Quiz: Clustering

Lecture series „Machine Learning“

Niels Landwehr

Research Group „Data Science“
Institute of Computer Science
University of Hildesheim

Quiz: K-Means

- Assume a 2D toy data set $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ where the three data points are given by

$$\mathbf{x}_1 = (0,0)$$
$$\mathbf{x}_2 = (1,0)$$
$$\mathbf{x}_3 = (4,0)$$

- Assume we run K-Means for K=2 on the data with initial cluster centers

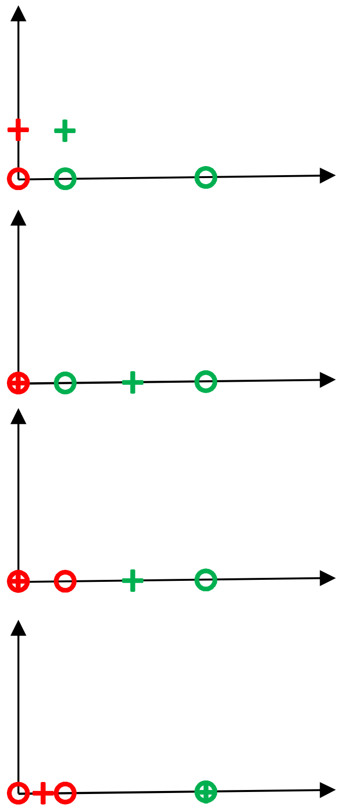
$$\boldsymbol{\mu}_1 = (0,1)$$
$$\boldsymbol{\mu}_2 = (1,1)$$

- What will be the final cluster centers after convergence?
  - $\boldsymbol{\mu}_1 = (1.5,0),\ \boldsymbol{\mu}_2 = (3,0)$
  - $\boldsymbol{\mu}_1 = (1,1),\ \boldsymbol{\mu}_2 = (2,1)$
  - $\boldsymbol{\mu}_1 = (0,0),\ \boldsymbol{\mu}_2 = (2.5,0)$
  - $\boldsymbol{\mu}_1 = (5/3,0),\ \boldsymbol{\mu}_2 = (5/3,0)$
  - $\boldsymbol{\mu}_1 = (0.5,0),\ \boldsymbol{\mu}_2 = (4,0)$

Solution: K-Means

- **Solution**: $\boldsymbol{\mu}_1 = (0.5, 0),\ \boldsymbol{\mu}_2 = (4, 0)$

- This is easiest to see if we just write down the points in a plot and see how the cluster assignments and cluster centers evolve (alternatively, could also compute)



Initial cluster centers with initial assignment

$\boldsymbol{\mu}_1 = (0,1)$
$\boldsymbol{\mu}_2 = (1,1)$

$\mathbf{x}_1 = (0,0)$
$\mathbf{x}_2 = (1,0)$
$\mathbf{x}_3 = (4,0)$

Recompute cluster centers ("Maximization")

Recompute assignments ("Expectation")

Recompute cluster centers ("Maximization")

Quiz: Multivariate Normal

- Assume a two-dimensional multivariate normal distribution with mean vector $\boldsymbol{\mu} = (0,0)$ and covariance matrix

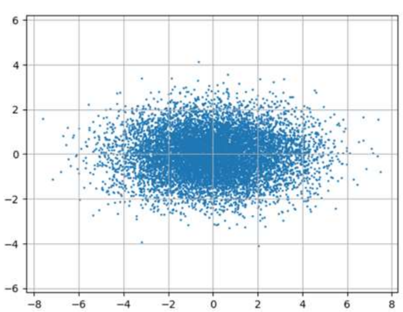
$$\Sigma = \begin{pmatrix} 1.5 & -1 \\ -1 & 1.5 \end{pmatrix}$$

  Note that the inverse covariance matrix is given by

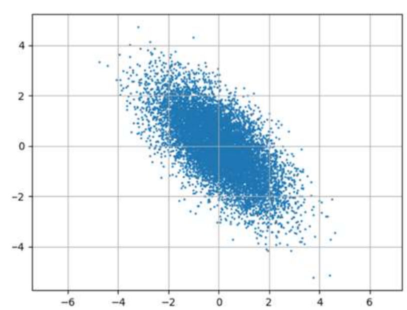
$$\Sigma^{-1} = \begin{pmatrix} 1.2 & 0.8 \\ 0.8 & 1.2 \end{pmatrix}$$

- If we sample 10000 points from this distribution, how will the point cloud look like?

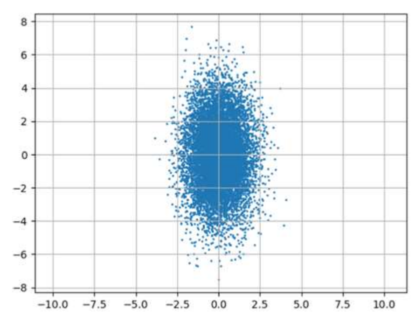
A



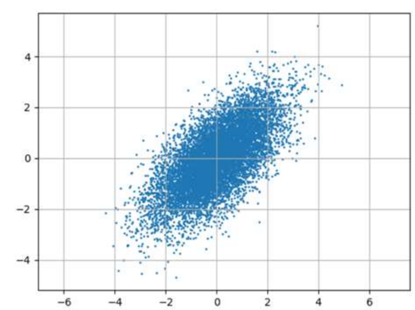
B



C

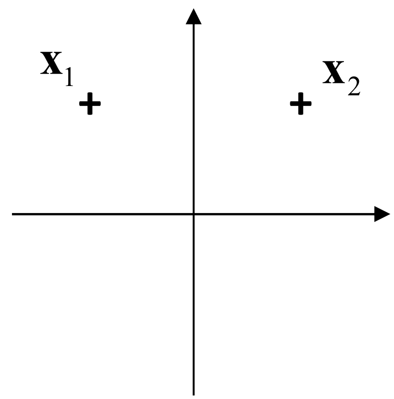


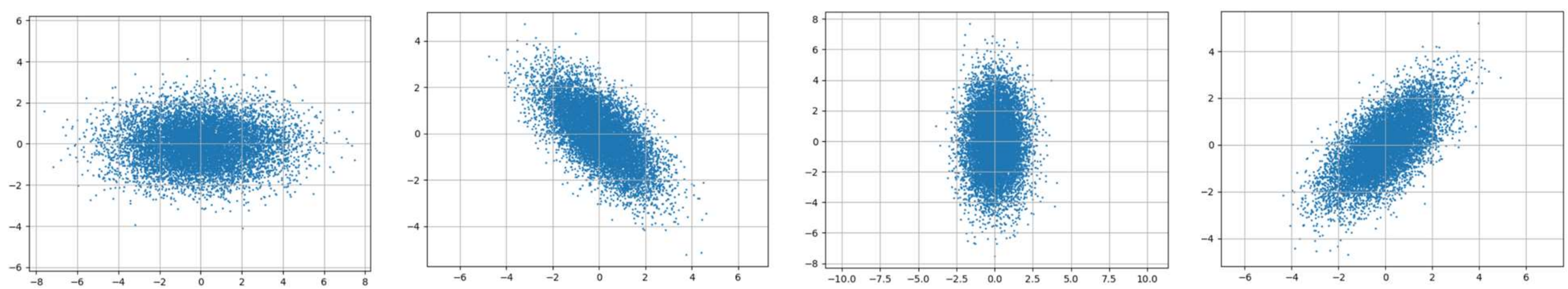
D



Solution: Multivariate Normal

- **Solution**: Figure B is correct.
- We can verify this by checking the density at $\mathbf{x}_1 = (-1,1)$ and $\mathbf{x}_2 = (1,1)$:





- If density at $\mathbf{x}_1$ is larger than density at $\mathbf{x}_2$, it must be Figure B

Solution: Multivariate Normal

- As the mean vector is zero, the density is

$$\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{Z} \exp\left(-\frac{1}{2} \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x} \right)$$

- Let first compute the quadratic forms:

$$\mathbf{x}_1{}^{\mathrm{T}} \Sigma^{-1} \mathbf{x}_1 = \begin{pmatrix} -1 & 1 \end{pmatrix} \begin{pmatrix} 1.2 & 0.8 \\ 0.8 & 1.2 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 & 1 \end{pmatrix} \begin{pmatrix} -0.4 \\ 0.4 \end{pmatrix} = 0.8$$

$$\mathbf{x}_2{}^{\mathrm{T}} \Sigma^{-1} \mathbf{x}_2 = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 1.2 & 0.8 \\ 0.8 & 1.2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 4$$

- Because the density is $\exp\left(-\frac{1}{2} \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x} \right)$, it is higher for $\mathbf{x}_1$ (note that Z is positive)

Reminder: please fill in the lecture evaluation form, you received a link by email