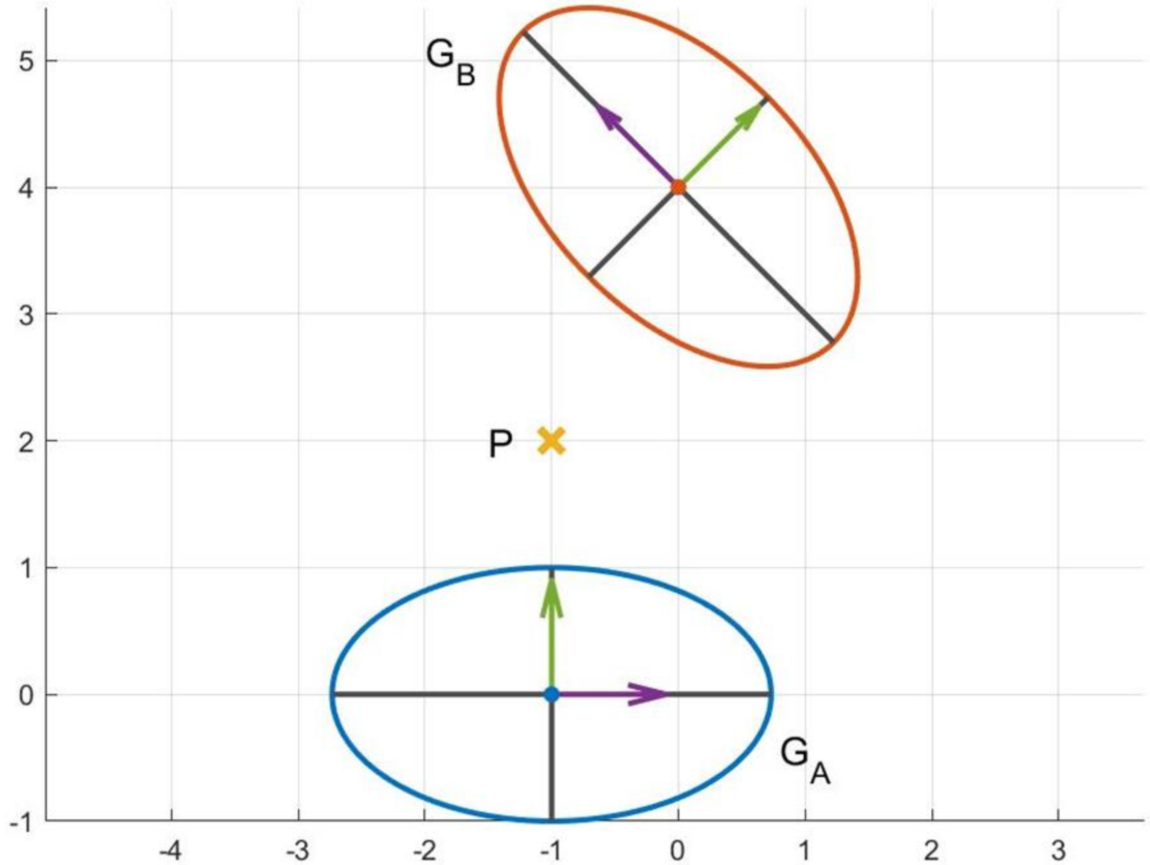
EXERCICE 1 (*Outline*)

$$\boldsymbol{\mu}_A = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \boldsymbol{\Sigma}_\mathrm{A} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\boldsymbol{\mu}_B = \begin{pmatrix} 0 \\ 4 \end{pmatrix}, \boldsymbol{\Sigma}_\mathrm{B} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

(a) Draw the ellipses corresponding to the Mahalanobis distances 1 of both Gaussian models.



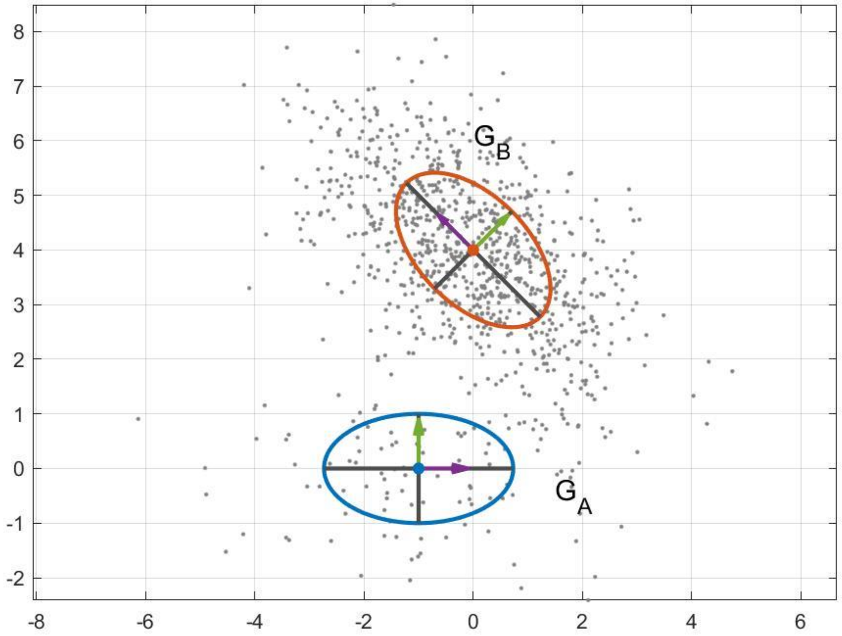
EXERCICE 1 (*Outline*)

$$\boldsymbol{\mu}_A = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \boldsymbol{\Sigma}_\mathrm{A} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$$

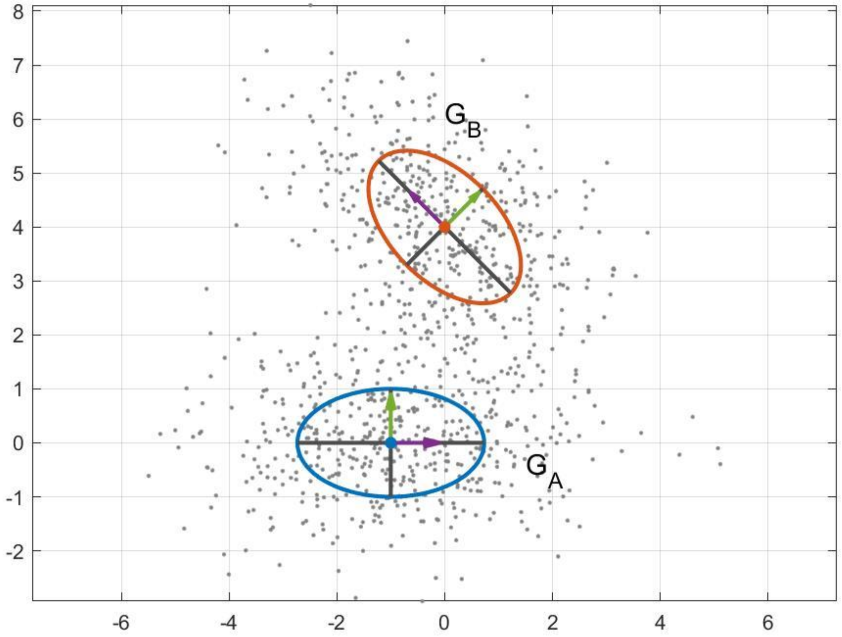
$$\boldsymbol{\mu}_B = \begin{pmatrix} 0 \\ 4 \end{pmatrix}, \boldsymbol{\Sigma}_\mathrm{B} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

(b) Draw 1000 random samples from a mixture of Gaussians G_A and G_B for different values of the mixing coefficients (**modeling point of view**).

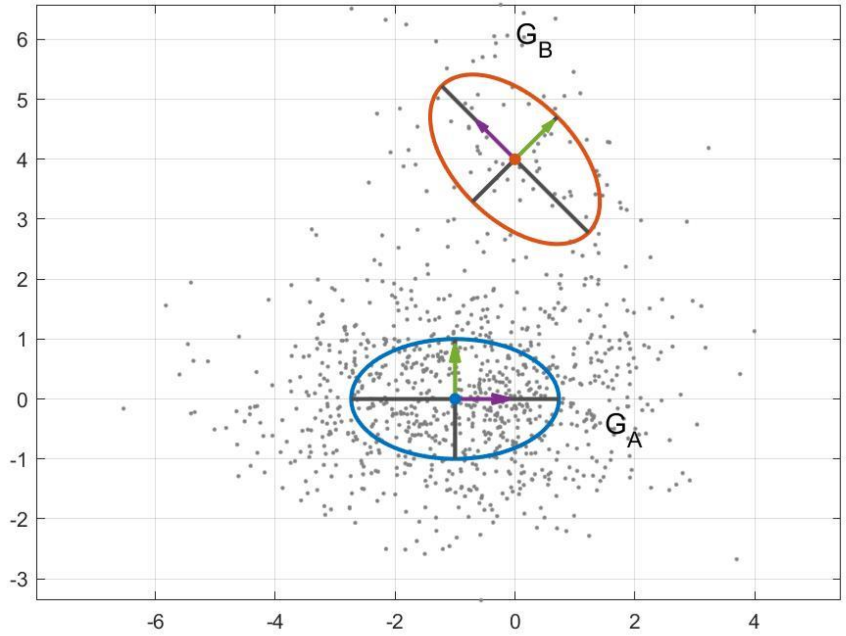
$\pi_A = 0.1$
$\pi_B = 0.9$



$\pi_A = 0.5$
$\pi_B = 0.5$



$\pi_A = 0.9$
$\pi_B = 0.1$

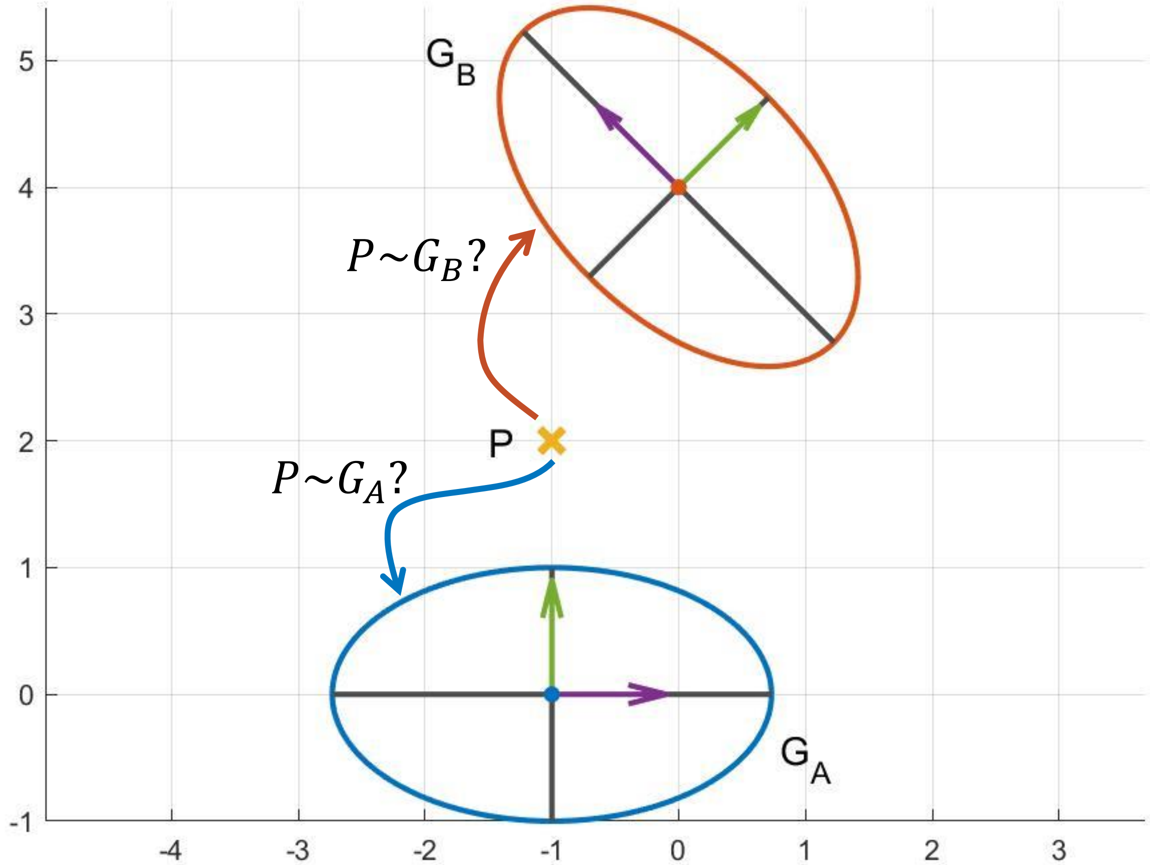


EXERCICE 1 (*Outline*)

$$\boldsymbol{\mu}_A = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \boldsymbol{\Sigma}_\mathrm{A} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\boldsymbol{\mu}_B = \begin{pmatrix} 0 \\ 4 \end{pmatrix}, \boldsymbol{\Sigma}_\mathrm{B} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

(c) Estimate the probability a posteriori of a point $P = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$ of being generated by G_A and G_B (**clustering point of view**) given mixing coefficients $\pi_A = 0.1$ and $\pi_A = 0.2$.

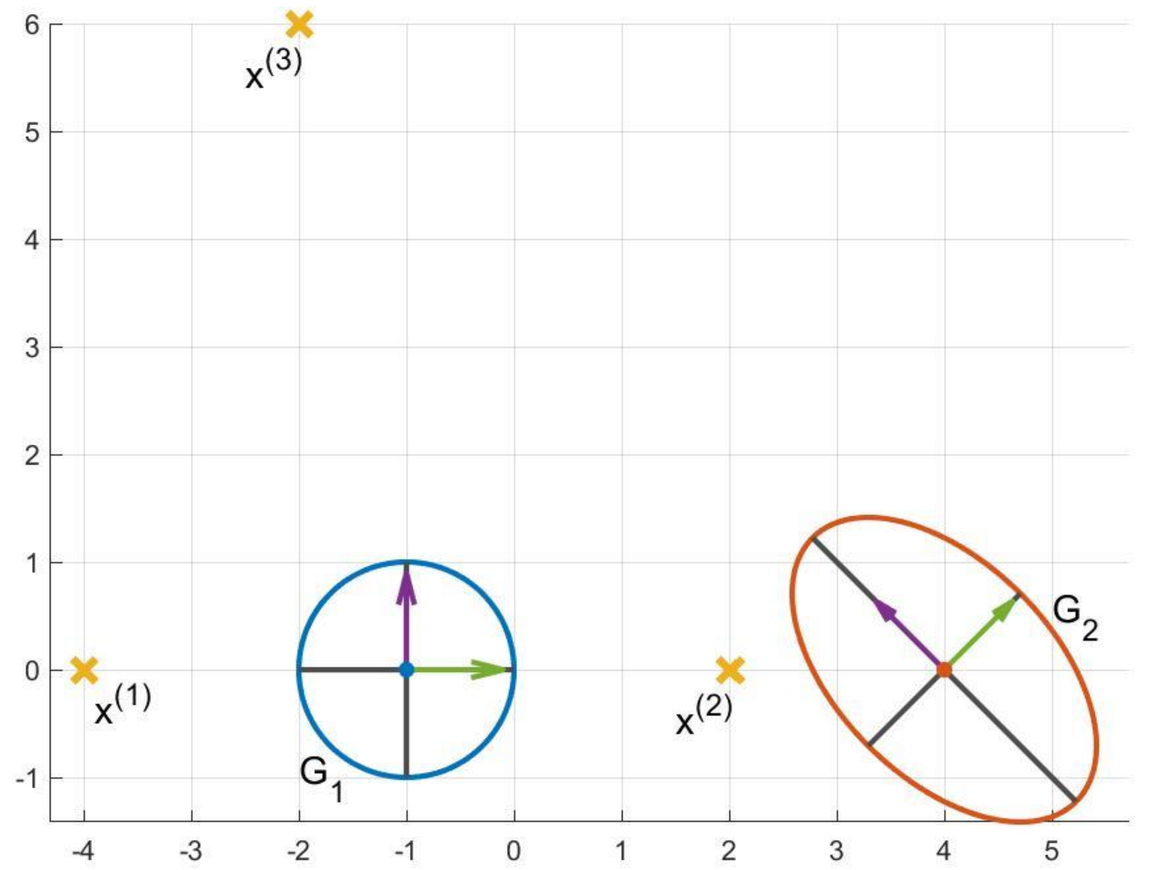


EXERCICE 2

Given the points $\left\{\boldsymbol{x}^{(1)} = \begin{pmatrix}-4\\0\end{pmatrix}, \boldsymbol{x}^{(2)} = \begin{pmatrix}2\\0\end{pmatrix}, \boldsymbol{x}^{(3)} = \begin{pmatrix}-2\\6\end{pmatrix}\right\} \subset \mathbb{R}^2$, derive an Expectation-Maximisation update for a Mixture of Gaussian with initial condition

$$\pi_1 = 0.8, \boldsymbol{\mu}_1 = \begin{pmatrix}-1\\0\end{pmatrix}, \boldsymbol{\Sigma}_1 = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix},$$

$$\pi_2 = 0.2, \boldsymbol{\mu}_2 = \begin{pmatrix}4\\0\end{pmatrix}, \boldsymbol{\Sigma}_2 = \begin{pmatrix}2 & -1\\-1 & 2\end{pmatrix}.$$



Be prepared to use Python! This will avoid you to waste time with the computations.

EXERCICE 2

$$\left\{\boldsymbol{x}^{(1)} = \begin{pmatrix}-4\\0\end{pmatrix}, \boldsymbol{x}^{(2)} = \begin{pmatrix}2\\0\end{pmatrix}, \boldsymbol{x}^{(3)} = \begin{pmatrix}-2\\6\end{pmatrix}\right\} \subset \mathbb{R}^2$$

$$\pi_1 = 0.8, \boldsymbol{\mu}_1 = \begin{pmatrix}-1\\0\end{pmatrix}, \quad \boldsymbol{\Sigma}_1 = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix},$$

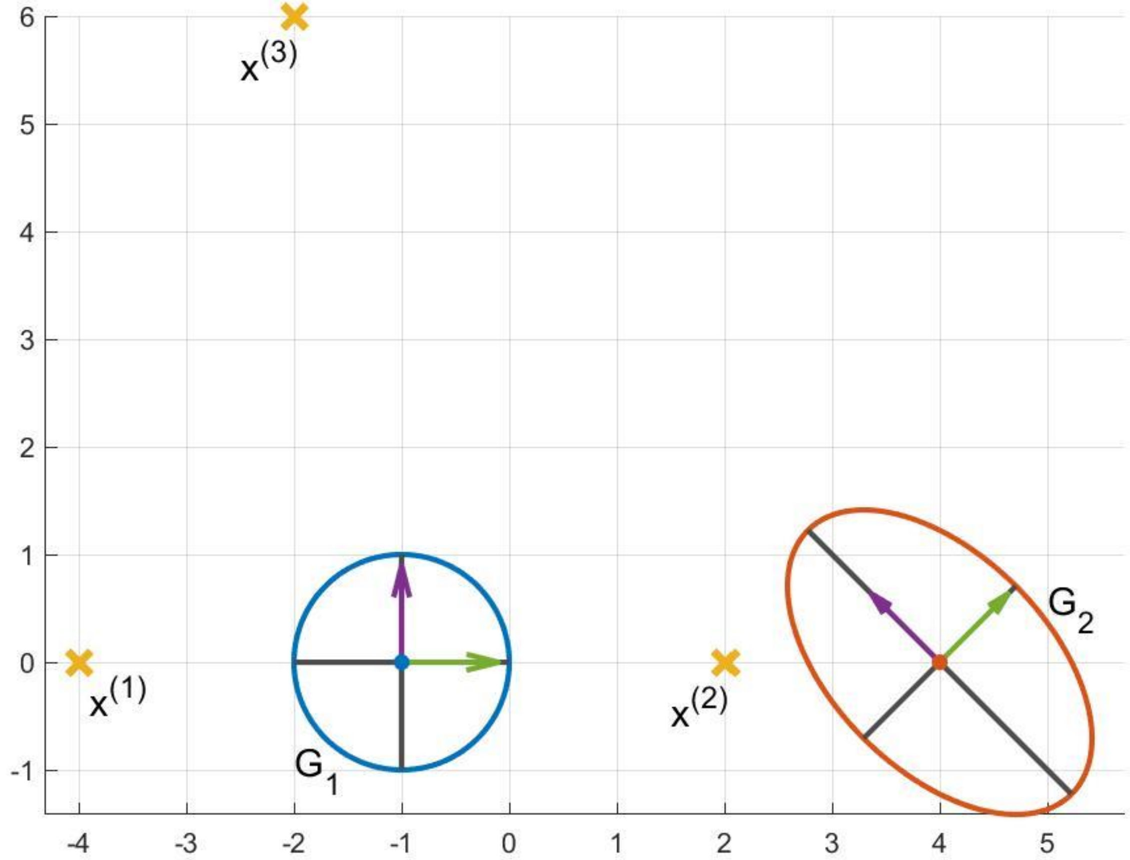
$$\pi_2 = 0.2, \boldsymbol{\mu}_2 = \begin{pmatrix}4\\0\end{pmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{pmatrix}2 & -1\\-1 & 2\end{pmatrix}.$$

(a) Draw the given points and the initial Gaussian models (ellipses). What do you expect the responsibilities to be? $r_1^{(1)} > r_2^{(1)}$ or otherwise? And what about $\boldsymbol{x}^{(2)}$ and $\boldsymbol{x}^{(3)}$?

To draw the Gaussian models: previous session (Gaussian models)

TIP

```
np.linalg.svd
```



EXERCICE 2

$$\left\{\boldsymbol{x}^{(1)} = \begin{pmatrix}-4\\0\end{pmatrix}, \boldsymbol{x}^{(2)} = \begin{pmatrix}2\\0\end{pmatrix}, \boldsymbol{x}^{(3)} = \begin{pmatrix}-2\\6\end{pmatrix}\right\} \subset \mathbb{R}^2$$

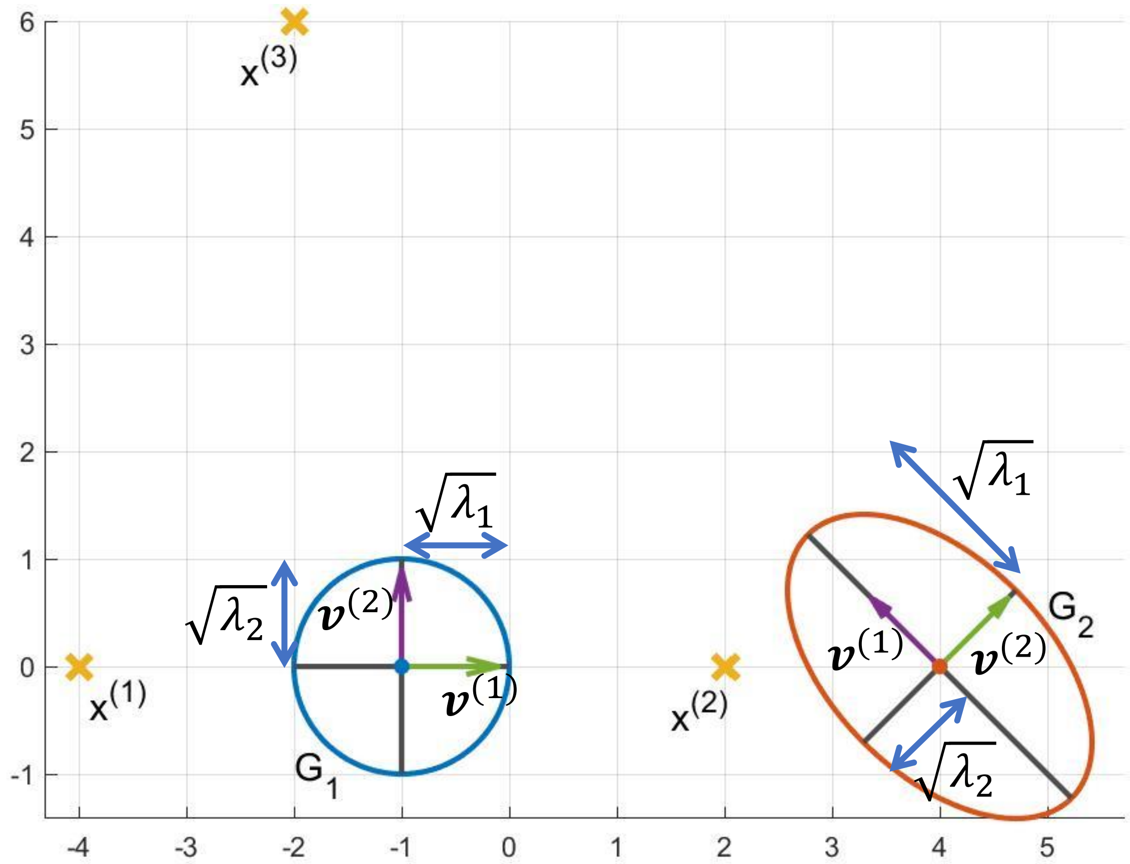
$$\pi_1 = 0.8, \boldsymbol{\mu}_1 = \begin{pmatrix}-1\\0\end{pmatrix}, \quad \boldsymbol{\Sigma}_1 = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix},$$

$$\pi_2 = 0.2, \boldsymbol{\mu}_2 = \begin{pmatrix}4\\0\end{pmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{pmatrix}2 & -1\\-1 & 2\end{pmatrix}.$$

(a) Draw the given points and the initial Gaussian models (ellipses). What do you expect the responsibilities to be? $r_1^{(1)} > r_2^{(1)}$ or otherwise? And what about $\boldsymbol{x}^{(2)}$ and $\boldsymbol{x}^{(3)}$?

$$\boldsymbol{\Sigma}_1 = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix} \xrightarrow{\text{SVD}} \begin{cases} \lambda_1 = 1 & \boldsymbol{v}^{(1)} = \begin{pmatrix}1\\0\end{pmatrix} \\ \lambda_2 = 1 & \boldsymbol{v}^{(2)} = \begin{pmatrix}0\\1\end{pmatrix} \end{cases}$$

$$\boldsymbol{\Sigma}_2 = \begin{pmatrix}2 & -1\\-1 & 2\end{pmatrix} \xrightarrow{\text{SVD}} \begin{cases} \lambda_1 = 3 & \boldsymbol{v}^{(1)} = \frac{\sqrt{2}}{2}\begin{pmatrix}-1\\1\end{pmatrix} \\ \lambda_2 = 1 & \boldsymbol{v}^{(2)} = \frac{\sqrt{2}}{2}\begin{pmatrix}1\\1\end{pmatrix} \end{cases}$$



EXERCICE 2

$$\left\{\boldsymbol{x}^{(1)} = \begin{pmatrix}-4\\0\end{pmatrix}, \boldsymbol{x}^{(2)} = \begin{pmatrix}2\\0\end{pmatrix}, \boldsymbol{x}^{(3)} = \begin{pmatrix}-2\\6\end{pmatrix}\right\} \subset \mathbb{R}^2$$

$$\pi_1 = 0.8, \boldsymbol{\mu}_1 = \begin{pmatrix}-1\\0\end{pmatrix}, \quad \boldsymbol{\Sigma}_1 = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix},$$

$$\pi_2 = 0.2, \boldsymbol{\mu}_2 = \begin{pmatrix}4\\0\end{pmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{pmatrix}2 & -1\\-1 & 2\end{pmatrix}.$$

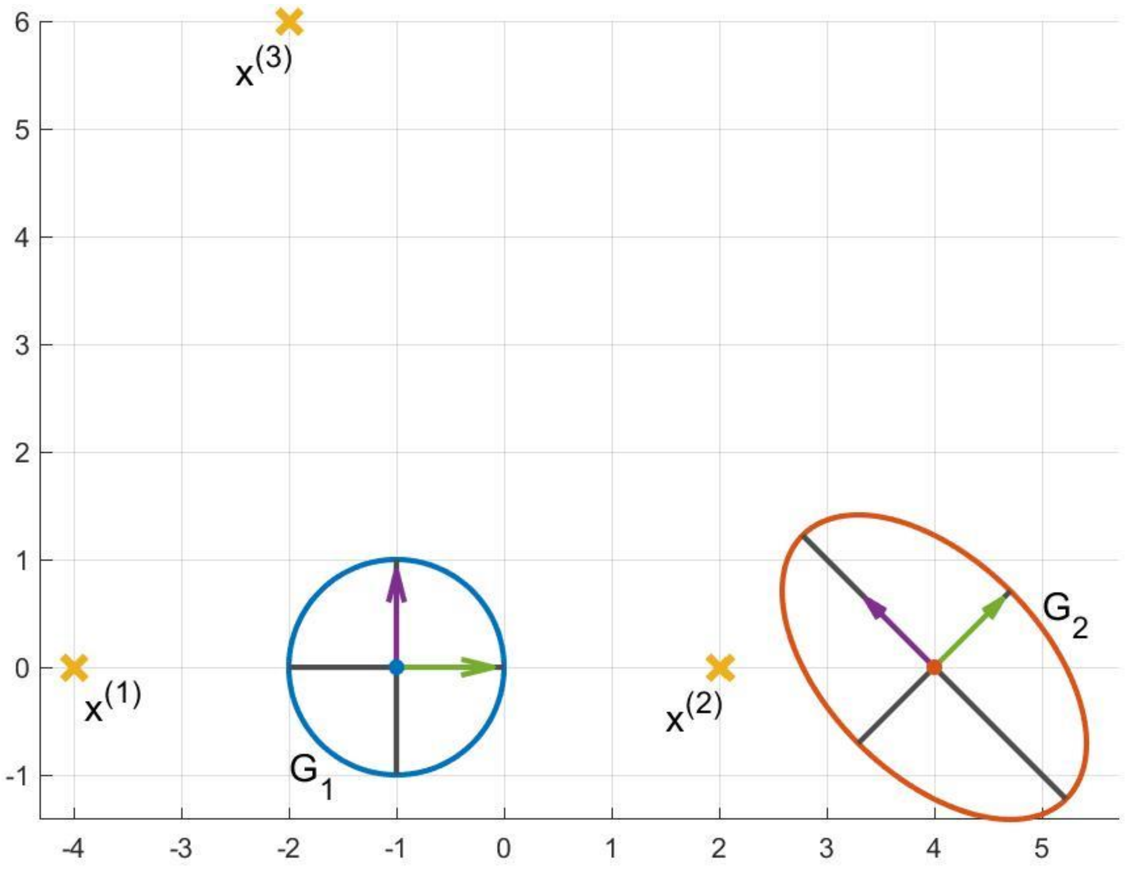
(a) Draw the given points and the initial Gaussian models (ellipses). What do you expect the responsibilities to be? $r_1^{(1)} > r_2^{(1)}$ or otherwise? And what about $\boldsymbol{x}^{(2)}$ and $\boldsymbol{x}^{(3)}$?

REMINDER

$k = 1, \ldots, K$ $\longleftarrow$ clusters / Gaussian models (in this case, $K = 2$)
$n = 1, \ldots, N$ $\longleftarrow$ data points (in this case, $N = 3$)

$$\boxed{r_k^{(n)} = P\left(k \middle| \boldsymbol{x}^{(n)}\right)}$$

The responsibilities are the probability of a given point $\boldsymbol{x}^{(n)}$ of being generated by component k.



EXERCICE 2

$$\left\{x^{(1)} = \begin{pmatrix}-4\\0\end{pmatrix}, x^{(2)} = \begin{pmatrix}2\\0\end{pmatrix}, x^{(3)} = \begin{pmatrix}-2\\6\end{pmatrix}\right\} \subset \mathbb{R}^2$$

$$\pi_1 = 0.8, \boldsymbol{\mu}_1 = \begin{pmatrix}-1\\0\end{pmatrix}, \quad \boldsymbol{\Sigma}_1 = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix},$$

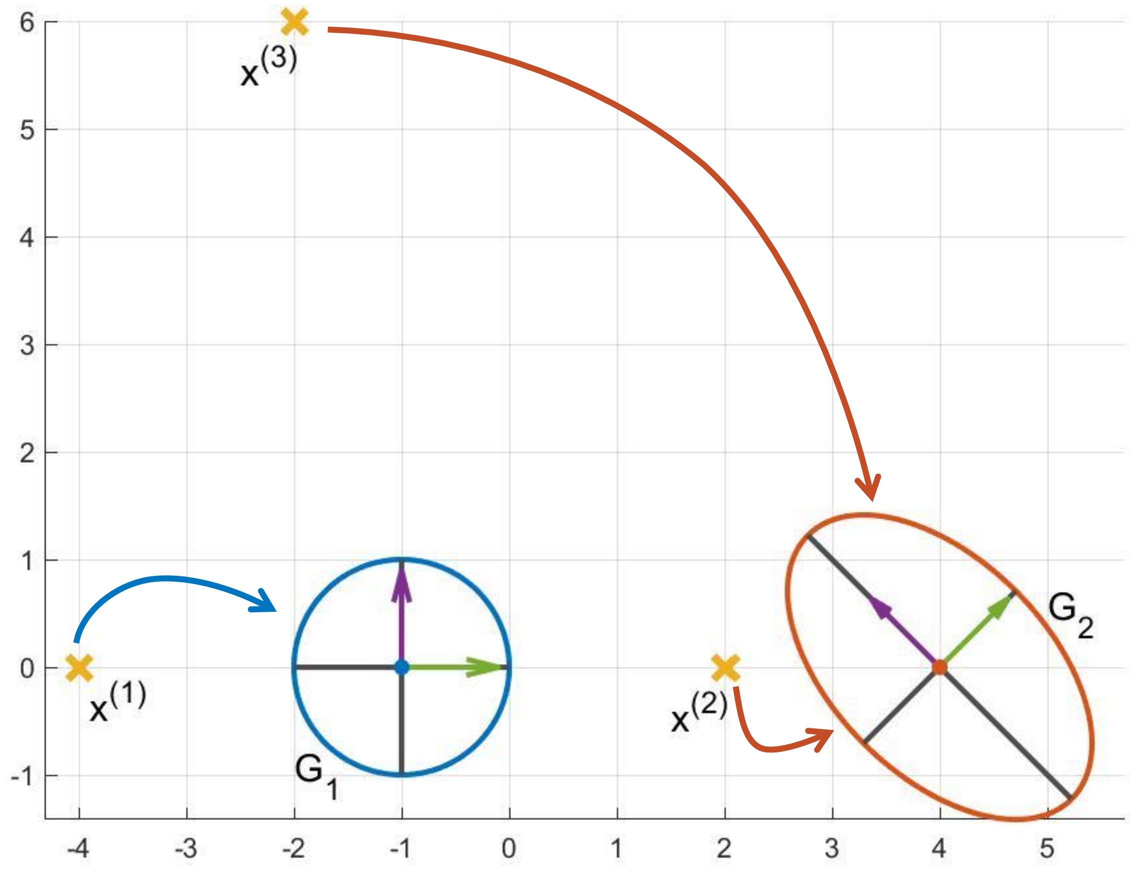
$$\pi_2 = 0.2, \boldsymbol{\mu}_2 = \begin{pmatrix}4\\0\end{pmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{pmatrix}2 & -1\\-1 & 2\end{pmatrix}.$$

(a) Draw the given points and the initial Gaussian models (ellipses). What do you expect the responsibilities to be? $r_1^{(1)} > r_2^{(1)}$ or otherwise? And what about $\boldsymbol{x}^{(2)}$ and $\boldsymbol{x}^{(3)}$?

$r_1^{(1)} > r_2^{(1)}$

$r_1^{(2)} < r_2^{(2)}$

$r_1^{(3)} < r_2^{(3)}$



EXERCICE 2

$$\left\{\boldsymbol{x}^{(1)} = \begin{pmatrix}-4\\0\end{pmatrix}, \boldsymbol{x}^{(2)} = \begin{pmatrix}2\\0\end{pmatrix}, \boldsymbol{x}^{(3)} = \begin{pmatrix}-2\\6\end{pmatrix}\right\} \subset \mathbb{R}^2$$

$$\pi_1 = 0.8, \boldsymbol{\mu}_1 = \begin{pmatrix}-1\\0\end{pmatrix}, \quad \boldsymbol{\Sigma}_1 = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix},$$

$$\pi_2 = 0.2, \boldsymbol{\mu}_2 = \begin{pmatrix}4\\0\end{pmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{pmatrix}2 & -1\\-1 & 2\end{pmatrix}.$$

(b) (Expectation step) Compute the responsibilities. Are they as you expected?

$$r_k^{(n)} = P\left(k \middle| \boldsymbol{x}^{(n)}\right) = \frac{\pi_k P\left(\boldsymbol{x}^{(n)} \middle| k\right)}{\sum_{l=1}^{K} \pi_l P\left(\boldsymbol{x}^{(n)} \middle| l\right)}$$

where $\left(\boldsymbol{x}^{(n)} \middle| l\right) \sim \mathcal{N}\left(\boldsymbol{x}^{(n)} \middle| \boldsymbol{\mu}_l, \boldsymbol{\Sigma}_l\right)$

$$\Longrightarrow \quad P\left(\boldsymbol{x}^{(n)} \middle| l\right) = (2\pi)\det(\boldsymbol{\Sigma}_l)^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\left(\boldsymbol{x}^{(n)} - \boldsymbol{\mu}_l\right)^{\mathrm{T}} \boldsymbol{\Sigma}_l^{-1}\left(\boldsymbol{x}^{(n)} - \boldsymbol{\mu}_l\right)\right]$$

TIP

```
from scipy.stats import multivariate_normal as multiNorm
px = multiNorm.pdf(x,mu,cov)
```

EXERCICE 2

$$\left\{ \boldsymbol{x}^{(1)} = \begin{pmatrix} -4 \\ 0 \end{pmatrix}, \boldsymbol{x}^{(2)} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \boldsymbol{x}^{(3)} = \begin{pmatrix} -2 \\ 6 \end{pmatrix} \right\} \subset \mathbb{R}^2$$

$$\pi_1 = 0.8, \boldsymbol{\mu}_1 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad \boldsymbol{\Sigma}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

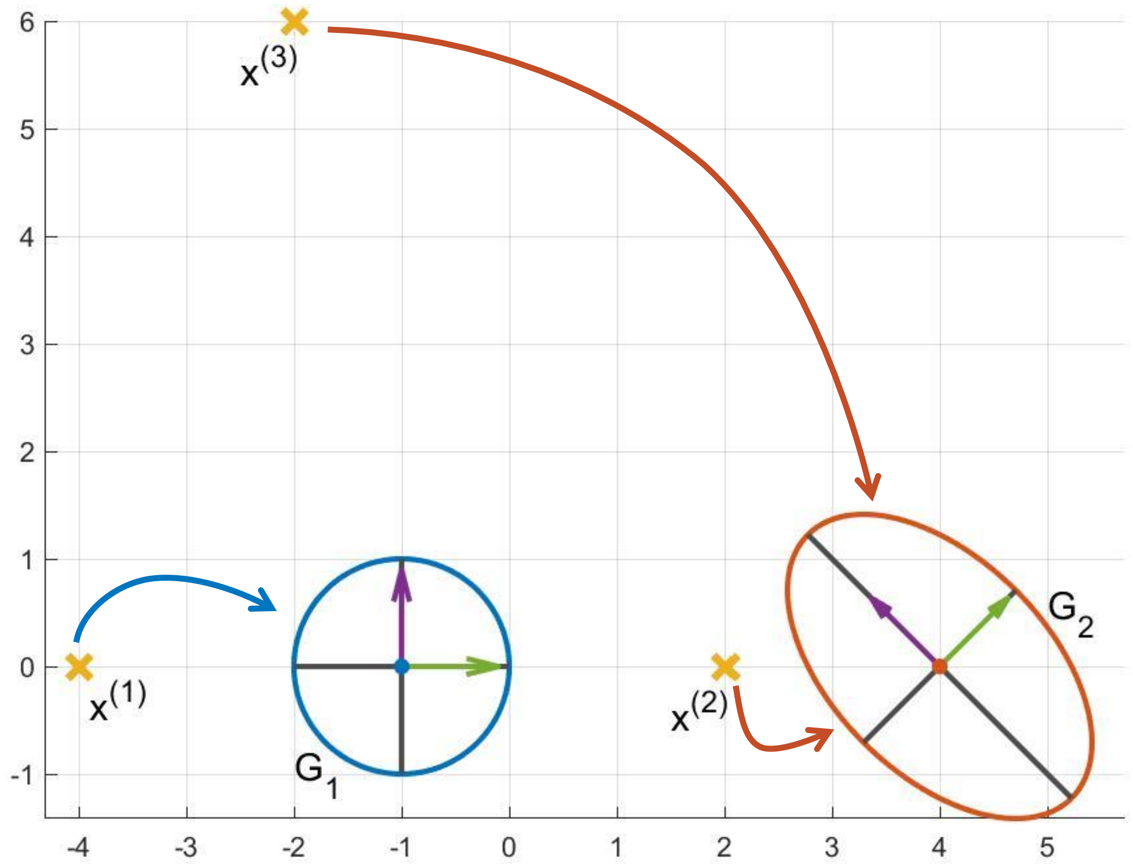
$$\pi_2 = 0.2, \boldsymbol{\mu}_2 = \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

(b) (Expectation step) Compute the responsibilities. Are they as you expected?

$$r_1^{(1)} = P(1|\boldsymbol{x}^{(1)}) = \cdots = 1$$
$$r_2^{(1)} = P(2|\boldsymbol{x}^{(1)}) = \cdots = 0$$

$$r_1^{(2)} = P(1|\boldsymbol{x}^{(2)}) = \cdots = 0.226$$
$$r_2^{(2)} = P(2|\boldsymbol{x}^{(2)}) = \cdots = 0.774$$

$$r_1^{(3)} = P(1|\boldsymbol{x}^{(3)}) = \cdots = 0.0103$$
$$r_2^{(3)} = P(2|\boldsymbol{x}^{(3)}) = \cdots = 0.9897$$



EXERCICE 2

$$\left\{\boldsymbol{x}^{(1)} = \begin{pmatrix}-4\\0\end{pmatrix}, \boldsymbol{x}^{(2)} = \begin{pmatrix}2\\0\end{pmatrix}, \boldsymbol{x}^{(3)} = \begin{pmatrix}-2\\6\end{pmatrix}\right\} \subset \mathbb{R}^2$$

$$\pi_1 = 0.8, \boldsymbol{\mu}_1 = \begin{pmatrix}-1\\0\end{pmatrix}, \quad \boldsymbol{\Sigma}_1 = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix},$$

$$\pi_2 = 0.2, \boldsymbol{\mu}_2 = \begin{pmatrix}4\\0\end{pmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{pmatrix}2 & -1\\-1 & 2\end{pmatrix}.$$

(c) (Maximisation step) Update the means, covariance matrices and mixing coefficients.

REMINDER

$k = 1, \dots, K \longleftarrow$ clusters / Gaussian models (in this case, $K = 2$)
$n = 1, \dots, N \longleftarrow$ data points (in this case, $N = 3$)

	MEANS UPDATE	COVARIANCE MATRICES UPDATE	MIXING COEFS UPDATE
$N_k = \sum_{n=1}^{N} r_k^{(n)}$	$\boldsymbol{\mu}_k = \frac{1}{N_k}\sum_{n=1}^{N} r_k^{(n)}\boldsymbol{x}^{(n)}$	$\boldsymbol{\Sigma}_k = \frac{1}{N_k}\sum_{n=1}^{N} r_k^{(n)}\left(\boldsymbol{x}^{(n)} - \boldsymbol{\mu}_k\right)\cdot\left(\boldsymbol{x}^{(n)} - \boldsymbol{\mu}_k\right)^{\mathrm{T}}$	$\pi_k = \frac{N_k}{N}$

EXERCICE 2

$$\left\{\boldsymbol{x}^{(1)} = \begin{pmatrix}-4\\0\end{pmatrix}, \boldsymbol{x}^{(2)} = \begin{pmatrix}2\\0\end{pmatrix}, \boldsymbol{x}^{(3)} = \begin{pmatrix}-2\\6\end{pmatrix}\right\} \subset \mathbb{R}^2$$

$$\pi_1 = 0.8, \boldsymbol{\mu}_1 = \begin{pmatrix}-1\\0\end{pmatrix}, \quad \boldsymbol{\Sigma}_1 = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix},$$

$$\pi_2 = 0.2, \boldsymbol{\mu}_2 = \begin{pmatrix}4\\0\end{pmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{pmatrix}2 & -1\\-1 & 2\end{pmatrix}.$$

(c) (Maximisation step) Update the means, covariance matrices and mixing coefficients.

$$N_k = \sum_{n=1}^{N} r_k^{(n)} \longrightarrow \begin{cases} N_1 = r_{11} + r_{21} + r_{31} = \cdots = 1.24 \\ N_2 = r_{12} + r_{22} + r_{32} = \cdots = 1.76 \end{cases}$$

MEANS

$$\boldsymbol{\mu}_k = \frac{1}{N_k}\sum_{n=1}^{N} r_k^{(n)}\boldsymbol{x}^{(n)} \longrightarrow \begin{cases} \boldsymbol{\mu}_1 = \dfrac{1}{N_1}\left(r_1^{(1)}\boldsymbol{x}^{(1)} + r_1^{(2)}\boldsymbol{x}^{(2)} + r_1^{(3)}\boldsymbol{x}^{(3)}\right) = \begin{pmatrix}-2.89\\0.05\end{pmatrix} \\ \boldsymbol{\mu}_2 = \dfrac{1}{N_2}\left(r_2^{(1)}\boldsymbol{x}^{(1)} + r_2^{(2)}\boldsymbol{x}^{(2)} + r_2^{(3)}\boldsymbol{x}^{(3)}\right) = \begin{pmatrix}-2.45\\3.37\end{pmatrix} \end{cases}$$

EXERCICE 2

$$\left\{\boldsymbol{x}^{(1)} = \begin{pmatrix}-4\\0\end{pmatrix}, \boldsymbol{x}^{(2)} = \begin{pmatrix}2\\0\end{pmatrix}, \boldsymbol{x}^{(3)} = \begin{pmatrix}-2\\6\end{pmatrix}\right\} \subset \mathbb{R}^2$$

$$\pi_1 = 0.8, \boldsymbol{\mu}_1 = \begin{pmatrix}-1\\0\end{pmatrix}, \quad \boldsymbol{\Sigma}_1 = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix},$$

$$\pi_2 = 0.2, \boldsymbol{\mu}_2 = \begin{pmatrix}4\\0\end{pmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{pmatrix}2 & -1\\-1 & 2\end{pmatrix}.$$

(c) (Maximisation step) Update the means, covariance matrices and mixing coefficients.

$$N_1 = 1.24 \qquad N_2 = 1.76$$

$$\boldsymbol{\mu}_1 = \begin{pmatrix}-2.89\\0.05\end{pmatrix} \qquad \boldsymbol{\mu}_2 = \begin{pmatrix}-2.45\\3.37\end{pmatrix}$$

COVARIANCE MATRICES

$$\boldsymbol{\Sigma}_k = \frac{1}{N_k}\sum_{n=1}^{N} r_k^{(n)}\left(\boldsymbol{x}^{(n)} - \boldsymbol{\mu}_k\right)\cdot\left(\boldsymbol{x}^{(n)} - \boldsymbol{\mu}_k\right)^{\mathrm{T}}$$

$$\boldsymbol{\Sigma}_1 = \frac{1}{N_1}\left(r_1^{(1)}\left(\boldsymbol{x}^{(1)} - \boldsymbol{\mu}_1\right)\cdot\left(\boldsymbol{x}^{(1)} - \boldsymbol{\mu}_1\right)^{\mathrm{T}} + r_1^{(2)}\left(\boldsymbol{x}^{(2)} - \boldsymbol{\mu}_1\right)\cdot\left(\boldsymbol{x}^{(2)} - \boldsymbol{\mu}_1\right)^{\mathrm{T}} + r_1^{(3)}\left(\boldsymbol{x}^{(3)} - \boldsymbol{\mu}_1\right)\cdot\left(\boldsymbol{x}^{(3)} - \boldsymbol{\mu}_1\right)^{\mathrm{T}}\right)$$

$$= \begin{pmatrix}5.37 & 0.04\\0.04 & 0.3\end{pmatrix}$$

$$\boldsymbol{\Sigma}_2 = \frac{1}{N_2}\left(r_2^{(1)}\left(\boldsymbol{x}^{(1)} - \boldsymbol{\mu}_2\right)\cdot\left(\boldsymbol{x}^{(1)} - \boldsymbol{\mu}_2\right)^{\mathrm{T}} + r_2^{(2)}\left(\boldsymbol{x}^{(2)} - \boldsymbol{\mu}_2\right)\cdot\left(\boldsymbol{x}^{(2)} - \boldsymbol{\mu}_2\right)^{\mathrm{T}} + r_2^{(3)}\left(\boldsymbol{x}^{(3)} - \boldsymbol{\mu}_2\right)\cdot\left(\boldsymbol{x}^{(3)} - \boldsymbol{\mu}_2\right)^{\mathrm{T}}\right)$$

$$= \begin{pmatrix}3.94 & -5.91\\-5.91 & 8.86\end{pmatrix}$$

EXERCICE 2

$$\left\{\boldsymbol{x}^{(1)} = \begin{pmatrix}-4\\0\end{pmatrix}, \boldsymbol{x}^{(2)} = \begin{pmatrix}2\\0\end{pmatrix}, \boldsymbol{x}^{(3)} = \begin{pmatrix}-2\\6\end{pmatrix}\right\} \subset \mathbb{R}^2$$

$$\pi_1 = 0.8, \boldsymbol{\mu}_1 = \begin{pmatrix}-1\\0\end{pmatrix}, \quad \boldsymbol{\Sigma}_1 = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix},$$

$$\pi_2 = 0.2, \boldsymbol{\mu}_2 = \begin{pmatrix}4\\0\end{pmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{pmatrix}2 & -1\\-1 & 2\end{pmatrix}.$$

(c) (Maximisation step) Update the means, covariance matrices and mixing coefficients.

$$N_1 = 1.24 \qquad N_2 = 1.76$$

MIXING COEFFICIENTS

$$\pi_k = \frac{N_k}{N} \longrightarrow \begin{cases} \pi_1 = \dfrac{N_1}{N} = \dfrac{1.24}{3} = 0.41 \\[2ex] \pi_2 = \dfrac{N_2}{N} = \dfrac{1.76}{3} = 0.59 \end{cases}$$

EXERCICE 2

$$\left\{\boldsymbol{x}^{(1)} = \begin{pmatrix} -4 \\ 0 \end{pmatrix}, \boldsymbol{x}^{(2)} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \boldsymbol{x}^{(3)} = \begin{pmatrix} -2 \\ 6 \end{pmatrix}\right\} \subset \mathbb{R}^2$$

$$\pi_1 = 0.41, \boldsymbol{\mu}_1 = \begin{pmatrix} -2.89 \\ 0.05 \end{pmatrix}, \quad \boldsymbol{\Sigma}_1 = \begin{pmatrix} 5.37 & 0.04 \\ 0.04 & 0.3 \end{pmatrix},$$

$$\pi_2 = 0.59, \boldsymbol{\mu}_2 = \begin{pmatrix} -2.45 \\ 3.37 \end{pmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{pmatrix} 3.94 & -5.91 \\ -5.91 & 8.86 \end{pmatrix}.$$

(d) (Convergence) Compute the log-likelihood at the current iteration.

REMINDER

$k = 1, \dots, K \longleftarrow$ clusters / Gaussian models (in this case, $K = 2$)
$n = 1, \dots, N \longleftarrow$ data points (in this case, $N = 3$)

$$\boxed{\ln P(X|\pi, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \sum_{n=1}^{N} \ln\left[\sum_{k=1}^{K} \pi_k P\left(\boldsymbol{x}^{(n)}\middle|k\right)\right]}$$

where $\left(\boldsymbol{x}^{(n)}\middle|k\right) \sim \mathcal{N}\left(\boldsymbol{x}^{(n)}\middle|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\right)$

$$\Longrightarrow \quad P\left(\boldsymbol{x}^{(n)}\middle|k\right) = (2\pi)\det(\boldsymbol{\Sigma}_k)^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\left(\boldsymbol{x}^{(n)} - \boldsymbol{\mu}_k\right)^{\mathrm{T}} \boldsymbol{\Sigma}_k^{-1}\left(\boldsymbol{x}^{(n)} - \boldsymbol{\mu}_k\right)\right]$$

TIP

```
from scipy.stats import multivariate_normal as multiNorm
px = multiNorm.pdf(x,mu,cov)
```

EXERCICE 2

$$\left\{\boldsymbol{x}^{(1)} = \binom{-4}{0}, \boldsymbol{x}^{(2)} = \binom{2}{0}, \boldsymbol{x}^{(3)} = \binom{-2}{6}\right\} \subset \mathbb{R}^2$$

$$\pi_1 = 0.41, \boldsymbol{\mu}_1 = \begin{pmatrix} -2.89 \\ 0.05 \end{pmatrix}, \quad \boldsymbol{\Sigma}_1 = \begin{pmatrix} 5.37 & 0.04 \\ 0.04 & 0.3 \end{pmatrix},$$

$$\pi_2 = 0.59, \boldsymbol{\mu}_2 = \begin{pmatrix} -2.45 \\ 3.37 \end{pmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{pmatrix} 3.94 & -5.91 \\ -5.91 & 8.86 \end{pmatrix}.$$

(d) (Convergence) Compute the log-likelihood at the current iteration.

$$\boxed{\ln P(X|\pi, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \sum_{n=1}^{N} \ln \left[\sum_{k=1}^{K} \pi_k P\left(\boldsymbol{x}^{(n)} \middle| k\right) \right]}$$

where $P\left(\boldsymbol{x}^{(n)} \middle| k\right) = (2\pi)\det(\boldsymbol{\Sigma}_k)^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\left(\boldsymbol{x}^{(n)} - \boldsymbol{\mu}_{\boldsymbol{k}}\right)^{\mathrm{T}} \boldsymbol{\Sigma}_{\boldsymbol{k}}^{-1}\left(\boldsymbol{x}^{(n)} - \boldsymbol{\mu}_{\boldsymbol{k}}\right)\right]$

$$\ln P(X|\pi, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \sum_{n=1}^{3} \ln \left[\sum_{k=1}^{2} \pi_k P\left(\boldsymbol{x}^{(n)} \middle| k\right) \right] = \cdots = 4.7254$$

EXERCICE 2

$$\left\{\boldsymbol{x}^{(1)} = \begin{pmatrix}-4\\0\end{pmatrix}, \boldsymbol{x}^{(2)} = \begin{pmatrix}2\\0\end{pmatrix}, \boldsymbol{x}^{(3)} = \begin{pmatrix}-2\\6\end{pmatrix}\right\} \subset \mathbb{R}^2$$

$$\pi_1 = 0.41, \boldsymbol{\mu}_1 = \begin{pmatrix}-2.89\\0.05\end{pmatrix}, \quad \boldsymbol{\Sigma}_1 = \begin{pmatrix}5.37 & 0.04\\0.04 & 0.3\end{pmatrix},$$

$$\pi_2 = 0.59, \boldsymbol{\mu}_2 = \begin{pmatrix}-2.45\\3.37\end{pmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{pmatrix}3.94 & -5.91\\-5.91 & 8.86\end{pmatrix}.$$

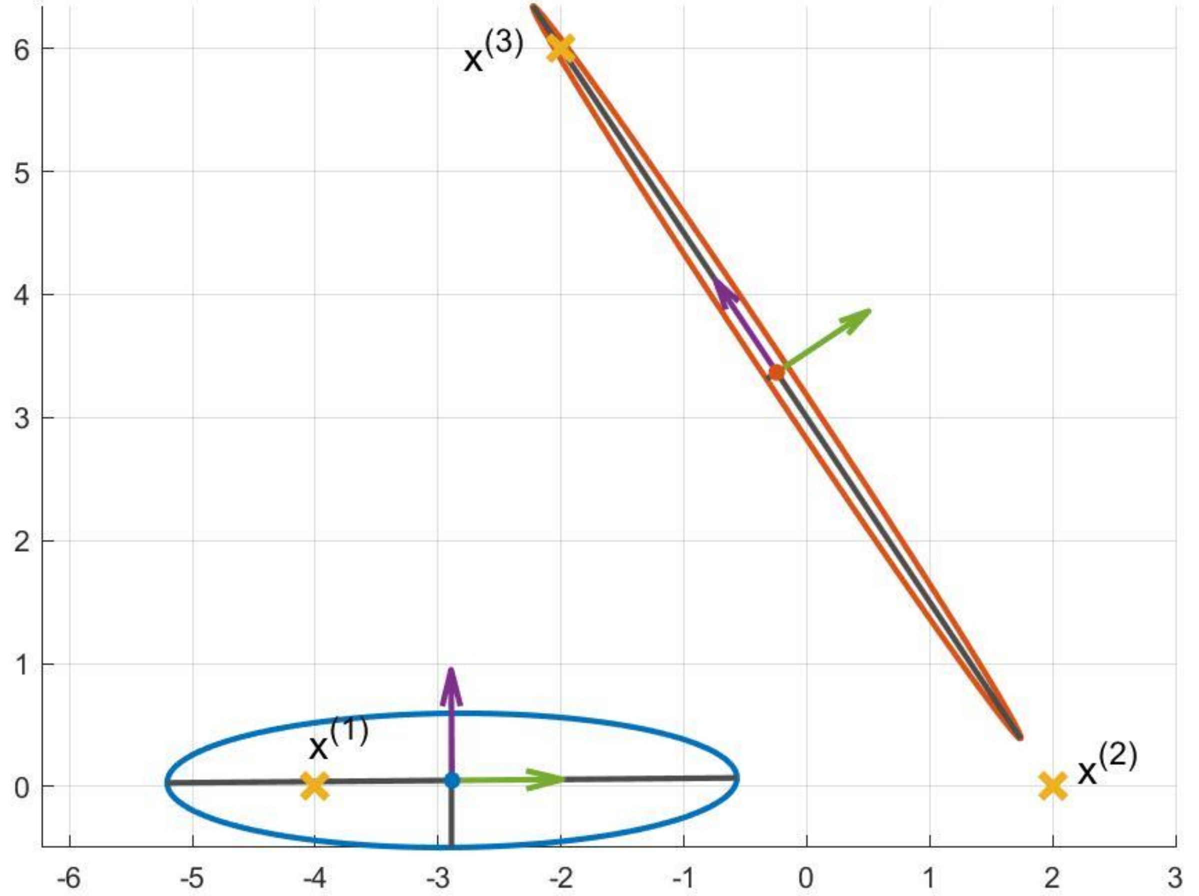
(e) Draw the updated Gaussian models. Do the updated Gaussian models look more accurate than the initial ones? What do you think the tendency of the EM algorithm is in this case?

TIP

```
np.linalg.svd
```

$$G_1 \begin{cases} \boldsymbol{\mu}_1 = \begin{pmatrix}-2.89\\0.05\end{pmatrix} \\ \boldsymbol{\Sigma}_1 = \begin{pmatrix}5.37 & 0.04\\0.04 & 0.3\end{pmatrix} \end{cases}$$

$$G_2 \begin{cases} \boldsymbol{\mu}_2 = \begin{pmatrix}-2.45\\3.37\end{pmatrix} \\ \boldsymbol{\Sigma}_2 = \begin{pmatrix}3.94 & -5.91\\-5.91 & 8.86\end{pmatrix} \end{cases}$$



EXERCICE 2

$$\left\{\boldsymbol{x}^{(1)} = \begin{pmatrix}-4\\0\end{pmatrix}, \boldsymbol{x}^{(2)} = \begin{pmatrix}2\\0\end{pmatrix}, \boldsymbol{x}^{(3)} = \begin{pmatrix}-2\\6\end{pmatrix}\right\} \subset \mathbb{R}^2$$

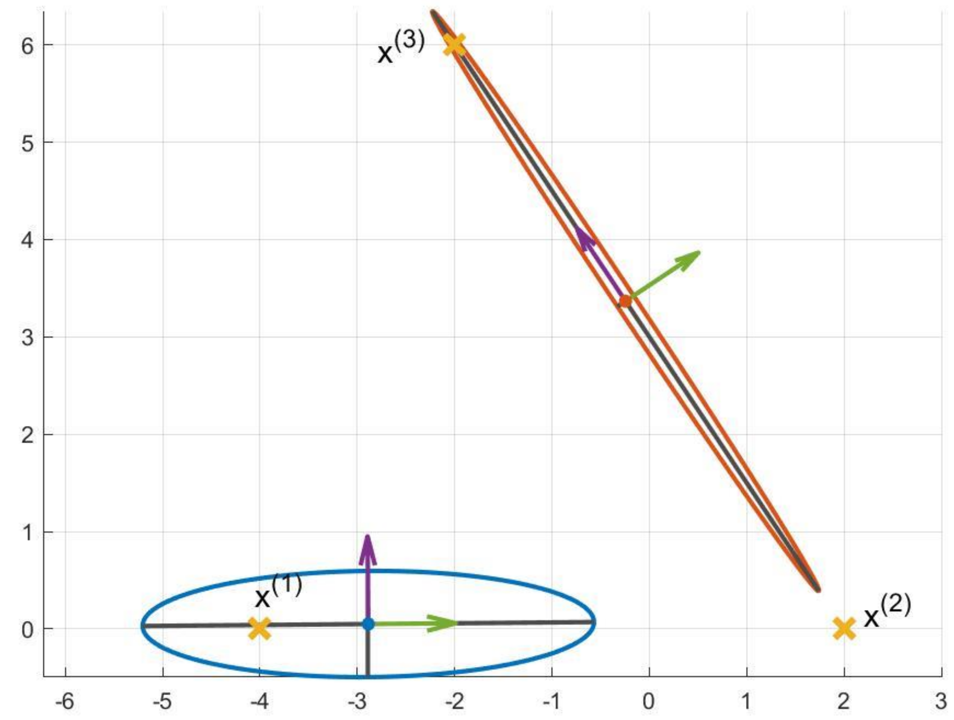
$$\pi_1 = 0.41, \boldsymbol{\mu}_1 = \begin{pmatrix}-2.89\\0.05\end{pmatrix}, \quad \boldsymbol{\Sigma}_1 = \begin{pmatrix}5.37 & 0.04\\0.04 & 0.3\end{pmatrix},$$

$$\pi_2 = 0.59, \boldsymbol{\mu}_2 = \begin{pmatrix}-2.45\\3.37\end{pmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{pmatrix}3.94 & -5.91\\-5.91 & 8.86\end{pmatrix}.$$

(e) Draw the updated Gaussian models. Do the updated Gaussian models look more accurate than the initial ones? What do you think the tendency of the EM algorithm is in this case?

$$G_1 \begin{cases} \boldsymbol{\mu}_1 = \begin{pmatrix}-2.89\\0.05\end{pmatrix} \\ \boldsymbol{\Sigma}_1 = \begin{pmatrix}5.37 & 0.04\\0.04 & 0.3\end{pmatrix} \xrightarrow{\text{SVD}} \begin{cases} \lambda_1 = 5.38 & \boldsymbol{v}^{(1)} = \begin{pmatrix}1\\0.009\end{pmatrix} \\ \lambda_2 = 0.3 & \boldsymbol{v}^{(2)} = \begin{pmatrix}-0.009\\1\end{pmatrix} \end{cases} \end{cases}$$

$$G_2 \begin{cases} \boldsymbol{\mu}_2 = \begin{pmatrix}-2.45\\3.37\end{pmatrix} \\ \boldsymbol{\Sigma}_2 = \begin{pmatrix}3.94 & -5.91\\-5.91 & 8.86\end{pmatrix} \xrightarrow{\text{SVD}} \begin{cases} \lambda_1 = 12.8 & \boldsymbol{v}^{(1)} = \begin{pmatrix}-0.56\\0.83\end{pmatrix} \\ \lambda_2 = 0 & \boldsymbol{v}^{(2)} = \begin{pmatrix}0.83\\0.56\end{pmatrix} \end{cases} \end{cases}$$



EXERCICE 2

$$\left\{\boldsymbol{x}^{(1)} = \begin{pmatrix}-4\\0\end{pmatrix}, \boldsymbol{x}^{(2)} = \begin{pmatrix}2\\0\end{pmatrix}, \boldsymbol{x}^{(3)} = \begin{pmatrix}-2\\6\end{pmatrix}\right\} \subset \mathbb{R}^2$$

$$\pi_1 = 0.41, \boldsymbol{\mu}_1 = \begin{pmatrix}-2.89\\0.05\end{pmatrix}, \quad \boldsymbol{\Sigma}_1 = \begin{pmatrix}5.37 & 0.04\\0.04 & 0.3\end{pmatrix},$$

$$\pi_2 = 0.59, \boldsymbol{\mu}_2 = \begin{pmatrix}-2.45\\3.37\end{pmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{pmatrix}3.94 & -5.91\\-5.91 & 8.86\end{pmatrix}.$$

(e) Draw the updated Gaussian models. Do the updated Gaussian models look more accurate than the initial ones? What do you think the tendency of the EM algorithm is in this case?