Instructor: Vishnu Boddeti
CSE 840: Computational Foundations in AI
Homework 2

Instructions:

- **Format:** Please submit a PDF file with typed answers (LaTeX or otherwise). Handwritten notes will not be graded.
- **Submission:** Only homework uploaded to Gradescope will be graded. Make sure to show all the steps of your derivations in order to receive full credit.
- **Integrity and Collaboration:** You are expected to work on the homework by yourself. You are not permitted to discuss them with anyone except the instructor. The homework that you hand in should be entirely your own work. You may be asked to demonstrate how you got any results that you report.
- **Clarifications:** If you have any questions, please look at Piazza first. Other students may have encountered the same problem and is solved already. If not, post your question there. We will respond as soon as possible.

1 [10 points] Prove that:

$$\langle x, y \rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + j\|x+jy\|^2 - j\|x-jy\|^2\right)$$

2 [10 points] Consider the following $n \times n$ matrix:

$$\mathbf{A} = \mathbf{I}_n - \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T,$$

where $\mathbf{X} \in \mathbb{R}^{n \times d}$ with $n > d$ and $\mathbf{I}_n$ is an $n \times n$ identity matrix. Note that $\mathbf{A}$ represents the residual error of linear regression models since

$$\mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{X}\mathbf{w} = \mathbf{y} - \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} = (\mathbf{I} - \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T)\mathbf{y}$$

Check whether $\mathbf{A}$ is symmetric, orthogonal, and idempotent (i.e., $\mathbf{A}^2 = \mathbf{A}$). Make sure you show the steps to verify each property separately.

3 [20 points]

1. Let $\mathbf{R}$ be a 2×2 rotation matrix at some counter-clockwise angle of θ. Compute $\mathbf{R}^8$ and provide a geometric interpretation of the matrix. For full credit, make sure you simplify the expression for $\mathbf{R}^8$ to its simplest form before making the interpretation. **Hint:** You may refer to the trigonometry identities available at `https://www2.clarku.edu/faculty/djoyce/trig/identities.html` to simplify your expression for $\mathbf{R}^8$.

2. Consider the vector $\mathbf{x} = \begin{bmatrix} 8 \\ 6 \end{bmatrix}$. Find the rotation angle θ needed to eliminate (zero out) the value in the second element of vector $\mathbf{x}$. **Hint:** Use Givens rotation and the fact that an orthogonal transform preserves the norm of a vector to answer this question. Show your calculations clearly to obtain full credit.

3. Consider the vector $\mathbf{x} = \begin{bmatrix} 8 \\ 6 \end{bmatrix}$. Find the unit normal vector $\hat{\mathbf{v}}$ of the reflection hyperplane that linearly transforms the vector to $\mathbf{x}' = \begin{bmatrix} 10 \\ 0 \end{bmatrix}$. **Hint:** use Householder reflection to solve for $\hat{\mathbf{v}}$.

4 [20 points] Let $\boldsymbol{\Sigma} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 3 & 2 \\ 0 & 2 & 3 \end{bmatrix}$ be the covariance matrix for some 3-dimensional dataset. Answer each of the following questions below. For full credit, you must show your intermediate steps clearly.

1. Compute $\|\boldsymbol{\Sigma}\|_1$, $\|\boldsymbol{\Sigma}\|_\infty$, $\|\boldsymbol{\Sigma}\|_F$ and $\|\boldsymbol{\Sigma}\|_2$.

2. Compute the inverse of $\boldsymbol{\Sigma}$. Make sure you show its determinant and cofactor matrix as well.

3. Determine whether $\boldsymbol{\Sigma}$ is a positive definite or semi-definite, negative definite or semi-definite, or indefinite matrix.

5 [20 points]

1. Consider the normal equation $\mathbf{X}^T\mathbf{X}\mathbf{w} = \mathbf{X}^T\mathbf{y}$ associated with the multiple linear regression formulation. Assuming $\mathbf{X} \in \mathbb{R}^{n\times d}$ is a full-rank matrix (where $n > d$), using singular value decomposition on $\mathbf{X}$, derive a closed-form solution for $\mathbf{w}$ in terms of $\mathbf{y}$ as well as the matrix of singular values and singular vectors of $\mathbf{X}$.

2. Derive the singular value decomposition of the following matrix:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \\ 2 & 1 \end{bmatrix}$$

   Show your derivation clearly.