ECE 537 Practice Exam 1

UNIVERSITY OF ILLINOIS
Department of Electrical and Computer Engineering

The real exam will be October 24, 2022 in class

- This is a closed-book exam.
- You are allowed to bring one 8.5x11 sheet of handwritten notes (front and back).
- No calculators are allowed. Please do not simplify explicit numerical expressions.
- There are 100 points in the exam. Points for each problem are specified by the problem number.

Name: _____

NetID: _____

Possibly Useful Charts and Formulas

Dynamic Time Warping

$$A_{i,k} = \max\left(A_{i+1,k}, A_{i,k+1}, a_{i,k} + A_{i+1,k+1}\right)$$

Linear Prediction

$$s[n] = Ge[n] + \sum_{m=1}^{p} a_m s[n-m] = h[n] * x[n]$$

$$H(z) = \frac{G}{1 - \sum_{m=1}^{N} a_m z^{-m}} = \frac{G}{\prod_{k=1}^{N}\left(1 - p_k z^{-1}\right)}$$

$$\mathcal{E} = \sum_{n=0}^{N-1} e^2[n] = \sum_{n=0}^{N-1}\left(s[n] - \sum_{m=1}^{p} a_m s[n-m]\right)^2$$

$$0 = \sum_{n=0}^{N-1}\left(s[n] - \sum_{m=1}^{p} a_m s[n-m]\right) s[n-k], \quad 1 \le k \le p$$

$$\vec{c} = \Phi\vec{a}$$

Hidden Markov Models

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

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

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

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

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

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

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

$$\bar{a}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{j=1}^{N}\sum_{t=1}^{T-1} \xi_t(i,j)}$$

$$\bar{U}_i = \frac{\sum_{t=1}^{T} \gamma_t(i)(\vec{o}_t - \vec{\mu}_i)(\vec{x}_t - \vec{\mu}_i)^T}{\sum_{t=1}^{T} \gamma_t(i)}$$

$$\bar{\mu}_i = \frac{\sum_{t=1}^{T} \gamma_t(i)\vec{o}_t}{\sum_{t=1}^{T} \gamma_t(i)}$$

1. (20 points) In (Velichko & Zagoruyko, 1970), the best time alignment was defined to be the one that maximized the sum of similarities. Consider, instead, an alignment that minimizes the sum of distances. Consider the problem of finding a sequence of alignments $I = [i(1), \ldots, i(T)]$ and $J = [j(1), \ldots, j(T)]$ where I, J, and T are chosen to maximize

$$D = \max_{i(t),j(t)\forall 1 \leq t} d\left(i(t-1), j(t-1), i(t), j(t)\right)$$

where

$$d\left(i(t-1), j(t-1), i(t), j(t)\right) = \begin{cases} C & i(t) = i(t-1), j(t) = j(t-1) + 1 \\ C & i(t) = i(t-1) + 1, j(t) = j(t-1) \\ d_{i(t),j(t)} & i(t) = i(t-1) + 1, j(t) = j(t-1) + 1 \\ \infty & \text{otherwise} \end{cases}$$

In other words, the cost of any horizontal or vertical move in the alignment matrix is a constant (C), the cost of a diagonal move is dependent on $i(t)$ and $j(t)$, and no other moves are possible. Suppose that the lengths of the two utterances are the same, i.e., $i(T) = j(T) = N$. What is the smallest value of C that is guaranteed to find an alignment with N diagonal moves?

Solution: Any diagonal move can be avoided by taking one horizontal move and one vertical move. The smallest value of C sufficient to guarantee that there are no vertical or horizontal moves is

$$C \geq \frac{1}{2} \max_{1 \leq i,j \leq N} d_{i,j}$$

2. (20 points) Suppose we have the following model of pitch prediction:

$$e[n] = d[n] - \beta_1 d[n-P+1] - \beta_2 d[n-P] - \beta_3 [n-P-1],$$

and we wish to find the values of $\vec{\beta} = [\beta_1, \beta_2, \beta_3]^T$ that minimizes

$$\mathcal{E} = \sum_{n=0}^{N-1} (e[n])^2$$

What value of the vector $\vec{\beta}$ minimizes $\mathcal{E}$? If you write your answer in terms of any other vectors, matrices, or covariance functions, be sure to define them.

Solution: By the principle of orthogonality, $\mathcal{E}$ is minimized if and only if

$$\sum_{n=0}^{N-1} e[n]d[n-P+1] = 0$$
$$\sum_{n=0}^{N-1} e[n]d[n-P] = 0$$
$$\sum_{n=0}^{N-1} e[n]d[n-P-1] = 0$$

Substituting in the definition of $e[n]$, we find that $\mathcal{E}$ is minimized by

$$\vec{\beta} = \Phi^{-1}\vec{c},$$

where

$$\Phi = \begin{bmatrix} \phi(P-1,P-1) & \phi(P-1,P) & \phi(P-1,P+1) \\ \phi(P,P-1) & \phi(P,P) & \phi(P,P+1) \\ \phi(P+1,P-1) & \phi(P+1,P) & \phi(P+1,P+1) \end{bmatrix}, \quad \vec{c} = \begin{bmatrix} \phi(P-1,0) \\ \phi(P,0) \\ \phi(P+1,0) \end{bmatrix},$$

and

$$\phi(m,k) = \sum_{n=0}^{N-1} d[n-m]d[n-k]$$

3. (20 points) An LPC-based speech coder generates every N-sample frame of a synthetic speech signal, $\hat{s}[n]$, from a synthetic quantized residual, $\hat{q}[n]$. There are two ways this may be done. First, the synthesis may be done using IIR filters, as shown in Equations (1) and (2):

$$\hat{d}[n] = \hat{q}[n] + \sum_{m=1}^{3} \beta_m \hat{d}[n - P + 2 - m] \tag{1}$$

$$\hat{s}[n] = \hat{d}[n] + \sum_{m=1}^{10} a_k \hat{s}[n - m] \tag{2}$$

Second, the synthesis may be performed using FIR filters, as shown in Equation (3):

$$\hat{s}[n] = \sum_{m=0}^{n} h[n - m]\hat{q}[m], \tag{3}$$

where

$$H(z) = \left(\frac{1}{1 - \sum_{m=1}^{3} \beta_m z^{-P+2-m}} \right) \left(\frac{1}{1 - \sum_{m=1}^{10} a_k z^{-m}} \right)$$

Suppose that most of the samples of $\hat{q}[n]$ are zero; only M samples per frame, on average, are non-zero, where $M \ll N$. Suppose that $h[n]$, a_m, and β_m have already been computed, and your only task now is to decide whether the FIR or the IIR synthesis method is more computationally efficient. For what values of M is the FIR synthesis more computationally efficient (fewer multiply operations) than the IIR synthesis method?

Solution: The IIR synthesis method requires 13 multiplications per sample: 3 to compute $\hat{d}[n]$, then 10 more to compute $\hat{s}[n]$. The FIR synthesis method requires us to multiply every nonzero sample of $\hat{q}[n]$ with, on average, $\frac{N}{2}$ samples of $h[n]$ (half a frame), thus it requires $MN/2$ multiplications per frame, or $M/2$ multiplications per sample. The FIR method is therefore more efficient if

$$M < 26$$

Note that this argument is only valid if $h[n]$ has already been computed.

4. (20 points) Consider a two-state HMM with parameters. Suppose that all of the parameters of this HMM have the value of 0.5. In other words, suppose that:

$$\begin{aligned} \pi_i &= P(q_1 = i) = 0.5 \quad \forall i \in \{1, 2\} \\ a_{i,j} &= P(q_t = j | q_{t-1} = i) = 0.5 \quad \forall i, j \in \{1, 2\} \\ b_j(k) &= P(o_t = k | q_t = j) = 0.5 \quad \forall j, k \end{aligned}$$

Define

$$\hat{\alpha}_t(i) = P(q_t = i | o_1, \ldots, o_t, \lambda)$$

Find the value of $\hat{\alpha}_{300}(1)$.

Solution: Using the scaled forward algorithm, we see that

$$\begin{aligned} \tilde{\alpha}_1(i) &= \pi_i b_i(o_1) = 0.25, \\ c_1 &= \tilde{\alpha}_1(1) + \tilde{\alpha}_1(2) = 0.5, \\ \hat{\alpha}_1(i) &= \frac{\tilde{\alpha}_1(i)}{c_1} = 0.5. \end{aligned}$$

Then,

$$\begin{aligned} \tilde{\alpha}_2(j) &= \sum_i \hat{\alpha}_1(i) a_{i,j} b_j(o_2) = 0.25 \\ c_2 &= \tilde{\alpha}_2(1) + \tilde{\alpha}_2(2) = 0.5, \\ \hat{\alpha}_2(i) &= \frac{\tilde{\alpha}_2(i)}{c_2} = 0.5. \end{aligned}$$

Following this logic, we find that $\hat{\alpha}_t(i) = 0.5$ for all i and t, therefore

$$\hat{\alpha}_{300}(1) = 0.5$$

5. (20 points) Recall that Baum's auxiliary can be written as

$$Q(\lambda, \bar{\lambda}) = \sum_{Q} P(Q|O, \lambda) \ln P(O, Q|\bar{\lambda}),$$

and that the part related to the observation pdf can be simplified to

$$Q_b(\lambda, \bar{\lambda}) = \sum_{t=1}^{T} \sum_{i=1}^{N} \gamma_t(i) \ln \bar{b}_i(o_t),$$

where the terms are defined as

$$\gamma_t(i) = P(q_t = i|O, \lambda)$$
$$\bar{b}_i(o_t) = P(o_t = o_t|q_t = i, \bar{\lambda})$$

Suppose that we have a sequence of scalar observations, $O = [o_1, \ldots, o_T]$, modeled by state-dependent Gaussian pdfs,

$$\bar{b}_i(o_t) = \frac{1}{\sqrt{2\pi\bar{\rho}_i}} e^{-\frac{(o_t - \bar{\mu}_i)^2}{2\bar{\rho}_i}},$$

where $\bar{\mu}_i$ and $\bar{\rho}_i$ are the state-dependent mean and variance, respectively. As you know, the variance terms in a Gaussian need to be non-negative. We can force $\bar{\rho}_i$ to be non-negative by maximing a Lagrangian term of the form

$$\mathcal{L}(\bar{\lambda}) = Q_b(\lambda, \bar{\lambda}) - \sum_{i=1}^{N} \kappa_i \bar{\rho}_i$$

Simplify $\mathcal{L}(\bar{\lambda})$ so that it is a function of only $\gamma_t(i)$, o_t, $\bar{\mu}_i$, $\bar{\rho}_i$, and κ_i for $1 \leq t \leq T$, $1 \leq i \leq N$.

Solution: This looks complicated, but after a bit of staring, we realize that it is only asking us to take the logarithm of a Gaussian. The result is

$$\mathcal{L}(\bar{\lambda}) = \sum_{t=1}^{T} \sum_{i=1}^{N} \gamma_t(i) \left(-\frac{1}{2} \ln(2\pi\rho_i) - \frac{(o_t - \bar{\mu}_i)^2}{2\bar{\rho}_i} \right) - \sum_{i=1}^{N} \kappa_i \bar{\rho}_i$$