Mathematics for Machine Learning (2025 Fall)
Assignment 02*

Chuang-Chieh Lin

Department of Computer Science and Engineering,
National Taiwan Ocean University

1. (20%) Diagonalize
$$\begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix} = \mathbf{P}\mathbf{D}\mathbf{P}^\top$$

   such that $\mathbf{P}$ consists of orthonormal column vectors.

2. (20%) Suppose that $\mathbf{y} = (y_1, y_2 \ldots, y_n) \in \mathbb{R}^n$ is the observation and $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ is a prediction from a model such that $\sum_{i=1}^n x_i = \sum_{j=1}^n y_j = 1$. What is $\dfrac{\partial}{\partial \mathbf{x}} \mathrm{D}_{\mathrm{KL}}(\mathbf{x}\|\mathbf{y})$?

3. (20%) Given $\mathbf{x} \in \mathbb{R}^n$, please compute $\frac{\mathrm{d}}{\mathrm{d}\mathbf{x}}(2\mathbf{x})$.

4. (20%) Compute the derivative $f'(x)$ of the (univariate) logistic sigmoid
$$f(x) = \frac{1}{1 + \exp(-x)}.$$

5. (20%) Given the formula
$$\frac{\partial \mathbf{x}^\top \mathbf{B} \mathbf{x}}{\partial \mathbf{x}} = \mathbf{x}^\top (\mathbf{B} + \mathbf{B}^\top)$$

   for a square matrix $\mathbf{B}$, compute the gradient

$$\frac{\partial}{\partial \mathbf{s}} \left((\mathbf{x} - \mathbf{A}\mathbf{s})^\top \mathbf{A}\mathbf{A}^\top (\mathbf{x} - \mathbf{A}\mathbf{s}) + \|\mathbf{s}\|^2 \right)$$

---

* List the required intermediate steps next to each problem. Note that any answers generated directly by AI are invalid for this assignment.