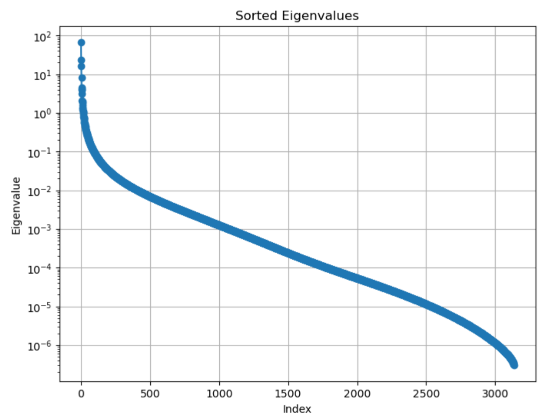
EECS 553 HW7

Lingqi Huang

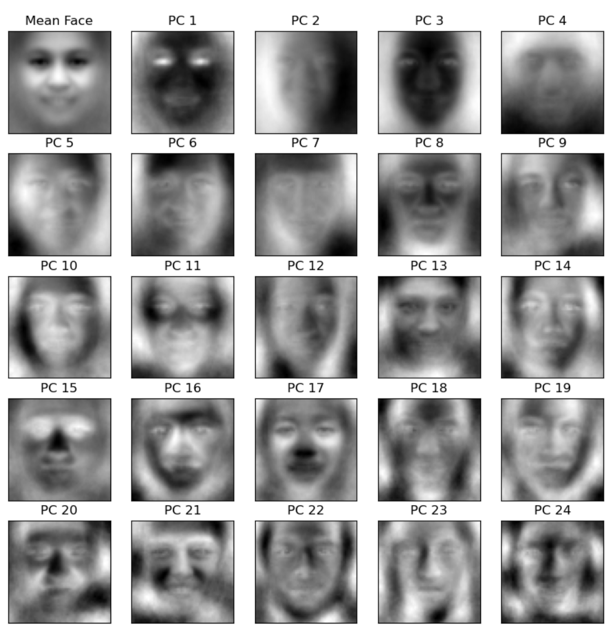
October 2024

1 Problem 1

Part(a): We notice that we need 143 principle components to explain 95% of variation, and we need 535 principle components to explain 99% variation. To explain 95% variation, the percentage of dimension reduced is about 95.4%, to explain 99% variation, the percentage of dimension reduced is about 82.94%.



Part(b): Notice that the first principle component captures the lighting variation of eyes, the third principle component may captures the face shape of a person. The 15th principle component may captures the light variation of nose.



Part(c): We can see that from part(b), the second principle component focus on the shadow or darkness around the face, and the 11-th principle component focus on the shadow or darkness around the eye, which is consistent with our below pictures, that the top 5 images with highest score for both principle components have significant characteristic that been addressed by these principle components.

Figure 1: Images with Top5 Score in PC2

Figure 2: Images with Top5 Score in PC11

2 Problem 2

Claim: The necessary and sufficient condition is $\lambda_k > \lambda_{k+1}$

" $\Rightarrow$ ": Given $A \in \mathcal{A}_k$. Suppose that the k-th largest eigenvalue $\lambda_k > \lambda_{k+1}$, we must have that the direction of their eigenvector is different, which means the subspace span by $[u_1, \dots, u_{k-1}, u_k]$ is different from the subspace span by $[u_1, \dots, u_{k-1}, u_{k+1}]$. This shows that the subspace $\langle A \rangle$ in PCA is unique.

" $\Leftarrow$ ": Now we prove this statement by contrapositive. If $\lambda_{k+1} = \lambda_k$, this means the subspace $\langle A \rangle$ could be either spanned by $[u_1, \dots, u_{k-1}, u_k]$ or $[u_1, \dots, u_{k-1}, u_{k+1}]$, where the eigenvector $u_k \neq u_{k+1}$. This contradicts with the fact that $\langle A \rangle$ is unique.

3 Problem 3

According to lecture notes page 2, we know that one solution to the optimization problem is that:

$$\boldsymbol{\mu} = \bar{\mathbf{x}}$$

$$\mathbf{A} = [\boldsymbol{u}_1, \dots, \boldsymbol{u}_k]$$
$$\theta_i = \mathbf{A}^T(\mathbf{x}_i - \bar{\mathbf{x}})$$

where

$$\mathbf{S} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T = \mathbf{U}\Lambda\mathbf{U}^T$$
$$\mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = I$$

and Λ is a diagonal matrix contains all eigenvalues $\lambda_1, \dots, \lambda_d$. Thus, it is same to find the minimum of the optimization problem of

$$\min_{\mathbf{A}} \sum_{i=1}^{n} ||\mathbf{x}_i - \bar{\mathbf{x}} - \mathbf{A}\mathbf{A}^T(\mathbf{x}_i - \bar{\mathbf{x}})||^2$$

Now we set $\tilde{\mathbf{x}} = \mathbf{x}_i - \bar{\mathbf{x}}$ that is normalized with mean of 0, the optimization problem becomes

$$\min_{\mathbf{A}} ||\tilde{\mathbf{X}} - \mathbf{A}\mathbf{A}^T\tilde{\mathbf{X}}||_F^2$$

According to page 7 of the lecture notes, given optimal $\mathbf{A}$, we find that

$$||\tilde{\mathbf{X}} - \mathbf{A}\mathbf{A}^T\tilde{\mathbf{X}}||_F^2 = tr(\tilde{\mathbf{X}}\tilde{\mathbf{X}}^T) - tr(\tilde{\mathbf{X}}\mathbf{A}\mathbf{A}^T\tilde{\mathbf{X}}^T)$$

Since we know that $\mathbf{S}$ is symmetric, so the trace of $\mathbf{S}$ is summation of its eigenvalues, and so

$$tr(\tilde{\mathbf{X}}\tilde{\mathbf{X}}^T) = n\sum_{i=1}^{d}\lambda_i$$

Also, by page 12 of lecture notes, we know that

$$tr(\tilde{X}\mathbf{A}\mathbf{A}^T\tilde{\mathbf{X}}^T) = n\sum_{i}^{k}\lambda_i$$

Thus, we conclude that

$$\min_{\boldsymbol{\mu},\mathbf{A},\{\theta_i\}}\sum_{i=1}^{n}||\mathbf{x}_i - \boldsymbol{\mu} - \mathbf{A}\theta_i||^2 = n\sum_{j=k+1}^{d}\lambda_j$$

This completes the proof.