

Deep Learning KU (708.220) WS22

Assignment 1: Maximum Likelihood Estimation

First part - derivation:

Consider a problem, in which the samples come from a 3-dimensional Gaussian distribution with mean vector $\boldsymbol{\mu}$ and a covariance matrix $\Sigma = \sigma^2 I_3$, where I_3 is the 3×3 identity matrix and $\sigma \in \mathbb{R}^+$.

Probability of data point vector \mathbf{x} equals:

$$p(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^3 |\Sigma|}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})}$$

Hint#1: $|\Sigma|$ is the determinant of Σ and equals σ^6 .

Data:

Your dataset is: $X = \langle \mathbf{x}^1, \dots, \mathbf{x}^M \rangle$ in which each data point \mathbf{x}^m is a 3-dimensional vector ($\mathbf{x}^m \in \mathbb{R}^3$). The samples are independent and identically distributed.

You will model and fit the distribution:

$$p(\mathbf{X}|\boldsymbol{\theta})$$

with parameters:

$$\boldsymbol{\theta} = \langle \boldsymbol{\mu}, \sigma \rangle$$

Your tasks:

- a) (3 points) : Write down the likelihood $p(\mathbf{x}^m|\boldsymbol{\theta})$ of a single data point \mathbf{x}^m .
Then, write down the likelihood $p(\mathbf{X}|\boldsymbol{\theta})$ of the whole data set \mathbf{X} .
Use this to derive the log-likelihood of the data set.
- b) (5 points) : For this model, derive the formula for the maximum-likelihood estimate of $\boldsymbol{\mu}$.
(*Hint#2: For easier derivation, calculate $(\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})$, taking into account $\Sigma = \sigma^2 I_3$*)

Second part - practical work:

You have a set of data about social capital. Download the data from:
<https://data.humdata.org/dataset/social-capital-atlas>

For the tasks below create the appropriate code in Python and discuss your experiences, findings and choices in your report.

Your tasks:

- c) (7 points) : First, choose one of the four datasets that you will work with. Then, choose three variables as the three dimensions of your 3-dimensional MLE problem above. Explain your choice and plot chosen data. Comment on the appropriateness of the assumption that the data follows a 3-dimensional Gaussian distribution. Show your arguments on appropriate plots.
- d) (5 points) : Using Python, code your MLE formula from (b) and calculate the maximum-likelihood estimate of μ .

Total: 20 points

Provide full derivations including intermediate steps. Present your results clearly, structured and legible. Submit your Python code and report at the TeachCenter. Do not zip the report together with the code.

Assignment details:

- *Assignment issued:* October 12th, 2022, 08:00
- *Deadline:* November 9th, 2022, 08:00
- *Solution Submission:* Upload to TeachCenter one PDF (report) and a single .py (code) file, separately as explained above (i.e. not in a zip file)
- *Presentation session:* November 16th, 2022
- *Rules:* There are no groups allowed for this task.
Copying of solutions, reports or parts of reports from other students is strictly forbidden.