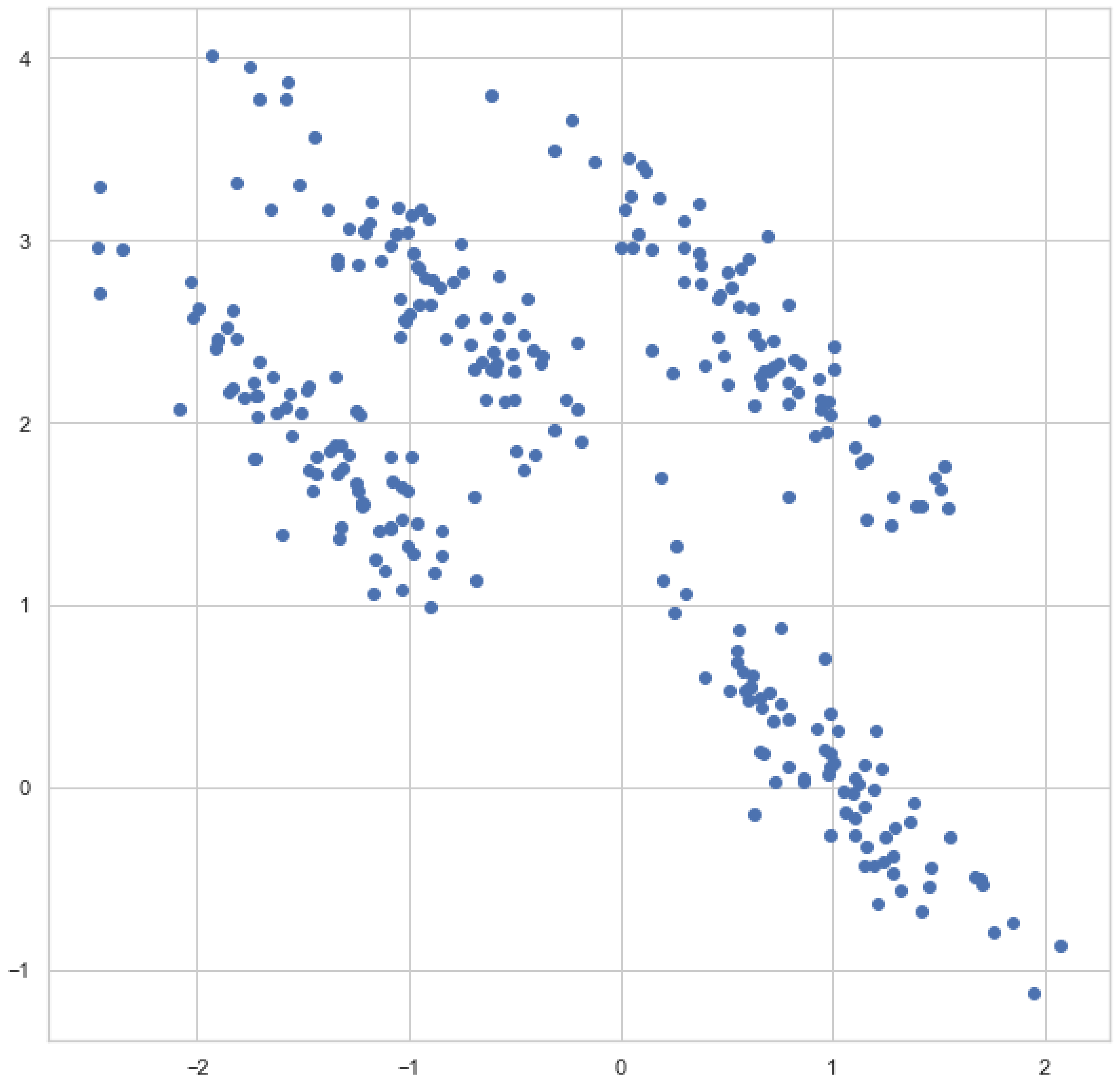
REVIEW

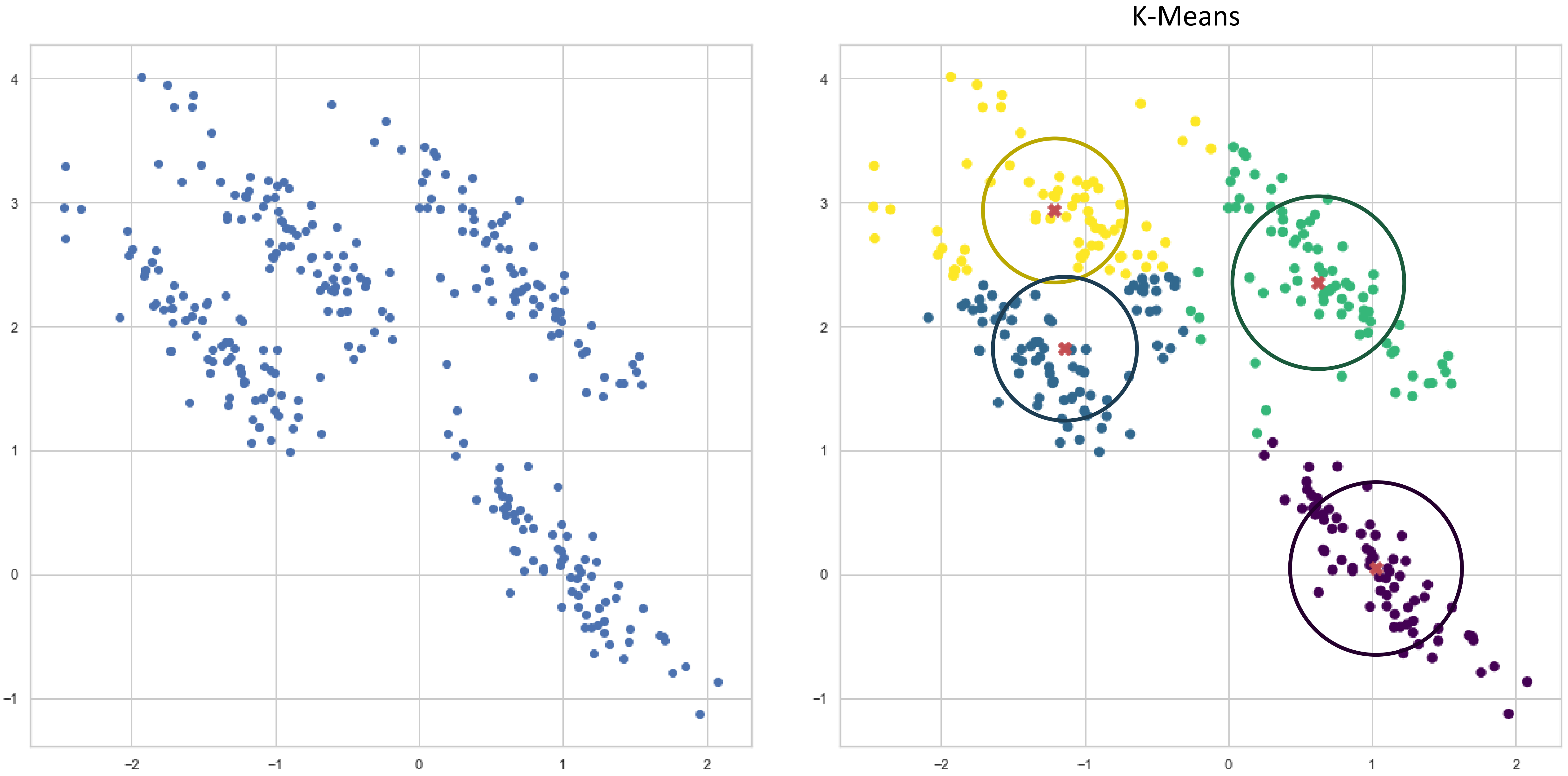
GAUSSIAN MODELS

REVIEW: GAUSSIAN MODELS

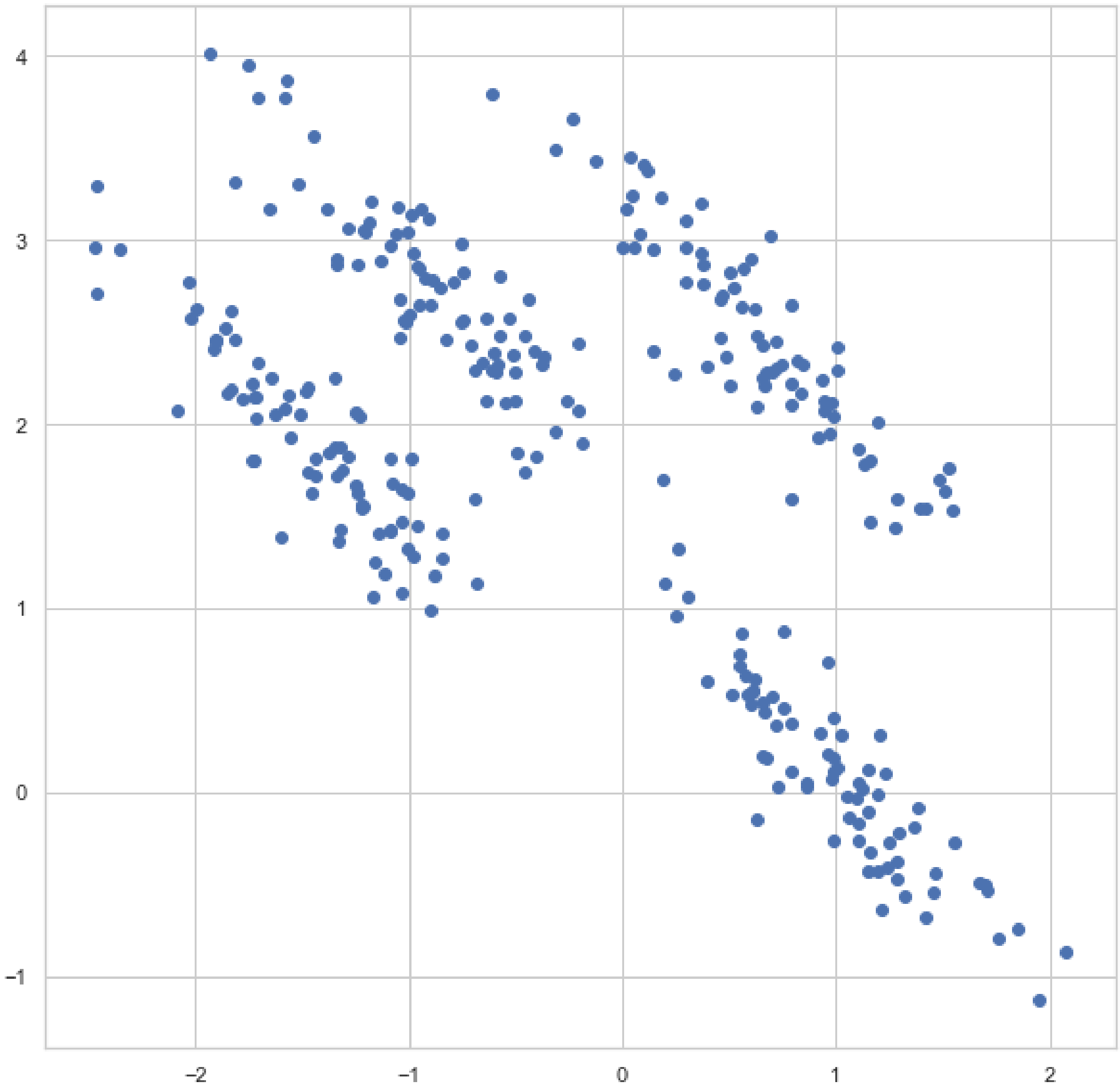
What if the dataset we want to clusterise looks like this?

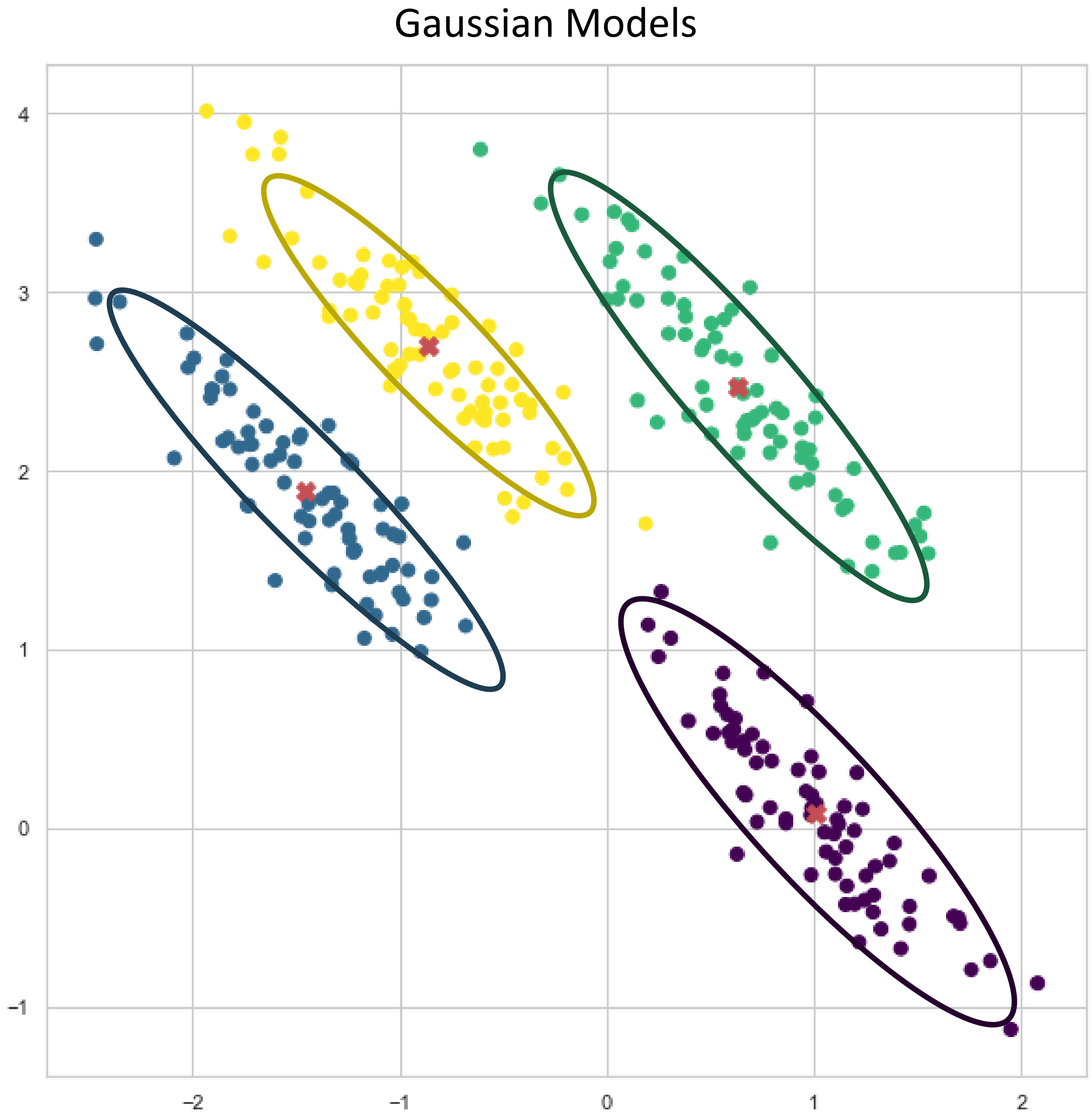


REVIEW: GAUSSIAN MODELS



REVIEW: GAUSSIAN MODELS





REVIEW: GAUSSIAN MODELS

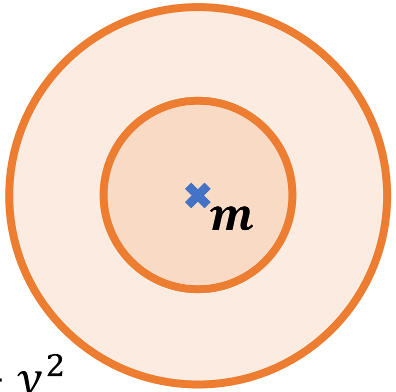
K-Means

- **Euclidean** distance

$$d(\boldsymbol{x},\boldsymbol{m}) = \sqrt{(x_1 - m_1)^2 + (x_2 - m_2)^2}$$

$$\Longrightarrow d^2 = (x_1 - m_1)^2 + (x_2 - m_2)^2$$

- The points at the same distance from the center form a **circle**.
- Its center is $\boldsymbol{m} = \sum_{i=1}^{N} \boldsymbol{x}^{(i)}$



$$k^2 = x^2 + y^2$$

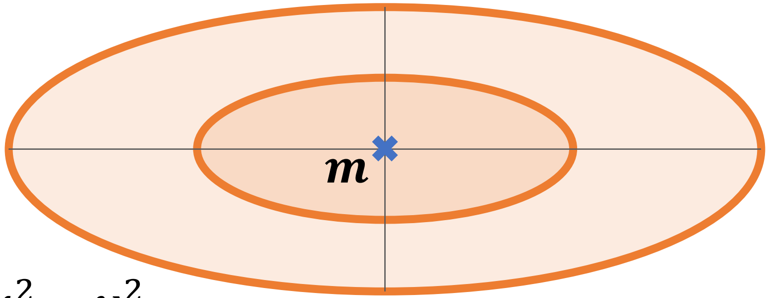
Gaussian model

- **Mahalanobis** distance

$$d(\boldsymbol{x},\boldsymbol{m}) = \sqrt{(\boldsymbol{x} - \boldsymbol{m})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{m})}$$

$$\Longrightarrow d^2 = (\boldsymbol{x} - \boldsymbol{m})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{m})$$

- The points at the same distance from the center form an **ellipsoid**.
- Its center is $\boldsymbol{m} = \sum_{i=1}^{N} \boldsymbol{x}^{(i)}$.
- Its principal directions are defined by the **eigenvectors** of $\boldsymbol{\Sigma}$.
- Its semi-axes are defined by the **eigenvalues** of $\boldsymbol{\Sigma}$.



$$k^2 = \frac{x^2}{a^2} + \frac{y^2}{b^2}$$

REVIEW: GAUSSIAN MODELS

Gaussian model

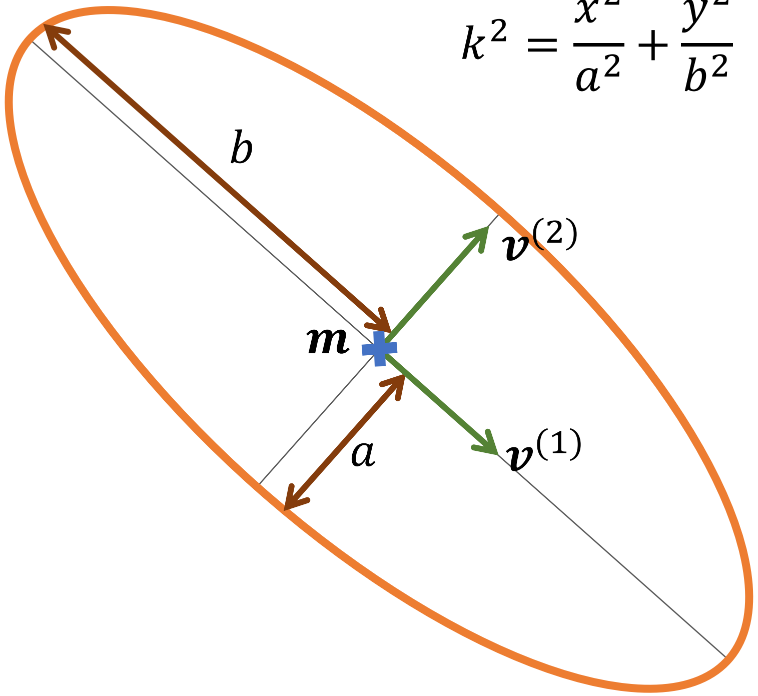
- **Mahalanobis** distance

$$d(\boldsymbol{x}, \boldsymbol{m}) = \sqrt{(\boldsymbol{x} - \boldsymbol{m})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{m})}$$
$$\Longrightarrow d^2 = (\boldsymbol{x} - \boldsymbol{m})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{m})$$

All the points at the same Mahalanobis distance from $\boldsymbol{m}$ have the same probability (Gaussian/Normal distribution).

- The points at the same distance from the center form an **ellipsoid**.
- Its center is $\boldsymbol{m} = \sum_{i=1}^{N} \boldsymbol{x}^{(i)}$.
- Its principal directions are defined by the **eigenvectors** of $\boldsymbol{\Sigma}$.
- Its semi-axes are defined by the **eigenvalues** of $\boldsymbol{\Sigma}$.

$$k^2 = \frac{x^2}{a^2} + \frac{y^2}{b^2}$$



REVIEW: GAUSSIAN MODELS

Gaussian model

All the points at the same Mahalanobis distance from $\boldsymbol{m}$ have the same probability (Gaussian/Normal distribution).

- **Mahalanobis** distance

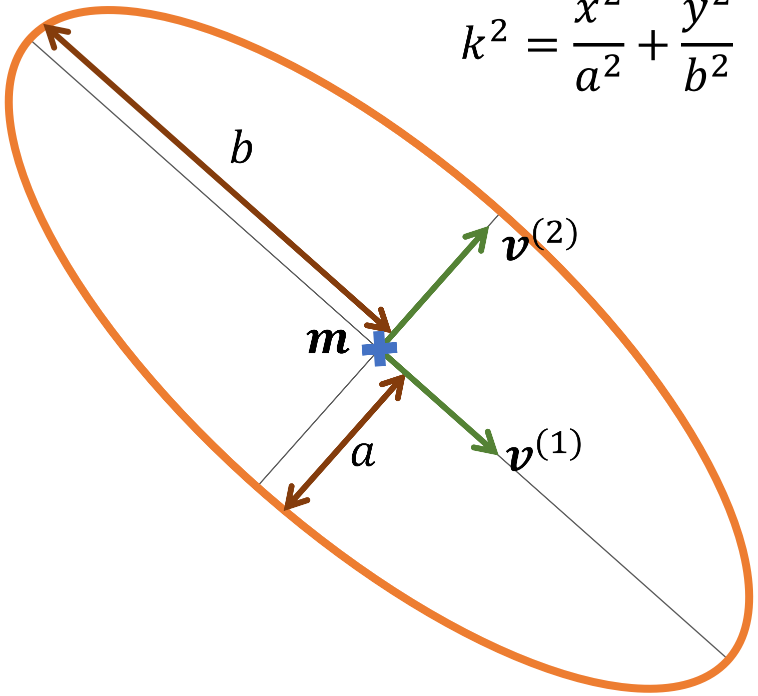
$$d(\boldsymbol{x}, \boldsymbol{m}) = \sqrt{(\boldsymbol{x} - \boldsymbol{m})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{m})}$$
$$\Longrightarrow d^2 = (\boldsymbol{x} - \boldsymbol{m})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{m})$$

- The points at the same distance from the center form an **ellipsoid**.

  same probability

- Its center is $\boldsymbol{m} = \sum_{i=1}^{N} \boldsymbol{x}^{(i)}$.
- Its principal directions are defined by the **eigenvectors** of $\boldsymbol{\Sigma}$.
- Its semi-axes are defined by the **eigenvalues** of $\boldsymbol{\Sigma}$.

$$k^2 = \frac{x^2}{a^2} + \frac{y^2}{b^2}$$



REVIEW: GAUSSIAN MODELS

Gaussian model

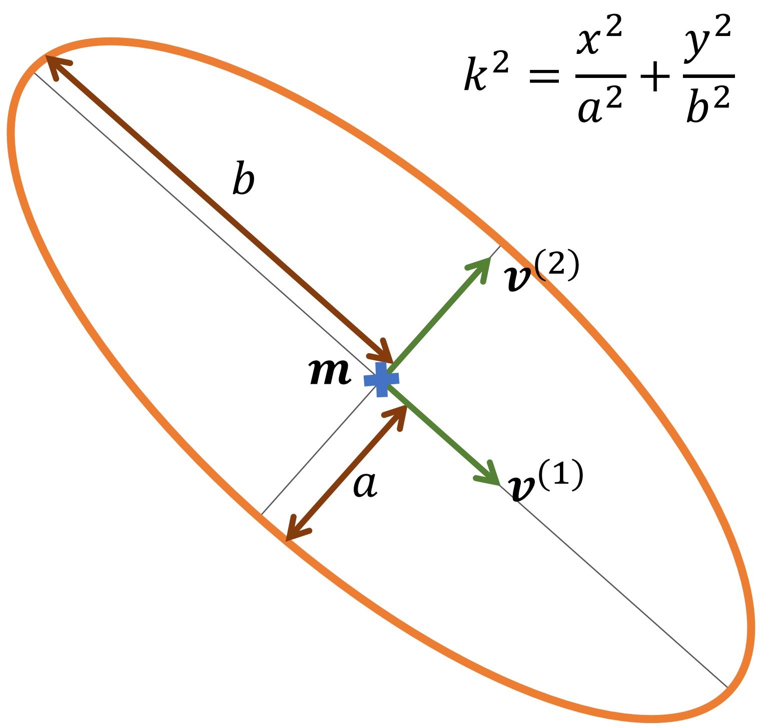
- **Mahalanobis** distance

$$d(\boldsymbol{x}, \boldsymbol{m}) = \sqrt{(\boldsymbol{x} - \boldsymbol{m})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{m})}$$
$$\Longrightarrow d^2 = (\boldsymbol{x} - \boldsymbol{m})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{m})$$

- The points at the same distance from the center form an **ellipsoid**.

Estimate parameters of Gaussian distribution from data.

- Its center is $\boldsymbol{m} = \sum_{i=1}^{N} \boldsymbol{x}^{(i)}$.
- Its principal directions are defined by the **eigenvectors** of $\boldsymbol{\Sigma}$.
- Its semi-axes are defined by the **eigenvalues** of $\boldsymbol{\Sigma}$.



REVIEW: GAUSSIAN MODELS

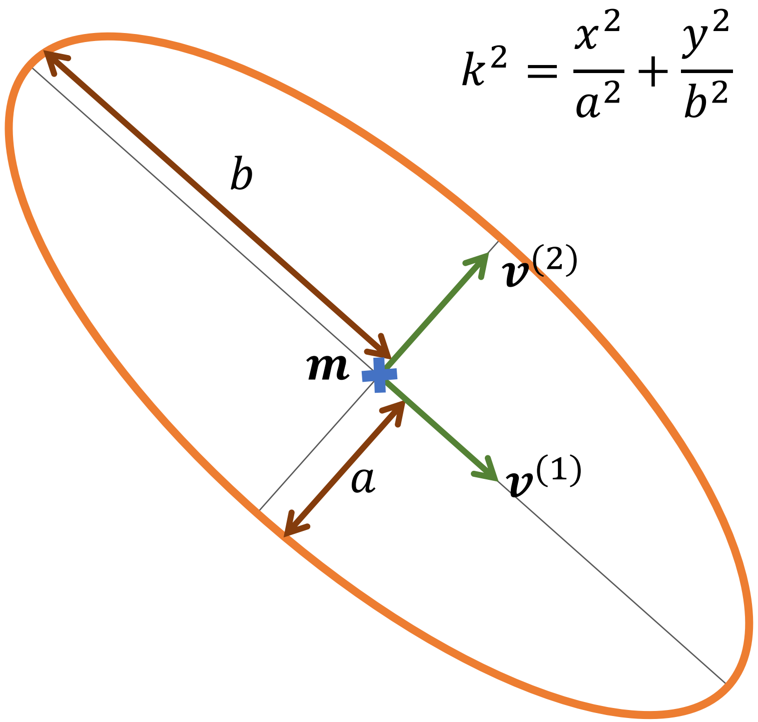
Gaussian model

- **Mahalanobis** distance

$$d(\boldsymbol{x}, \boldsymbol{m}) = \sqrt{(\boldsymbol{x} - \boldsymbol{m})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{m})}$$
$$\Longrightarrow d^2 = (\boldsymbol{x} - \boldsymbol{m})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{m})$$

- The points at the same distance from the center form an **ellipsoid**.
- Its center is $\boldsymbol{m} = \sum_{i=1}^{N} \boldsymbol{x}^{(i)}$.
- Its principal directions are defined by the **eigenvectors** of $\boldsymbol{\Sigma}$.
- Its semi-axes are defined by the **eigenvalues** of $\boldsymbol{\Sigma}$.

$$k^2 = \frac{x^2}{a^2} + \frac{y^2}{b^2}$$



The **eigenvectors** give the orientation of the ellipse.

The **eigenvalues** give variance of the data in the directions given by the eigenvectors.