**Machine Learning - 2020-2021**
Universitat Pompeu Fabra

# Problems 2: Gaussian Models

## Goal

The goal of this practice is to learn to estimate a Gaussian (model) from a given set of points. Also, you will learn to understand and graphically interpret the Mahalanobis distance and the ellipses given by the estimated Gaussian. It is important to understand the applications estimating a Gaussian from a set of points: modelling and clustering.

## Needed Concepts

- Basis of a vector space and change of basis
  If $\mathcal{V} = \{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}\}$ is a basis of $\mathbb{R}^2$ (in cartesian coordinates), then

$$\mathbf{P} = \begin{pmatrix} \vdots & \vdots \\ \mathbf{v}^{(1)} & \mathbf{v}^{(2)} \\ \vdots & \vdots \end{pmatrix}$$

  is the matrix that changes a vector $\mathbf{x}_{<\mathcal{V}>}$ written in coordinates of the basis $\mathcal{V}$ to $\mathbf{x}_{<\mathcal{E}>}$ in cartesian coordinates, i.e., $\mathbf{x}_{<\mathcal{E}>} = \mathbf{P}\ \mathbf{x}_{<\mathcal{V}>}$. We will denote as ${}_{<\mathcal{V}>}\mathbf{M}_{<\mathcal{E}>}$ the matrix that changes from basis $\mathcal{V}$ to basis $\mathcal{E}$.
  *Reminder*: if $\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}\}$ are orthonormal, then $\mathbf{P}^{-1} = \mathbf{P}^T$.

- Singular Value Decomposition (SVD): used to diagonalise non-square matrices $A \in \mathbb{R}^{n\times m}$.

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^T$$

  where $\mathbf{U} \in \mathbb{R}^{n\times n}$, $\mathbf{D} \in \mathbb{R}^{n\times m}$ is a diagonal matrix and $\mathbf{V} \in \mathbb{R}^{m\times m}$.
  *Reminder*: The columns of $\mathbf{U}$ and $\mathbf{V}$ define new **orthonormal** bases in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively.

- Gaussian distribution (or multivariate normal distribution)

$$PDF : f(\mathbf{x}) = (2\pi)^{-\frac{k}{2}}\det(\mathbf{\Sigma})^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\mathbf{x}-\mu)^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\mu)\right]$$

  where $k$ is the dimension of $\mathbf{x}$.

## Exercises

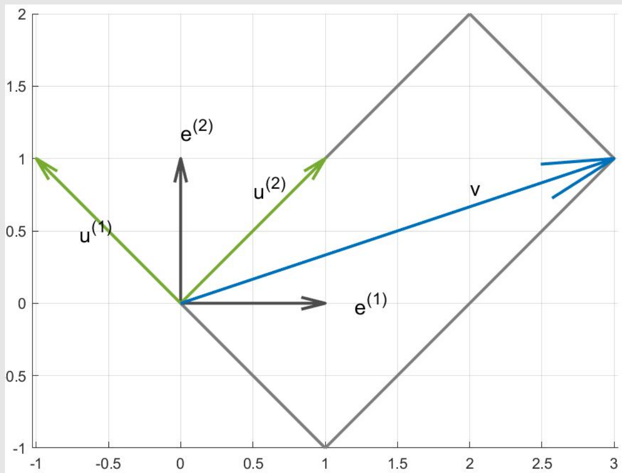
**REFRESH: Change of basis and system of coordinates**

1. (a) Consider $\mathbf{v} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \in \mathbb{R}^2$ written in coordinates of the standard basis of $\mathbb{R}^2$, $\mathcal{E} = \{\mathbf{e}^{(1)}, \mathbf{e}^{(2)}\}$. Find the coordinates of $\mathbf{v}$ in the basis $\mathcal{U} = \{\mathbf{u}^{(1)}, \mathbf{u}^{(2)}\}$ where $\mathbf{u}^{(1)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ and $\mathbf{u}^{(2)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Draw an scheme of that change of basis.
   (b) Consider the standard system of coordinates of $\mathbb{R}^2$, $\mathcal{E} = \{O; \mathbf{e}^{(1)}, \mathbf{e}^{(2)}\}$ and a new system of coordinates $\mathcal{U} = \{P; \mathbf{u}^{(1)}, \mathbf{u}^{(2)}\}$, where $P = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$ and $\mathbf{u}^{(1)}, \mathbf{u}^{(2)}$ are the same as in the previous exercise. Find the coordinates of $A = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ in the new system of coordinates $\mathcal{U}$. Draw an scheme of that change of coordinates.

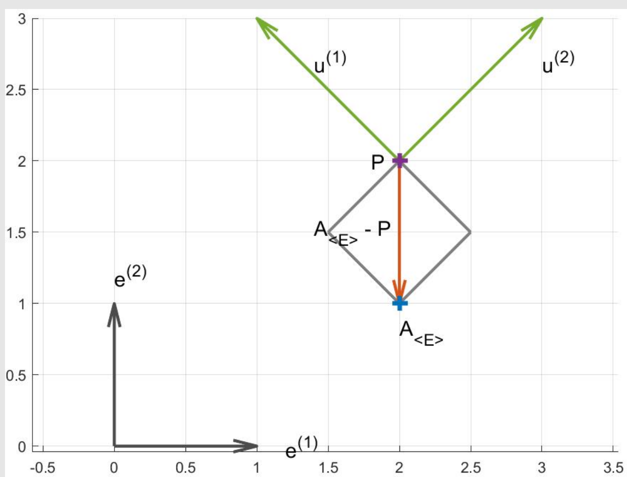
**Solution:**

(a) $$_{<\mathcal{E}>}\mathbf{M}_{<\mathcal{U}>} = \begin{pmatrix} \vdots & \vdots \\ \mathbf{e}^{(1)}_{<\mathcal{U}>} & \mathbf{e}^{(2)}_{<\mathcal{U}>} \\ \vdots & \vdots \end{pmatrix} = \begin{pmatrix} \vdots & \vdots \\ \mathbf{u}^{(1)}_{<\mathcal{E}>} & \mathbf{u}^{(2)}_{<\mathcal{E}>} \\ \vdots & \vdots \end{pmatrix}^{-1} = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}^{-1} = \frac{1}{2}\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}.$$

$$\implies \mathbf{v}_{<\mathcal{U}>} = {}_{<\mathcal{E}>}\mathbf{M}_{<\mathcal{U}>}\ \mathbf{v}_{<\mathcal{E}>} = \frac{1}{2}\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}.$$




(b) $$A_{<\mathcal{U}>} = {}_{<\mathcal{E}>}\mathbf{M}_{<\mathcal{V}>}\ (A_{<\mathcal{E}>} - P) = \frac{1}{2}\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}\left(\begin{pmatrix} 2 \\ 1 \end{pmatrix} - \begin{pmatrix} 2 \\ 2 \end{pmatrix}\right) = \frac{1}{2}\begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

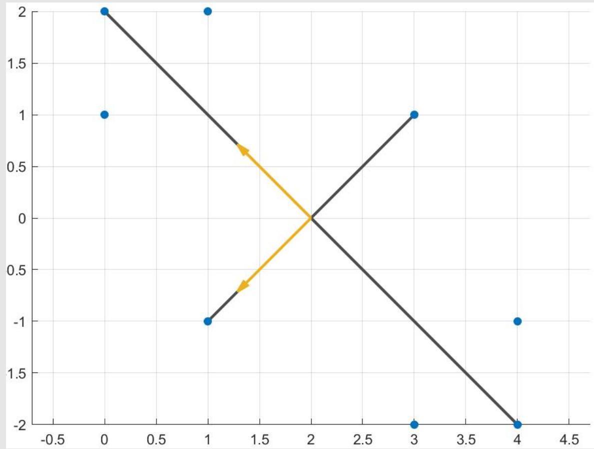



## Gaussian models

2. Let the set of points $\left\{ \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ -2 \end{pmatrix}, \begin{pmatrix} 4 \\ -1 \end{pmatrix}, \begin{pmatrix} 4 \\ -2 \end{pmatrix} \right\} \subset \mathbb{R}^2$ be driven from a Gaussian distribution. Estimate the parameters of such Gaussian distribution. To do so,

   (a) Draw the set of points. Intuitively, draw the principal directions of the covariance matrix given by the Gaussian model.

**Solution:**



(b) Compute the mean and the covariance matrix of the given set of points.

**Solution:**

$\mu = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$

$X = \begin{pmatrix} 0 & 0 & 1 & 1 & 3 & 3 & 4 & 4 \\ 2 & 1 & 2 & -1 & 1 & -2 & -1 & -2 \end{pmatrix} \implies X_{\text{c}} = X - \mu = \begin{pmatrix} -2 & -2 & -1 & -1 & 1 & 1 & 2 & 2 \\ 2 & 1 & 2 & -1 & 1 & -2 & -1 & -2 \end{pmatrix}$

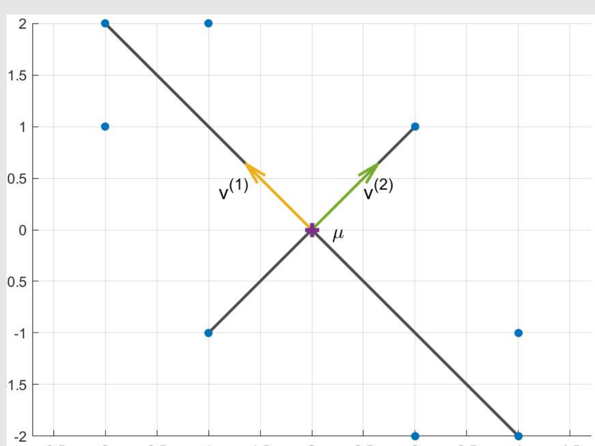
$\implies \mathbf{\Sigma} = \dfrac{1}{8-1} X X^T = \dfrac{1}{7} \begin{pmatrix} 20 & -14 \\ -14 & 20 \end{pmatrix}$

(c) Compute the principal directions of the covariance matrix. Draw the resulting new coordinate system and check if they correspond to the directions you sketched in (1).

**Solution:**

$$p_{\mathbf{\Sigma}}(\lambda) = \det \begin{pmatrix} \frac{20}{7} - \lambda & \frac{-14}{7} \\ \frac{-14}{7} & \frac{20}{7} - \lambda \end{pmatrix} = \left( \lambda^2 - \frac{40}{7}\lambda + \frac{204}{49} \right) = \left( \lambda - \frac{6}{7} \right) \left( \lambda - \frac{34}{7} \right) \implies \begin{cases} \lambda_1 = 34/7 \\ \lambda_2 = 6/7 \end{cases}$$

- $\lambda_1 = 34/7 \implies \mathbf{v}^{(1)} = \text{Ker} \begin{pmatrix} 20-34 & -14 \\ -14 & 20-34 \end{pmatrix} = \text{Ker} \begin{pmatrix} -14 & -14 \\ -14 & -14 \end{pmatrix} \implies \mathbf{v}^{(1)} = \frac{\sqrt{2}}{2} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$
- $\lambda_2 = 6/7 \implies \mathbf{v}^{(2)} = \text{Ker} \begin{pmatrix} 20-6 & -14 \\ -14 & 20-6 \end{pmatrix} = \text{Ker} \begin{pmatrix} 14 & -14 \\ -14 & 14 \end{pmatrix} \implies \mathbf{v}^{(2)} = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$



(d) Compute the coordinates of the points $A = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $B = \begin{pmatrix} 4 \\ 0 \end{pmatrix}$ in the system of coordinates adapted to the Gaussian.

**Solution:**

$$A_{<\mathcal{V}>} = {}_{<\mathcal{E}>}\mathbf{M}_{<\mathcal{V}>}(A - \mu) = \frac{\sqrt{2}}{2}\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}^T \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 2 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} \sqrt{2} \\ -\sqrt{2} \end{pmatrix}$$

$$B_{<\mathcal{V}>} = {}_{<\mathcal{E}>}\mathbf{M}_{<\mathcal{V}>}(B - \mu) = \frac{\sqrt{2}}{2}\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}^T \left(\begin{pmatrix} 4 \\ 0 \end{pmatrix} - \begin{pmatrix} 2 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} -\sqrt{2} \\ \sqrt{2} \end{pmatrix}$$

(e) Write the PDF of the resulting Gaussian model in the new coordinate system and the standard coordinate system.

**Solution:**

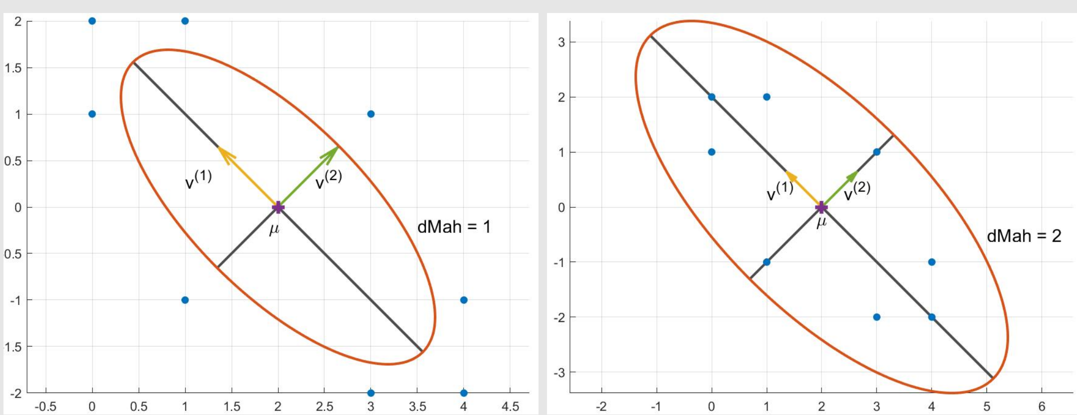
OLD: $f(\mathbf{x}) = (2\pi)^{-1}\left(\frac{34 \cdot 6}{7^2}\right)^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\left(\mathbf{x} - \begin{pmatrix} 2 \\ 0 \end{pmatrix}\right)^T \begin{pmatrix} 20/7 & -2 \\ -2 & 20/7 \end{pmatrix}^{-1} \left(\mathbf{x} - \begin{pmatrix} 2 \\ 0 \end{pmatrix}\right)\right]$

NEW:

$$\begin{aligned} f(\mathbf{x}) = (2\pi)^{-1}\left(\frac{34 \cdot 6}{7^2}\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\mathbf{x}^T \begin{pmatrix} 34/7 & 0 \\ 0 & 6/7 \end{pmatrix}^{-1} \mathbf{x}\right) = \\ = (2\pi)^{-1}\left(\frac{34 \cdot 6}{7^2}\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\mathbf{x}^T \begin{pmatrix} 7/34 & 0 \\ 0 & 7/6 \end{pmatrix} \mathbf{x}\right) \\ = (2\pi)^{-1}\left(\frac{34 \cdot 6}{7^2}\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\left(\frac{7}{34}{x_1}^2 + \frac{7}{6}{x_2}^2\right)\right) \end{aligned}$$

(f) Draw the ellipses defined by the points with Mahalanobis distances equal to 1 and 2 (without computing it).

**Solution:**




(g) Write the equation of the ellipse defined by the points at Mahalanobis distance 3 in the new coordinate system and the cartesian coordinate system.

**Solution:**

OLD

$$\sqrt{(\mathbf{x}-\mu)^T\,\Sigma^{-1}\,(\mathbf{x}-\mu)} = 3 \implies 9 = (\mathbf{x}-\mu)^T\,\Sigma^{-1}\,(\mathbf{x}-\mu)$$

$$\implies 9 = \begin{pmatrix} x_1-2 & x_2 \end{pmatrix} \begin{pmatrix} 20/7 & -2 \\ -2 & 20/7 \end{pmatrix}^{-1} \begin{pmatrix} x_1-2 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1-2 & x_2 \end{pmatrix} \frac{7}{102} \begin{pmatrix} 10 & 7 \\ 7 & 10 \end{pmatrix} \begin{pmatrix} x_1-2 \\ x_2 \end{pmatrix}$$

$$\implies 9 = \frac{7}{102}\left((x_1-2)(10(x_1-2)+7x_2) + x_2(7(x_1-2)+10x_2)\right)$$

$$\implies 9 = \frac{7}{51}\left(5x_1^2 + 7x_1x_2 - 20x_1 + 5x_2^2 - 14x_2 + 20\right)$$

NEW

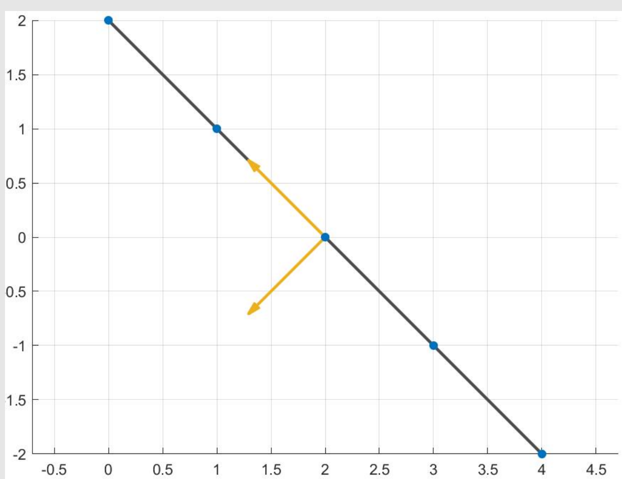
$$9 = (\mathbf{x}-\mu)^T\,\Sigma^{-1}\,(\mathbf{x}-\mu) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 34/7 & 0 \\ 0 & 6/7 \end{pmatrix}^{-1} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 7/34 & 0 \\ 0 & 7/6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\implies 9 = \frac{x_1^2}{34/7} + \frac{x_2^2}{6/7}$$

3. Repeat the previous exercise (Ex. 3) with the set of points $\left\{ \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \begin{pmatrix} 4 \\ -2 \end{pmatrix} \right\} \subset \mathbb{R}^2$. What problems do you encounter?

**Solution:**

(a)




(b) $\mu = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$, $\mathbf{\Sigma} = \frac{1}{2}\begin{pmatrix} 5 & -5 \\ -5 & 5 \end{pmatrix}$

(c) $\lambda_1 = 5, \boxed{\lambda_2 = 0}$ with $\mathbf{v}^{(1)} = \frac{\sqrt{2}}{2}\begin{pmatrix} -1 \\ 1 \end{pmatrix}, \mathbf{v}^{(2)} = \frac{\sqrt{2}}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

(d) Same.

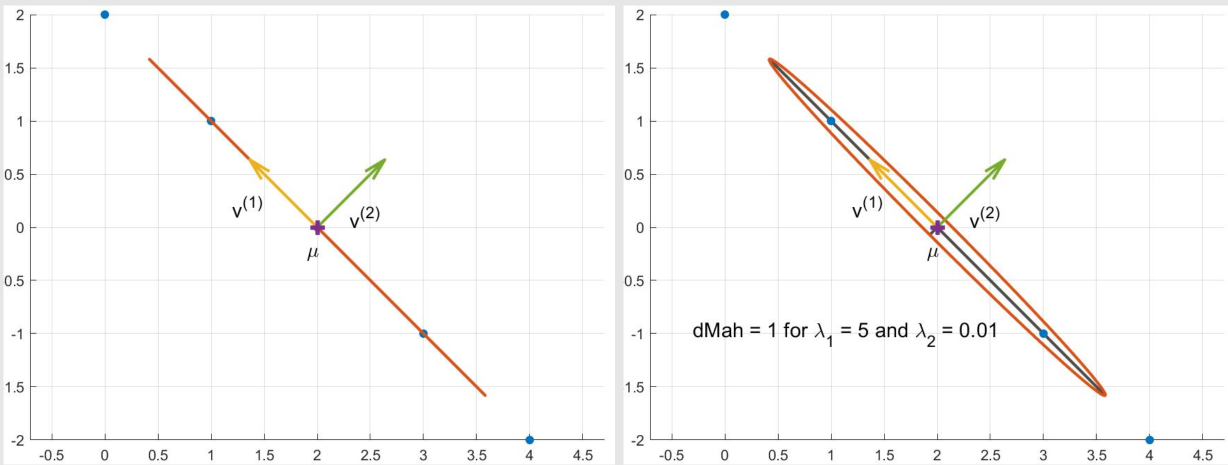
(e) PROBLEM: The covariance matrix is not invertible in this case since $\lambda_2 = 0$. Thus, we set $\lambda_2 := 0.3$.

$\implies \lambda_1 = 5, \boxed{\lambda_2 := 0.3}$ with $\mathbf{v}^{(1)} = \frac{\sqrt{2}}{2}(-1,1), \mathbf{v}^{(2)} = \frac{\sqrt{2}}{2}(1,1)$

Thus, in the new coordinate system:

$$f(\mathbf{x}) = (2\pi)^{-1}(5\cdot 0.3)^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\mathbf{x}^T\begin{pmatrix}5 & 0\\ 0 & 0.3\end{pmatrix}^{-1}\mathbf{x}\right)$$
$$= (2\pi)^{-1}(5\cdot 0.3)^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\mathbf{x}^T\begin{pmatrix}1/5 & 0\\ 0 & 10/3\end{pmatrix}\mathbf{x}\right)$$

(f)




(g) In the new coordinate system:

$$9 = (\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu) = \begin{pmatrix}x_1 & x_2\end{pmatrix}\begin{pmatrix}5 & 0\\ 0 & 0.3\end{pmatrix}^{-1}\begin{pmatrix}x_1\\ x_2\end{pmatrix} = \begin{pmatrix}x_1 & x_2\end{pmatrix}\begin{pmatrix}1/5 & 0\\ 0 & 10/3\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}$$
$$\implies 9 = \frac{x_1^2}{5} + \frac{x_2^2}{0.3}$$

4. *(Exam 2019-2020)* Consider two different Gaussian models, $G_1$ and $G_2$ with means and covariances:

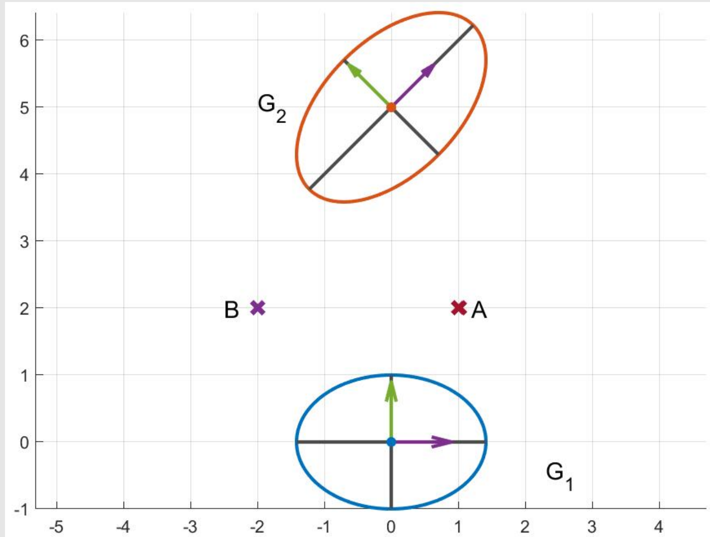
$$\mu_1 = \begin{pmatrix}0\\ 0\end{pmatrix}, \mathbf{\Sigma}_1 = \begin{pmatrix}2 & 0\\ 0 & 1\end{pmatrix},$$
$$\mu_2 = \begin{pmatrix}0\\ 5\end{pmatrix}, \mathbf{\Sigma}_2 = \begin{pmatrix}2 & 1\\ 1 & 2\end{pmatrix}.$$

(a) Which Gaussian gives larger probability the points $A = \begin{pmatrix}1\\ 2\end{pmatrix}$ and $B = \begin{pmatrix}-2\\ 2\end{pmatrix}$? First, make a guess by studying graphically both Gaussians. Then, check your answer with computations.

**Solution:**

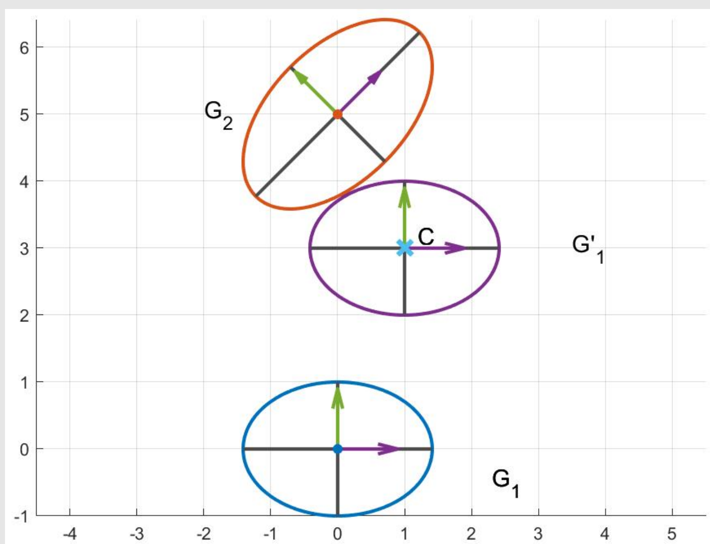
$$\mathbf{\Sigma}_1^{-1} = \begin{pmatrix}1/2 & 0\\ 0 & 1\end{pmatrix},\quad \mathbf{\Sigma}_2^{-1} = \frac{1}{3}\begin{pmatrix}2 & -1\\ -1 & 2\end{pmatrix},$$

$P(A|G_1) = (2\pi)^{-1}\det(\mathbf{\Sigma}_1)^{-\frac{1}{2}}\exp\left[-\frac{1}{2}(A-\mu_1)^T\mathbf{\Sigma}_1^{-1}(A-\mu_1)\right] = \cdots = 0.0119$
$P(A|G_2) = \cdots = 0.0012$
$P(B|G_1) = \cdots = 0.0056$
$P(B|G_2) = \cdots = 0.0089$
$\implies P(A|G_1) > P(A|G_2) \quad \& \quad P(B|G_1) < P(B|G_2)$

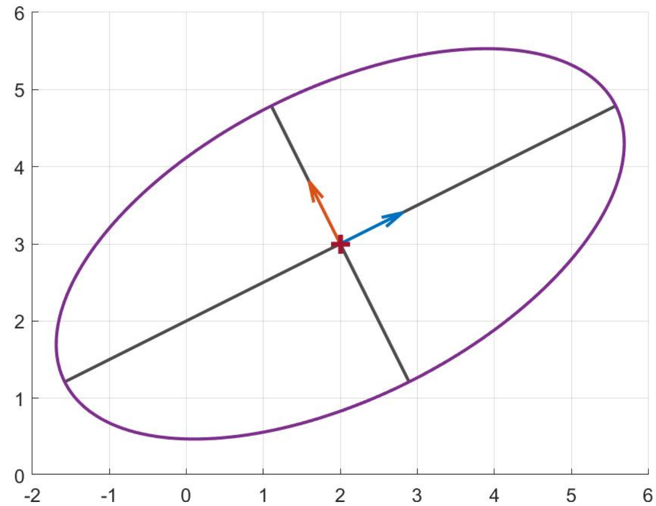



(b) Find a new system of coordinates for $G_1$ (basis and center) such that the point $C = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ is classified as 1 and not 2, i.e, that it is better explained by $G_1$ than $G_2$. Compute the new Gaussian parameters.

**Solution:**
For example, take $\mu_1' = C$ and the same basis as $G_1$.



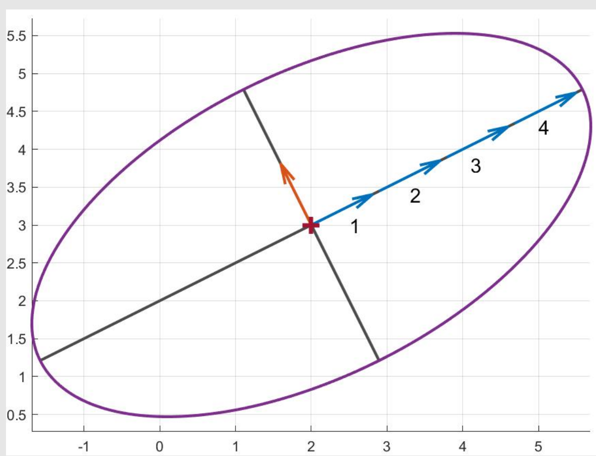

5. *(Exam 2019-2020)* Consider a Gaussian distribution $G$ whose ellipse of Mahalanobis distance equal to 1 is (best seen in colour)




(a) Find the mean, the eigenvectors and eigenvalues that determine this Gaussian distribution.

**Solution:**
The mean of the distribution is the centre of the ellipse, thus, $\mu = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$.

The first eigenvector is the one with highest eigenvalue, thus the blue vector. Clearly, it is the unitary vector with the same direction as the vector $\begin{pmatrix} 2 \\ 1 \end{pmatrix} \implies \mathbf{u}^{(1)} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix}$. Its eigenvalue can be determined by finding how many of the blue vector we can allocate between the mean and the ellipse $\implies 4 = a = \sqrt{\lambda_1} \implies \lambda_1 = 16$.




The second eigenvector can be found in the same way as the first, but there is another way to do so. Since we are working on $\mathbb{R}^2$, there are only two unitary vectors that are orthogonal with $\mathbf{u}^{(1)}$: the orange vector (let it be $\mathbf{u}^{(2)}$) and $-\mathbf{u}^{(2)}$. Thus, $\mathbf{u}^{(2)}$ can be the unitary vector in the same direction as $\begin{pmatrix} -1 \\ 2 \end{pmatrix}$ or in the direction as $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$, since both satisfy

$$(-1, 2) \cdot \mathbf{u}^{(1)} = 0,$$

$$(1, -2) \cdot \mathbf{u}^{(1)} = 0.$$

But, looking at the figure we determine that $\mathbf{u}^{(2)} = \frac{1}{\sqrt{5}} \begin{pmatrix} -1 \\ 2 \end{pmatrix}$, with $2 = b = \sqrt{\lambda_2} \implies \lambda_2 = 4$.

(b) Calculate the covariance matrix $\mathbf{\Sigma}$.

**Solution:** $\mathbf{\Sigma}$ can be computed in three different ways:

- $\mathbf{\Sigma} = \mathbf{UDU}^\top = \frac{1}{\sqrt{5}}\begin{pmatrix}2 & -1\\ 1 & 2\end{pmatrix}\begin{pmatrix}16 & 0\\ 0 & 4\end{pmatrix}\frac{1}{\sqrt{5}}\begin{pmatrix}2 & 1\\ -1 & 2\end{pmatrix} = \frac{1}{5}\begin{pmatrix}68 & 24\\ 24 & 32\end{pmatrix}$
- $\mathbf{\Sigma} = 16\mathbf{u}_1 * \mathbf{u}_1^\top + 4\mathbf{u}_2 * \mathbf{u}_2^\top$
- $\mathbf{\Sigma}$ is a symmetric matrix with 3 unknows: $\mathbf{\Sigma} = \left(\begin{smallmatrix}a & b\\ b & c\end{smallmatrix}\right)$ and we know the eigenvectors and eigenvalues. Using $\left(\begin{smallmatrix}a-16 & b\\ b & c-16\end{smallmatrix}\right)\mathbf{u}_1 = \mathbf{0}$ and $\left(\begin{smallmatrix}a-4 & b\\ b & c-4\end{smallmatrix}\right)\mathbf{u}_2 = \mathbf{0}$. Then $\mathbf{\Sigma}$ is obtained after solving the resulting four linear equations.

(c) Indicate how you can use this to estimate the $\mathbf{\Sigma}^{-1}$. Which are the values of its eigenvalues?

**Solution:**
$\mathbf{\Sigma}^{-1} = (\mathbf{UDU}^\top)^{-1} = (\mathbf{DU}^\top)^{-1}\mathbf{U}^{-1} = (\mathbf{U}^\top)^{-1}\mathbf{D}^{-1}\mathbf{U}^{-1}$
But, since $\mathbf{U}$ is orthonormal, $\mathbf{U}^{-1} = \mathbf{U}^\top \implies \mathbf{\Sigma}^{-1} = \mathbf{UD}^{-1}\mathbf{U}^\top$, which has the form of an SVD.
Besides, since $\mathbf{D} = \text{diag}(\lambda_1, \lambda_2)$ is a diagonal matrix, $\mathbf{D}^{-1} = \text{diag}\left(\frac{1}{\lambda_1}, \frac{1}{\lambda_2}\right)$.
Therefore, the eigenvalues of $\mathbf{\Sigma}^{-1}$ are $\lambda_1' = \frac{1}{16}, \lambda_2' = \frac{1}{4}$, and $\mathbf{\Sigma}^{-1} = \mathbf{UD}^{-1}\mathbf{U}^\top = \frac{1}{80}\begin{pmatrix}8 & -6\\ -17 & 9\end{pmatrix}$.

(d) Assume this Gaussian $G$ has mean $\mu = (5, 8)$ and covariance $\mathbf{\Sigma} = \left(\begin{smallmatrix}5 & 1\\ 1 & 2\end{smallmatrix}\right)$ and that it models distribution of the mark of the final exam (first variable) and mark of the continuous evaluation (second variable). What are the mean and variance of the continuous evaluation?

**Solution:** Let $X$ be the mark of the final exam and $Y$ the mark of the continuous evaluation, then

$$\mu = \begin{pmatrix}\mu_X\\ \mu_Y\end{pmatrix}$$

and

$$\mathbf{\Sigma} = \begin{pmatrix}\text{var}(X) & \text{cov}(X,Y)\\ \text{cov}(Y,X) & \text{var}(Y)\end{pmatrix}.$$

Therefore, $\mu_Y = 8$ and $\text{var}(Y) = 2$.

6. *(In Jupyter Notebook)*

(a) Write a Python function that, given a mean and a covariance (in $\mathbb{R}^2$), returns the principal directions of the Gaussian and their variances, and plots the points at Mahalanobis distance $k$ (where $k$ is given as input).

(b) Write a Python function that, given a set of points, estimates the Gaussian model from which they are drawn and plots the points, together with several (of choice) ellipses of Mahalanobis distances and the principal directions.

(c) Write a Python function that, given a Gaussian model, i.e, $\mu, \mathbf{\Sigma}$, draws $n$ samples from that distribution. *TIP*: Use the Numpy function *np.random.multivariate_normal*.