This document has some practice questions to help you prepare for the midterm.

# 1 Derivatives of Matrices and Vectors

For the following questions, and unless otherwise specified, we assume the following,

- A variable in bold font in lower caps, for example $\mathbf{x}$ , represents a column-vector with $n$ elements.

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

- A variable in bold font in upper caps, for example, $\mathbf{A}$, represents a matrix with $n$-rows and $m$-columns.

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & & a_{ij} & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

1. Is the following equality true? (Do the algebra to show it)

$$\frac{\partial\, \mathbf{a}^\top \mathbf{x}}{\partial\, \mathbf{x}} = \frac{\partial\, \mathbf{x}^\top \mathbf{a}}{\partial\, \mathbf{x}}$$

2. What is the following derivative?

$$\frac{\partial\, \mathbf{x}^\top \mathbf{A}\, \mathbf{b}}{\partial\, \mathbf{x}},$$

where $\mathbf{A}$ is a diagonal square matrix with the following form,

$$\mathbf{A} = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & & a_i & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}.$$

3. Compute the inverse of the following matrix,

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & -2 \\ 2 & -3 & -5 \\ -1 & 3 & 5 \end{pmatrix}$$

# 2 Matrix inverse

1. Prove that the matrix $\mathbf{A}$ has an inverse $\mathbf{B}$,

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}$$

# 3 Linear Regression

In class we studied the least square problem (LSP) where the goal was to solve the following problem

$$\mathbf{w}^* = \arg\min_{\mathbf{w}} \underbrace{\sum_i^N \left(y_i - (w_0 + w_1 x_i)\right)^2}_{\mathcal{L}(\mathbf{w})} = \arg\min_{\mathbf{w}} \underbrace{\left(\mathbf{Y} - \mathbf{X}\,\mathbf{w}\right)^\top \left(\mathbf{Y} - \mathbf{X}\,\mathbf{w}\right)}_{\mathcal{L}(\mathbf{w})},$$

where $\mathbf{Y}$ and $\mathbf{X}$ are the matrices with the training data, if you are not familiar with this, please revise the notes. In class, we show that this problem has an exact solution when the Jacobian of $\mathcal{L}$ with respect to $\mathbf{w}$ is equal to zero, leading to,

$$\mathbf{w}^* = \left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{Y}\,.$$

Before we continue please revise the notes on the least square problem and the use of matrix notation for $\mathcal{L}(\mathbf{w})$ and the gradient $\frac{\partial \mathcal{L}(\mathbf{w})}{\partial \mathbf{w}}$.

The main objective of this exercise is to redo the same methodology but for the **weighted least square problem** (WLSP), where the error for each point is weighted by a positive scalar value $\lambda_i$. Similarly to the LSP, we can define an error function $\mathcal{L}(\mathbf{w})$ for the WLSP,

$$\mathbf{w}^* = \arg\min_{\mathbf{w}} \sum_i^N \lambda_i \left(y_i - (w_0 + w_1 x_i)\right)^2$$

Things to do,

1. Rewrite $\mathcal{L}(\mathbf{w})$ for the WLSP in matrix form.
2. Derive the expression for $\frac{\partial \mathcal{L}(\mathbf{w})}{\partial \mathbf{w}}$.
3. Set $\frac{\partial \mathcal{L}(\mathbf{w})}{\partial \mathbf{w}} = \mathbf{0}$ and solve for $\mathbf{w}$.

• I encourage you to consult Chapter 2 of `http://matrixcookbook.com`, specially subsection 2.4.

After completing the previous steps and finding the equation for $\mathbf{w}*$, you can use the following data to test it.

$$\mathbf{X} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} 2.1 \\ 3.9 \\ 5.8 \\ 8.2 \end{bmatrix}, \quad [\lambda_1, \lambda_2, \lambda_3, \lambda_4] = [1.0, 2.0, 5.0, 10.0]$$

$\mathbf{X}$ already includes the bias term. For this data, the weighted least squares estimates for $\mathbf{w}^*$ are:

$$w_0 \approx 1.26, \quad w_1 \approx 1.70.$$

The End.