

Chapter 1

Vecteurs Gaussiens

1.1 Loi normale

On commence par rappeler certaines propriétés de la loi normale.

Définition 1.1. Soient $\mu \in \mathbb{R}$ et $\sigma > 0$ deux paramètres. On appelle variable aléatoire de loi normale, ou gaussienne, de paramètres μ et σ toute v.a. X de densité : pour tout $x \in \mathbb{R}$

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

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

Bien noter que le deuxième paramètre dans la notation précédente est σ^2 la variance (et pas σ l'écart type).

Les propriétés qui suivent sont bien connues retrouver leurs preuves ne doit pas poser de problème au lecteur.

Propriété 1.2. Si $X \sim \mathcal{N}(\mu, \sigma^2)$ alors

1. $\mathbb{E}(X) = \mu$ et $\mathbf{Var}(X) = \sigma^2$.
2. La v.a. $Z = (X - \mu)/\sigma$ suit une loi $\mathcal{N}(0, 1)$ qui est appelée loi normale centrée réduite, ou standard.
3. On pourra toujours écrire $X = \mu + \sigma Z$ avec $Z \sim \mathcal{N}(0, 1)$.

Le résultat qui suit donne une idée du poids des queues de distribution (ensemble de valeurs éloignées de la moyenne) de la loi normale standard.

Proposition 1.3. Queue de distribution de la loi $\mathcal{N}(0, 1)$.

Soit $Z \sim \mathcal{N}(0, 1)$. Lorsque t tends vers l'infini on a l'équivalent

$$\mathbb{P}(Z \geq t) \sim \frac{1}{t\sqrt{2\pi}} e^{-t^2/2}$$

La preuve de cet équivalent est une conséquence d'un encadrement que l'on verra en exercice.

Par voie de conséquence : si $X \sim \mathcal{N}(\mu, \sigma^2)$ on a

$$\mathbb{P}(|X - \mu| \geq t) \sim \frac{2\sigma}{t\sqrt{2\pi}} e^{-t^2/(2\sigma^2)}$$

1.2 Fonction caractéristique

On rappelle que la fonction caractéristique d'une v.a. X est une fonction à valeurs dans \mathbb{C} telle que $\forall t \in \mathbb{R}$ on a $\Psi_X(t) = \mathbb{E}(e^{itX})$. Cette fonction existe toujours et caractérise la loi de la variable aléatoire au même titre que la fonction de répartition (ou la densité si la v.a. est absolument continue).

C'est une fonction uniformément continue sur \mathbb{R} et telle que $\Psi_X(0) = 1$. De plus elle est bornée : $|\Psi_X(t)| \leq 1$.

Lorsque X est une v.a. à densité f_X sa fonction caractéristique correspond à un type de transformée de Fourier de f_X

$$\int_{\mathbb{R}} e^{itx} f_X(x) dx$$

De plus si X admet une fonction génératrice des moments $M_X(t)$, i.e. type de transformée de Laplace de f_X ,

$$\int_{\mathbb{R}} e^{tx} f_X(x) dx$$

qui n'existe pas toujours, alors on a $\forall t \in \mathbb{R}$

$$M_X(t) = \Psi_X(-it) = \mathbb{E}(e^{tX})$$

Le calcul de la fonction caractéristique fait intervenir le calcul intégral dans \mathbb{C} .

Remarque : On observera qu'en "retournant" la formule précédente on obtient

$$\Psi_X(t) = M_X(it)$$

Ce qui permet en pratique (pour les cas où la FGM existe) de trouver les fonctions caractéristiques sans passer par de l'intégration complexe.

Proposition 1.4. *On a*

1. Soit $Z \sim \mathcal{N}(0, 1)$ pour tout $t \in \mathbb{R}$ on a

$$\Psi_Z(t) := \mathbb{E}[e^{itZ}] = e^{-t^2/2}$$

2. Soit $X \sim \mathcal{N}(\mu, \sigma^2)$ sa fonction caractéristique est donnée par

$$\forall t \in \mathbb{R}, \Psi_X(t) := \mathbb{E}[e^{itX}] = \exp\left(it\mu - \frac{\sigma^2 t^2}{2}\right)$$

Preuve. Commençons donc par calculer la fonction génératrice des moments : On a pour tout $t \in \mathbb{R}$

$$M_Z(t) = \mathbb{E}[e^{tZ}] = \int_{\mathbb{R}} e^{tx} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx = e^{t^2/2} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-t)^2}{2}\right) dx = e^{t^2/2}$$

d'où la fonction caractéristique par la remarque ci dessus

$$\Psi_Z(t) = M_Z(it) = e^{-t^2/2}$$

Par ailleurs en écrivant $X = \mu + \sigma Z$ avec $Z \sim \mathcal{N}(0, 1)$ on a

$$\mathbb{E}[e^{itX}] = \mathbb{E}[e^{it(\mu + \sigma Z)}] = e^{it\mu} \mathbb{E}[e^{it\sigma Z}] = \exp\left(it\mu - \frac{\sigma^2 t^2}{2}\right)$$

Remarque : en utilisant l'intégration dans le plan complexe on peut montrer que $\forall \zeta \in \mathbb{C}$ on a $\mathbb{E}[e^{\zeta Z}] = e^{\zeta^2/2}$. □

1.3 Stabilité par convergence en loi

On rappelle que la convergence en loi est équivalente à la convergence des fonctions caractéristiques vers une fonction limite Ψ , continue en 0 qui est alors la fonction caractéristique de la loi limite.

Un résultat remarquable des lois Gaussienne est :

Proposition 1.5. *Soit $(X_n)_{n \geq 1}$ une suite de v.a. de lois respectives $\mathcal{N}(\mu_n, \sigma_n^2)$ alors les 2 propriétés suivantes sont équivalentes :*

1. *La suite (X_n) converge en loi.*
2. $\lim_{n \rightarrow +\infty} \mu_n = \mu \in \mathbb{R}$ et $\lim_{n \rightarrow +\infty} \sigma_n^2 = \sigma^2 \in \mathbb{R}^+$.

et dans ces cas la loi limite est $\mathcal{N}(\mu, \sigma^2)$.

Preuve. Clairement 2 implique 1. En effet par continuité de la fonction exp on a

$$\lim_{n \rightarrow +\infty} \Psi_{X_n}(u) = \exp\left(iu\mu - \frac{\sigma^2 u^2}{2}\right) = \Psi(u)$$

qui est bien une fonction continue en $u = 0$ et qui de plus est la FC d'une v.a. de loi $\mathcal{N}(\mu, \sigma^2)$.

Montrer que 1 implique 2 est un peu plus laborieux :

Par hypothèse on a $\Psi_{X_n}(u) \rightarrow \Psi(u)$ où Ψ est une fonction continue en 0. On remarque au passage que comme pour tout n on a $\Psi_{X_n}(0) = 1$ alors on a $\Psi(0) = 1$. Par la convergence des fonctions on a la convergence des modules :

$$|\Psi_{X_n}(u)| = \exp\left(-\frac{u^2 \sigma_n^2}{2}\right) \rightarrow |\Psi(u)| \text{ quand } n \rightarrow +\infty$$

on en déduit que la limite de σ_n^2 existe, on la note σ^2 et on a

$$|\Psi(u)| = \exp\left(-\frac{u^2 \sigma^2}{2}\right)$$

Pour la partie imaginaire on a $\forall u \in \mathbb{R}$

$$\exp(iu\mu_n) \rightarrow \Psi(u) \exp\left(\frac{u^2 \sigma^2}{2}\right) := \varphi(u) \text{ quand } n \rightarrow +\infty \quad (1.1)$$

Par convergence dominée on a $\forall r > 0$

$$\lim_{n \rightarrow +\infty} \int_0^r \exp(iu\mu_n) du = \int_0^r \varphi(u) du$$

On va distinguer deux cas suivant que 0 n'est pas ou est valeur d'adhérence de la suite $(\mu_n)_{n \geq 1}$.

1er cas : Si 0 n'est pas valeur d'adhérence alors par intégration on a $\forall r > 0$

$$\int_0^r \exp(iu\mu_n) du = \frac{1}{i\mu_n} [\exp(ir\mu_n) - 1]$$

ou de manière équivalente

$$\mu_n = \frac{\exp(ir\mu_n) - 1}{i \int_0^r \exp(iu\mu_n) du} \quad (1.2)$$

on a que $\varphi(0) = \Psi(0) \cdot 1 = 1$ et comme φ est continue en 0 donc on peut trouver un $r_0 > 0$ tel que

$$\int_0^{r_0} \varphi(u) du \neq 0 \quad (1.3)$$

comme

$$\lim_{n \rightarrow +\infty} \exp(ir_0\mu_n) = \varphi(r_0)$$

on déduit de (1.1), (1.2) et (1.3) que μ_n converge vers

$$\mu := \frac{\varphi(r_0) - 1}{i \int_0^{r_0} \varphi(u) du}.$$

2ème cas : A présent si 0 est valeur d'adhérence de $(\mu_n)_{n \geq 0}$ alors il existe une sous suite $(\mu_{n_k})_{k \geq 0}$, avec n_k tendant vers $+\infty$ telle que μ_{n_k} converge vers 0, par suite

$$\forall u \in \mathbb{R}, \lim_{k \rightarrow +\infty} \exp(iu\mu_{n_k}) = 1$$

donc $\varphi(u) = 1, \forall u \in \mathbb{R}$ et par conséquent

$$\lim_{n \rightarrow +\infty} \exp(iu\mu_n) = \varphi(u) = 1, \forall u \in \mathbb{R} \quad (1.4)$$

Si on suppose que μ_n ait une autre valeur d'adhérence $\gamma \neq 0$ alors il existe une autre sous suite $(\mu_{n_j})_{j \in \mathbb{N}}$, avec n_j tendant vers $+\infty$ telle que

$$\lim_{j \rightarrow +\infty} \mu_{n_j} = \gamma$$

et $\forall r > 0$

$$\int_0^r \exp(iu\mu_{n_j}) du = \frac{1}{i\mu_{n_j}} [\exp(ir\mu_{n_j}) - 1] \xrightarrow{(1.4)} 0 \text{ quand } j \rightarrow +\infty$$

qui entraîne que $\forall r > 0$

$$\int_0^r \varphi(u) du = 0$$

ce qui contredit $\varphi(0) = 1$ et φ continue. Cela veut donc dire que 0 est alors l'unique valeur d'adhérence et donc

$$\lim_{n \rightarrow +\infty} \mu_n = 0 = \mu.$$

□

Exercice :

Trouvez des contre-exemples au résultat précédent lorsque la loi n'est pas gaussienne.

1.4 Vecteurs gaussiens

Notation : dans ce qui suit on supposera que si X est un vecteur de \mathbb{R}^d il s'agit d'un vecteur colonne. Le vecteur ligne s'écrira X^T .

Définition 1.6. Un vecteur aléatoire X à valeurs dans \mathbb{R}^d est un vecteur gaussien si et seulement si $\forall a \in \mathbb{R}^d$ la v.a.

$$\langle a, X \rangle = a^T \cdot X = \sum_{i=1}^d a_i X_i$$

est de loi normale (éventuellement dégénérée).

Proposition 1.7. Un vecteur aléatoire dont toutes les composantes sont gaussiennes et indépendantes est un vecteur gaussien.

Preuve. Soient donc X un vecteur aléatoire dont les composantes sont gaussiennes et indépendantes et $a \in \mathbb{R}^d$ on a $\forall u \in \mathbb{R}$

$$\mathbb{E} \left[e^{iu \sum_{i=1}^d a_i X_i} \right] = \prod_{i=1}^d \mathbb{E} \left[e^{iu a_i X_i} \right] = \prod_{i=1}^d \exp \left(iu a_i \mu_i - \frac{\sigma_i^2 a_i^2 u^2}{2} \right) = \exp \left(iu \sum_{i=1}^d a_i \mu_i - \frac{u^2 \sum_{i=1}^d \sigma_i^2 a_i^2}{2} \right)$$

on en déduit que la v.a. $\sum_{i=1}^d a_i X_i$ est de loi $\mathcal{N}(\sum_{i=1}^d a_i \mu_i, \sum_{i=1}^d a_i^2 \sigma_i^2)$. □

□

Exercice :

Soient X_1 de loi $\mathcal{N}(0, 1)$, U une v.a. indépendante de X_1 de loi $\mathbb{P}(U = -1) = \mathbb{P}(U = 1) = \frac{1}{2}$ et enfin $X_2 = UX_1$.

1. Montrer que X_2 est gaussienne;
2. Calculer $\mathbf{Cov}(X_1, X_2)$;
3. Calculer $\Psi_{X_1+X_2}$ et en déduire que $X_1 + X_2$ n'est pas gaussien.
4. Conclure.

Définition 1.8. Etant donné un vecteur aléatoire X à valeurs dans \mathbb{R}^d , on définit sa matrice de covariance comme la matrice $K_X = (k_{ij})_{1 \leq i, j \leq d}$, carré $d \times d$ telle que

$$k_{ij} = \mathbf{Cov}(X_i, X_j)$$

Elle rassemble toutes les covariances des composantes du vecteur X et sur sa diagonale on a les variances des composantes.

Il est bien entendu que pour que cette matrice existe il faut que toutes les composantes soient de carré intégrable. Sauf indication contraire on supposera cette hypothèse toujours satisfaite dans ce qui suit.

Cette matrice est symétrique semi-définie positive (i.e. toutes ses valeurs propres sont positives ou nulles ou de façon équivalente $\forall u \in \mathbb{R}^d$ on a $u^T \cdot K_X \cdot u \geq 0$).

On peut également définir une matrice de covariance de deux vecteurs aléatoires :

Définition 1.9. Etant donnés X un vecteur aléatoire de \mathbb{R}^d et Y un vecteur aléatoire de \mathbb{R}^n on définit la matrice de covariance de X et Y comme la matrice $\mathbf{COV}(X, Y) \in \mathcal{M}_{d \times n}$ telle que

$$\mathbf{COV}(X, Y)_{ij} = \mathbf{Cov}(X_i, Y_j) \quad \forall 1 \leq i \leq d, 1 \leq j \leq n$$

que l'on note également

$$\mathbf{COV}(X, Y) = \mathbb{E}[(X - \mathbb{E}(X)) \cdot (Y - \mathbb{E}(Y))^T] = \mathbb{E}(X \cdot Y^T) - \mathbb{E}(X) \cdot \mathbb{E}(Y)^T$$

Avec les notations précédente on a $\mathbf{COV}(X, X) = K_X$.

Lemme 1.10. Si X est un vecteur de \mathbb{R}^d de moyenne μ et de matrice de covariance K_X et si A est une matrice réelle de $\mathcal{M}_{m \times d}$ alors le vecteur aléatoire $A \cdot X$ de \mathbb{R}^m a pour moyenne $A \cdot \mu$ et pour matrice de covariance

$$K_{A \cdot X} = A \cdot K_X \cdot A^T \in \mathcal{M}_{m \times m}.$$

Preuve. Par linéarité de l'espérance

$$\mathbb{E}(A \cdot X) = A \cdot \mathbb{E}(X) = A \cdot \mu$$

et

$$\mathbb{E}[(A \cdot X - A \cdot \mu) \cdot (A \cdot X - A \cdot \mu)^T] = \mathbb{E}[A \cdot (X - \mu) \cdot (X - \mu)^T \cdot A^T] = A \cdot K_X \cdot A^T$$

□

Proposition 1.11. Fonction caractéristique d'un vecteur gaussien.

Un vecteur aléatoire X de \mathbb{R}^d est un vecteur gaussien si et seulement si il existe $\mu \in \mathbb{R}^d$ et $\Sigma^2 \in \mathcal{M}_{d \times d}$ symétrique semi-définie positive tels que $\forall u \in \mathbb{R}^d$

$$\Psi_X(u) := \mathbb{E}(e^{iu^T \cdot X}) = \exp\left(iu^T \cdot \mu - \frac{u^T \Sigma^2 u}{2}\right)$$

Remarques : On déduit de ce résultat que si X et Y sont gaussiens tels que $\mu = \mathbb{E}(X) = \mathbb{E}(Y)$ et $\Sigma^2 = K_X = K_Y$ alors $X = Y$. Cette proposition peut également servir de définition d'un vecteur gaussien. L'hypothèse Σ^2 symétrique semi-définie positive autorise le cas dégénéré (i.e. $\det(\Sigma^2) = 0$).

Preuve. Condition suffisante : si X est gaussien alors ses composantes sont toutes gaussiennes de plus X est de carré intégrable et μ et Σ^2 sont bien définis. Soit $u \in \mathbb{R}^d$ alors par le lemme précédent la variable aléatoire $u^T \cdot X$ a pour moyenne $\mu_{u^T \cdot X} = u^T \cdot \mu$ et variance $K_{u^T \cdot X} = u^T \cdot \Sigma^2 \cdot u$ et est de loi gaussienne. De plus par la Proposition 1.4 on a

$$\Psi_X(u) := \mathbb{E}[e^{iu^T \cdot X}] = \Psi_{u^T \cdot X}(1) = \exp\left(iu^T \cdot \mu - \frac{u^T \cdot \Sigma^2 \cdot u}{2}\right).$$

Condition nécessaire : Ceci découle des égalités $\forall t \in \mathbb{R}$ et $\forall a \in \mathbb{R}^d$

$$\begin{aligned} \Psi_{a^T \cdot X}(t) &= \mathbb{E}[\exp(it a^T \cdot X)] = \Psi_X(ta) \\ &= \exp\left(it a^T \cdot \mu - \frac{(ta)^T \cdot \Sigma^2 \cdot (ta)}{2}\right) \\ &= \exp\left(it(a^T \cdot \mu) - t^2 \frac{a^T \cdot \Sigma^2 \cdot a}{2}\right) \end{aligned}$$

qui montre que $a^T \cdot X$ est gaussien.

□

Proposition 1.12. Soit $(X^T, Y^T)^T$ un vecteur gaussien. Il y a équivalence entre

1. $\mathbf{COV}(X, Y) = \mathbf{0}$ (matrice nulle)
2. Les vecteurs aléatoires X et Y sont indépendants.

Remarque : On retiendra donc que dès que l'on sait qu'un vecteur est gaussien alors ses composantes sont indépendantes si et seulement si leur covariance est nulle. Ceci explique l'intérêt de travailler avec des vecteurs gaussiens.

Preuve. 2. entraîne 1. est immédiat par définition de la covariance.

Pour montrer que 1. entraîne 2. on écrit

$$\Psi_{(X,Y)}(u,v) = \mathbb{E} \left[\exp(i(u^T, v^T) \cdot (X, Y)) \right] = \exp \left[i(u^T, v^T) \cdot \mathbb{E}(X, Y) - \frac{1}{2}(u^T, v^T) \cdot K_{(X,Y)} \cdot (u, v) \right]$$

or

$$(u^T, v^T) \cdot \mathbb{E}(X, Y) = u^T \cdot \mu_X + v^T \cdot \mu_Y$$

et

$$K_{(X,Y)} = \begin{pmatrix} K_X & 0 \\ 0 & K_Y \end{pmatrix}$$

d'où

$$(u^T, v^T) \cdot K_{(X,Y)}(u, v) = u^T \cdot K_X \cdot u + v^T \cdot K_Y \cdot v$$

donc $\forall u, v$

$$\Psi_{(X,Y)}(u, v) = \Psi_X(u) \Psi_Y(v)$$

ce qui établit l'indépendance. □

Proposition 1.13. Soient X un vecteur gaussien de \mathbb{R}^d de paramètres $\mu \in \mathbb{R}^d$ et $\Sigma^2 \in \mathcal{M}_{d \times d}$. On considère M une matrice $\mathcal{M}_{n \times d}$ et b un vecteur de \mathbb{R}^n . Alors le vecteur aléatoire de \mathbb{R}^n défini par

$$M \cdot X + b$$

est gaussien de paramètres $M \cdot \mu + b$ et $M \cdot \Sigma^2 \cdot M^T$.

Remarque : On retiendra que toute transformation affine d'un vecteur gaussien est un vecteur gaussien.

Preuve. Il suffit de calculer sa fonction caractéristique : $\forall u \in \mathbb{R}^n$

$$\begin{aligned} \Psi_{M \cdot X + b}(u) &= \mathbb{E}[\exp(iu^T \cdot (M \cdot X + b))] \\ &= e^{iu^T \cdot b} \mathbb{E}[\exp(iu^T \cdot M \cdot X)] = e^{iu^T \cdot b} \mathbb{E}[\exp(i(M^T \cdot u)^T \cdot X)] \\ &= e^{iu^T \cdot b} \Psi_X(M^T \cdot u) \\ &= e^{iu^T \cdot b} \exp \left(i(M^T \cdot u)^T \cdot \mu - \frac{(M^T \cdot u)^T \cdot \Sigma^2 \cdot M^T \cdot u}{2} \right) \\ &= \exp \left(iu^T \cdot (M \cdot \mu + b) - \frac{u^T \cdot M \cdot \Sigma^2 \cdot M^T \cdot u}{2} \right) \end{aligned}$$

□

Le résultat qui suit est une conséquence d'un théorème d'algèbre matricielle.

Théorème 1.14. (admis, conséquence du Théorème de Schur.).

1. Soit Σ^2 une matrice symétrique semi-définie positive. Alors il existe une matrice symétrique notée Σ vérifiant $\Sigma \cdot \Sigma = \Sigma^2$.
2. Lorsque Σ^2 est non dégénérée alors Σ est inversible d'inverse $\Sigma^{-1} = ((\Sigma^2)^{-1})^{1/2}$.

Proposition 1.15. Densité d'un vecteur gaussien non dégénéré.

Soient X un vecteur gaussien de \mathbb{R}^d de paramètres $\mu \in \mathbb{R}^d$ et $\Sigma^2 \in \mathcal{M}_{d \times d}$. On suppose $\det(\Sigma^2) \neq 0$ (cas non dégénéré). Alors X possède une densité de probabilité f_X sur \mathbb{R}^d définie par : $\forall x \in \mathbb{R}^d$

$$f_X(x) = \frac{1}{(2\pi)^{d/2} (\det(\Sigma^2))^{1/2}} \exp \left[-\frac{(x - \mu)^T \cdot (\Sigma^2)^{-1} \cdot (x - \mu)}{2} \right]$$

Preuve. Soit Y un vecteur gaussien de \mathbb{R}^d de moyenne $\mathbf{0}$ et de matrice de covariance l'identité $I_d \in \mathcal{M}_{d \times d}$. Donc par la Proposition 1.12 ses coordonnées sont indépendantes et sa densité jointe s'écrit donc $\forall y \in \mathbb{R}^d$

$$f_Y(y) = \frac{1}{(2\pi)^{d/2}} \exp \left(-\frac{\|y\|^2}{2} \right)$$

On applique alors la Proposition 1.13 (et le théorème précédent) avec $M = \Sigma$ et $b = \mu$ alors X et $\Sigma \cdot Y + \mu$ ont même loi, ce qui implique que pour toute fonction g mesurable bornée on a

$$\mathbb{E}[g(X)] = \mathbb{E}[g(\Sigma \cdot Y + \mu)] = \int_{\mathbb{R}^d} g(\Sigma \cdot y + \mu) \frac{1}{(2\pi)^{d/2}} \exp \left(-\frac{\|y\|^2}{2} \right) dy$$

On pose alors $x = \Sigma \cdot y + \mu = \Phi(y)$. Le Jacobien de cette transformation (qui est un \mathcal{C}^1 difféomorphisme sur \mathbb{R}) est

$$J_{\Phi^{-1}}(x) = \det(\Sigma^{-1}) = (\det(\Sigma^2))^{-1/2}$$

d'où

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}^d} g(x) \frac{1}{(2\pi)^{d/2} (\det(\Sigma^2))^{1/2}} \exp\left(-\frac{\|\Sigma^{-1} \cdot (x - \mu)\|^2}{2}\right) dx$$

et on conclut en observant que

$$\|\Sigma^{-1} \cdot (x - \mu)\|^2 = ((x - \mu)^T \cdot \Sigma^{-1}) \cdot (\Sigma^{-1} \cdot (x - \mu)) = (x - \mu)^T \cdot (\Sigma^2)^{-1} \cdot (x - \mu)$$

et en identifiant la densité sous l'intégrale. □

1.5 Simulation des lois Multinormales

Nous venons de voir que les lois gaussiennes $\mathcal{N}_d(\mu, \Sigma^2)$ sont caractérisées par leur vecteur moyenne $\mu \in \mathbb{R}^d$ et leur matrice de covariance $\Sigma^2 \in \mathcal{M}_{d \times d}$ (symétrique semi-définie positive).

On a vu que si $Z \sim \mathcal{N}_d(\mathbf{0}, I_d)$ alors le vecteur

$$X = A \cdot Z + \mu$$

est de loi $\mathcal{N}_d(\mu, A \cdot A^T)$.

Par ailleurs on peut facilement simuler des variables aléatoires normales centrées réduites indépendantes Z_1, \dots, Z_d . Le vecteur $Z = (Z_1, \dots, Z_d)^T$ est gaussien de loi $\mathcal{N}_d(\mathbf{0}, I_d)$.

Pour simuler un vecteur gaussien de loi $\mathcal{N}_d(\mu, \Sigma^2)$ il suffit donc de trouver explicitement une matrice $A = (a_{ij})_{1 \leq i, j \leq d}$ telle que $A \cdot A^T = \Sigma^2$.

1.5.1 Factorisation de Cholesky

Parmi l'ensemble de telles matrices celle qui sont triangulaire inférieures sont plus pratiques. En effet le calcul de $A \cdot Z + \mu$ donne

$$\begin{aligned} X_1 &= a_{11}Z_1 + \mu_1 \\ X_2 &= a_{21}Z_1 + a_{22}Z_2 + \mu_2 \\ &\vdots \\ X_d &= a_{d1}Z_1 + a_{d2}Z_2 + \dots + a_{dd}Z_d + \mu_d \end{aligned}$$

Définition 1.16. *Factorisation de Cholesky*

Une représentation de Σ^2 en $A \cdot A^T$ avec A triangulaire inférieure est appelée factorisation de Cholesky.

Propriété 1.17. Si Σ^2 est définie positive alors Σ^2 admet une factorisation de Cholesky notée A . De plus A est unique au changement de signe près.

Exemple : Soient $\sigma_1 > 0$, $\sigma_2 > 0$ et $|\rho| \leq 1$ des paramètres constants et soit

$$\Sigma^2 = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}$$

alors on peut prendre

$$A = \begin{pmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sqrt{1-\rho^2} \sigma_2 \end{pmatrix}$$

et on peut représenter $X \sim \mathcal{N}_2(\mu, \Sigma^2)$ par

$$\begin{aligned} X_1 &= \sigma_1 Z_1 + \mu_1 \\ X_2 &= \rho\sigma_2 Z_1 + \sqrt{1-\rho^2} \sigma_2 Z_2 + \mu_2 \end{aligned}$$

où $Z = (Z_1, Z_2)^T \sim \mathcal{N}_2(\mathbf{0}, I_2)$.

En dimension d en résolvant $A \cdot A^T = \Sigma^2 = (\sigma_{ij}^2)_{1 \leq i, j \leq d}$ il vient

$$\begin{array}{ccccccc} a_{11}^2 = \sigma_{11}^2 & a_{11}a_{21} = \sigma_{12}^2 & \cdots & a_{11}a_{d1} = \sigma_{1d}^2 \\ a_{21}a_{11} = \sigma_{21}^2 & a_{21}^2 + a_{22}^2 = \sigma_{22}^2 & \cdots & \\ \vdots & & & \\ a_{d1}a_{11} = \sigma_{d1}^2 & & \cdots & a_{d1}^2 + \cdots + a_{dd}^2 = \sigma_{dd}^2 \end{array}$$

Formellement : pour tout $1 \leq i, j \leq d$

$$\sigma_{ij}^2 = \sum_{k=1}^j a_{ik}a_{jk}$$

On obtient ainsi pour $1 \leq j < i$

$$a_{ij} = \frac{\sigma_{ij}^2 - \sum_{k=1}^{j-1} a_{ik}a_{jk}}{a_{jj}}$$

et

$$a_{ii} = \sqrt{\sigma_{ii}^2 - \sum_{k=1}^{i-1} a_{ik}^2}$$

Exercice : Ecrire un algorithme de factorisation.

1.5.2 Factorisation en composantes principales

Comme Σ^2 est une matrice symétrique semi-définie positive alors Σ^2 est diagonalisable et a d valeurs propres $\lambda_i \geq 0$ et elle admet un ensemble de vecteurs propres associés v_i orthonormés vérifiant

$$v_i^T \cdot v_i = 1, \quad v_i^T \cdot v_j = 0, \quad \forall i \neq j \in \{1, \dots, d\}$$

et

$$\Sigma v_i = \lambda_i v_i$$

On note V la matrice orthonormale dont les colonnes sont v_1, \dots, v_d et on note Λ la matrice diagonale des valeurs propres associées.

Alors en choisissant $A = V\Lambda^{1/2}$ on obtient bien $A \cdot A^T = V \cdot \Lambda \cdot V^T = \Sigma^2$. Si de plus on suppose Σ^2 définie positive alors $\forall i$ on a $\lambda_i > 0$ et on a $A^{-1} = \Lambda^{-1/2} \cdot V^T$.