Gaussian curves in bivariate case and covariance matrix.

We often remember that covariance matrix is defined as

$$\Sigma_{ii} = cov(X_i, X_i) = Var(X_i)$$

$$\forall i,j \in [\![1,n]\!] \quad \Sigma_{ij} = cov(X_i, X_j) = cov(X_j, X_i)$$

but knowing that : $X = (X_1, \dots, X_n)^T$

$$\boxed{\Sigma = \mathbb{E}\left((X - \mathbb{E}(X))(X - \mathbb{E}(X))^T\right)}$$

... is much more practical to immediately see that Σ is in $\underbrace{S_n^{++}(\mathbb{R})}_{\text{Symmetric positive definite}}$.

In fact: $\forall w \in \mathbb{R}^n \setminus \{0\}$,

$$w^T \Sigma w = \mathbb{E}\left(\| (X - \mathbb{E}(X))^T w \|^2\right)$$

If we had $w^T \Sigma w = 0$ for a $w \neq 0 \in \mathbb{R}^n$, then

$\forall \omega \in \Omega$, $(X(\omega) - \mathbb{E}(X))^T w = 0$

impossible

$$\boxed{\Sigma \in S_n^{++}(\mathbb{R})}$$

According to the spectral theorem, $Sp(\Sigma) \subset \mathbb{R}^+ \setminus \{0\}$

Spectral decomposition of Σ : $\exists Q \in O_n(\mathbb{R})$ orthogonal matrix such that : $\Sigma = Q \, Diag(\lambda_i) \, Q^T$

such that : $(Q_{i,:})^T$ is the unit eigenvector associated with λ_i

so $\Sigma^{-1} = Q^T \, \mathrm{Diag}(\lambda_i^{-1}) \, Q$

$$= \underbrace{Q^T \sqrt{\mathrm{Diag}(\lambda_i^{-1})}\, Q}_{:= C} \, Q^T \sqrt{\mathrm{Diag}(\lambda_i^{-1})}\, Q$$

$\Sigma^{-1} = C^2$ where C is symmetric positive definite

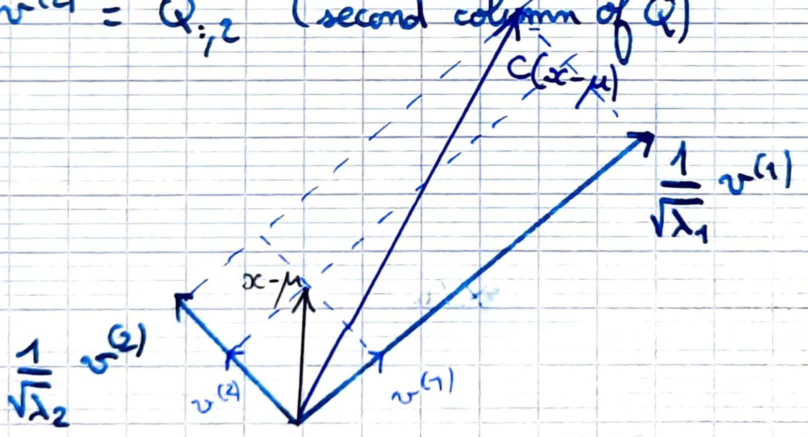
The PDF of Gaussian multivariate is :

$$\mathcal{N}(x;\mu;\Sigma) = \sqrt{\frac{1}{(2\pi)^n \det(\Sigma)}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right)$$

$$= K \exp\left(-\frac{1}{2}\|C(x-\mu)\|^2\right)$$

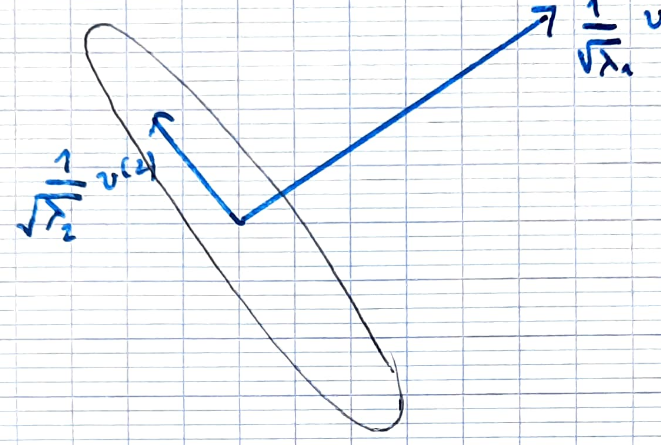
<u>Study of n=2</u> Let $v^{(1)} = Q_{:,1}$ (first column of Q)

$v^{(2)} = Q_{:,2}$ (second column of Q)



The smaller λ_1, the larger $\frac{1}{\sqrt{\lambda_1}}$, the larger the component of $C(x-\mu)$ along $v^{(1)}$ and the larger $\frac{1}{2}\|C(x-\mu)\|^2$. Therefore : the smaller λ_1, the smaller $\mathcal{N}(x;\mu;\Sigma)$.

Hence the observation in the notebook: if λ_i is big, the halo is extended in the direction of the respective $v^{(i)}$ (which is - be careful - an eigenvector of Σ^{-1}, not Σ).



Now let's do a little bit more: let's compute the equations of the isocurves. For instance the black curve above is an isocurve:

$$\| C(x-\mu) \|^2 = R^2$$

$$\langle C(x-\mu), C(x-\mu) \rangle = R^2$$

$$(x-\mu)^T C^T C (x-\mu) = R^2$$

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

<u>Study of $n=2$</u) $\quad \Sigma = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \quad \Sigma^{-1} = \frac{1}{ac-b^2} \begin{pmatrix} c & -b \\ -b & a \end{pmatrix}$

$\mu = (\mu_1 \ \mu_2)^T$

$$(x-\mu)^T \Sigma^{-1} (x-\mu) = (x-\mu)^T \frac{1}{ac-b^2} \begin{pmatrix} c(x_1-\mu_1) - b(x_2-\mu_2) \\ -b(x_1-\mu_1) + a(x_2-\mu_2) \end{pmatrix}$$

$$= \frac{1}{ac-b^2}\Big[(x_1-\mu_1)\big(c(x_1-\mu_1)-b(x_2-\mu_2)\big)+(x_2-\mu_2)\big(-b(x_1-\mu_1)+a(x_2-\mu_2)\big)\Big]$$

$$= \frac{1}{ac-b^2}\Big[c(x_1-\mu_1)^2-b(x_1-\mu_1)(x_2-\mu_2)-b(x_1-\mu_1)(x_2-\mu_2)+a(x_2-\mu_2)^2\Big]$$

$$R^2 = \frac{1}{ac-b^2}\Big[c(x_1-\mu_1)^2-2b(x_1-\mu_1)(x_2-\mu_2)+a(x_2-\mu_2)^2\Big]$$

$$R^2 = \frac{1}{c(ac-b^2)}\Big[c^2(x_1-\mu_1)^2-2bc(x_1-\mu_1)(x_2-\mu_2)+b^2(x_2-\mu_2)^2+\lambda_1\lambda_2(x_2-\mu_2)^2\Big]$$

$$R^2 = \frac{1}{\lambda_1\lambda_2 c}\Big[\big(c(x_1-\mu_1)-b(x_2-\mu_2)\big)^2+\lambda_1\lambda_2(x_2-\mu_2)^2\Big]$$

$$\begin{cases} \mathrm{Tr}(\Sigma) = a+c = \lambda_1+\lambda_2 \\ \det(\Sigma) = ac-b^2 = \lambda_1\lambda_2 \end{cases}$$