EE141L: Lab 8
Eigenfaces

November 16, 2018

1 Introduction

In this Lab you will use eigenvectors to analyze a set of face images.

1.1 Covariance matrix: definition and intuition

Consider a set of images in column vector form $\mathbf{x}_1, \cdots, \mathbf{x}_n$, all $\mathbf{x}_i \in \mathbb{R}^m$. The average among all images is defined as

$$\bar{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i.$$

Now let us introduce the basic idea of an outer product. To keep the discussion simple, let $\mathbf{x}$ be a vector in $\mathbb{R}^2$. Then the outer product obtained from $\mathbf{x}$ is:

$$\mathbf{M} = \mathbf{x}\mathbf{x}^\top = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\begin{bmatrix} x_1 & x_2 \end{bmatrix} = \begin{bmatrix} x_1\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} & x_2\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\end{bmatrix} = \begin{bmatrix} x_1^2 & x_1x_2 \\ x_1x_2 & x_2^2 \end{bmatrix},$$

where we obtain a 2×2 matrix from the vector $\mathbf{x}$. Now let us focus on the properties of this matrix. First, observe that $\text{rank}(\mathbf{M}) = 1$. Then, note that $\mathbf{x}$ and $\mathbf{x}_0$ are eigenvectors of $\mathbf{x}\mathbf{x}^\top$, where $\mathbf{x}_0$ is a vector orthogonal to $\mathbf{x}$, i.e., $\mathbf{x}^\top\mathbf{x}_0 = 0$. This can be seen by writing for an arbitrary $\mathbf{y}$

$$\mathbf{M}\mathbf{y} = \mathbf{x}\mathbf{x}^\top\mathbf{y} = (\mathbf{x}^\top\mathbf{y})\mathbf{x},$$

where the first term $\mathbf{x}^\top\mathbf{y}$ is a scalar (inner product between $\mathbf{x}$ and $\mathbf{y}$. Then, noting that $\mathbf{x}^\top\mathbf{x} = |\mathbf{x}|^2$ and $\mathbf{x}^\top\mathbf{x}_0 = 0$, we can write that:

$$\mathbf{M}\mathbf{x} = |\mathbf{x}|^2\mathbf{x} \quad \mathbf{M}\mathbf{x}_0 = \mathbf{0} = 0\mathbf{x}_0$$

This is consistent with our observation that the matrix does not have full rank (it has a zero eigenvalue). Note also that the same idea extends to dimensions greater than $m > 2$. That is, each outer product matrix will have rank 1 (all columns are the same up to a scale factor), one eigenvector will be $\mathbf{x}$, with multiplicity one, and there will be multiple eigenvectors with eigenvalue 0, with multiplicity $m - 1$. The eigenvectors corresponding to eigenvalue 0 will be in the orthogonal complement of span($\mathbf{x}$)., which has dimension $m-1$, since span($\mathbf{x}$) has dimension 1.

Then, the covariance matrix of a set of images is defined as

$$\mathbf{C} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^\top. \tag{1}$$

Note that $\mathbf{C}$ is an average of outer products and thus it is a symmetric $m \times m$ matrix (because all $m \times m$ outer products are symmetric by construction). Moreover, the rank 1 outer product corresponding to $\mathbf{x}_i - \bar{\mathbf{x}}$ has the larger (and only non-zero) eigenvalue, with eigenvector equal to $\mathbf{x}_i - \bar{\mathbf{x}}$. The key intuition behind the covariance matrix is that because it is the average of a series of rank one matrices, its own eigendecomposition will (loosely speaking) capture an average of the eigendecompositions of each of the rank 1 terms. For example, assume that there are many vectors in our dataset that have a direction similar to $\mathbf{x}_i - \bar{\mathbf{x}}$, then the first eigenvector of C will be close to $\mathbf{x}_i - \bar{\mathbf{x}}$ as well.

Problem 1. Implement a MATLAB function that given a set of m dimensional vectors, stored in a matrix $\mathbf{X}$ of size $m \times n$, returns its covariance matrix. For simplicity, I suggest dividing the process into three steps

1. Compute $\bar{\mathbf{x}}$.
2. Remove mean from all data points, i.e. for $i = 1, \cdots, n$ compute $\mathbf{y}_i = \mathbf{x}_i - \bar{\mathbf{x}}$
3. Compute covariance matrix as

$$\frac{1}{n}\sum_{i=1}^{n}\mathbf{y}_i\mathbf{y}_i^\top. \tag{2}$$

1.2 Eigenvectors

Since $\mathbf{C}$ is symmetric, then it has a set of orthonormal eigenvectors, $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m$, that satisfy

$$\mathbf{u}_i^\top \mathbf{u}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

and a set of eigenvalues $\lambda_1 \geq \lambda_2, \cdots, \geq \lambda_m$ that satisfy $\mathbf{C}\mathbf{u}_i = \lambda_i \mathbf{u}_i$.

2 Computing eigenfaces

2.1 Dataset

We will use the Yale Face Database [1], which consistes of 165 images of size 64×64. There are 15 subjects, and for each subject 11 images with different expressions and accessories.

[1] P. Belhumeur, J. Hespanha, D. Kriegman, Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection, IEEE Transactions on Pattern Analysis and Machine Intelligence, July 1997, pp. 711-720.

2.2 Eigenfaces for all images

Problem 2. After loading the Yale dataset, each column of the matrix faces contains a vectorized image, and the vector subject contains numbers from 1 to 15, that correspond to subject labels, meaning faces(:,i) is an image of subject subject(i)

1. Complete the code display_all_subjects.m to generate a $64 \times 64 * 15$ matrix that contains one image per subject. Put this image in your report.

2. Compute the covariance matrix of all faces (columns of faces). Compute the 10 eigenvectors of the covariance matrix associated with the 10 largest eigenvalues, for this use the matlab function eigs, DO NOT compute all eigenvectors, see Matlab documentation on how to use function eigs.

3. Each eigenvector is a 64^2 dimensional vector. Reshape them into 64×64 matrices. These images are called **Eigenfaces**. Put all 10 images in a 64×640 image, and put that in your report. To create this image you can modify the code from Problem 2 part 1.

4. Repeat parts 2 and 3, but now use the eigenvectors with the next 10 largest eigenvalues, i.e. $\lambda_{11}, \lambda_{12}, \cdots, \lambda_{20}$. Plot the largest 20 eigenvalues, and discuss what happens with the eigenfaces as the eigenvalues decrease.

2.3 Eigenfaces for pairs of subjects

Pick a random pair of subjects, denote them by subject a and subject b. There are 210 possible combinations, so please be creative and do not pick $a = 1, b = 2$.

Problem 3.

1. Compute covariance matrices for the images of each subject, call them $\mathbf{C}_a$ and $\mathbf{C}_b$.

2. For each subject, compute eigenvectors corresponding to largest 5 eigenvalues, and display corresponding eigenfaces (all in one image as in problem 2).

3. Compute the covariance matrix for images of both subjects together, call it $\mathbf{C}_{ab}$. Compute its eigenvectors corresponding to its 5 largest eigenvalues, and display corresponding eigenfaces.

4. Discuss differences between eigenfaces of subject a, eigenfaces of subject b, and eigenfaces of both subjects together.