CMSE 820 HW3

This HW is due on Sep 29st at 11:59 pm.

Question 1: For the statistical view of PCA, finish the proof for the theorem of Principal Components of a Random Vector assuming there is no repeated eigenvalues.

Question 2: Given $\mathbf{y} \in \mathbb{R}^n$ and $\mathbf{X} = I \in \mathbb{R}^{n\times n}$, derive the analytical Lasso solution step by step solving the

$$\arg\min \|\mathbf{y} - \mathbf{X}^T\beta\|^2 + \lambda\|\beta\|_1.$$

Question 3: Let $A \in \mathbb{R}^{m\times n}$ be a matrix.

a. The nuclear norm of A is defined as $f(A) = \|A\|_* = \sum_{i=1}^{\min(m,n)} \sigma_i(A)$, where $\sigma_i(A)$ is the ith largest singular value of A. Given the nuclear norm is a norm (I know it is ...weird to state that, but remember the ℓ_0 is not a norm). Show that it is also a convex function.

b. The subdifferential of the nuclear norm at A is define by

$$\partial\|A\|_* = \left\{G \in \mathbb{R}^{m\times n} : \|B\|_* \geq \|A\|_* + \mathrm{Tr}\left[(B-A)^T G\right],\ \text{for all } B \in \mathbb{R}^{m\times n}\right\}.$$

Then for $G \in \partial\|A\|_*$ prove that

(1) $\|A\|_* = \mathrm{Tr}(G^T A)$.

(2) $\|G\|_*^* = 1$ where

$$\|G\|_*^* = \max_{\|B\|_*\leq 1} \mathrm{Tr}(B^T G).$$

Here, $\|\cdot\|_*^*$ is actually the dual norm to $\|\cdot\|_*$.

c. Given that the subdifferential of the nuclear norm has the following expression

$$\partial\|A\|_* = \{UV^T + W : U^T W = 0, WV = 0, \|W\|_2 \leq 1, W \in \mathbb{R}^{m\times n}\}$$

where $A = U\Sigma V^T$ is singular-value decomposed (into $U \in \mathbb{R}^{m\times r}, \Sigma \in \mathbb{R}^{r\times r}$, and $V \in \mathbb{R}^{n\times r}$) and $\|\cdot\|_2$ is the spectral norm which is the maximum singular value of a matrix. Show that the optimal solution of

$$\min_A \frac{1}{2}\|X - A\|_F^2 + \lambda\|A\|_*$$

is given by $A = \mathcal{D}_\lambda(X) = U_x\mathcal{S}_\lambda(\Sigma_x)V_x^T$, where D_λ is called the singular-value thresholding operator and $X = U_x\Sigma_x V_x^T$.

Question 4: Face recognition using PCA. In this exercise you will use a small subset of the Yale B dataset, which contains photos of ten individuals under various illumination conditions. Specifically, you will use only images from the first three individuals under ten different illumination conditions. Download the file YaleB-Dataset.zip. This file contains the image database along with the MATLAB function loadimage.m. Decompress the file and type help loadimage at the MATLAB prompt to see how to use this function.

a. Write your own code for PCA (you can use any program language).

b. Apply PCA with $d = 2$ to all 10 images from individual

   1. Plot the mean face μ and the first two eigenfaces u_1 and u_2. What do you observe?

   2. Plot $\mu + y_i \mathbf{u}_i$ for $y_i = -\sigma_i, -0.8\sigma_i, -0.6\sigma_i, \ldots, 0.6\sigma_i, 0.8\sigma_i, \sigma_i$ with $i \in \{1, 2\}$. Here, σ_i is the standard deviation of $y[i]$ (the ith principal component). What do the first two principal components capture?

. Repeat for individuals 2 and 3.