Data Science I

Overall Points: / 100

Exercise 6: Linear Algebra & Linear Regression

Submission Deadline: January 12 2026, 07:00 UTC
University of Oldenburg
Winter 2025/2026
Instructors: Jannik Schröder, Wolfram "Wolle" Wingerath

Submitted by: <your names here>

Part 1: Linear Algebra

/ 50

1.) Give a pair of square matrices A and B such that:

a) $AB = BA$ **(it commutes)**

Solution:

< your solution here >

b) $AB \neq BA$ **(it does not commute)**

Solution:

< your solution here >

2.) Are the matrices MM^T **nd** M^TM **square and symmetric? Please explain.**

Solution:

< your solution here >

3.) Let A be a feature matrix, w be a weight vector, and b be the target vector, so that matrix equation $Aw = b$ holds. Please explain how to compute w with linear algebra. Does it make a difference whether A is a square and invertible matrix? Why or why not?

Solution:

< your solution here >

4.) Please implement a function that computes the result of multiplying two matrices. Compare the speed of a library function for matrix multiplication to your own implementation.

a) How much faster is the library on products of random n × n matricies, as a function of n as n gets large?

Solution:

< your solution here >

b) What about the product of an n × m and m × n matrix, where n << m?

Solution:

< your solution here >

c) Now make sure that your implementation calculates C = A · B by first transposing B internally, so that all dot products are computed along rows of the matrices. Does that improve performance compared with an implementation that does not transpose B first? By how much?

Solution:

< your solution here >

Part 2: Linear Regression

/ 50

5.) Construct an example on $n \geq 6$ points where the optimal regression line is $y = x$, even though none of the input points lie directly on this line. Please visualize your example with a plot.

Solution:

< your solution here >

6.) Suppose we fit a regression line to predict the shelf life of an apple based on its weight. For a particular apple, we predict the shelf life to be 4.6 days. The apples residual is −0.6 days. Did we over or under estimate the shelf-life of the apple? Please explain your reasoning.

Solution:

< your solution here >

7.) Suppose we want to find the best-fitting function $f(x) = y = w^2x + wx$. How can we use linear regression to find the best value of w? Please explain.

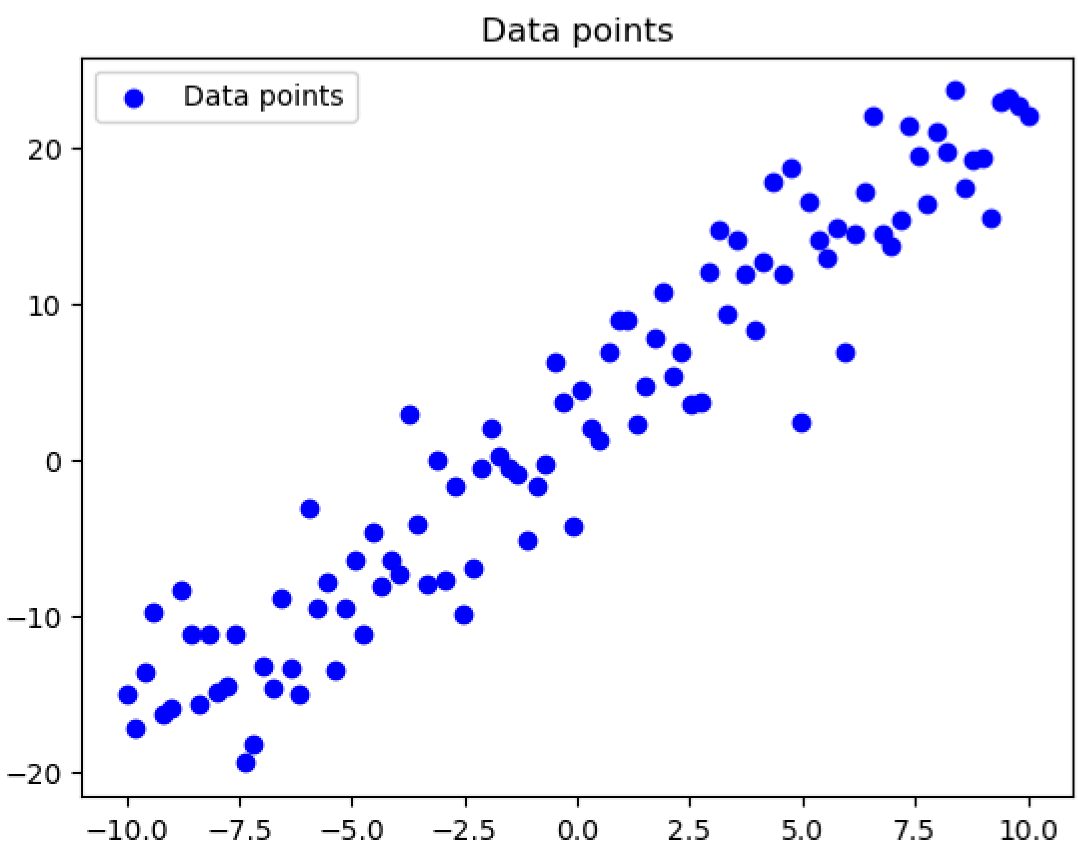
Solution:

< your solution here >

8.) Consider the following data set:

In [32…

```
import numpy as np
import matplotlib.pyplot as plt

np.random.seed(42)

# Parameters of the line
m = 2
b = 3

# Number of data points
n = 100

# Generate data points
x = np.linspace(-10, 10, n)
y = m * x + b + np.random.normal(0, 4, n)

# Plot the data points
plt.scatter(x, y, color='blue', label='Data points')
plt.title('Data points')
plt.legend()
plt.show()
```



Assume that the data points have been generated by a linear function $y = mx + b$ with some random noise added to the y values. Please implement the gradient

descent algorithm to estimate the parameters m (slope) and b (intercept) of the linear function via the following steps.

1. **Plot the provided dataset on a scatter plot.**
2. **Implement the gradient descent algorithm. Use the mean squared error as the cost function.**
3. **For each iteration of the gradient descent algorithm, plot the line with the current estimate of m and b on the scatter plot with the data points. You should end up with a plot where lines of different colors represent the estimated line of best fit at different stages of the algorithm. This will help to visualize how the line of best fit improves with each iteration.**

Solution:

< your solution here >

Finally: Submission

Save your notebook and submit it (as both **notebook and PDF file**). And please don't forget to ...

- ... choose a **file name** according to convention (see Exercise Sheet 1, but please **add your group name as a suffix** like `_group01`) and to
- ... include the **execution output** in your submission!