

AI Course

Chapter 2. Quiz

For students

©2023 SAMSUNG. All rights reserved.

Samsung Electronics Corporate Citizenship Office holds the copyright of this document.

This document is a literary property protected by copyright law so reprint and reproduction without permission are prohibited.

To use this document other than the curriculum of Samsung Innovation Campus, you must receive written consent from copyright holder.

Samsung Innovation Campus

1. Create examples of zero vector, one vector, square matrix, diagonal matrix, identity matrix, and symmetric matrix one by one, and represent vectors and matrices with NumPy.

2. The shares of three companies A, B, and C are 1 million won, 800,000 won, and 500,000 won, respectively. We want to find the amount required to purchase 3, 4, and 5 shares of these stocks, respectively.
 - 1) Express the stocks' price and quantity as p vector and n vector, respectively and coded with NumPy.
 - 2) The amount required to purchase stocks is expressed by multiplication, and the value is calculated by Numpy operation.

3. When the following code is executed, all data of the MNIST numeric image is converted into vectors to create a single NumPy matrix X. Use this matrix to solve the following problem.

```
from sklearn.datasets import load_digits
X=load_digits().data
```

- 1) Find the similarity between the first image and the tenth image using dot product.

 - 2) Find the similarity for a combination of all images using the dot product, how would it be efficient to implement it (hint: using matrices and multiplication of matrices)
-
4. Calculate the following inverse matrix.

$$\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}^{-1}$$

5. The Boston house price problem is a problem of predicting the housing price of each town in Boston, USA using features such as the crime rate and air pollution in the area. It can be imported from the `load_boston` function. Find the weight vector x when the Boston house price problem is solved with the linear prediction model $Ax=\hat{b}$. Matrix and vector data can be obtained as follows.

Here, to simplify the problem, we limited the input data to crime rate (CRIM), air quality (NOX), number of rooms (RM), and age (AGE), and only four data were used.

Run the code below to check whether the magnitude or sign of the weight vector obtained from running the program is consistent the common notion. In order to find it, interpret the printed output for all the factors: CRIM, NOX, RM and AGE.

※ Write the interpreted output like *"the house price is in inverse proportion to the crime rate (CRIM)."*

```
import pandas as pd
import numpy as np

def load_boston():
    data_url = "http://lib.stat.cmu.edu/datasets/boston"
    raw_df = pd.read_csv(data_url, sep="\s+", skiprows=22, header=None)
    data = np.hstack([raw_df.values[::2, :], raw_df.values[1::2, :2]])
    target = raw_df.values[1::2, 2]
    return {'data': data, 'target': target}

boston=load_boston()
X=boston['data']
y=boston['target']
A=X[:4, [0,4,5,6]] # 'CRIM', 'NOX', 'RM', 'AGE'
b=y[:4]
```

6. Find the weight vector w when the Boston house price problem is solved with the linear prediction model $X_w = \hat{y}$ by the least-squares method. Matrix and vector data can be obtained as follows.
 ※ This question is related to Question 5.

```
import pandas as pd
import numpy as np

def load_boston():
    data_url = "http://lib.stat.cmu.edu/datasets/boston"
    raw_df = pd.read_csv(data_url, sep="\s+", skiprows=22, header=None)
    data = np.hstack([raw_df.values[::2, :], raw_df.values[1::2, :2]])
    target = raw_df.values[1::2, 2]
    return {'data': data, 'target': target}

boston=load_boston()
X=boston['data']
y=boston['target']
```

The meaning of each column of matrix X is as follows.

- CRIM: crime rate
 - INDUS: Non-retail commercial area ratio
 - NOX: Nitric Oxide Concentration
 - RM: Number of rooms per house
 - LSTAT: Proportion of the lower class of the population
 - B: Proportion of black people in the population
 - PTRATIO: Student/Teacher Ratio
 - ZN: Percentage of residential areas exceeding 25,000 square feet
 - CHAS: 1 if located on the Charles River border, 0 otherwise
 - AGE: Percentage of houses built before 1940
 - RAD: Distance to radial highway
 - DIS: Weighted average distance to 5 Boston Job Centers
 - TAX: property tax rate
- 1) Run the program above to check whether the magnitude or sign of the weight vector obtained from running the program is consistent the common notion. In order to find it, interpret the printed output for all the factors suggested above.
 ※ Write the interpreted output like *"the house price is in inverse proportion to the crime rate (CRIM)"*.
 - 2) Explain how the result differs from the value obtained in Question 5.

7. The ratings given by three users a, b, and c to 4 movies are expressed as vectors as follows.

$$a = \begin{pmatrix} 4 \\ 5 \\ 2 \\ 2 \end{pmatrix}, b = \begin{pmatrix} 4 \\ 0 \\ 2 \\ 0 \end{pmatrix}, c = \begin{pmatrix} 2 \\ 2 \\ 0 \\ 1 \end{pmatrix}$$

- 1) Find the Euclidean distance between a, b, and c. Which two users are the closest? And which two users are farthest apart?
- 2) Find the cosine distance between a, b, and c. Which two users are the closest? And which two users are farthest apart?

8. Find the eigenvalues of the following matrix using the characteristic equation.

$$D = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

9. The $\log 2$ value is about 0.69, and the $\log 3$ value is about 1.10. At this time, find the $\log 12$ value.

10. Find the inverse function of the logistic function.

$$y = \sigma(x) = \frac{1}{1 + \exp(-x)}$$

11. Differentiate the following functions. In this equation, k, a, and b are constants, not variables.

A. $f(x) = x^3 - 1$

B. $f(x) = \log(x^2 - 3k)$

C. $f(x) = \exp(ax^b)$

12. Find the first/second partial derivatives $f_x, f_y, f_{xx}, f_{xy}, f_{yx}, f_{yy}$ for the following function.

$$f(x, y) = \exp(x^2 + 2y^2)$$

13. Find the derivative that differentiates the following function using SymPy. In this expression, k, a, and b are constants, not variables.

A. $f(x) = x^3 - 1$

B. $f(x) = \log(x^2 - 3k)$

C. $f(x) = \exp(ax^b)$

14. Find the first/second partial derivatives $f_x, f_y, f_{xx}, f_{xy}, f_{yx}, f_{yy}$ for the following function using SymPy.

$$f(x, y) = \exp(x^2 + 2y^2)$$