UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
Department of Electrical and Computer Engineering

ECE 417 Principles of Signal Analysis
Spring 2014

EXAM 1

Tuesday, February 25, 2014

- This is a CLOSED BOOK exam.
- There are a total of 100 points in the exam. Each problem specifies its point total. Plan your work accordingly.
- You must SHOW YOUR WORK to get full credit.

Problem	Score
1	
2	
3	
4	
5	
6	
Total	

Name: SOLUTION

Problem 1 (16 points)

A particular dataset has the scatter matrix $S = \sum_{k=1}^{n}(\vec{x}_k - \vec{m})(\vec{x}_k - \vec{m})^T$, whose first two eigenvectors are $\vec{v}_1$ and $\vec{v}_2$, characterized by eigenvalues $\lambda_1 = 450$ and $\lambda_2 = 150$. Define the transform $\vec{y}_k = [\vec{v}_1, \vec{v}_2]^T(\vec{x}_k - \vec{m})$. Define the 2×2 matrix

$$Q = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} = \sum_{k=1}^{n} \vec{y}_k \vec{y}_k^T$$

Find the numerical values of the elements q_{11}, q_{12}, q_{21}, and q_{22} of matrix Q.

$$Q = \sum_{k=1}^{n} V^T(\vec{x}_k - \vec{m})(\vec{x}_k - \vec{m})^T V$$

$$= V^T S V = \Lambda = \begin{bmatrix} 450 & 0 \\ 0 & 150 \end{bmatrix}$$

Problem 2 (16 points)

A particular dataset has six data vectors, given by

$$\{\vec{x}_1,\ldots,\vec{x}_6\} = \left\{ \begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}-1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\-1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}, \begin{bmatrix}0\\0\\-1\end{bmatrix} \right\}$$

By calling `randn` in matlab, you generate a 3×2 random projection matrix V, given by

$$V = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \\ v_{31} & v_{32} \end{bmatrix}$$

Using this random projection matrix, you compute the transformed feature vectors $\vec{y}_k = V^T\vec{x}_k$. The total energy of the transformed dataset can be written as

$$E = \sum_{k=1}^{6} \vec{y}_k^T \vec{y}_k$$

Find the value of E in terms of the random projection matrix elements v_{ij}.

$$Y = V^T \underline{X} = \begin{bmatrix} v_{11} & -v_{11} & v_{21} & -v_{21} & v_{31} & -v_{31} \\ v_{12} & -v_{12} & v_{22} & -v_{22} & v_{32} & -v_{32} \end{bmatrix}$$

$$\boxed{\begin{aligned} E &= \sum_{k=1}^{6} \|\vec{y}_k\|^2 = 2\sum_{i=1}^{3}\sum_{j=1}^{2} v_{ij}^2 \\ &= 2\left(v_{11}^2 + v_{12}^2 + v_{21}^2 + v_{22}^2 + v_{31}^2 + v_{32}^2\right) \end{aligned}}$$

Problem 3 (16 points)

A 200×200 sunset image is bright on the bottom, and dark on top, thus the pixel in the i^{th} row and j^{th} column has intensity $A[i,j] = 200 - i$. Suppose that this formula extends past the edge of the image, and also extends to non-integer pixel locations, so you can assume that it holds for every real-valued coordinate value (i,j).

Suppose your goal is to recognize sunset images that have been taken by people laying upside-down on the beach. You decide to train the classifier by rotating the image $A[i,j]$ to every possible angle, thus creating the training images

$$B_k[i,j] = A[i\cos\theta_k - j\sin\theta_k, i\sin\theta_k + j\cos\theta_k], \quad \theta_k = \frac{2\pi k}{100}, \quad 0 \le k \le 99$$

Your next step is to reshape each 200×200 image $B_k[i,j]$ into a vector of raw pixel intensities, $\vec{x}_k$, then to compute the dataset mean, $\vec{m} = \frac{1}{100}\sum_{k=0}^{99}\vec{x}_k$.

(a) What is the length of the vector $\vec{m}$?

$$\text{length}(\vec{m}) = \text{length}(\vec{x}_k) = 200 \times 200 = \boxed{40{,}000}$$

(b) What is the numerical value of $\vec{m}$? Provide enough information to specify the value of every element of the vector.

$$M[i,j] = \frac{1}{100}\sum_{k=0}^{99} B_k[i,j]$$

$$= \frac{1}{100}\sum_{k=0}^{99} A\left[i\cos\frac{2\pi k}{100} - j\sin\frac{2\pi k}{100}, i\sin\theta_k + j\cos\theta_k\right]$$

$$= \frac{1}{100}\sum_{k=0}^{99}\left(200 - i\cos\left(\frac{2\pi k}{100}\right) + j\sin\left(\frac{2\pi k}{100}\right)\right)$$

$$= 200 \quad \text{BECAUSE} \quad \sum_{k=0}^{99}\cos\left(\frac{2\pi k}{100}\right) = 0$$

$$\sum_{k=0}^{99}\sin\left(\frac{2\pi k}{100}\right) = 0$$

$$\boxed{\vec{m} = \begin{bmatrix} 200 \\ 200 \\ \vdots \\ \end{bmatrix} \Bigg\} \text{40,000 DIMENSIONS}}$$

Problem 4 (16 points)

Suppose you have a 1000-sample audio waveform, $x[n]$, such that $x[n] \neq 0$ for $0 \leq n \leq 999$. You want to chop this waveform into 200-sample frames, with 10% overlap between frames. How many nonzero samples are there in the last frame?

(10% overlap) × (200 samples) = 20 sample overlap

200 − 20 = 180 new samples / frame

$(N-1) \times 180 \geq 1000 - 200$

$N - 1 \geq \frac{800}{180}$

$N - 1 = 5$

$N = 6$

200 + 5 · 180 = 1100 . SAMPLES TOTAL

NONZERO IN LAST FRAME = 200 − 100 = $\boxed{100}$

Problem 5 (16 points)

An audio signal has the following 256-point DFT:

$$X[k] = \begin{cases} 1 & k = 16, 240 \\ 0.1 & \text{otherwise} \end{cases}$$

Find its cepstrum.

$$\log|X[k]| = \begin{cases} 0 & k=16, 240 \\ \log(0.1) & \text{else} \end{cases}$$

$$= \log(0.1) - \log(0.1)\,\delta[k-16] - \log(0.1)\,\delta[k-240]$$

$$c[n] = \frac{1}{256}\sum_{k=0}^{255} \log|X[k]|\, e^{j2\pi kn/256}$$

$$= \log(0.1)\,\delta[n] - \frac{1}{256}\log(0.1)\, e^{j\pi n/8} - \frac{1}{256}\log(0.1)\, e^{-j\pi n/8}$$

$$\boxed{= \log(0.1)\,\delta[n] - \frac{1}{128}\log(0.1)\cos\left(\frac{\pi n}{8}\right)}$$

$$\boxed{= -\log(10)\,\delta[n] + \frac{\log(10)}{128}\cos\left(\frac{\pi n}{8}\right)}$$

Problem 6 (20 points)

Suppose you have a database with an infinite number of class 2 samples, but only two samples from class 1. Thus the training class labels, c_k, are given by

$$\{c_1, c_2, c_3, \ldots, c_k, \ldots\} = \{1, 1, 2, \ldots, 2, \ldots\} \tag{1}$$

The corresponding training vectors are given by

$$\{\vec{x}_1, \vec{x}_2, \vec{x}_3, \ldots, \vec{x}_k, \ldots\} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \alpha_3 \end{bmatrix}, \ldots \begin{bmatrix} 0 \\ \alpha_k \end{bmatrix}, \ldots \right\} \tag{2}$$

where the numbers α_k cover all of the rational numbers, $-\infty < \alpha_k < \infty$, thus for most practical purposes, the samples from class 2 cover the $x_1 = 0$ axis.

(a) A particular test token is given by $\vec{x} = [a, b]^T$, and is classified by a nearest-neighbor (NN) classifier using the training database specified in Eqs. 1 and 2. What values of $[a, b]^T$ are classified into class 2 by the NN classifier?

$$\min_{\alpha} \left\| \begin{bmatrix} a \\ b \end{bmatrix} - \begin{bmatrix} 0 \\ \alpha \end{bmatrix} \right\|^2 = a^2$$

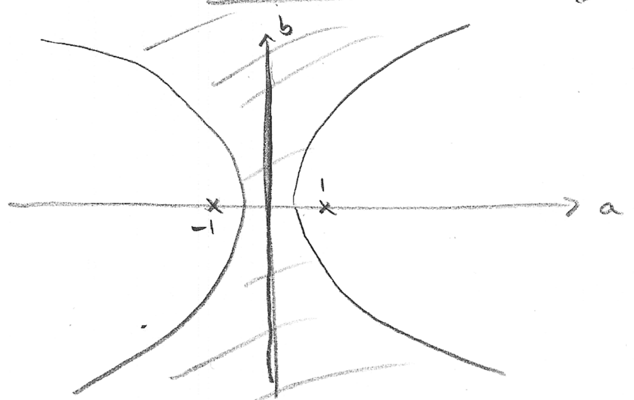
$$\min_{\pm 1} \left\| \begin{bmatrix} a \\ b \end{bmatrix} - \begin{bmatrix} \pm 1 \\ 0 \end{bmatrix} \right\|^2 = \min\left((a-1)^2 + b^2,\ (a+1)^2 + b^2\right)$$

CLASS 2 IFF $\cancel{a^2} < \min\left(\cancel{a^2} - 2a + 1 + b^2,\ \cancel{a^2} + 2a + 1 + b^2\right)$

$$0 < \min\left(1 + b^2 - 2a,\ 1 + b^2 + 2a\right)$$

$$\max(2a, -2a) < 1 + b^2$$

$$\boxed{1 + b^2 > 2|a|} \qquad \boxed{-\left(\frac{1+b^2}{2}\right) < a < \left(\frac{1+b^2}{2}\right)}$$



(b) Same question as part (a), but now using a KNN classifier with $K = 3$. For what values of $\vec{x} = [a, b]^T$ will KNN classify $\vec{x}$ into class 2?

KNN will $\boxed{\text{always}}$ choose class 2.
Of the 3 nearest neighbors, at least 2 are on the b-axis.