Universität Zürich
Institut für Informatik

HS 2021

Student Name: Youngseo Kim
Matrikel-Nr: 20-744-470

Numerical Methods in Informatics - Exercise 5

Hand out: 29.11.2021 - Due to: 12.12.2021

Please upload your solutions to the Olat system.

Theory

5.1 Least Squares

a) (2 Min, 2 Points) Comparison to Eigendecomposition (e.g. in approximation)
Please name two advantages and two disadvantages, Least Squares has in applications compared to an Eigendecomposition.

<u>Advantages</u>

- To minimise the error, the least squares problem requires to find an $\hat{x}$ that makes $\|b - A\hat{x}\|$ as small as possible. However, with Eigendecomposition, approximating an eigenvalue requires an iterative calculation.

- We can find a solution for an $m \times n$ matrix using the least squares method. However, Eigendecomposition can be applied only for an $n \times n$ matrix.

Disadvantages

- Least Squares is sensitive to outliers in the data used to fit a model. This makes model validation, especially with respect to outliers, critical to obtaining sound answers to the questions.
- For inherently nonlinear processes, it becomes increasingly difficult to find a linear model that fits the data as well as the data increases. It will result in an extreme linear model, which means that linear models may not be effective for extra-polating the results of a process for which data cannot be collected in the region of interest.

b) (8 Min, 8 Points)

Please calculate the best fitting line appromximing the points P (also shown in Figure 1) using least squares:

$$P = \{(0, -0.2), (1, 1.3), (2, 1.6), (3, 3.2)\}$$

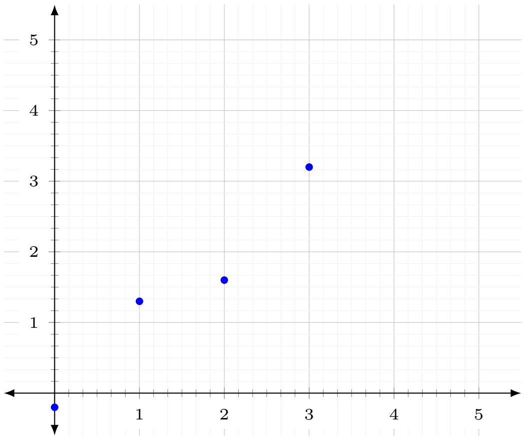

Figure 1: Noisy points

$$X = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \quad y = \begin{bmatrix} -0.2 \\ 1.3 \\ 1.6 \\ 3.2 \end{bmatrix}$$

$\Leftrightarrow$ For the least-squares solution of $X\beta = y$, obtain the normal equations:

$$X^T X \beta = X^T y$$

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 1+1+1+1 & 0+1+2+3 \\ 0+1+2+3 & 0+1+4+9 \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ 6 & 14 \end{bmatrix}$$

$$X^T y = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} -0.2 \\ 1.3 \\ 1.6 \\ 3.2 \end{bmatrix} = \begin{bmatrix} \overset{2.7}{-0.2+1.3+1.6+3.2} \\ \underset{4.5}{0+1.3+3.2+9.6} \end{bmatrix} = \begin{bmatrix} 5.9 \\ 14.1 \end{bmatrix}$$

The normal equations are:

$$\begin{bmatrix} 4 & 6 \\ 6 & 14 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 5.9 \\ 14.1 \end{bmatrix}$$

(continue)

Hence $\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ 6 & 14 \end{bmatrix}^{-1} \begin{bmatrix} 5.9 \\ 14.1 \end{bmatrix}$

$$= \underbrace{\frac{1}{56-36} \begin{bmatrix} 14 & -6 \\ -6 & 4 \end{bmatrix}} \begin{bmatrix} 5.9 \\ 14.1 \end{bmatrix}$$

$$\begin{bmatrix} \cancel{14}^{7}/\cancel{20}_{10} & -\cancel{6}^{3}/\cancel{20}_{10} \\ -6/20 & \cancel{4}_{1}/\cancel{20}_{5} \end{bmatrix} = \begin{bmatrix} 0.7 & -0.3 \\ -0.3 & 0.2 \end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} 0.7 & -0.3 \\ -0.3 & 0.2 \end{bmatrix} \begin{bmatrix} 5.9 \\ 14.1 \end{bmatrix} = \begin{bmatrix} \overset{4.13}{0.7 \cdot 5.9} - \overset{-4.23}{0.3 \cdot 14.1} \\ \underset{-1.77}{-0.3 \cdot 5.9} + \underset{2.82}{0.2 \cdot 14.1} \end{bmatrix} \qquad \begin{array}{r} 2.82 \\ -1.77 \\ \hline 1.05 \end{array}$$

$$= \begin{bmatrix} -0.1 \\ 1.05 \end{bmatrix} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$$

Thus, the least-squares line has the equation

$\therefore\ y = -0.1 + 1.05x$

Practice

5.2 Least Squares

a) (100 Min, 5 Points) Linear Approximation via Least Squares

Please implement the method `linearLSQ(x, y)` in `backend.py`, which performs line fitting (as described in the last exercise) via least squares.

b) (100 Min, 5 Points) Orthonormalization

Please implement the function `orthonormalize(sourceBase)` in `backend.py`. The function is supposed to create an orthonormal basis from the given `sourceBase`, spanning the same vectorspace.

Handing in:

Please only include your `backend.py` in your hand in.