1 Deriving the second principal component

1. Let

\[ J(v_2, z_2) = \frac{1}{n} \sum_{i=1}^{n} (x_i - z_1 v_1 - z_2 v_2)^T (x_i - z_1 v_1 - z_2 v_2) \]  

Show that \( \frac{\partial J}{\partial z_2} = 0 \) yields \( z_2 = v_2^T x_i \).

2. Show that the value of \( v_2 \) that minimizes

\[ \tilde{J}(v_2) = -v_2^T C v_2 + \lambda_2 (v_2^T v_2 - 1) + \lambda_{12} (v_2^T v_1 - 0) \]

is given by the eigenvector of \( C \) with the second largest eigenvalue. Hint: recall that \( Cv_1 = \lambda_1 v_1 \) and

\[ \frac{\partial x^T A x}{\partial x} = (A + A^T)x \]

2 Deriving the residual error for PCA

1. Prove that

\[ ||x_i - \sum_{j=1}^{K} z_{ij} v_j||^2 = x_i^T x_i - \sum_{j=1}^{K} v_j^T x_i x_i^T v_j \]

Hint: first consider the case \( K = 2 \). Use the fact that \( v_j^T v_j = 1 \) and \( v_j^T v_k = 0 \) for \( k \neq j \). Also, recall \( z_{ij} = x_i^T v_j \).

2. Now show that

\[ J_K = \frac{1}{n} \sum_{i=1}^{n} \left( x_i^T x_i - \sum_{j=1}^{K} v_j^T x_i x_i^T v_j \right) = \frac{1}{n} \sum_{i=1}^{n} x_i^T x_i - \sum_{j=1}^{K} \lambda_j \]

Hint: recall \( v_j^T C v_j = \lambda_j v_j^T v_j = \lambda_j \).

3. If \( K = d \) there is no truncation, so \( J_d = 0 \). Use this to show that the error from only using \( K < d \) terms is given by

\[ J_K = \sum_{j=K+1}^{d} \lambda_j \]

Hint: partition the sum \( \sum_{j=1}^{d} \lambda_j \) into \( \sum_{j=1}^{K} \lambda_j \) and \( \sum_{j=K+1}^{d} \lambda_j \).
3 Derivation of Fisher’s linear discriminant

Show that the maximum of

\[ J(w) = \frac{w^T S_B w}{w^T S_W w} \]  

is given by

\[ S_B w = \lambda S_W w \]  

where

\[ \lambda = \frac{w^T S_B w}{w^T S_W w} \]  

Hint: recall that the derivative of a ratio of two scalars is given by

\[ \frac{d}{dx} \frac{f(x)}{g(x)} = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2} \]  

where \( f' = \frac{d}{dx} f(x) \) and \( g' = \frac{d}{dx} g(x) \). Also, recall that

\[ \frac{d}{dx} x^T A x = (A + A^T)x \]  

4 PCA and nearest neighbors for face recognition (Matlab)

Download the files olivettifaces.mat and Code.zip from the MLABA web page. Fill in the parts of faceRecStub.m marked with question marks, and thus reproduce Figure 1 and Figure 2. Turn in your code and plots.
Figure 2: Misclassification rate vs number of PCA dimensions.