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

ECE 417 Multimedia Signal Processing
Fall 2017

EXAM 2

Thursday, October 26, 2017

- This is a CLOSED BOOK exam. You may use one sheet (front and back) of hand-written notes.
- No calculators are permitted. You need not simplify explicit numerical expressions.
- There are a total of 40 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	
Total	

Name: ____

Possibly Useful Formulas

Fourier Transforms

$$x[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} X(\omega)e^{j\omega n}d\omega \leftrightarrow X(\omega) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n}$$

$$x[n] = \frac{1}{N}\sum_{k=0}^{N-1} X[k]e^{j2\pi kn/N} \leftrightarrow X[k] = \sum_{n=0}^{N-1} x[n]e^{-j2\pi kn/N}$$

$$h[n] = \frac{\sin\omega_c n}{\pi n} \leftrightarrow H(\omega) = \begin{cases} 1 & |\omega| < \omega_c \\ 0 & \text{otherwise} \end{cases}$$

$$u[n] - u[n-N] \leftrightarrow e^{-j\frac{\omega(N-1)}{2}}\frac{\sin(\omega N/2)}{\sin(\omega/2)}$$

$$\delta[n] \leftrightarrow 1$$

$$e^{j\alpha n} \leftrightarrow 2\pi\delta(\omega-\alpha)$$

$$\sum_{\ell=-\infty}^{\infty}\delta[n-\ell T_0] \leftrightarrow \left(\frac{2\pi}{T_0}\right)\sum_{k=1}^{T_0-1}\delta\left(\omega - \frac{2\pi k}{T_0}\right)$$

Gaussians, Mahalanobis, and PCA

$$\mathcal{N}\left(\vec{x};\vec{\mu},\Sigma\right) = \frac{1}{(2\pi)^{D/2}|\Sigma|^{1/2}}e^{-\frac{1}{2}(\vec{x}-\vec{\mu})^T\Sigma^{-1}(\vec{x}-\vec{\mu})}$$

$$\begin{aligned} \Sigma &= U\Lambda U^T \\ \Sigma^{-1} &= U\Lambda^{-1}U^T \\ U^T\Sigma U &= \Lambda \\ U^TU &= I \end{aligned}$$

$$d_\Sigma^2(\vec{x},\vec{\mu}) = (\vec{x}-\vec{\mu})^T\Sigma^{-1}(\vec{x}-\vec{\mu}) = \vec{y}^T\Lambda^{-1}\vec{y}$$

$$\vec{y} = U^T(\vec{x}-\vec{\mu})$$

Problem 1 (10 points)

A speech signal can be modeled as an excitation passed through a filter, $s[n] = h[n] * e[n]$. A reasonable (very) simplified model of voiced speech might use

$$e[n] = \delta[n] + \sum_{m=-\infty}^{\infty} \delta[n - mT_0], \quad h[n] = \left(A_1 p_1^n + A_1^*(p_1^*)^n\right) u[n]$$

where p_1 is the complex first formant frequency, and A_1 is some appropriate constant. For purposes of this problem, define the cepstrum to be the inverse DTFT of the log DTFT:

$$\hat{s}[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln S(\omega) e^{j\omega n} d\omega$$

You are given the cepstrum of $h[n]$:

$$\hat{h}[n] = \left(\frac{p_1^n}{n} + \frac{(p_1^*)^n}{n}\right) u[n]$$

and you may assume that

$$\ln\left(1 + \frac{2\pi}{T_0} \sum_{k=1}^{T_0-1} \delta\left(\omega - \frac{2\pi k}{T_0}\right)\right) \approx \ln\left(\frac{2\pi}{T_0}\right) + \sum_{k=1}^{T_0-1} \delta\left(\omega - \frac{2\pi k}{T_0}\right)$$

Under these assumptions, find the cepstrum $\hat{s}[n]$ of the speech signal, in terms of the parameters T_0 and p_1.

Problem 2 (20 points)

A speech signal, $x[n]$, has been sampled at F_s samples/second. Suppose the signal is only N samples long, so we can compute its N-sample DFT, $X[k]$. The mel spectrum is then defined to be

$$\tilde{X}[m] = \begin{cases} \sum_{k=0}^{N/2} W_m[k] \cdot |X[k]| & 1 \leq m \leq M \\ 0 & \text{otherwise} \end{cases}$$

where the filters $W_m[k]$ are uniformly spaced in a mel-frequency scale. Suppose that the speech signal is known to be $x[n] = h[n] * e[n]$, where $e[n]$ has mel spectrum $\tilde{E}[m]$, and $h[n]$ has mel spectrum $\tilde{H}[m]$. Define the mel-frequency cepstral coefficients to be

$$\tilde{x}[n] = \text{DFT}^{-1}\left\{\ln\left(\tilde{X}[m] + \tilde{X}[2M+2-m]\right)\right\}$$

where the DFT has a length of $N = 2M + 2$. Prove that $\tilde{x}[n]$ is a symmetric real-valued sequence.

Problem 3 (10 points)

Three scalar random variables A, V, and Y are jointly distributed as

$$p_{Y|V}(y|v) = \frac{1}{3}C\left(y_k = y\right), \quad p_{A|Y}(a|y) = \mathcal{N}(a;\mu_y,1) \tag{1}$$

where $\mathcal{N}(a;\mu_y,1)$ is a scalar Gaussian with unit variance, and with class-dependent means

$$\mu_0 = -1, \quad \mu_1 = 1 \tag{2}$$

The operator $C(y_k = y)$ is a nearest-neighbor count operator: it finds three training samples v_k with the three smallest values of $(v_k - v)^2$, then counts how many of those samples have the label y. The training samples are

$$\{v_1,\ldots,v_9\} = [-2,-1,0,1,2,3], \quad \{y_1,\ldots,y_9\} = [0,0,0,1,1,1] \tag{3}$$

(a) Sketch $p_{Y|V}(1|v)$ as a function of v.

(b) Based on Eqs. 1 through 3 on the previous page, design a classifier $\hat{y}(a,v)$ such that

$$\hat{y}(a,v) = \arg\max_y \ln\left(p_{A|Y}(a|y)p_{Y|V}(y|v)\right)$$

Draw a two-dimensional space whose axes are v and a; show the region $-5 \le v, a \le 5$. In the two-dimensional space, draw the decision boundary between the class $\hat{y}(a,v) = 0$ and the class $\hat{y}(a,v) = 1$.

Problem 4 (10 points)

Suppose you have a scalar random variable X, with training examples $[x_1, x_2, x_3, x_4] = [-1, 0, 1, 2]$. You want to try to cluster these into two clusters. You have initial estimates of the cluster centroids, as $\mu_0 = 0$, $\mu_1 = 1$.

(a) Perform one iteration of K-means clustering: assign each token to a cluster, then recompute the new cluster centroids. What are the new cluster centroids?

(b) Assume that part (a) never happened. Instead, perform one iteration of EM training. You have initial parameter estimates $\mu_0 = 0$, $\mu_1 = 1$, $\sigma_0^2 = \sigma_1^2 = 1$, and $c_0 = c_1 = 0.5$. Perform one iteration of EM training. What are the new cluster centroids? Give numerical values for the new values of μ_0 and μ_1; in order to do so, assume that the unit normal Gaussian pdf has the following values: $\mathcal{N}(0; 0, 1) \approx \frac{2}{5}$, $\mathcal{N}(1; 0, 1) \approx \frac{1}{4}$, $\mathcal{N}(2; 0, 1) \approx \frac{1}{20}$, and $\mathcal{N}(x; 0, 1) \approx 0$ for $x \geq 3$. Note: since you didn't bring a calculator to this exam, feel free to leave your answer in the form of a sum of fractions.