Homework Set 2, CPSC 8420, Fall 2024

Your Name

Due 10/28/2024, Monday, 11:59PM EST

Problem 1

For PCA, from the perspective of maximizing variance (assume the data is already self-centered)

- show that the first column of $\mathbf{U}$, where $[\mathbf{U}, \mathbf{S}] = svd(\mathbf{X}^T\mathbf{X})$ will maximize $\|\mathbf{X}\boldsymbol{\phi}\|_2^2$, $s.t.$ $\|\boldsymbol{\phi}\|_2 = 1$. (Note: you need prove why it is optimal than any other reasonable combinations of $\mathbf{U}_i$, say $\hat{\boldsymbol{\phi}} = 0.8 * \mathbf{U}(:,1) + 0.6 * \mathbf{U}(:,2)$ which also satisfies $\|\hat{\boldsymbol{\phi}}\|_2 = 1$.)
- show that the solution is not unique, say if $\boldsymbol{\phi}$ is the optimal solution, so is $-\boldsymbol{\phi}$.
- show that first r columns of $\mathbf{U}$, where $[\mathbf{U}, \mathbf{S}] = svd(\mathbf{X}^T\mathbf{X})$ maximize $\|\mathbf{X}\mathbf{W}\|_F^2$, $s.t.$ $\mathbf{W}^T\mathbf{W} = \mathbf{I}_r$.
- Assume the singular values are all different in $\mathbf{S}$, then how many possible different $\mathbf{W}$'s will maximize the objective above?

Problem 2

Given matrix $\mathbf{X} \in \mathbb{R}^{n \times p}$ (assume each column is centered already), where n denotes sample size while p feature size. To conduct PCA, we need find eigenvectors to the largest eigenvalues of $\mathbf{X}^T\mathbf{X}$, where usually the complexity is $\mathcal{O}(p^3)$. Apparently when $n \ll p$, this is not economic when p is large. Please consider conducting PCA based on $\mathbf{X}\mathbf{X}^T$ and obtain the eigenvectors for $\mathbf{X}^T\mathbf{X}$ accordingly and use experiment to demonstrate the acceleration.

Problem 3

Let $\theta^* \in \mathbb{R}^d$ be the ground truth linear model parameter and $\mathbf{X} \in \mathbb{R}^{N\times d}$ be the observing matrix and each column of $\mathbf{X}$ is independent. Assume the linear model is $\mathbf{y} = \mathbf{X}\theta^* + \epsilon$ where ϵ follows $Gaussian(0, \sigma^2\mathbf{I})$. Assume $\hat{\theta} = \arg\min_{\theta} \|\mathbf{X}\theta - \mathbf{y}\|^2$.

- Please show that $\mathbf{X}^T\mathbf{X}$ is invertible.
- Show that $MSE(\theta^*, \hat{\theta}) := E_\epsilon\{\|\theta^* - \hat{\theta}\|^2\} = \sigma^2 trace((\mathbf{X}^T\mathbf{X})^{-1})$
- Show that as N increases, MSE decreases. (hint: make use of 'Woodbury matrix identity')