Assignment 1

CMPS 392

Linear Algebra

Q1

Prove that $\boldsymbol{x}^\top \boldsymbol{B}\boldsymbol{x} = 0$ if $\boldsymbol{B}$ is the anti-symmetric part of a matrix A.

Q2

What is the gradient of $\mathrm{tr}(\boldsymbol{A})$:

$$\nabla_{\boldsymbol{A}} \mathrm{tr}(\boldsymbol{A}) = ?$$

Q3

What is the gradient of the sigmoid function:

$$\sigma(\boldsymbol{x}) = \frac{1}{1 + \exp(-\boldsymbol{w}^\top \boldsymbol{x})}$$

$$\nabla_{\boldsymbol{x}} \sigma(\boldsymbol{x}) = ?$$

Q4

What is the following gradient:

$$\nabla_{\boldsymbol{x}} \frac{\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}}{||\boldsymbol{x}||_2^2}$$

assuming $\boldsymbol{A}$ is symmetric.

Q5

Is $\boldsymbol{A}^\top \boldsymbol{A}$ always:

- positive definite ?
- symmetric ?
- positive semi-definite ?

Explain.

Q6

Prove that:

$$\operatorname{tr}(\boldsymbol{ABC}) = \operatorname{tr}(\boldsymbol{CAB})$$

Q7

We know that :

- $\operatorname{tr}(\boldsymbol{A}) = \sum_i \lambda_i$
- $\det(\boldsymbol{A}) = \prod_i \lambda_i$

Prove it for the special case of a symmetric matrix $\boldsymbol{A}$.

Q8

(a) Is a diagonalizable matrix always non singular?

(b) Is a non-singular matrix always diagonalizable?

Give counter examples.

Q9

What is the image of the unit circle under multiplication by the matrix:

$$\boldsymbol{A} = \begin{bmatrix} 3 & 2 \\ 2 & 4 \end{bmatrix}$$

Write $\boldsymbol{A}$ under the form $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^\top$, then track how a point $\boldsymbol{x}$ from the circle moves after multiplication by $\boldsymbol{U}^\top$, then by $\boldsymbol{\Lambda}$, and finally by $\boldsymbol{U}$.

Q10

What is the SVD decomposition of the following matrix:

$$\boldsymbol{A} = \begin{bmatrix} 3 & 2 \\ 2 & 4 \\ 1 & 2 \end{bmatrix}$$

(write under the form : $\boldsymbol{U\Sigma V}^\top$ where $\boldsymbol{U}$ and $\boldsymbol{V}$ are orthogonal, and $\boldsymbol{\Sigma}$ is diagonal).

Programming Assignment

In this exercise, we will use linear regression to predict the weight of a fish based on five other attributes: $Length1$, $Length2$, $Length3$, $Height$ and $Width$ (You can also include *Species* which is a categorical variable in the way you see convenient if you like):

1. Download the data file "fish.csv" from Moodle.
2. The dataset has 159 datapoints. Let's divide them into 100 datapoints for training and 59 datapoints for testing.
3. Compute a linear regression model over the training data solely.
4. Compute the mean squared error (MSE) for the training data, and the MSE for the testing data, respectively.

Next, we will try to increase the capacity of the model by enriching the representation of the data. Let's add $Length1^2$, $Length2^2$, $Length3^2$, $Height^2$ and $Width^2$ to the set of features of our model. Repeat the experiment and report the new training MSE and the new testing MSE.

Next we can add $Length1^3$, $Length2^3$, $Length3^3$, $Height^3$ and $Width^3$ and so on. Each time you increase the capacity of your model compute the training MSE and the testing MSE.

What are your observations? What is the best capacity for minimizing the testing MSE? What is the difference between optimization and machine learning?