EECS 553 Homework 7 Solution (FA24)

1. PCA Eigenfaces

a. Grading Rubrics:

- 1.5 pts for correct pattern (as long as the trend is similar, do not have to match values) in the eigenvalue plots.
- 1.5 pts for correct number of PCs to achieve the desired percentages of explained variances.
- 0 pts if no effort or entirely wrong.

143 (within ± 2) PCs are needed to represent 95% of the total variation.

535 (within ± 2) PCs are needed to represent 99% of the total variation.

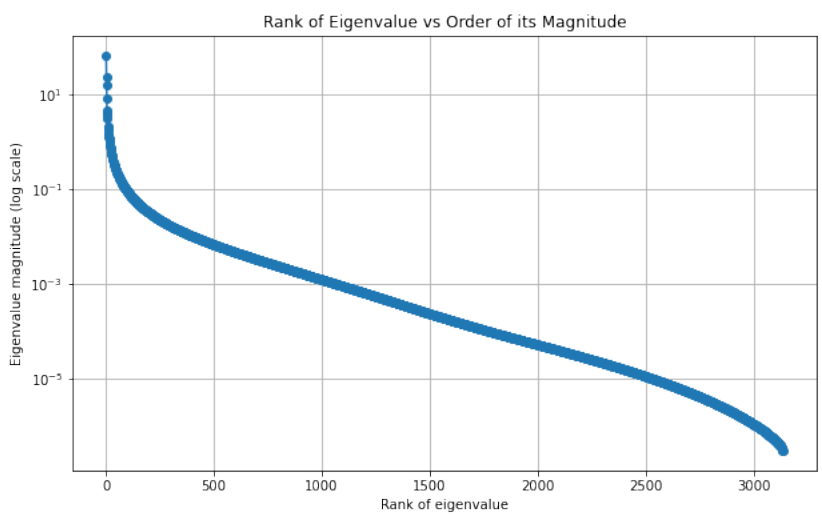

Figure 1: 1a Eigenvalues on semilog plot

b. Grading Rubrics:

- Full points for generally correct patterns in the eigenface plots (do not have to match perfectly).
- 0 pts if no effort or entirely wrong.

Some of the principal components capture background lighting (first row, second image),

c. Grading Rubrics:

- 1.5 pts for each set (for 2nd PC and 11th PC) of original images plotted. Note: only five images are required for each set.
- 0 pts if no effort or entirely wrong.

2. Uniqueness of PCA Subspace

Let the eigenvectors of the covariance matrix be $\{\boldsymbol{u}_i\}_{i=1}^d$ with corresponding eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$.

A necessary and sufficient condition is that $\lambda_k \neq \lambda_{k+1}$. $\lambda_k \neq \lambda_{k+1} \implies$ the subspace $\langle A \rangle$ is uniquely spanned by $\{\boldsymbol{u}_i\}_{i=1}^k$. If $\lambda_k = \lambda_{k+1}$, $\langle \boldsymbol{A} \rangle$ can be either spanned by $\{\boldsymbol{u}_i\}_{i=1}^k$ or $\{\boldsymbol{u}_i\}_{i=1}^{k-1} \cup \{\boldsymbol{u}_{k+1}\}$.

3. PCA Optimal Object Function Value

Let $\bar{\boldsymbol{x}}$ be the mean of features vectors and $\boldsymbol{X}$ be the matrix whose columns are centered feature vectors $\boldsymbol{x}_i - \bar{\boldsymbol{x}}$. Let $S = \boldsymbol{X}^T\boldsymbol{X}$ denote the covariance matrix. Let $\boldsymbol{A}_*$ whose columns are the first k eigenvectors of S.

$$
\begin{aligned}
\min_{\boldsymbol{\mu},\boldsymbol{A},\{\boldsymbol{\theta}_i\}} \sum_{i=1}^n \|\boldsymbol{x}_i - \boldsymbol{\mu} - \boldsymbol{A}\boldsymbol{\theta}_i\|^2 &= \sum_{i=1}^n \|\boldsymbol{x}_i - \bar{\boldsymbol{x}} - \boldsymbol{A}_*\boldsymbol{A}_*^T(\boldsymbol{x}_i - \bar{\boldsymbol{x}})\|_2^2 \\
&= \|\boldsymbol{X} - \boldsymbol{A}_*\boldsymbol{A}_*^T\boldsymbol{X}\|_F^2 \\
&= tr(\boldsymbol{X}^T\boldsymbol{X}) - ntr(\boldsymbol{A}_*^T\boldsymbol{S}\boldsymbol{A}_*) \\
&= tr(\boldsymbol{X}\boldsymbol{X}^T) - ntr(\boldsymbol{A}_*^T\boldsymbol{S}\boldsymbol{A}_*) \\
&= n\sum_{i=1}^d \lambda_i - n\sum_{i=1}^k \lambda_i \\
&= n\sum_{i=k+1}^d \lambda_i,
\end{aligned}
$$

where the first equality follows the PCA lecture notes, the fourth equality is a direct consequence of the property of trace operator.