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

# ECE 417 Multimedia Signal Processing
Spring 2021

## EXAM 2

Tuesday, November 2, 2021

- This is a CLOSED BOOK exam.
- You are permitted one sheet of handwritten notes, 8.5x11.
- Calculators and computers are not permitted.
- If you're taking the exam online, you will need to have your webcam turned on. The exam will be posted to zoom at exactly 9:30am; you will need to photograph and upload your answers to Gradescope by exactly 11:00am.
- 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.
- The second page is a formula sheet.
- Do not simplify explicit numerical expressions. The expression "$e^{-5}\cos(3)$" is a MUCH better answer than "-0.00667".

Name: ________________________________________________

**Gaussians and GMMs**

$$p_{\vec{X}}(\vec{x}) = \frac{1}{(2\pi)^{D/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(\vec{x}-\vec{\mu})^T\Sigma^{-1}(\vec{x}-\vec{\mu})}$$

$$p_{\vec{X}}(\vec{x}) = \sum_{k=0}^{K-1} c_k \mathcal{N}(\vec{x}|\vec{\mu}_k, \Sigma_k)$$

**Principal Component Analysis**

$$(n-1)\Sigma = V\Lambda V^T, \quad \frac{1}{n-1}\Lambda = V^T\Sigma V, \quad V^TV = VV^T = I$$

$$\sum_{d=1}^{D}\sigma_d^2 = \frac{1}{n-1}\text{trace}\left(X^TX\right) = \frac{1}{n-1}\text{trace}\left(Y^TY\right) = \frac{1}{n-1}\sum_{d=1}^{D}\lambda_d$$

$$\Lambda = V^T(X^TX)V = U^T(XX^T)U$$

**Expectation Maximization**

$$Q(\Theta, \hat{\Theta}) = E\left[\ln p(\mathcal{D}_v, \mathcal{D}_h|\Theta)\,\middle|\,\mathcal{D}_v, \hat{\Theta}\right]$$

**Hidden Markov Model**

$$\alpha_t(j) = \sum_{i=1}^{N}\alpha_{t-1}(i)a_{ij}b_j(\vec{x}_t), \quad 1 \le j \le N,\ 2 \le t \le T$$

$$\beta_t(i) = \sum_{j=1}^{N} a_{ij}b_j(\vec{x}_{t+1})\beta_{t+1}(j), \quad 1 \le i \le N,\ 1 \le t \le T-1$$

$$\tilde{\alpha}_t(j) = \sum_{i=1}^{N}\hat{\alpha}_{t-1}(i)a_{ij}b_j(\vec{x}_t)$$

$$g_t = \sum_{j=1}^{N}\tilde{\alpha}_t(j)$$

$$\hat{\alpha}_t(j) = \frac{1}{g_t}\tilde{\alpha}_t(j)$$

$$\gamma_t(i) = \frac{\alpha_t(i)\beta_t(i)}{\sum_{k=1}^{N}\alpha_t(k)\beta_t(k)}$$

$$\xi_t(i,j) = \frac{\alpha_t(i)a_{ij}b_j(\vec{x}_{t+1})\beta_{t+1}(j)}{\sum_{k=1}^{N}\sum_{\ell=1}^{N}\alpha_t(k)a_{k\ell}b_\ell(\vec{x}_{t+1})\beta_{t+1}(\ell)}$$

$$a'_{ij} = \frac{\sum_{t=1}^{T-1}\xi_t(i,j)}{\sum_{j=1}^{N}\sum_{t=1}^{T-1}\xi_t(i,j)}$$

$$\Sigma'_i = \frac{\sum_{t=1}^{T}\gamma_t(i)(\vec{x}_t-\vec{\mu}_i)(\vec{x}_t-\vec{\mu}_i)^T}{\sum_{t=1}^{T}\gamma_t(i)}$$

$$\vec{\mu}'_i = \frac{\sum_{t=1}^{T}\gamma_t(i)\vec{x}_t}{\sum_{t=1}^{T}\gamma_t(i)}$$

1. (15 points) Suppose $\vec{X} = [X_1, X_2]^T$ is a Gaussian random variable with mean and covariance matrix given by

$$\vec{\mu} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$$

Sketch the set of points such that $p_{\vec{X}}(\vec{x}) = \frac{1}{4\pi}e^{-\frac{1}{2}}$, where $p_{\vec{X}}(\vec{x})$ is the pdf of $\vec{X}$. Clearly label at least four of the points included in this set.

2. (20 points) The binary random variable $Y$ has the following prior distribution:

$$p_Y(0) = a, \quad p_Y(1) = 1 - a$$

The random vector $\vec{X}$ depends on $Y$, with the conditionally Gaussian pdf $p_{\vec{X}|Y}(\vec{x}|y) = \mathcal{N}(\vec{x}; \vec{\mu}_y, \Sigma_y)$, where the mean vectors and covariance matrices are given by

$$\vec{\mu}_0 = \begin{bmatrix} b \\ c \end{bmatrix}, \quad \vec{\mu}_1 = \begin{bmatrix} d \\ e \end{bmatrix}, \quad \Sigma_0 = \Sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Given a sample measurement of $\vec{X} = \vec{x}$, it's possible to infer the value of $Y = \hat{y}$ with minimum probability of error using the following decision rule:

$$\hat{y} = \begin{cases} 1 & \text{if } \vec{w}^T \vec{x} + \beta > 0 \\ 0 & \text{otherwise} \end{cases}$$

Find $\vec{w}$ and $\beta$ in terms of the constants $a, b, c, d, e$.

3. (30 points) A particular unlabeled dataset, $\mathcal{D} = \{\vec{x}_1, \ldots, \vec{x}_n\}$ has been centered so that the sample mean is $\vec{\mu} = [0, 0]^T$. The sample covariance matrix, $\Sigma$ has the following value, and the following eigenvalue decomposition:

$$\Sigma = \left[\begin{array}{cc} 4 & -2 \\ -2 & 5 \end{array}\right] = \left[\begin{array}{cc} -0.79 & 0.62 \\ -0.62 & -0.79 \end{array}\right] \left[\begin{array}{cc} 2.44 & 0 \\ 0 & 6.56 \end{array}\right] \left[\begin{array}{cc} -0.79 & -0.62 \\ 0.62 & -0.79 \end{array}\right]$$

Suppose you want to find a unit-length vector $\vec{v}$ that makes the quantity $\mathcal{J}$, defined in the following equation, as large as possible:

$$\vec{v} = \arg\max \mathcal{J} \quad s.t. \ \|\vec{v}\| = 1 \ \text{ and } \ \mathcal{J} = \frac{\sum_{i=1}^{n} (\vec{v}^T \vec{x}_i)^2}{\sum_{i=1}^{n} \|\vec{x}_i\|^2}$$

(a) What is the numerical value of $\vec{v}$ that maximizes $\mathcal{J}$? You may leave your answer as an explicit function of numerical quantities, if you wish.

(b) What is the maximum achievable numerical value of $\mathcal{J}$? You may leave your answer as an explicit function of numerical quantities, if you wish.

(c) The gram matrix is an $n \times n$ matrix, $G$, whose $(i, j)^{\text{th}}$ element is

$$G[i, j] = \vec{x}_i^T \vec{x}_j$$

In terms of $n$, what are the eigenvalues of $G$?

4. (20 points) An HMM has the parameters $\Lambda = \{\pi_i, a_{i,j}, b_j(\vec{x}_t) : 1 \le i,j \le N, 1 \le t \le T\}$, with the standard definitions:

$$\pi_i = p(q_1 = i)$$
$$a_{i,j} = p(q_t = j | q_{t-1} = i)$$
$$b_j(\vec{x}_t) = p(\vec{x} = \vec{x}_t | q_t = j),$$

where $q_t$ is the state index at time $t$. Suppose you have software available that will compute the forward, backward, scaled forward, and/or scaled backward algorithm for you, and will therefore provide you with any or all of the following quantities, for any values of $1 \le i,j \le N$ and $1 \le t \le T$:

$$\alpha_t(i) = p(\vec{x}_1, \ldots, \vec{x}_t, q_t = i | \Lambda)$$
$$\beta_t(i) = p(\vec{x}_{t+1}, \ldots, \vec{x}_T | q_t = i, \Lambda)$$
$$\hat{\alpha}_t(i) = p(q_t = i | \vec{x}_1, \ldots, \vec{x}_t, \Lambda)$$
$$g_t = p(\vec{x}_t | \vec{x}_1, \ldots, \vec{x}_{t-1}, \Lambda)$$
$$\hat{\beta}_t(i) = \beta_t(i) / \max_j \beta_t(j)$$

In terms of $\pi_i, a_{ij}, b_j(\vec{x}_t), \alpha_t(i), \beta_t(i), \hat{\alpha}_t(i), g_t$, and/or $\hat{\beta}_t(i)$, find a formula for the following quantity, assuming that $T$ is much larger than 19:

$$p(q_{16} = 4, q_{17} = 5 | \vec{x}_1, \ldots, \vec{x}_{18}, \Lambda)$$

5. (15 points) A second-order HMM is like a standard HMM, except that the state at each time step depends on the two preceding states. The parameters are $\Lambda = \{\pi_{i,j}, a_{i,j,k}, b_k(\vec{x}_t) : 1 \leq i,j,k \leq N, 1 \leq t \leq T\}$, with the definitions:

$$\begin{aligned} \pi_{i,j} &= p(q_1 = i, q_2 = j) \\ a_{i,j,k} &= p(q_t = k | q_{t-2} = i, q_{t-1} - j) \\ b_k(\vec{x}_t) &= p(\vec{x} = \vec{x}_t | q_t = k), \end{aligned}$$

where $q_t$ is the state index at time $t$. Suppose you have software available that will compute the forward and backward algorithms for you, and will therefore provide you with the following quantities, for any values of $1 \leq i,j \leq N$ and $2 \leq t \leq T$:

$$\begin{aligned} \alpha_t(i,j) &= p(\vec{x}_1, \ldots, \vec{x}_t, q_{t-1} = i, q_t = j | \Lambda) \\ \beta_t(i,j) &= p(\vec{x}_{t+1}, \ldots, \vec{x}_T | q_{t-1} = i, q_t = j, \Lambda) \end{aligned}$$

In terms of $\pi_{i,j}, a_{i,j,k}, b_k(\vec{x}_t), \alpha_t(i,j)$, and/or $\beta_t(i,j)$, find the following expected value:

$$\mathbb{E}\left[\#\text{ times, } t, \text{ for which } q_{t-2} = i, q_{t-1} = j, q_t = k | \vec{x}_1, \ldots, \vec{x}_T, \Lambda\right]$$