Introduction aux Processus Gaussiens

Application aux données spatio-temporelles

Clément Lejeune

Thales Services Numériques,
AD/AD3/IA

May 29, 2024

Plan

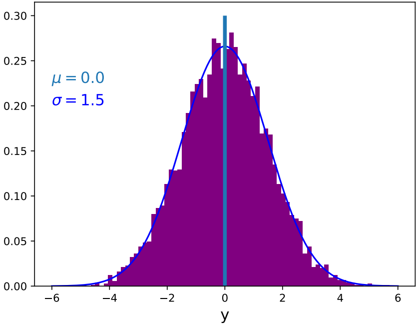
1. Gaussien: vecteur vs. processus

Gaussien: vecteur vs. processus

Loi Gaussienne unidiemensionnelle:

$$y \sim \mathcal{N}(m, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-m)^2}{2\sigma^2}}$$

1. m: espérance (aka moyenne) de y
2. $\sigma > 0$: écart-type



Loi Gaussienne multidimensionnelle (vecteur Gaussien): Distribution *jointe* d'un vecteur d-dimensionnel dont les *marginales* sont Gaussiennes (unidimensionnelles).

$$\mathbf{y} := [y_1, \ldots, y_d]^\top \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\sqrt{(2\pi)^d \det|\boldsymbol{\Sigma}|}} e^{-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})}$$

1. $\boldsymbol{\mu} \in \mathbb{R}^d$: *vecteur moyen* $\implies$ μ_j: moyenne de la Gaussienne y_j
2. $\boldsymbol{\Sigma} \in \mathbb{R}^{d \times d}$ définie positive[1]: *matrice de covariance*

[1] i.e. $\mathbf{a}^\top \boldsymbol{\Sigma} \mathbf{a} > 0$ (donc symmétrique)

Gaussien: vecteur vs. processus

Loi Gaussienne multidimensionnelle (vecteur Gaussien): Distribution *jointe* d'un vecteur d-dimensionnel dont les *marginales* sont Gaussiennes (unidimensionnelles).

$$\mathbf{y} := [y_1, \ldots, y_d]^\top \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\sqrt{(2\pi)^d \det |\boldsymbol{\Sigma}|}} e^{-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})}$$

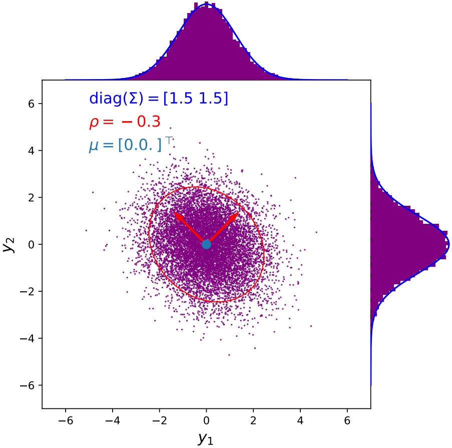
1. $\boldsymbol{\mu} \in \mathbb{R}^d$: *vecteur moyen* $\implies$ μ_j: moyenne de la Gaussienne y_j
2. $\boldsymbol{\Sigma} \in \mathbb{R}^{d \times d}$ définie positive[1]: *matrice de covariance*

$$\boldsymbol{\Sigma} = \begin{pmatrix} \Sigma_1^2 & \Sigma_{12} & \cdots & \Sigma_{1d} \\ \Sigma_{21} & \ddots & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{d1} & \cdots & \cdots & \Sigma_d^2 \end{pmatrix} \implies \begin{cases} \Sigma_j : \text{écart-type de } y_j \\ \Sigma_{ij} : \text{covariance entre } y_i \text{ et } y_j \end{cases}$$

[1] i.e. $\mathbf{a}^\top \boldsymbol{\Sigma} \mathbf{a} > 0$ (donc symmétrique)

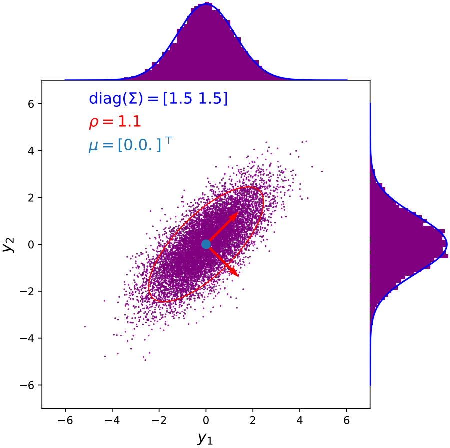
Cas $d = 2$:

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} \Sigma_1 & \Sigma_{12} \\ \Sigma_{21} & \Sigma_2 \end{bmatrix} = \begin{bmatrix} 1.5 & \rho = -0.3 \\ \rho = -0.3 & 1.5 \end{bmatrix}$$



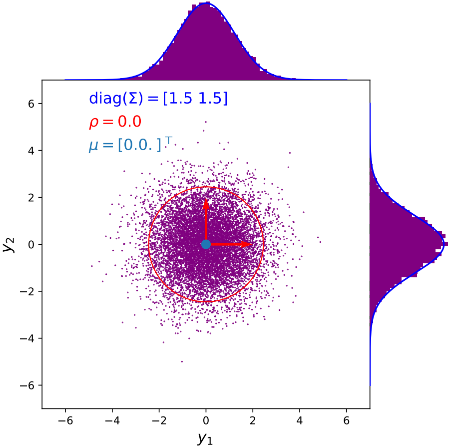
Cas $d = 2$:

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} \Sigma_1 & \Sigma_{12} \\ \Sigma_{21} & \Sigma_2 \end{bmatrix} = \begin{bmatrix} 1.5 & \rho = 1.1 \\ \rho = 1.1 & 1.5 \end{bmatrix}$$



Cas $d = 2$:

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} \Sigma_1 & \Sigma_{12} \\ \Sigma_{21} & \Sigma_2 \end{bmatrix} = \begin{bmatrix} 1.5 & \rho = 0 \\ \rho = 0 & 1.5 \end{bmatrix}$$



Cas $d = 2$:

$$\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 0 \\ \dots \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} 1.5 & 0.99 & 0.98 & 0.96 & 0.94 \\ 0.99 & 1.5 & 0.99 & 0.98 & 0.96 \\ 0.98 & 0.99 & 1.5 & 0.99 & 0.98 \\ 0.96 & 0.98 & 0.99 & 1.5 & 0.99 \\ 0.94 & 0.96 & 0.98 & 0.99 & 1.5 \end{bmatrix}$$

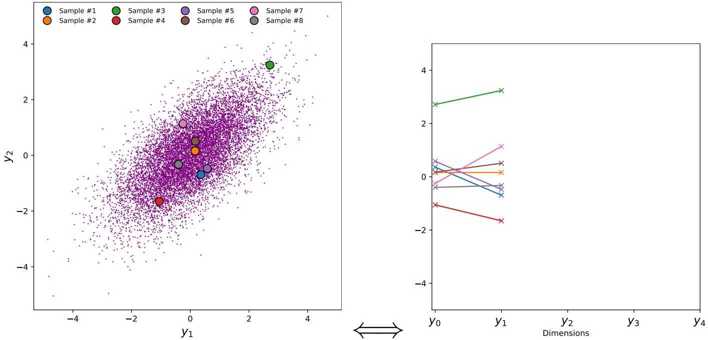


Figure: Dimensions $j = 1, 2$. Gauche: 10^4 $+8$ échantillons. Droite: 8 *mêmes* échantillons.

Cas $d = 5$:

$$\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 0 \\ \ldots \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} 1.5 & 0.99 & 0.98 & 0.96 & 0.94 \\ 0.99 & 1.5 & 0.99 & 0.98 & 0.96 \\ 0.98 & 0.99 & 1.5 & 0.99 & 0.98 \\ 0.96 & 0.98 & 0.99 & 1.5 & 0.99 \\ 0.94 & 0.96 & 0.98 & 0.99 & 1.5 \end{bmatrix}$$

Chaque courbe est un tirage de $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

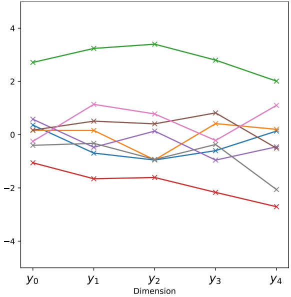


Figure: 8 échantillons, dimensions $j = 1, \ldots, 5$

Cas $d = 50$:

$$\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 0 \\ \dots \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} 1.5 & 0.998 & 0.993 & 0.986 & 0.975 & 0.962 & 0.946 & 0.927 & \cdots \\ 0.998 & 1.5 & 0.998 & 0.993 & 0.986 & 0.975 & 0.962 & 0.946 & \cdots \\ 0.993 & 0.998 & 1.5 & 0.998 & 0.993 & 0.986 & 0.975 & 0.962 & \cdots \\ 0.986 & 0.993 & 0.998 & 1.5 & 0.998 & 0.993 & 0.986 & 0.975 & \cdots \\ 0.975 & 0.986 & 0.993 & 0.998 & 1.5 & 0.998 & 0.993 & 0.986 & \cdots \\ 0.962 & 0.975 & 0.986 & 0.993 & 0.998 & 1.5 & 0.998 & 0.993 & \cdots \\ 0.946 & 0.962 & 0.975 & 0.986 & 0.993 & 0.998 & 1.5 & 0.998 & \cdots \\ 0.927 & 0.946 & 0.962 & 0.975 & 0.986 & 0.993 & 0.998 & 1.5 & \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 1.5 \end{bmatrix}$$

Chaque courbe est un tirage de $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

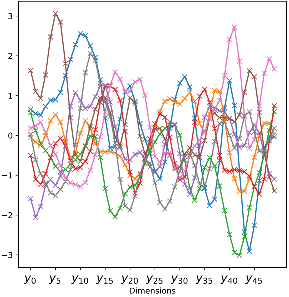


Figure: 8 échantillons, dimensions $j = 1, \dots, 50$

Cas $d = \infty$: chaque courbe est une *collection infinie* d'échantillons Gaussiens.

1. $\boldsymbol{\mu}$: vecteur de taille infinie
2. $\boldsymbol{\Sigma}$: une fonction de covariance