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. What is the following derivative?

$$\frac{\partial \mathbf{b}^\top \mathbf{A}(\mathbf{x} - \mathbf{s})}{\partial \mathbf{x}},$$

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

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

4. Compute the inverse of the following matrix,

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

5. What is the Jacobian of the following vectorial function $F(\mathbf{x})$,

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} F_1(x_1, x_2, x_3, x_4) \\ F_2(x_1, x_2, x_3, x_4) \\ F_3(x_1, x_2, x_3, x_4) \\ F_4(x_1, x_2, x_3, x_4) \end{bmatrix} = \begin{bmatrix} x_1^2 + x_2 x_3 - \sin(x_4) \\ e^{x_1} + x_3^2 \\ x_2 x_4 + \cos(x_1) \\ x_1 x_2 x_3 + \ln(x_4 + 2) \end{bmatrix}.$$

6. What is the following derivative?

$$\frac{\partial\, \mathbf{b}^\top \mathbf{A}(\mathbf{c} - \mathbf{s}(\mathbf{x}))}{\partial\, \mathbf{x}},$$

where $\mathbf{s}(\mathbf{x})$ is,

$$\mathbf{s}(\mathbf{x}) = \begin{bmatrix} \sin(x_1) - x_2^2 \\ \cos(x_2) + x_1^2 \end{bmatrix}.$$

For this question, you can consider $\mathbf{x}^\top = [x_1, x_2]$, $\mathbf{A}$ is a (2×2) diagonal square matrix, and $\mathbf{c}$ and $\mathbf{b}$ are vectors independent of $\mathbf{x}$.

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

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

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 you 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 λ_i. Similarly to the LSP, we can define an error function $\mathcal{L}(\mathbf{w})$ for the WLSP,

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

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 $\mathbf{w}$.
4. Can we use $\frac{\partial \mathcal{L}(\mathbf{w})}{\partial \mathbf{w}}$ for gradient descent?

• 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:

$$w_0 \approx -0.35, \quad w_1 \approx 2.12.$$

4 Eigenvalue problem

1. What are the eigenvalues of the matrix $\mathbf{A}$,

$$\mathbf{A} = \begin{bmatrix} -1 & 1 \\ 1 & 3 \end{bmatrix}$$

2. What are the eigenvalues and eigenvectors of the matrix $\mathbf{A}$,

$$\mathbf{A} = \begin{bmatrix} 6 & 0 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix}$$

3. (This one is slightly challenging) Prove that the trace of a matrix is equal to the sum of the eigenvalues, $\mathrm{Tr}[\mathbf{A}] = \sum_i^n \lambda_i$, where $\{\lambda_i\}_i^n$ are the eigenvalues of $\mathbf{A}$.
   To do so you need to consider two things,
   - $\mathbf{A}$ can be rewritten in terms of the eigenbasis.
   - The trace of matrix multiplication is cyclic; meaning, $\mathrm{Tr}[\mathbf{A}\,\mathbf{B}\,\mathbf{C}] = \mathrm{Tr}[\mathbf{B}\,\mathbf{C}\,\mathbf{A}] = \mathrm{Tr}[\mathbf{C}\,\mathbf{A}\,\mathbf{B}]$.

4.1 Linear coupled ordinary differential equations

1. What is the most general solution for the following coupled linear ODEs,

$$\frac{d\,\mathbf{x}}{dt} = \mathbf{A}\,\mathbf{x} = \begin{bmatrix} 1 & 4 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

2. Given the following reactions, $\mathrm{A} \xrightarrow{\mathrm{k_1}} \mathrm{B} \xrightarrow{\mathrm{k_2}} \mathrm{C}$. These give us the following first-order reaction rates,

$$\begin{aligned}\frac{d[\mathrm{A}]}{dt} &= -k_1[\mathrm{A}] \\ \frac{d[\mathrm{B}]}{dt} &= k_1[\mathrm{A}] - k_2[\mathrm{B}] \\ \frac{d[\mathrm{C}]}{dt} &= k_2[\mathrm{B}]\end{aligned}$$

- Solve these set of coupled linear equations
- State the eigenvalues and eigenvectors.
- Plot the values of [A](t), [B(t)], and [C(t)] over time, you can consider that at initial time, $[\mathrm{B}(\mathrm{t}_0)] = [\mathrm{C}(\mathrm{t}_0)] = 0$ and $[\mathrm{A}(\mathrm{t}_0)] = [\mathrm{A}]_0$

The End.