C) = (B \otimes C')(I_m \otimes \Sigma)(B' \otimes C) = BB' \otimes C'\Sigma C$.
On déduit que

$$Vec(Y') \sim \mathcal{N}_{nm}((B1_m) \otimes (C'\mu), BB' \otimes C'\Sigma C).$$

On a $E(X) = 1_m \otimes \mu'$, alors

$$E(Y) = E(BXC) = BE(X)C = B(1_m \otimes \mu')C = (B1_m) \otimes (C'\mu)'.$$

Par conséquent

$$Y \sim \mathcal{MN}_{m,n}((B1_m) \otimes (C'\mu)', BB' \otimes C'\Sigma C).$$

On peut écrire

$$Y \sim \mathcal{MN}_{m,n}((B1_m) \otimes (C'\mu)', BB' \otimes C'\Sigma C) \Leftrightarrow Vec(Y') \sim \mathcal{N}_{nm}((B1_m) \otimes (C'\mu), BB' \otimes C'\Sigma C).$$

ii) Si $\Sigma = I_n$, on a

$$X \sim \mathcal{MN}_{m,n}(1_m \otimes \mu', I_m \otimes I_n) \Leftrightarrow Vec(X') \sim \mathcal{N}_{nm}(1_m \otimes \mu, I_m \otimes I_n).$$

Donc

$$Y \sim \mathcal{MN}_{m,n}((B1_m) \otimes (C'\mu)', BB' \otimes C'C) \Leftrightarrow Vec(Y') \sim \mathcal{N}_{nm}((B1_m) \otimes (C'\mu), BB' \otimes C'C).$$

Définition 2.3.2.
On dit qu'une variable $X(m \times n)$ suit une loi normale matricielle $\mathcal{MN}_{m,n}(M, \Omega)$ si sa matrice de covariance est définie positive et elle se factorise en produit de Kronecker de deux matrices définies positives $\Omega = \Psi \otimes \Sigma$. Ces deux matrices sont la matrice de covariance des lignes $\Psi(m \times m)$ et la matrice de covariance des colonnes $\Sigma(n \times n)$.
Alors $Vec(X')$ suit une loi normale vectorielle $\mathcal{N}_{nm}(Vec(M'), \Psi \otimes \Sigma)$.
On écrit simplement

$$X \sim \mathcal{MN}_{m,n}(M, \Psi \otimes \Sigma) \Leftrightarrow Vec(X') \sim \mathcal{N}_{nm}(Vec(M'), \Psi \otimes \Sigma).$$

Dans ce cas, la variable X possède une fonction de densité par rapport à la mesure de Lebesgue et elle est donnée dans le théorème suivant :

Théorème 2.3.1.
Si $X \sim \mathcal{MN}_{m,n}(M, \Psi \otimes \Sigma)$, où $\Psi(m \times m)$ et $\Sigma(n \times n)$ sont des matrices définies positifs, alors X admet une fonction de densité par rapport à la mesure de Lebesgue sur $\mathcal{M}_{m,n}(\mathbb{R})$ donnée par

$$f_X(x) = \frac{1}{(2\pi)^{\frac{mn}{2}} (\det \Psi)^{\frac{n}{2}} (\det \Sigma)^{\frac{m}{2}}} \exp\left(-\frac{1}{2} \operatorname{tr}\left(\Psi^{-1}(x-M)\Sigma^{-1}(x-M)'\right)\right), \quad \forall\, x \in \mathcal{M}_{m,n}(\mathbb{R}).$$

Démonstration. On a $X \sim \mathcal{MN}_{m,n}(M, \Psi \otimes \Sigma)$, donc $Vec(X') \sim \mathcal{N}_{nm}(Vec(M'), \Psi \otimes \Sigma)$.
Soit $Y = Vec(X')$ et $\mu = Vec(M')$, alors $Y \sim \mathcal{N}_{nm}(\mu, \Psi \otimes \Sigma)$.
La fonction de densité de Y est :

$$f_Y(y) = \frac{1}{(2\pi)^{\frac{mn}{2}} \det(\Psi \otimes \Sigma)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(y-\mu)'(\Psi \otimes \Sigma)^{-1}(y-\mu)\right), \quad \forall\, y \in \mathbb{R}^{nm}.$$

D'après la Propriété 2) de la Proposition(2.1.2), on a

$$(\Psi \otimes \Sigma)^{-1} = \Psi^{-1} \otimes \Sigma^{-1}$$

et en utilisant la Propriété 4 de la proposition(2.1.2), on trouve que

$$\det(\Psi \otimes \Sigma)^{\frac{1}{2}} = (\det \Psi)^{\frac{n}{2}} (\det \Sigma)^{\frac{m}{2}}.$$

Par conséquent

$$f_Y(y) = \frac{1}{(2\pi)^{\frac{mn}{2}} (\det \Psi)^{\frac{n}{2}} (\det \Sigma)^{\frac{m}{2}}} \exp\left(-\frac{1}{2}\left(Vec\left((x-\mu)'\right)\right)' \left(\Psi^{-1} \otimes \Sigma^{-1}\right) \left(Vec\left((x-\mu)'\right)\right)\right).$$

Nous écrivons maintenant l'identité de la propriété *iii)* du Lemme (2.1.6) sous la forme particulière

$$\operatorname{tr}(DZ'EZ) = (Vec(Z))' (D' \otimes E) Vec(Z).$$

Nous prenons $Z=(x-\mu)^{'}$, $D=D'=\Psi^{-1}$ et $E=\Sigma^{-1}$. Alors

$$\left(Vec\left((x-\mu)^{'}\right)\right)^{'}\left(\Psi^{-1}\otimes\Sigma^{-1}\right)Vec\left((x-\mu)^{'}\right)= \text{ tr } \left(\Psi^{-1}(x-\mu)\Sigma^{-1}(x-\mu)\right)^{'}.$$

D'où

$$f_X(x)=\frac{1}{(2\pi)^{\frac{mn}{2}}(\det\Psi)^{\frac{n}{2}}(\det\Sigma)^{\frac{m}{2}}} \quad \exp\left(-\frac{1}{2}\text{ tr } \left(\Psi^{-1}(x-\mu)\Sigma^{-1}(x-\mu)^{'}\right)\right).$$

□

Remarque 2.3.3. *Si $m=n=1$, la loi de X n'est que la loi normale réelle.*

Exemple 2.3.3. *Si $m=n=2$ et $X=\begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22}\end{pmatrix}\sim\mathcal{MN}_{2,2}(M,\Psi\otimes\Sigma)$, avec*

$$M=\begin{pmatrix} 0 & 0 \\ 0 & 0\end{pmatrix} \; et \;\; \Psi=\Sigma=I_2=\begin{pmatrix} 1 & 0 \\ 0 & 1\end{pmatrix}.$$

On a : $\det(\Psi)=\det(\Sigma)=1$ *et* $\Psi^{-1}=\Sigma^{-1}=I_2=\begin{pmatrix} 1 & 0 \\ 0 & 1\end{pmatrix}$.

On a aussi :

$$\begin{aligned}\Psi^{-1}(x-\mu)\Sigma^{-1}(x-\mu)^{'} &= \begin{pmatrix} 1 & 0 \\ 0 & 1\end{pmatrix}\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22}\end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1\end{pmatrix}\begin{pmatrix} x_{11} & x_{21} \\ x_{12} & x_{22}\end{pmatrix}\\ &= \begin{pmatrix} x_{11}^2+x_{12}^2 & x_{11}x_{21}+x_{12}x_{22} \\ x_{21}x_{11}+x_{22}x_{12} & x_{21}^2+x_{22}^2\end{pmatrix}\end{aligned}$$

D'où

$$\text{tr } (\Psi^{-1}(x-\mu)\Sigma^{-1}(x-\mu)^{'})=x_{11}^2+x_{12}^2+x_{21}^2+x_{22}^2.$$

La fonction de densité de X est donnée par :

$$f_X(x)=\frac{1}{4\pi^2}e^{-\frac{1}{2}(x_{11}^2+x_{12}^2+x_{21}^2+x_{22}^2)}$$

et sa représentation graphique est donnée par la FIGURE 2.1 suivante :

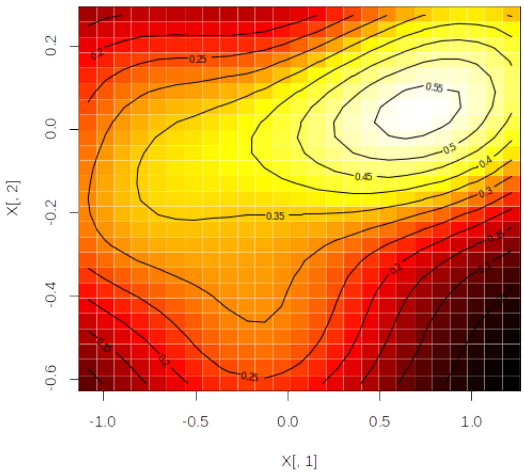

FIGURE 2.1 – La densité de la loi normal matricielle

2.3.2 Propriétés de la loi normale matricielle

Dans cette partie, nous présentons les propriétés les plus importantes de la loi normale matricielle (Voir [7],[1] et [13]). Nous commençons par la fonction caractéristique.

Théorème 2.3.2.
Si $X \sim \mathcal{MN}_{m,n}(M, \Psi \otimes \Sigma)$, *alors la fonction caractéristique de* X *est*

$$\varphi_X(Z) = \exp\left(\operatorname{tr} \left(iZ'M - \frac{1}{2}(Z'\Psi Z\Sigma) \right) \right).$$

Démonstration.
La fonction caractéristique de X est définie par

$$\begin{aligned} \varphi_X(Z) &= \mathbb{E}\left(\exp\left(\operatorname{tr}\, (iZ'X)\right)\right) \\ &= \mathbb{E}\left(\exp\left(i\left(Vec(Z')\right)' Vec(X')\right)\right) \qquad \forall \Updownarrow, \backslash\mathbb{R}. \end{aligned}$$

On a $X \sim \mathcal{MN}_{m,n}(M, \Psi \otimes \Sigma)$, donc $Vec(X') \sim \mathcal{N}_{nm}(Vec(M'), \Psi \otimes \Sigma)$.
En utilisant la formule de la fonction caractéristique d'un vecteur gaussien définie dans le Théorème (1.3.3), nous obtenons :

$$\varphi_X(Z) = \exp\left(i\left(Vec(Z')\right)' Vec(M') - \frac{1}{2}\left(Vec(Z')\right)'(\Psi \otimes \Sigma)\left(Vec(Z')\right)\right)$$

et d'aprés la propriété $iii)$ du Lemme (2.1.6), la fonction caractéristique de X est donnée par

$$\varphi_X(Z) = \exp\left(\operatorname{tr} \left(iZ'M - \frac{1}{2}(Z'\Psi Z\Sigma) \right) \right).$$

$\square$

Théorème 2.3.3.
Soit $X \sim \mathcal{MN}_{m,n}(M, \Psi \otimes \Sigma)$, $B(n \times t)$ et $D(n \times s)$ deux matrices non aléatoires. Alors XB et XD sont indépendantes si et seulement si $B^{'}\Sigma D = 0$.

Démonstration.
Sans perte de généralités, on pose que $M = 0$. La matrice de covariance entre XB et XD est donnée par :

$$\begin{aligned} Cov(XB, XD) &= Cov(Vec((XB)^{'}), Vec((XD)^{'})) \\ &= E(Vec((XB)^{'})(Vec((XD)^{'}))^{'}). \end{aligned}$$

En utilisant la propriété $i)$ du lemme (2.1.6) nous obtenons

$$\begin{aligned} Cov(XB, XD) &= E((I_m \otimes B^{'})Vec(X^{'})(Vec(X^{'}))^{'}(I_m \otimes D)) \\ &= (I_m \otimes B^{'})E(Vec(X^{'})(Vec(X^{'}))^{'})(I_m \otimes D) \\ &= (I_m \otimes B^{'})Var(Vec(X^{'}))(I_m \otimes D). \end{aligned}$$

On a $X \sim \mathcal{MN}_{m,n}(M, \Psi \otimes \Sigma)$ alors $Vec(X^{'}) \sim \mathcal{N}_{nm}(Vec(M^{'}), \Psi \otimes \Sigma)$ d'où

$$\begin{aligned} Cov(XB, XD) &= (I_m \otimes B^{'})(\Psi \otimes \Sigma)(I_m \otimes D) \\ &= (I_m\Psi \otimes B^{'}\Sigma)(I_m \otimes D) \\ &= \Psi \otimes (B^{'}\Sigma D). \end{aligned}$$

Donc $Cov(XB, XD) = 0$ si et seulement si $B^{'}\Sigma D = 0$. $\square$

Dans cette partie, nous calculons les moments du produit de Kronecker de deux matrices gaussiennes.

Théorème 2.3.4.
Soit $Z = (X \quad Y)$ une matrice aléatoire où $X \in \mathcal{M}_{m,n}(\mathbb{R})$ et $Y \in \mathcal{M}_{m,p}(\mathbb{R})$. Supposons

$$E(Z) = (E(X) \quad E(Y)) = (M \quad N)$$

et

$$Var(Vec(Z)) = \Omega = \begin{pmatrix} \Omega_{XX} & \Omega_{XY} \\ \Omega_{YX} & \Omega_{YY} \end{pmatrix},$$

alors

$$E(X \otimes Y) = \sum_{i=1}^{p}\sum_{j=1}^{m}(E_{ij}^{'} \otimes I_m)\Omega_{YX}(I_p \otimes E_{ij}^{'}) + M \otimes N.$$

Démonstration.

En utilisant le Lemme (2.1.9), on trouve que

$$Vec(X \otimes Y) = (I_m \otimes K_{pm} \otimes I_m)E(Vec(X) \otimes Vec(Y))$$

Comme l'opérateur Vec est linéaire on peut donc calculer

$$\begin{aligned} Vec(E(X \otimes Y)) &= E(Vec(X \otimes Y)) \\ &= E((I_n \otimes K_{pm} \otimes I_m)(Vec(X) \otimes Vec(Y))) \\ &= (I_n \otimes K_{pm} \otimes I_m)E(Vec(X) \otimes Vec(Y)). \end{aligned}$$

D'aprés la propriété *ii*) de la Proposition (2.1.4), on a :

$$\begin{aligned} E(Vec(X) \otimes Vec(Y)) &= E(Vec(Vec(Y)(Vec(X)))^{'}) \\ &= Vec(E(Vec(Y)(Vec(X)))^{'})) \\ &= Vec(Cov(Vec(Y), Vec(X)) + E(Vec(Y))E((Vec(X))^{'})) \\ &= Vec(\Omega_{YX} + Vec(N)(Vec(M))^{'}) \\ &= Vec(\Omega_{YX}) + Vec(M) \otimes Vec(N). \end{aligned}$$

Alors

$$Vec(E(X \otimes Y)) = (I_n \otimes K_{pm} \otimes I_m)(Vec(\Omega_{YX}) + Vec(M) \otimes Vec(N)).$$

En utilisant le Lemme(2.1.9), on trouve que

$$Vec(E(X \otimes Y)) = (I_n \otimes K_{pm} \otimes I_m)Vec(\Omega_{YX}) + Vec(M) \otimes Vec(N).$$

D'après la définition de la matrice de commutation

$$K_{pm} = \sum_{i=1}^{p} \sum_{j=1}^{m} (E_{ij} \otimes E_{ij}^{'})$$

et le Lemme(2.1.5), on a

$$\begin{aligned} Vec(E(X \otimes Y)) &= \sum_{i=1}^{p} \sum_{j=1}^{m} (I_n \otimes E_{ij} \otimes E_{ij}^{'} \otimes I_m)Vec(\Omega_{YX}) + Vec(M) \otimes Vec(N) \\ &= Vec\left(\sum_{i=1}^{p} \sum_{j=1}^{m} (E_{ij}^{'} \otimes I_m)\Omega_{YX}(I_n \otimes E_{ij}^{'})\right) + Vec(M \otimes N). \end{aligned}$$

Alors

$$E(X \otimes Y) = \sum_{i=1}^{p} \sum_{j=1}^{m} (E_{ij}^{'} \otimes I_m)\Omega_{YX}(I_n \otimes E_{ij}^{'}) + M \otimes N.$$

$\square$

Corollaire 2.3.2. *On peut déduire que*

$$E(X \otimes X) = \sum_{i=1}^{p} \sum_{j=1}^{m} (E_{ij}^{'} \otimes I_m) \Omega_{XX} (I_n \otimes E_{ij}^{'}) + M \otimes M.$$

Proposition 2.3.3.

Si la matrice de covariance de X est de la forme

$$\Omega_{XX} = \Psi \otimes \Sigma,$$

où $\Psi(n \times n)$ et $\Sigma(m \times m)$ deux matrices symétriques définies positives. Alors

$$E(X \otimes X) = Vec(\Sigma)(Vec(\Psi))^{'} + M \otimes M.$$

Démonstration.

D'après le corollaire précédent, on a

$$\begin{aligned} E(X \otimes X) &= \sum_{i=1}^{p} \sum_{j=1}^{m} (E_{ij}^{'} \otimes I_m) \Omega_{XX} (I_n \otimes E_{ij}^{'}) + M \otimes M \\ &= \sum_{i=1}^{p} \sum_{j=1}^{m} (E_{ij}^{'} \otimes I_m)(\Psi \otimes \Sigma)(I_n \otimes E_{ij}^{'}) + M \otimes M \\ &= \sum_{i=1}^{p} \sum_{j=1}^{m} (E_{ij}^{'} \Psi \otimes \Sigma E_{ij}^{'}) + M \otimes M. \end{aligned}$$

En remplaçant E_{ij} par $e_i e_j^{'}$, on obtient

$$\begin{aligned} E(X \otimes X) &= \sum_{i=1}^{p} \sum_{j=1}^{m} (e_j e_i^{'} \Psi \otimes \Sigma e_j e_i^{'}) + M \otimes M \\ &= \sum_{i=1}^{p} \sum_{j=1}^{m} (e_j \Psi_{i.} \otimes \Sigma_{.j} e_i^{'}) + M \otimes M \\ &= \sum_{i=1}^{p} (e_j \otimes \Sigma_{.j}) \sum_{j=1}^{m} (\Psi_{i.} \otimes e_i^{'}) + M \otimes M. \end{aligned}$$

En utilisant la propriété $ii)$ de la Proposition(2.1.9),on trouve que

$$E(X \otimes X) = \sum_{i=1}^{p} Vec(\Sigma_{.j} e_j^{'}) \sum_{j=1}^{m} (Vec(e_i \Psi_{i.}))^{'} + M \otimes M = Vec(\Sigma)(Vec(\Psi))^{'} + M \otimes M.$$

□

Théorème 2.3.5.

Soient $Z = (X \quad Y)$ une matrice aléatoire où $X \in \mathcal{M}_{m,n}(\mathbb{R})$ et $Y \in \mathcal{M}_{m,p}(\mathbb{R})$.
Supposons que

$$Vec(Z) \sim \mathcal{N}_{mn+mp}\left(\begin{pmatrix}\mu_X \\ \mu_Y\end{pmatrix}, \Omega\right),$$

où $\mu_X = Vec\,(E(X)) = Vec\,(M)$, $\mu_Y = Vec\,(E(Y)) = Vec\,(N)$ et $\Omega = \begin{pmatrix}\Omega_{XX} & \Omega_{XY} \\ \Omega_{YX} & \Omega_{YY}\end{pmatrix}$. Alors

$$\begin{aligned}
Var(Vec(X \otimes Y)) = {} & (I_n \otimes K_{pm} \otimes I_m)(\Omega_{XX} \otimes \Omega_{YY} + \Omega_{XX} \otimes \mu_Y \mu_Y^{'} + \mu_X \mu_X^{'} \otimes \Omega_{YY})(I_n \otimes K_{pm} \otimes I_m) \\
& + (K_{np} \otimes K_{mm})(I_p \otimes K_{nm} \otimes I_m)(\Omega_{YX} \otimes \Omega_{XY} + \Omega_{YX} \otimes \mu_X \mu_Y^{'} + \mu_Y \mu_X^{'} \otimes \Omega_{XY}) \\
& \times (I_n \otimes K_{mp} \otimes I_m).
\end{aligned}$$

Pour démontrer le théorème, nous avons besoin de quelques résultats des moments du produit de Kronecker de deux vecteurs gaussiens donnés dans le théorème suivant :

Théorème 2.3.6.

Soient deux vecteurs gaussiens $X \sim \mathcal{N}_m(\mu, \Omega_{XX})$ et $Y \sim \mathcal{N}_n(v, \Omega_{YY})$.
Supposons que

$$Z = \begin{pmatrix}X \\ Y\end{pmatrix} \sim \mathcal{N}_{m+n}(W, \Omega),$$

avec

$$W = E(Z) = \begin{pmatrix}E(X) \\ E(Y)\end{pmatrix} = \begin{pmatrix}\mu \\ v\end{pmatrix}$$

et

$$Var(Z) = \begin{pmatrix}\Omega_{XX} & \Omega_{XY} \\ \Omega_{YX} & \Omega_{YY}\end{pmatrix} = \begin{bmatrix}Var(X) & Cov(X,Y) \\ Cov(Y,X) & Var(Y)\end{bmatrix}.$$

Alors

$$E(X \otimes Y) = Vec(\Omega_{YX}) + \mu_X \otimes \mu_Y,$$

$$E(X \otimes X) = Vec(\Omega_{XX}) + \mu_X \otimes \mu_X$$

et

$$\begin{aligned}
Var(X \otimes Y) = {} & \Omega_{XX} \otimes \Omega_{YY} + \Omega_{XX} \otimes \mu_Y \mu_Y^{'} + \mu_X \mu_X^{'} \otimes \Omega_{YY} \\
& + K_{mn}(\Omega_{YX} \otimes \Omega_{XY} + \Omega_{YX} \otimes \mu_X \mu_Y^{'} + \mu_Y \mu_X^{'} \otimes \Omega_{XY}).
\end{aligned}$$

Démonstration.

Soient $x = Vec(X)$, $y = Vec(Y)$ et $z = Vec(Z)$. Alors

$$\begin{aligned} Vec(X \otimes Y) &= (I_n \otimes K_{pm} \otimes I_m)(Vec(X) \otimes Vec(Y)) \\ &= (I_n \otimes K_{pm} \otimes I_m)(x \otimes y). \end{aligned}$$

Par conséquent

$$\begin{aligned} Var(Vec(X \otimes Y)) &= Var(I_n \otimes K_{pm} \otimes I_m)(x \otimes y) \\ &= (I_n \otimes K_{pm} \otimes I_m) Var(x \otimes y)(I_n \otimes K_{pm} \otimes I_m)^{'} \\ &= (I_n \otimes K_{pm} \otimes I_m) Var(x \otimes y)(I_n \otimes K_{mp} \otimes I_m) \end{aligned}$$

En utilisant le théorème (2.3.6), on trouve que

$$\begin{aligned} Var(Vec(X \otimes Y)) = (I_n \otimes K_{pm} \otimes I_m)(&\Omega_{XX} \otimes \Omega_{YY} + \Omega_{XX} \otimes \mu_Y \mu_Y^{'} + \mu_X \mu_X^{'} \otimes \Omega_{YY} \\ &+ K_{mn,mp}(\Omega_{YX} \otimes \Omega_{XY} + \Omega_{YX} \otimes \mu_X \mu_Y^{'} + \mu_Y \mu_X^{'} \otimes \Omega_{XY}))(I_n \otimes K_{mp} \otimes I_m). \end{aligned}$$

En utilisant la propriété de la matrice de commutation de grande dimension

$$K_{mn,mp} = (I_n \otimes K_{mp} \otimes I_m)(K_{pn} \otimes K_{mm})(I_p \otimes K_{nm} \otimes I_m),$$

on trouve que

$$\begin{aligned} Var(Vec(X \otimes Y)) = &(I_n \otimes K_{pm} \otimes I_m)(\Omega_{XX} \otimes \Omega_{YY} + \Omega_{XX} \otimes \mu_Y \mu_Y^{'} + \mu_X \mu_X^{'} \otimes \Omega_{YY})(I_n \otimes K_{pm} \otimes I_m) \\ &+ (K_{np} \otimes K_{mm})(I_p \otimes K_{nm} \otimes I_m)(\Omega_{YX} \otimes \Omega_{XY} + \Omega_{YX} \otimes \mu_X \mu_Y^{'} + \mu_Y \mu_X^{'} \otimes \Omega_{XY}) \\ &\times (I_n \otimes K_{mp} \otimes I_m). \end{aligned}$$

□

Corollaire 2.3.4. *On déduit que :*

$$\begin{aligned} Var(Vec(X \otimes X)) = &(I_{m^2n^2} + K_{nn} \otimes K_{mm})(I_n \otimes K_{nm} \otimes I_m) \\ &\times (\Omega_{XX} \otimes \Omega_{XX} + \Omega_{XX} \otimes \mu_X \mu_X^{'} + \mu_X \mu_X^{'} \otimes \Omega_{XX})(I_n \otimes K_{mn} \otimes I_m). \end{aligned}$$

Proposition 2.3.5.

Si la matrice de covariance de X est de la forme

$$\Omega_{XX} = \Psi \otimes \Sigma,$$

où $\Psi(n \times n)$ et $\Sigma(m \times m)$ deux matrices symétriques définies positives. Alors

$$\begin{aligned} Var(Vec(X \otimes X)) = &(I_{m^2n^2} + K_{nn} \otimes K_{mm})(\Psi \otimes \Psi \otimes \Sigma \otimes \Sigma + \Psi \otimes (K_{nm} \otimes I_m)(\Sigma \otimes u_X u_X^{'}) \\ &\times (K_{mn} \otimes I_m) + (I_n \otimes K_{nm})(u_X u_X^{'} \otimes \Psi)(I_n \otimes K_{mn}) \otimes \Sigma). \end{aligned}$$

2.3.3 Lois marginales et conditionnelles

Théorème 2.3.7.

Soient $X \sim \mathcal{MN}_{m,n}(M, \Psi \otimes \Sigma)$, M et Ψ sont partionnées de la façon suivante :

$$X = \begin{array}{cc} \begin{pmatrix} X_{1r} \\ X_{2r} \end{pmatrix} & \begin{array}{l} s \\ m\text{-}s \end{array} \end{array} \quad , \; M = \begin{array}{cc} \begin{pmatrix} M_{1r} \\ M_{2r} \end{pmatrix} & \begin{array}{l} s \\ m\text{-}s \end{array} \end{array} \quad et \quad \Psi = \begin{array}{cc} \begin{pmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{21} & \Psi_{22} \end{pmatrix} & \begin{array}{l} s \\ m\text{-}s \end{array} \\ \begin{array}{cc} s & m\text{-}s \end{array} & \end{array}$$

alors

i) $X_{1r} \sim \mathcal{MN}_{s,n}(M_{1r}, \Psi_{11} \otimes \Sigma)$,

ii) $X_{2r}|X_{1r} \sim \mathcal{MN}_{m-s,n}(M_{2r} + \Psi_{21}\Psi_{11}^{-1}(X_{1r} - M_{1r}), \Psi_{22.1} \otimes \Sigma)$,
avec $\Psi_{22.1} = \Psi_{22} - \Psi_{21}\Psi_{11}^{-1}\Psi_{12}$.

Démonstration.

Pour démonter le résultat, il faut tout d'abord calculer l'inverse de $\Psi(m \times m)$ qui est une matrice symétrique définie positive partitionnée de la façon suivante :

Soient

$$\Psi = \begin{pmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{21} & \Psi_{22} \end{pmatrix}$$

où $\Psi_{11} \in \mathcal{M}_{s,s}(\mathbb{R}), \Psi_{12} \in \mathcal{M}_{s,m-s}(\mathbb{R}), \Psi_{21} \in \mathcal{M}_{m-s,s}(\mathbb{R}), \Psi_{22} \in \mathcal{M}_{m-s,m-s}(\mathbb{R})\, et\, \Psi_{12} = (\Psi_{21})'$.

et

$$\Psi^{-1} = \begin{pmatrix} \Psi^{11} & \Psi^{12} \\ \Psi^{21} & \Psi^{22} \end{pmatrix}$$

D'après la définition de l'inverse d'une matrice

$$\Psi\Psi^{-1} = \Psi^{-1}\Psi = I_m = \begin{pmatrix} I_s & 0 \\ 0 & I_{m-s} \end{pmatrix},$$

Nous avons :

$$\begin{cases} \Psi_{11}\Psi^{11} + \Psi_{21}\Psi^{21} &= I_s \\ \Psi_{21}\Psi^{11} + \Psi_{22}\Psi^{21} &= 0 \\ \Psi_{11}\Psi^{12} + \Psi_{12}\Psi^{22} &= 0 \\ \Psi_{21}\Psi^{12} + \Psi_{22}\Psi^{22} &= I_{m-s} \end{cases} \quad \text{et} \quad \begin{cases} \Psi^{11}\Psi_{11} + \Psi^{12}\Psi_{21} &= I_s \\ \Psi^{11}\Psi_{12} + \Psi^{12}\Psi_{22} &= 0 \\ \Psi^{21}\Psi_{11} + \Psi^{22}\Psi_{21} &= 0 \\ \Psi^{21}\Psi_{12} + \Psi^{22}\Psi_{22} &= I_{m-s} \end{cases}$$

Après des simplifications, on trouve que

$$\left\{\begin{array}{lcl}
\Psi^{11} & = & (\Psi_{11} - \Psi_{12}\,\Psi_{22}^{-1}\,\Psi_{21})^{-1} \\
\Psi^{22} & = & (\Psi_{22} - \Psi_{21}\,\Psi_{11}^{-1}\,\Psi_{12})^{-1} = \Psi_{22.1}^{-1} \\
\Psi^{12} & = & -\Psi_{11}^{-1}\,\Psi_{12}\,\Psi_{22.1}^{-1} \\
\Psi^{21} & = & -\Psi_{22.1}^{-1}\,\Psi_{21}\,\Psi_{11}^{-1} = (\Psi^{12})'
\end{array}\right.$$

De ce fait, pour tout $x \in \mathcal{M}_{m,n}(\mathbb{R})$

$$\begin{array}{rcl}
(x-M)'\Psi^{-1}(x-M) & = & \Big((x_{1r}-M_{1r})'(x_{2r}-M_{2r})'\Big)\begin{pmatrix}\Psi^{11} & \Psi^{12}\\ \Psi^{21} & \Psi^{22}\end{pmatrix}\begin{pmatrix}x_{1r}-M_{1r}\\ x_{2r}-M_{2r}\end{pmatrix} \\
& = & (x_{1r}-M_{1r})'\Psi^{11}(x_{1r}-M_{1r}) + (x_{1r}-M_{1r})'\Psi^{12}(x_{2r}-M_{2r}) \\
& + & (x_{2r}-M_{2r})'\Psi^{21}(x_{1r}-M_{1r}) + (x_{2r}-M_{2r})'\Psi^{22}(x_{2r}-M_{2r}) \\
& = & (x_{1r}-M_{1r})'(\Psi^{11}-\Psi^{12}(\Psi^{22})^{-1}\Psi^{21})(x_{1r}-M_{1r}) + (x_{1r}-M_{1r})' \\
& \times & \Psi^{12}(\Psi^{22})^{-1}\Psi^{21}(x_{1r}-M_{1r}) + (x_{1r}-M_{1r})'\Psi^{12}(x_{2r}-M_{2r}) \\
& + & (x_{2r}-M_{2r})'\Psi^{21}(x_{1r}-M_{1r}) + (x_{2r}-M_{2r})'\Psi^{22}(x_{2r}-M_{2r}) \\
& = & (x_{1r}-M_{1r})'\Psi_{11}^{-1}(x_{1r}-M_{1r}) + (x_{2r}-M_{2r}-\Psi_{21}\Psi_{11}^{-1}(x_{1r}-M_{1r}))' \\
& \times & \Psi_{22.1}^{-1}(x_{2r}-M_{2r}-\Psi_{21}\Psi_{11}^{-1}(x_{1r}-M_{1r})).
\end{array}$$

La fonction de densité d'une variable X est

$$\begin{array}{rcl}
f_X(x) & = & \dfrac{1}{(2\pi)^{\frac{mn}{2}}(\det\Psi)^{\frac{n}{2}}(\det\Sigma)^{\frac{m}{2}}}\ \exp\left(-\dfrac{1}{2}\,\mathrm{tr}\,\left(\Psi^{-1}(x-M)\Sigma^{-1}(x-M)'\right)\right) \\
& = & \dfrac{1}{(2\pi)^{\frac{mn}{2}}(\det\Psi)^{\frac{n}{2}}(\det\Sigma)^{\frac{m}{2}}}\ \exp\left(-\dfrac{1}{2}\,\mathrm{tr}\,\left((x-M)'\Psi^{-1}(x-M)\Sigma^{-1}\right)\right).
\end{array}$$

En utilisant la propriété suivante (voir [12])

$$\det(\Psi) = \det(\Psi_{11})\det(\Psi_{22.1}),$$

nous trouvons que

$$\begin{array}{l}
f_X(x) = \dfrac{1}{(2\pi)^{\frac{sn}{2}}(\det\Psi_{11})^{\frac{n}{2}}(\det\Sigma)^{\frac{s}{2}}}\ \exp\left(-\dfrac{1}{2}\,\mathrm{tr}\,\left(\Psi_{11}^{-1}(x_{1r}-M_{1r})\Sigma^{-1}(x_{1r}-M_{1r})'\right)\right) \\
\times\ \dfrac{\exp\left(-\frac{1}{2}\,\mathrm{tr}\,\left((x_{2r}-M_{2r}-\Psi_{21}\Psi_{11}^{-1}(x_{1r}-M_{1r}))'\Psi_{22.1}^{-1}(x_{2r}-M_{2r}-\Psi_{21}\Psi_{11}^{-1}(x_{1r}-M_{1r}))\Sigma^{-1}\right)\right)}{(2\pi)^{\frac{(m-s)n}{2}}(\det\Psi_{22.1})^{\frac{n}{2}}(\det\Sigma)^{\frac{m-s}{2}}}.
\end{array}$$

Donc

i) $X_{1r} \sim \mathcal{MN}_{s,n}(M_{1r}, \Psi_{11} \otimes \Sigma)$

ii) $X_{2r}|X_{1r} = x_{1r} \sim \mathcal{MN}_{m-s,n}(M_{2r} + \Psi_{21}\Psi_{11}^{-1}(x_{1r} - M_{1r}), \Psi_{22.1} \otimes \Sigma)$

$\square$

Corollaire 2.3.6.
Si X_{1r} et X_{2r} sont indépendantes, alors

$$X_{1r} \sim \mathcal{MN}_{s,n}(M_{1r}, \Psi_{11} \otimes \Sigma)$$

$$X_{2r} \sim \mathcal{MN}_{m-s,n}(M_{2r}, \Psi_{22} \otimes \Sigma)$$

Exemple 2.3.4. *Si $m = 3, n = 2$ et $s = 1$. On a*

$$X = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \\ X_{31} & X_{32} \end{pmatrix} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix},$$

avec $X_1(1 \times 2)$ est vecteur ligne et $X_2 \in M_{2,2}(\mathbb{R})$ et

$$M = \begin{pmatrix} \mu_{11} & \mu_{12} \\ \mu_{21} & \mu_{22} \\ \mu_{31} & \mu_{32} \end{pmatrix} = \begin{pmatrix} M_1 \\ M_2 \end{pmatrix},$$

avec $M_1(1 \times 2)$ est vecteur ligne et $M_2 \in M_{2,2}(\mathbb{R})$.
La matrice de covariance de ligne Ψ est une matrice définie positive de dimension (3×3) donnée par :

$$\Psi = \begin{pmatrix} \psi_{11} & \psi_{12} & \psi_{13} \\ \psi_{21} & \psi_{22} & \psi_{23} \\ \psi_{31} & \psi_{32} & \psi_{33} \end{pmatrix} = \begin{pmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{21} & \Psi_{22} \end{pmatrix},$$

avec $\Psi_{11} = \psi_{11} \in \mathbb{R}^{*+}$, $\Psi_{12} = (\psi_{12} \;\; \psi_{13})$ est un vecteur ligne, $\Psi_{21} = \Psi_{12}'$ et $\Psi_{22} = \begin{pmatrix} \psi_{22} & \psi_{23} \\ \psi_{32} & \psi_{33} \end{pmatrix}$ est une matrice symétrique définie positive de dimension (2×2).
La fonction de densité de variable X est :

$$\begin{aligned} f_X(x) &= \frac{1}{(2\pi)^3 \det(\Psi) \det(\Sigma)^{\frac{3}{2}}} \exp\left(-\frac{1}{2} \operatorname{tr} \left((x-M)'\Psi^{-1}(x-M)\Sigma^{-1}\right)\right) \\ &= \frac{1}{(2\pi) \det(\Psi_{11}) \det(\Sigma)^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \operatorname{tr} \left((x_1-M_1)'\Psi_{11}^{-1}(x_1-M_1)\Sigma^{-1}\right)\right) \\ &\times \frac{\exp\left(-\frac{1}{2} \operatorname{tr} \left((x_2 - M_2 - \Psi_{21}\Psi_{22}^{-1}(x_1-M_1))'\Psi_{22.1}^{-1}(x_2 - M_2 - \Psi_{21}\Psi_{11}^{-1}(x_1-M_1))\,\Sigma^{-1}\right)\right)}{(2\pi)^2 \det(\Psi_{22.1}) \det(\Sigma)} \end{aligned}$$

En utilisant les propriétés iii) du Lemme (2.1.6) et i) de la Proposition (2.1.4), on trouve que

$$
\begin{aligned}
f_X(x) &= \frac{1}{(2\pi)\det(\Psi_{11})\det(\Sigma)^{\frac{1}{2}}}\exp\left(-\frac{1}{2\Psi_{11}}\left((x_1-M_1)\Sigma^{-1}(x_1-M_1)^{'}\right)\right)\\
&\times \frac{\exp\left(-\frac{1}{2}\operatorname{tr}\left((x_2-M_2-\Psi_{21}\Psi_{22}^{-1}(x_1-M_1))'\Psi_{22.1}^{-1}(x_2-M_2-\Psi_{21}\Psi_{11}^{-1}(x_1-M_1))\,\Sigma^{-1}\right)\right)}{(2\pi)^2\det(\Psi_{22.1})\det(\Sigma)}.
\end{aligned}
$$

Donc

i) $X_1 \sim \mathcal{N}_2(M_1, \Psi_{11}\Sigma)$,

ii) $X_2|X_1 = x_1 \sim \mathcal{MN}_{2,2}(M_2 + \Psi_{21}\Psi_{11}^{-1}(x_1 - M_1), \Psi_{22.1} \otimes \Sigma)$.

Conclusion Générale

Dans la littérature sur l'analyse multidimensionnelle les matrice aléatoires reste encore un sujet d'actualité soit ce qui concerne la modélisation, ou bien l'estimation des différent paramètres. Ce travail est juste une introduction aux modèles aléatoires matricielles, qui sont nécessaires dans la modélisation de big data , où plusieurs échantillons sont extrais en même temps.

Nous avons donné quelques outils de base de la théorie des vecteurs aléatoires et nous avons étudié les matrices aléatoires dont la loi est normale multivariée. De plus, nous avons identifié les moments aléatoires (la moyenne, variance et covariance) et la fonction caractéristique qui jouent un rôle déterminant dans l'étude de la fonction de densité du modèle.

Nous introduisons également des outils matriciels (le produit de Kroneker, l'opérateur Vec et la matrice de commutation) qui facilitent l'étude d'une matrice aléatoire pour trouver la loi normal matriciel , sa forme et ses propriétés les plus importantes.

I. Annexe

Quelques résultats utiles d'algèbre linéaire

Définition

• Une matrice carrée A est dite symétrique si $A' = A$.

• Une matrice carrée A de $\mathcal{M}_{n,n}(\mathbb{K})$ est dite inversible, s'il existe une matrice B de $\mathcal{M}_{n,n}(\mathbb{K})$ telle que : $AB = I_n$ et $BA = I_n$. La matrice B est alors appelée la matrice inverse de A et on écrit $B = A^{-1}$.

• Une matrice Σ est inversible si $\det \Sigma \neq 0$.
Si Σ est inversible alors $\Sigma^{-1} = \frac{1}{\det(\Sigma)}\, com(\Sigma)^t$,
où $C = com(\Sigma)$ est la matrice de coefficients avec $C_{ij} = (-1)^{i+j} \det \Sigma_{ij}$

• Une matrice symétrique A d'ordre n est semi-définie positive si une de ces propriétés est vérifiée :

1) $\forall x \in \mathbb{R}^n, \quad x'Ax \geq 0$.
2) Les valeurs propres de la matrices sont positives ou nulles.
3) Les mineurs principaux sont positives ou nulles.

et définie positive : pour A est une matrice symétrique.

• Une matrice symétrique A d'ordre n est définie positive si une de ces propriétés est vérifiée :

1) $\forall x \in \mathbb{R}^n, \quad x'Ax > 0$.
2) Les valeurs propres de la matrices sont positives.
3) Les mineurs principaux sont positives.

II. Code de R

Chapitre 1 : Loi normale vectorielle

```
#library :charger et joindre des packages complémentaires
library(mvtnorm)
library(lattice)
# N2(c(2,1),matrix(c(3,0.5,0.5,3),byrow=T,nrow=2))
var <- matrix(c(3,0.5,0.5,3),byrow=T,nrow=2)
#Le code suivant explique comment appliquer la fonction expand.grid pour renvoyer chaque combinaison possible de deux variables factorielles.
g <- expand.grid(x = seq(-2,2,0.05), y = seq(-2,2,0.05))
#Le package dans R permettant d'obtenir la densité, les quantiles ou de générer des réalisations de lois normales multivariées est le package mvtnorm.
g$z <- dmvnorm(x=cbind(g$x,g$y),mean=c(2,1), sigma=var, log=FALSE)
wireframe( z ~ x * y,data = g,xlab=" ",ylab=" ",zlab=" ")
# N2(c(0,0),I_2)
g <- expand.grid(x = seq(-2,2,0.05), y = seq(-2,2,0.05))
g$z <- dmvnorm(x=cbind(g$x,g$y),mean=c(0,0), sigma=diag(2),log=FALSE)
wireframe(z ~ x * y,data=g,xlab=" ",ylab=" ",zlab=" ")
```

Chapitre 2 : Loi normale matricielle

```
library(LaplacesDemon)
N <- 10
K <- 4
Psi <- as.positive.definite(matrix(rnorm(N*N),N,N))
Sigma <- as.positive.definite(matrix(rnorm(K*K),K,K))
#Ces commandes fournissent la densité et la génération de nombres aléatoires pour la distribution normale de la matrice.
x <- dmatrixnorm(matrix(0,N,K), matrix(0,N,K), Psi, Sigma)
```

```
X <- rmatrixnorm(matrix(0,N,K), Psi, Sigma)
# Tracer graphiquement la fonction de densité de loi normale matricielle
joint.density.plot(X[,1], X[,2], color=TRUE)
```

Bibliographie

[1] **Arashib,A ; Tabatabaeya,S.M.M**. (2010). *On Conditional Applications of Matrix Variate Normal Distribution.*

[2] **Bogaert,P**. (2006). *Probabilités pour scientifiques et ingénieurs, Paris, De Boeck.*

[3] **Brez,H ; Douglas,R.G**. (2012). *Matrix variate Distributions.Jeffrey, University of Newcastle upon Tyne (Founding Editor)*

[4] **Craig,A**. (2011). *The Distributions of Random Matrix Theory and their Applications.Department of Mathematics UC Davis.*

[5] **Dauxois, J.Y**. (2014). Cours de Probabilités.

[6] **Deheuvels,P**. (2012). *Cours d'analyse statistique multivariée.*

[7] **Ghazal,A ;Neudecker,H**. (2000). *On second-order and fourth-order moments of jointly distributed random matrices : a survey.*

[8] **Gupta,A.K ; Nagar,D** (1999). *Matrix variate distributions.*

[9] **Henderson,H. V ; Searle,S.R** (1981). *The vec-permutation Matrix, the vec operator and Kroneker product : a review. Linar and Multilinear Algebra.* 9(4), 271-288.

[10] **Johson,R ;Wichern,D**. (2017). *Applied multivariate statistical analysis.*

[11] **Magnus,J.R ; Neudecker, H** (1986). *Symmetry, 0-1 matrices and Jacobians : A review. Econometric Theory.* 2(2), 157-190.

[12] **Muirhead,R**. (1982). *Aspects of multivariate statistical theory.*

[13] **Neudecker,H**. (1987). *Fourth-Order Properties of Normally Distributed Random Matrices.Department of Econometric University of Amsterdam.*

[14] **Rosenblatt,M**.*An Identity for the Wishart Distribution with Applications*.University of California, San Diego*

[15] **Pishro,N. H** . (2014). *Introduction to Probability, Statistics, and Random Processes.*

[16] **Rush,J.J**. (2013). Vecteurs gaussiens, préparation à l'agrégation Bordeaux 1.

[17] **Saporta,G**. (2006). *Probabilités, analyses des données et statistiques.*