Math 521 Homework 4

Due on Tuesday, April 17, 2018

Kristin Holmbeck

Contents

List of Figures

Theory

1.

Consider the two eigenvector problems

$$C_x\boldsymbol{u} = \lambda_x\boldsymbol{u}$$

and

$$C_s\boldsymbol{v} = \lambda_s\boldsymbol{v}$$

where the matrices are related by $C_x = C_x + \alpha I$, where α is a real number and I is the usual identity matrix. Show that if $\boldsymbol{u}$ is an eigenvector of C_x, then it is also an eigenvector of C_s associated with eigenvalue $\lambda_s = \lambda_x - \alpha$.

$$\begin{aligned}
C_x\boldsymbol{u} &= \lambda_x\boldsymbol{u} = C_s\boldsymbol{u} + \alpha I\boldsymbol{u} \\
C_s\boldsymbol{u} &= \lambda_x\boldsymbol{u} - \alpha I\boldsymbol{u} \\
C_s\boldsymbol{u} &= (\lambda_x - \alpha)\boldsymbol{u} \\
\text{but} \quad C_s\boldsymbol{v} &= \lambda_s\boldsymbol{v} \\
\implies \lambda_s &= \lambda_x - \alpha \qquad \text{is an eigenvalue with associated eigenvector } \boldsymbol{u}
\end{aligned}$$

2.

Let $A \in \mathbb{R}^{m\times n}$. Show that the matrix M defined as

$$M = \alpha^2 I + AA^T, \quad \alpha \neq 0 \in \mathbb{R}$$

is nonsingular, where $I = I_m$ and α is a nonzero real number.

For the trivial case, suppose $A = 0_{m\times n}$. Then $M = \alpha^2 I$ which is clearly nonsingular since $\alpha \neq 0$.

Next, note that $AA^T = (AA^T)^T$ since $(AA^T)^T = (A^T)^T A^T = AA^T$. This implies that M is symmetric.

3.

Show that the between-class scatter matrix, S_B, in the multi-class *Fisher Discriminant Analysis* is given by

$$S_B = \sum_{i=1}^{M} n_i(\boldsymbol{m}_i - \boldsymbol{m})(\boldsymbol{m}_i - \boldsymbol{m})^T,$$

where M is the total number of distinct classes, n_i is the number of data points in class i, $\boldsymbol{m}_i$ is the class mean of the i^{th} class, and m is the mean across all n data points. You may use the facts that

$$S_T = S_B + S_W, \quad S_W = \sum_{i=1}^{M}\sum_{x\in D_i}(x - \boldsymbol{m}_i)(x - \boldsymbol{m}_i)^T, \quad \text{and} \quad S_T = \sum_{i=1}^{n}(x_i - \boldsymbol{m})(x_i - \boldsymbol{m})^T$$

Given what we know above, and the fact that the ith class mean $\boldsymbol{m}_i = \frac{1}{n_i}\sum_{x\in D_i}\boldsymbol{x}$ and the mean $\boldsymbol{m}$ across all data points is given by $\boldsymbol{m} = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{x}_i$, we will start with evaluating $S_B = S_T - S_W$.

Furthermore, note that $\forall \boldsymbol{x}_i \in X$ (the entire data set), $\boldsymbol{x}_i \in \bigcup_{j=1}^{M} D_j$, the union of all distinct classes, and also $\boldsymbol{m}_i n_i = \sum_{x\in D_i}\boldsymbol{x}$.

$$
\begin{aligned}
S_T - S_W &= \sum_{i=1}^{n}(x_i-\boldsymbol{m})(x_i-\boldsymbol{m})^T - \sum_{i=1}^{M}\sum_{x\in D_i}(x-\boldsymbol{m}_i)(x-\boldsymbol{m}_i)^T \\
&= \sum_{i=1}^{M}\sum_{x\in D_i}\left[(x-\boldsymbol{m})(x-\boldsymbol{m})^T - (x-\boldsymbol{m}_i)(x-\boldsymbol{m}_i)^T\right] \\
&= \sum_{i=1}^{M}\sum_{x\in D_i}\left[(xx^T - \boldsymbol{m}x^T - x\boldsymbol{m}^T + \boldsymbol{m}\boldsymbol{m}^T) - (xx^T - \boldsymbol{m}_i x^T - x\boldsymbol{m}_i{}^T + \boldsymbol{m}_i\boldsymbol{m}_i{}^T)\right] \\
&= \sum_{i=1}^{M}\sum_{x\in D_i}\left[(\boldsymbol{m}\boldsymbol{m}^T - \boldsymbol{m}x^T - x\boldsymbol{m}^T) - (\boldsymbol{m}_i\boldsymbol{m}_i{}^T - \boldsymbol{m}_i x^T - x\boldsymbol{m}_i{}^T)\right] \\
&= \sum_{i=1}^{M}\left[n_i\boldsymbol{m}\boldsymbol{m}^T - n_i\boldsymbol{m}_i\boldsymbol{m}_i{}^T + \sum_{x\in D_i}(\boldsymbol{m}_i x^T + x\boldsymbol{m}_i{}^T - \boldsymbol{m}x^T - x\boldsymbol{m}^T)\right] \\
&= \sum_{i=1}^{M}\left[n_i\boldsymbol{m}\boldsymbol{m}^T - n_i\boldsymbol{m}_i\boldsymbol{m}_i{}^T + \boldsymbol{m}_i\sum_{x\in D_i}x^T + \boldsymbol{m}_i{}^T\sum_{x\in D_i}x - \boldsymbol{m}\sum_{x\in D_i}x^T - \boldsymbol{m}^T\sum_{x\in D_i}x\right] \\
&= \sum_{i=1}^{M}\left[n_i\boldsymbol{m}\boldsymbol{m}^T - n_i\boldsymbol{m}_i\boldsymbol{m}_i{}^T + \boldsymbol{m}_i n_i\boldsymbol{m}_i{}^T + \boldsymbol{m}_i{}^T n_i\boldsymbol{m}_i - \boldsymbol{m}n_i\boldsymbol{m}_i{}^T - \boldsymbol{m}^T n_i\boldsymbol{m}_i\right] \\
&= \sum_{i=1}^{M} n_i\left[\boldsymbol{m}\boldsymbol{m}^T + \boldsymbol{m}_i{}^T\boldsymbol{m}_i - \boldsymbol{m}\boldsymbol{m}_i{}^T - \boldsymbol{m}^T\boldsymbol{m}_i\right] \\
&= \sum_{i=1}^{M} n_i(\boldsymbol{m}-\boldsymbol{m}_i)(\boldsymbol{m}-\boldsymbol{m}_i)^T
\end{aligned}
$$

Computing

1.

This project concerns the application of the KL procedure for incomplete data [3]. Let the complete data set be translation-invariant:

$$f(x_m, t_\mu) = \frac{1}{N}\sum_{k=1}^{N}\frac{1}{k}\sin[k(x_m - t_\mu)],$$

where $m = 1, \ldots, M$, with M dimension of the ambient space (size of the spatial grid), and $\mu = 1, \ldots, P$, with P the number of points in the ensemble.

Let $x_m = \dfrac{2\pi(m-1)}{M}$ and $t_\mu = \dfrac{2\pi(\mu-1)}{P}$.

Select an ensemble of masks $\{\boldsymbol{m}^{(\mu)}\}$, $\mu = 1, \ldots, P$. where 10% of the indices are selected to be zero for each mask. Each pattern in the incomplete ensemble may be written as

$$\tilde{\boldsymbol{x}}^{(\mu)} = \boldsymbol{m}^{(\mu)} . \boldsymbol{f}^{(\mu)},$$

where $\left(\boldsymbol{f}^{(\mu)}\right)_m = \frac{1}{N} \sum_{k=1}^{N} \frac{1}{k} \sin[k(x_m - t_\mu)]$. Let $P = M = 64$ and $N = 3$.

a. Compute the eigenvectors of this ensemble using the gappy algorithm [3].

b. Plot the eigenvalues as a function of the iteration, and continue until they converge.

c. Plot your final eigenfunctions corresponding to the 10 largest eigenvalues.

d. Plot the element $\tilde{\boldsymbol{x}}^{(1)}$ and the vector $\tilde{\boldsymbol{x}}_D$ repaired according to Equation

$$\tilde{\boldsymbol{x}} \approx \tilde{\boldsymbol{x}}_D = \sum_{n=1}^{D} \tilde{a}_n \phi^{(n)}. \tag{1}$$

Determine the value of D that provides the best approximation to the original non-gappy pattern vector.

Code

Gram-Schmidt

References

[1] Chang, Jen-Mei. *Matrix Methods for Geometric Data Analysis and Recognition*. 2014.

[2] P. N. Belhumeur, J. P. Hespanha and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711-720, Jul 1997.

[3] R. Everson and L. Sirovich. The karhunen-loeve transform for incomplete data. *J. Opt. Soc. Am., A*, 12(8):16571664, 1995.