ELEC 378 – Spring 2023
Homework 6

Due: Friday February 24, 5PM

1 Optimization for Best Orthogonal Projection

Let $\mathbf{X}$ denoted the centered $n \times p$ data matrix, and suppose $0 < q < p$. Using Lagrange multipliers, prove that the PCA objective

$$\max_{\mathbf{u}_1,\ldots,\mathbf{u}_q} \sum_{i=1}^{q} \mathbf{u}_i^\intercal \mathbf{X}^\intercal \mathbf{X} \mathbf{u}_i \text{ subject to } \left\{\|\mathbf{u}_i\|_2^2 = 1\right\}_{i=1}^{q}$$

is maximized when $\mathbf{u}_i$ is chosen to be the i^{th} eigenvector of the covariance matrix $\mathbf{X}^\intercal\mathbf{X}$.

2 Pseudoinverse via SVD

Let $\mathbf{X}^\dagger = (\mathbf{X}^{\text{H}}\mathbf{X})^{-1}\mathbf{X}^{\text{H}}$ denote the Moore-Penrose pseudoinverse of the matrix $\mathbf{X}$. Prove that $\mathbf{X}^\dagger = \mathbf{V}\mathbf{\Sigma}^\dagger\mathbf{U}^{\text{H}}$, where $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\text{H}}$ and $\mathbf{\Sigma}^\dagger$ is obtained by taking the reciprocal of the non-zero entries of $\mathbf{\Sigma}$ and leaving the zeros in place.

3 Predicting House Sale Price with Linear Regression

Download `train.csv` from the following Kaggle competition:
`https://www.kaggle.com/competitions/house-prices-advanced-regression-techniques/data`

a) Which features can be used to predict the sale price using linear regression? Construct the data matrix $\mathbf{X}$, containing the appropriate features of each home for linear regression, and label vector $\mathbf{y}$, containing the sale price of each home. What are the dimensions n and p?

b) Use linear regression to predict the sale price $\mathbf{y}$ from the features $\mathbf{X}$. Are you able to make "good" prediction? Why or why not?

c) Which features are important in predicting a sale price according to your learned parameters $\mathbf{w}^*$? Which features are not important?

Submission Instructions

Every student must submit their work in PDF format, providing intermediate and final results as well as any necessary code. Submit your homework on Gradescope.

Collaboration Policy

Collaboration both inside and outside class is encouraged. You may talk to other students for general ideas and concepts, but individual write-ups must be done independently.

Plagiarism

Plagiarism of any form will not be tolerated. You are expected to credit all sources explicitly.