Projet Maths

Matthias ZDRAVKOVIC, Ameen MOHD FAIRUZ

Mai 2023

1 Développements mathématiques

1. a) On part de la loi normale bidimensionnelle :

$$f_Z(z) = \frac{1}{2\pi\sqrt{\det \Sigma}} exp(-\frac{1}{2}{}^t(z-\mu)\Sigma^{-1}(z-\mu))$$

On pose

$$\Sigma = U \cdot \Lambda \cdot U^T$$

$$\Leftrightarrow \det \Sigma = \det U \cdot \det \Lambda \cdot \det(U^T)$$

$$(\det U)^2 = 1 \text{ car c'est une matrice orthogonale}$$

$$\Leftrightarrow \det \Sigma = \det \Lambda$$

avec $\Sigma \in M_2(\mathbb{R})$ matrice symétrique positive définie, $U \in M_2(\mathbb{R})$ matrice orthogonale et $\Lambda \in M_2(\mathbb{R})$ matrice diagonale à coefficients positifs.
On pose

$$U = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

et

$$\Lambda = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$$

avec

$$\Sigma^{-1} = U \cdot \Lambda^{-1} \cdot U^{-1}$$

$$\Leftrightarrow \Sigma^{-1} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{b} \end{pmatrix} \cdot \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}^{-1}$$

$$\Leftrightarrow \Sigma^{-1} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{b} \end{pmatrix} \cdot \frac{1}{\cos^2\theta + \sin^2\theta} \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

$$\Leftrightarrow \Sigma^{-1} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{b} \end{pmatrix} \cdot \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

$$\Leftrightarrow \Sigma^{-1} = \begin{pmatrix} \frac{b\cos^2\theta + a\sin^2\theta}{ab} & \frac{b\sin\theta\cos\theta - a\sin\theta\cos\theta}{ab} \\ \frac{b\sin\theta\cos\theta - a\sin\theta\cos\theta}{ab} & \frac{b\sin^2\theta + a\sin^2\theta}{ab} \end{pmatrix}$$

On injecte dans la formule de $f_Z(z)$:

$$f_Z(z) = \frac{1}{2\pi\sqrt{ab}} exp(-\frac{1}{2}\begin{pmatrix} x-\mu_1 & y-\mu_2 \end{pmatrix} \cdot \begin{pmatrix} \frac{b\cos^2\theta + a\sin^2\theta}{ab} & \frac{b\sin\theta\cos\theta - a\sin\theta\cos\theta}{ab} \\ \frac{b\sin\theta\cos\theta - a\sin\theta\cos\theta}{ab} & \frac{b\sin^2\theta + a\sin^2\theta}{ab} \end{pmatrix} \begin{pmatrix} x-\mu_1 \\ y-\mu_2 \end{pmatrix})$$

$$\Leftrightarrow f_Z(z) = \frac{1}{2\pi\sqrt{ab}} exp(-\frac{1}{2ab}\begin{pmatrix} x-\mu_1 & y-\mu_2 \end{pmatrix} \cdot \begin{pmatrix} b\cos^2\theta + a\sin^2\theta & b\sin\theta\cos\theta - a\sin\theta\cos\theta \\ b\sin\theta\cos\theta - a\sin\theta\cos\theta & b\sin^2\theta + a\sin^2\theta \end{pmatrix} \cdot \begin{pmatrix} x-\mu_1 \\ y-\mu_2 \end{pmatrix})$$

Calculons le produit matriciel A dans l'exponentielle :

$$A = \begin{pmatrix} x-\mu_1 & y-\mu_2 \end{pmatrix} \cdot \begin{pmatrix} b\cos^2\theta + a\sin^2\theta & b\sin\theta\cos\theta - a\sin\theta\cos\theta \\ b\sin\theta\cos\theta - a\sin\theta\cos\theta & b\sin^2\theta + a\sin^2\theta \end{pmatrix} \cdot \begin{pmatrix} x-\mu_1 \\ y-\mu_2 \end{pmatrix}$$

$$\Leftrightarrow A = ((a\sin^2\theta + b\cos^2\theta)(x-\mu_1) + \cos\theta\sin\theta(b-a)(y-\mu_2))(x-\mu_1) \\ + ((b\sin^2\theta + a\cos^2\theta)(y-\mu_2) + \cos\theta\sin\theta(b-a)(x-\mu_1))(y-\mu_2)$$

$$\Leftrightarrow A = (a\sin^2\theta + b\cos^2\theta)(x-\mu_1)^2 + \cos\theta\sin\theta(b-a)(x-\mu_1)(y-\mu_2) \\ + (b\sin^2\theta + a\cos^2\theta)(y-\mu_2)^2 + \cos\theta\sin\theta(b-a)(x-\mu_1)(y-\mu_2)$$

$$\Leftrightarrow A = (a\sin^2\theta + b\cos^2\theta)(x-\mu_1)^2 + 2\cos\theta\sin\theta(b-a)(x-\mu_1)(y-\mu_2) \\ + (b\sin^2\theta + a\cos^2\theta)(y-\mu_2)^2$$

$$\Leftrightarrow A = a\sin^2\theta(x-\mu_1)^2 + b\cos^2\theta(x-\mu_1)^2 + 2\cos\theta\sin\theta(b-a)(x-\mu_1)(y-\mu_2) \\ + b\sin^2\theta(y-\mu_2)^2 + a\cos^2\theta(y-\mu_2)^2$$

$$\Leftrightarrow A = a(\sin^2\theta(x-\mu_1)^2 + \cos^2\theta(y-\mu_2)^2) + 2\cos\theta\sin\theta(b-a)(x-\mu_1)(y-\mu_2) \\ + b(\sin^2\theta(y-\mu_2)^2 + \cos^2\theta(x-\mu_1)^2)$$

$$\Leftrightarrow A = a((x-\mu_1)^2\sin^2\theta + (y-\mu_2)^2\cos^2\theta - 2\cos\theta\sin\theta(x-\mu_1)(y-\mu_2)) \\ + b((y-\mu_2)^2\sin^2\theta + (x-\mu_1)^2\cos^2\theta + 2\cos\theta\sin\theta(x-\mu_1)(y-\mu_2))$$

$$\Leftrightarrow A = a((x-\mu_1)\sin\theta - (y-\mu_2)\cos\theta)^2 + b((x-\mu_1)\cos\theta + (y-\mu_2)\sin\theta)^2$$

On obtient donc :

$$f_Z(z) = \frac{1}{2\pi\sqrt{ab}} exp(-\frac{((x-\mu_1)\sin\theta - (y-\mu_2)\cos\theta)^2}{2b} - \frac{((x-\mu_1)\cos\theta + (y-\mu_2)\sin\theta)^2}{2a}) = K$$

$$\Leftrightarrow \frac{((x-\mu_1)\sin\theta - (y-\mu_2)\cos\theta)^2}{2b} + \frac{((x-\mu_1)\cos\theta + (y-\mu_2)\sin\theta)^2}{2a} = -\ln(K2\pi\sqrt{ab})$$

$$\Leftrightarrow \frac{((x-\mu_1)\sin\theta - (y-\mu_2)\cos\theta)^2}{2b\ln(\frac{1}{K2\pi\sqrt{ab}})} + \frac{((x-\mu_1)\cos\theta + (y-\mu_2)\sin\theta)^2}{2a\ln(\frac{1}{K2\pi\sqrt{ab}})} = 1$$

Ici, le centre de l'ellipse est donné par $\mu = (\mu_1, \mu_2)$, $\sqrt{2a\log(\frac{1}{2\pi K\sqrt{ab}})}$ est la demi-longueur de l'axe principal et $\sqrt{2b\log(\frac{1}{2\pi K\sqrt{ab}})}$ la demi-longueur de l'axe secondaire, K est la constante de normalisation et θ est l'angle de rotation de l'ellipse.

1. b) Pour calculer la probabilité qu'un point tiré selon la loi Z appartienne à la surface interne S_k délimitée par l'ellipse d'isodensité K, on intègre la densité de Z par la surface de l'ellipse:

$$P(Z \in S_k) = \int \frac{1}{\sqrt{2\pi^2 det\Sigma}} exp(-\frac{1}{2}{}^t(z-\mu)\Sigma^{-1}(z-\mu))dz$$

$$\Leftrightarrow \int\int_{S_k} \frac{1}{2\pi\sqrt{ab}} exp(-\frac{[(x-\mu_1)cos\theta + (y-\mu_2)sin\theta]^2}{2a} - \frac{[(x-\mu_1)sin\theta - (y-\mu_2)cos\theta]^2}{2b})dxdy$$

Pour simplifier les calculs, on fait un changement de variable:

$$\begin{cases} x_1 = \frac{(x-\mu_1)cos\theta + (y-\mu_2)sin\theta}{\sqrt{2a}} \\ x_2 = \frac{(x-\mu_1)sin\theta - (y-\mu_2)cos\theta}{\sqrt{2a}} \end{cases}$$

$$\Leftrightarrow \begin{cases} x = \sqrt{2a}cos\theta x_1 + \sqrt{2b}sin\theta y_1 + \mu_1 \\ y = \sqrt{2a}sin\theta x_1 - \sqrt{2b}cos\theta y_1 + \mu_2 \end{cases}$$

$$\Leftrightarrow \frac{\partial x}{\partial x_1} = \sqrt{2a}cos\theta \ ; \ \frac{\partial x}{\partial y_1} = \sqrt{2b}sin\theta$$

$$\frac{\partial y}{\partial x_1} = \sqrt{2a}sin\theta \ ; \ \frac{\partial y}{\partial y_1} = -\sqrt{2b}cos\theta$$

En changeant (x, y) à (x_1, y_1), nous avons maintenant:

$$\int\int_{S_k} \frac{1}{2\pi\sqrt{ab}} exp(-x_1^2 - y_1^2) \begin{vmatrix} \frac{\partial x}{\partial x_1} & \frac{\partial x}{\partial y_1} \\ \frac{\partial y}{\partial x_1} & \frac{\partial y}{\partial y_1} \end{vmatrix} dx_1 dy_1$$

$$\Leftrightarrow \int\int_{S_k} \frac{1}{2\pi\sqrt{ab}} exp(-x_1^2 - y_1^2) \left| -\sqrt{4ab}cos^2\theta - \sqrt{4ab}sin^2\theta \right| dx_1 dy_1$$

$$\Leftrightarrow \int\int_{S_k} \frac{2\sqrt{ab}}{2\pi\sqrt{ab}} exp(-x_1^2 - y_1^2) dx_1 dy_1$$

$$\Leftrightarrow \int\int_{S_k} \frac{1}{\pi} exp(-x_1^2 - y_1^2) dx_1 dy_1$$

Changement en base polaire: on utilise $x_1 = rcos\varphi$ et $y_1 = rsin\varphi$

$$\int\int_{S_k} \frac{1}{\pi} exp(-r^2) r dr d\varphi$$

$$\Leftrightarrow \frac{1}{\pi} \int_0^r \int_0^{2\pi} re^{-r} d\varphi dr$$

$$\Leftrightarrow \int_0^r 2re^{-r} dr$$

Or, ici il faut utiliser $r =$ l'équation de l'ellipse en base (r, φ)
Pour cela, nous reprenons l'équation de l'ellipse trouvée à la question 1.a)

$$\frac{((x-\mu_1)\sin\theta - (y-\mu_2)\cos\theta)^2}{2b\ln(\frac{1}{K2\pi\sqrt{ab}})} + \frac{((x-\mu_1)\cos\theta + (y-\mu_2)\sin\theta)^2}{2a\ln(\frac{1}{K2\pi\sqrt{ab}})} = 1$$

$$\Leftrightarrow \frac{x_1^2}{\ln(\frac{1}{K2\pi\sqrt{ab}})} + \frac{y_1^2}{\ln(\frac{1}{K2\pi\sqrt{ab}})} = 1$$

$$\Leftrightarrow \frac{r^2}{\ln(\frac{1}{K2\pi\sqrt{ab}})} = 1$$

$$\Leftrightarrow r = \sqrt{\ln(\frac{1}{K2\pi\sqrt{ab}})}$$

Nous utilisons cette équation de r comme borne superieure de notre intégrale:

$$\int_0^{\sqrt{\ln(\frac{1}{K2\pi\sqrt{ab}})}} 2re^{-r} dr$$

$$\Leftrightarrow \left[-e^{r^2}\right]_0^{\sqrt{\ln(\frac{1}{K2\pi\sqrt{ab}})}}$$

$$\Leftrightarrow -e^{\sqrt{\ln(\frac{1}{K2\pi\sqrt{ab}})}} + e^0$$

Nous trouvons donc:

$$p = P(Z \in S_k)$$

$$\Leftrightarrow p = 1 - 2\pi K\sqrt{ab}$$

2. Intuitivement, on a envie de montrer que les estimateurs de μ et Σ sont :

$$\mu = \begin{pmatrix} \frac{\sum_{i=1}^{N} Xi}{N} \\ \frac{\sum_{i=1}^{N} Yi}{N} \end{pmatrix} \Leftrightarrow \mu = \begin{pmatrix} \bar{X} \\ \bar{Y} \end{pmatrix} = \frac{\sum_{i=1}^{N} z_i}{N}$$

$$\Sigma = \begin{pmatrix} \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_1)^2 & \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_1)(y_i - \mu_2) \\ \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_1)(y_i - \mu_2) & \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \mu_2)^2 \end{pmatrix} = \frac{1}{N-1}\sum_{i=1}^{N}(z_i - \mu)(z_i - \mu)^T$$

Montrons le grâce à la fonction log-vraisemblance. Pour commencer, on va trouver l'estimateur de μ en maximisant la fonction log-vraisemblance :

$$\mu \to \ln \mathcal{L}(\mu|z_i)$$

$$\Leftrightarrow \ln \mathcal{L}(\mu|z_i) = \ln \prod_{i=1}^{N} \frac{1}{2\pi\sqrt{|\Sigma|}} exp(-\frac{1}{2}(z_i - \mu)^T \Sigma^{-1}(z_i - \mu))$$

$$= \sum_{i=1}^{N} (\ln(\frac{1}{2\pi\sqrt{|\Sigma|}} exp(-\frac{1}{2}(z_i - \mu)^T \Sigma^{-1}(z_i - \mu))))$$

$$= \sum_{i=1}^{N} (\ln \frac{1}{2\pi\sqrt{|\Sigma|}}) - \frac{1}{2}\sum_{i=1}^{N}((z_i - \mu)^T \Sigma^{-1}(z_i - \mu))$$

$$= \sum_{i=1}^{N} (\ln \frac{1}{2\pi\sqrt{|\Sigma|}}) - \frac{1}{2}\sum_{i=1}^{N}(z_i^T \Sigma^{-1} z_i - z_i^T \Sigma^{-1}\mu - \mu^T \Sigma^{-1} z_i + \mu^T \Sigma^{-1}\mu)$$

$$= \sum_{i=1}^{N} (\ln \frac{1}{2\pi\sqrt{|\Sigma|}}) - \frac{1}{2}\sum_{i=1}^{N} z_i^T \Sigma^{-1} z_i + \frac{1}{2}\sum_{i=1}^{N} z_i^T \Sigma^{-1}\mu + \frac{1}{2}\sum_{i=1}^{N} \mu^T \Sigma^{-1} z_i - \frac{1}{2}\sum_{i=1}^{N} \mu^T \Sigma^{-1}\mu$$

On dérive par rapport à μ et on cherche le maximum. En sachant que pour n'importe quel vecteur v de taille 2 et matrice M de taille 2x2, on a[1,2] :

$$\frac{\partial(v^T M v)}{\partial v} = (M + M^T)v$$

$$\frac{\partial v^T M}{\partial v} = M^T$$

On utilise ces formules pour la dérivation matricielle :

$$\frac{\partial \ln \mathcal{L}(\mu|z_i)}{\partial \mu} = \frac{1}{2}(\sum_{i=1}^{N} z_i^T \Sigma^{-1} + \sum_{i=1}^{N}(\Sigma^{-1} z_i)^T - \sum_{i=1}^{N}(\Sigma^{-1} + (\Sigma^{-1})^T) \cdot \mu) = 0$$

[1] https://www.di.ens.fr/ fbach/courses/fall2009/formulaire.pdf
[2] https://en.wikipedia.org/wiki/Matrixcalculus

$$\Leftrightarrow \sum_{i=1}^{N}(\Sigma^{-1}+(\Sigma^{-1})^T)\cdot\mu = \sum_{i=1}^{N} z_i^T\Sigma^{-1} + \sum_{i=1}^{N}(z_i^T(\Sigma^{-1})^T)$$

$$\Leftrightarrow \sum_{i=1}^{N}\mu = \frac{\sum_{i=1}^{N} z_i^T\Sigma^{-1} + \sum_{i=1}^{N}(z_i^T(\Sigma^{-1})^T)}{\sum_{i=1}^{N}(\Sigma^{-1}+(\Sigma^{-1})^T)}$$

$$\Leftrightarrow \sum_{i=1}^{N}\mu = \sum_{i=1}^{N} z_i^T\frac{(\sum_{i=1}^{N}\Sigma^{-1}+((\Sigma^{-1})^T))}{\sum_{i=1}^{N}(\Sigma^{-1}+(\Sigma^{-1})^T)}$$

$$\Leftrightarrow N\cdot\hat{\mu} = \sum_{i=1}^{N} z_i^T$$

$$\Leftrightarrow \hat{\mu} = \frac{\sum_{i=1}^{N} z_i^T}{N}$$

Or $\frac{\sum_{i=1}^{N} z_i^T}{N} = \frac{\sum_{i=1}^{N} z_i}{N}$ car c'est une somme d'un vecteur colonne. On retrouve donc bien l'estimateur de μ. On fait la même chose pour Σ :

$$\Sigma \to \ln\mathcal{L}(\Sigma|z_i)$$

$$\Leftrightarrow \ln\mathcal{L}(\Sigma|z_i) = \ln\prod_{i=1}^{N}\frac{1}{2\pi\sqrt{|\Sigma|}}exp(-\frac{1}{2}(z_i-\mu)^T\Sigma^{-1}(z_i-\mu))$$

$$= \sum_{i=1}^{N}(\ln\frac{1}{2\pi\sqrt{|\Sigma|}}exp(-\frac{1}{2}(z_i-\mu)^T\Sigma^{-1}(z_i-\mu)))$$

$$= \sum_{i=1}^{N}\ln\frac{1}{2\pi\sqrt{|\Sigma|}} + \sum_{i=1}^{N}(-\frac{1}{2}(z_i-\mu)^T\Sigma^{-1}(z_i-\mu))$$

On dérive par rapport à Σ et on cherche le maximum.

On a aussi une autre formule de dérivation matricielle[3] :

$$\frac{\partial \det\Sigma}{\partial\Sigma} = \det(\Sigma)(\Sigma^{-1})^T$$

$$\frac{\partial a^T\Sigma^T b}{\partial\Sigma} = ab^T$$

avec a et b pas des fonctions de Σ.

$$\frac{\partial\ln\mathcal{L}(\Sigma|z_i)}{\partial\Sigma} = \sum_{i=1}^{N}(2\pi\sqrt{\det\Sigma}\cdot\frac{\partial}{\partial\Sigma}(\frac{1}{2\pi\sqrt{\det\Sigma}})) - \frac{1}{2}\sum_{i=1}^{N}(z_i-\mu)(z_i-\mu)^T = 0$$

[3]https://www.quora.com/What-is-the-differentiation-of-the-determinant-of-a-matrix-to-the-matrix-itself-and-why

$$\frac{\partial \ln \mathcal{L}(\Sigma|z_i)}{\partial \Sigma} = \sum_{i=1}^{N} (\sqrt{\det \Sigma} \cdot (-\frac{1}{2 \det \Sigma \sqrt{\det \Sigma}}) \cdot \det(\Sigma)(\Sigma^{-1})^T) - \frac{1}{2} \sum_{i=1}^{N} (z_i - \mu)(z_i - \mu)^T = 0$$

$$\frac{\partial \ln \mathcal{L}(\Sigma|z_i)}{\partial \Sigma} = -\frac{1}{2} \sum_{i=1}^{N} \cdot (\Sigma^{-1})^T - \frac{1}{2} \sum_{i=1}^{N} (z_i - \mu)(z_i - \mu)^T = 0$$

Or on sait que Σ est une matrice de rotation, donc $\Sigma^T = \Sigma^{-1}$.

$$\Leftrightarrow \frac{\partial \ln \mathcal{L}(\Sigma|z_i)}{\partial \Sigma} = -\frac{1}{2} \sum_{i=1}^{N} \Sigma - \frac{1}{2} \sum_{i=1}^{N} (z_i - \mu)(z_i - \mu)^T = 0$$

$$\Leftrightarrow \sum_{i=1}^{N} \Sigma = \sum_{i=1}^{N} (z_i - \mu)(z_i - \mu)^T$$

$$\Leftrightarrow N \cdot \Sigma = \sum_{i=1}^{N} (z_i - \mu)(z_i - \mu)^T$$

$$\Leftrightarrow \hat{\Sigma} = \frac{\sum_{i=1}^{N} (z_i - \mu)(z_i - \mu)^T}{N}$$