CHEM 3PC3

Assignment 2

November 24, 2025

Name: ____________________

Student Number: ____________________

SUBMISSION INSTRUCTIONS — READ CAREFULLY

To receive full credit, you must follow these steps:

1. Answer all the questions.
2. Write your solutions on new, separate pages. Do not write your solutions in the margins of this paper.
3. At the top of each solution page, clearly write the corresponding question number (e.g., **"Question 5"**). If you use more than one page for a question, write the question number on each page (e.g., **"Question 5 (Page 1 of 2)"**).
4. **If you cannot solve a question, still attach a page with the question number and write "Blank" or "No Answer"** to indicate you attempted it.
5. **If you are unsure of a complete answer, still attempt the question:** attach a page with the question number and write down any relevant thoughts, formulas, or initial steps. Partial credit may be awarded for demonstrated effort and correct reasoning, whereas a blank answer will receive no credit.
6. Show all your work clearly and legibly. Unorganized or illegible work may not receive credit.

1 Problems

Problem 1

In regularized linear regression, we encounter quadratic forms. Consider:

$$y = \sum_{i=1}^{n} \frac{1}{\lambda_i} x_i^2 \tag{1}$$

where $\lambda_i > 0$ are constants and $\mathbf{x} \in \mathbb{R}^n$ with components x_i.

1. Express y in a vector-matrix-vector multiplication format.
2. Using $y = \mathbf{x}^\top \boldsymbol{\Lambda} \mathbf{x}$, compute:
   (a) $\dfrac{\partial y}{\partial \mathbf{x}}$
   (b) $\dfrac{\partial y}{\partial \boldsymbol{\Lambda}}$

Hint:

1. Use dimensional analysis to find the shape of $\boldsymbol{\Lambda}$.

Problem 2

Given,

$$f(x) = \|\,\mathbf{x}\,\|_1 + \|\,\mathbf{x}\,\|_2^2 + \mathrm{ReLU}(\mathbf{w}^\top\,\mathbf{x}) + \mathbf{x}^\top\,\mathbf{A}\,\mathbf{x} \tag{2}$$

where $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} = (x_1, \dots, x_n)^\top$. Calculate

$$\frac{df(\mathbf{a})}{d\,\mathbf{x}} \tag{3}$$

for

$$\mathbf{a} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}.$$

For the term $\mathbf{x}^\top\,\mathbf{A}\,\mathbf{x}$, do not use the identity $\frac{d}{dx}(\mathbf{x}^\top\,\mathbf{A}\,\mathbf{x}) = (\mathbf{A} + \mathbf{A}^\top)\,\mathbf{x}$. Instead, explicitly expand the matrix product.

Hints:

1. Apply the definition of the L1-norm $\|\,\mathbf{x}\,\|_1$, the L2-norm $\|\,\mathbf{x}\,\|_2^2$ and the $\mathrm{ReLU}(x)$ function before taking the derivative.
2. The derivative of the L1-norm depends on the sign of x, so $\frac{d\|x\|_1}{dx_i} = \mathrm{sign}(x_i)$.
3. Calculate the operation $\mathbf{x}^\top\,\mathbf{A}\,\mathbf{x}$ and then take the derivative.

Problem 3

The **Log-Regularized Ridge Regression** (RRP-log) modifies the least squares problem by adding a penalty term to prevent large weight values. The error function is:

$$\mathcal{L}(\mathbf{w}) = \frac{1}{2n}\sum_{i=1}^{n}\left(\mathbf{w}^\top\boldsymbol{\phi}(x_i) - \hat{y}_i\right)^2 + \ln(1+\lambda)\sum_{j=1}^{d} w_j^2 \tag{4}$$

Let's consider a polynomial of order 2 linear model, meaning, we will represent for each point $\boldsymbol{\phi}(x_i)$ as,

$$\boldsymbol{\phi}(x_i)^\top = [1, x, x^2]. \tag{5}$$

Where,

- $\mathbf{w} = [w_1, w_2, \cdots, w_d]^\top$ are the model parameters
- $\lambda > 0$ is the regularization hyperparameter
- d is the number of parameters, n is the number of training samples.

To do,

1. Rewrite $\mathcal{L}(\mathbf{w})$ in matrix form.
2. Derive $\dfrac{\partial \mathcal{L}(\mathbf{w})}{\partial \mathbf{w}}$ in matrix form.

3. Set $\dfrac{\partial \mathcal{L}(\mathbf{w})}{\partial \mathbf{w}} = \mathbf{0}$ and solve for $\mathbf{w}$ to find $\mathbf{w}^*$. **Show all matrix inversion steps !**

4. After you derive the equations for $\mathbf{w}^*$, use the following data to verify your derivation. Consider the following 3 points,

$$\begin{bmatrix} x_1 = 1 \\ x_2 = 2 \\ x_3 = 3 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} y_1 = 1 \\ y_2 = 4 \\ y_3 = 9 \end{bmatrix}$$

   For $\lambda = 0.1$, you should get that $w_1 = -0.047, w_2 = 0.176$, and $w_3 = 0.937$.

5. How different is $\mathbf{w}^*$ if we set $\lambda = 0$? Find the explicit expression of $\mathbf{w}^*$ with $\lambda = 0$.

Problem 4

In class we studied the concept of diagonalizable matrices. A **square matrix** $\mathbf{A} \in \mathbb{R}^{n\times n}$ is said to be diagonalizable if it can be written in the form

$$\mathbf{A} = \mathbf{P}\,\mathbf{\Lambda}\,\mathbf{P}^{-1}, \tag{6}$$

or equivalently,

$$\mathbf{\Lambda} = \mathbf{P}^{-1}\,\mathbf{A}\,\mathbf{P}, \tag{7}$$

where $\mathbf{P}$ is the matrix whose columns are the eigenvectors of $\mathbf{A}$, $\mathbf{P}^{-1}$ is its inverse, and $\mathbf{\Lambda}$ is a diagonal matrix whose diagonal entries are the eigenvalues of $\mathbf{A}$. Given the matrices:

$$\mathbf{A}_1 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad \mathbf{A}_2 = \begin{pmatrix} 2 & 0 & 2 \\ -1 & 2 & 1 \\ 0 & 1 & 4 \end{pmatrix}, \quad \mathbf{A}_3 = \begin{pmatrix} 2 & 0 & 2 & 0 \\ -1 & 2 & 1 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}.$$

1. Calculate the determinant of each matrix $\mathbf{A}_i$ for $i = 1, 2, 3$.
2. For the matrices $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$ where the determinant exists, does this guarantee they are diagonalizable?
3. Find all eigenvalues of each matrix $\mathbf{A}_i$ for $i = 1, 2, 3$.
4. Compute an eigenvector associated with each eigenvalue.
5. Compute the $\mathbf{P}$ matrix for each case and calculate its determinant.
6. What does the determinant of $\mathbf{P}$ indicate about diagonalizability?
7. Explicitly write the diagonal matrix $\mathbf{\Lambda}$ for each matrix if it exists.