E9 261 – Speech Information Processing

Final Examination (Marks = 100)

2-5PM, June 4, 2021

Instructions:

- Final examination is open lecture notes.
- Write clearly and legibly.
- To get credit, you need to give detailed explanations of your answers.
- Cheating or violating academic integrity will incur 0 points.

1. **Discrete cepstral coefficients**
   Consider a real discrete-time sequence $x[n]$ with DTFT $X(e^{j\omega})$. Suppose the value of $X(e^{j\omega})$ is known at L distinct frequencies $\omega_l,\ \ 1 \le l \le L$. Find the discrete cepstral coefficients $c_n,\ -p \le n \le p$ to model $\log|X(e^{j\omega})|$ such that the following mean squared error is minimized:

   $$\epsilon = \sum_{l=1}^{L} \left|\log|X(e^{j\omega_l})| - \mathrm{DTFT}(c_n)\right|^2$$

   Clearly write each step in deriving the solution. (**Points 15**)

2. **Levinson's Recursion**

   (a) Consider the auto-correlation of a sequence as follows:

   $$R[k] = (24/5) \times 2^{-|k|} - (27/10) \times 3^{-|k|}.$$

   Solve second-order optimal linear prediction co-efficients using Levinson's recursion formulae. Find the corresponding minimum error signal energy. (**Points 5**)

   (b) Do the partial correlation coefficients k_1 and k_2 in part (a) have magnitudes less than 1? Is it true even for p-th order linear prediction, when $p > 2$? Use Levinson recursion equation to argue. (**Points 5**)

(c) Can you propose an alternative way to solve the coefficients in part (a) without using Levinson's recursion? If yes, between the Levinson's recursion and the alternative method which one has computational advantage? Can you quantify the advantage for solving p-th order linear prediction coefficients? (**Points 5**)

3. **FIR Cepstrum!**
Is the following statement true (justify your answer)? – The complex cepstrum of the following sequence is FIR.

$$x[n] = \begin{cases} (-1)^{n/2}\frac{\alpha^n}{(n/2)!}, & n > 0 \text{ and } n \text{ even} \\ 1 & n = 0 \\ 0 & n < 0 \text{ or } n \text{ odd} \end{cases}$$

Hint: $e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$ (**Points 15**)

4. **Partial derivative of STFT magnitude and phase**
Consider a modified short-time Fourier transform (STFT) defined for a continuous-time signal $x(t)$ with analysis window $h(t)$ as follows:

$$X^h(t,\omega) = e^{\mathrm{j}\frac{\omega t}{2}} \int_{-\infty}^{+\infty} x(u)h(t-u)e^{-\mathrm{j}\omega u}\mathrm{d}u = M_x^h(t,\omega)e^{\mathrm{j}\Phi_x^h(t,\omega)}$$

where $M_x^h(t,\omega)$ and $\Phi_x^h(t,\omega)$ are magnitude and phase respectively. As a part of spectrogram re-assignment, one often has to compute the partial derivative of magnitude and phase. Show that the partial derivative of magnitude and phase with respect to t and ω can be easily computed as follows:

(a) $\frac{\partial}{\partial t}\log(M_x^h(t,\omega)) = \mathrm{Re}\left(\frac{X^{Dh}(t,\omega)}{X^h(t,\omega)}\right)$, $\quad \frac{\partial}{\partial t}\Phi_x^h(t,\omega) = \mathrm{Im}\left(\frac{X^{Dh}(t,\omega)}{X^h(t,\omega)}\right) + \frac{\omega}{2}$

(b) $\frac{\partial}{\partial \omega}\log(M_x^h(t,\omega)) = -\mathrm{Im}\left(\frac{X^{Th}(t,\omega)}{X^h(t,\omega)}\right)$, $\quad \frac{\partial}{\partial \omega}\Phi_x^h(t,\omega) = \mathrm{Re}\left(\frac{X^{Dh}(t,\omega)}{X^h(t,\omega)}\right) - \frac{t}{2}$

where $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ indicate the real and imaginary part of a complex number and $Dh(t) = \frac{\mathrm{d}}{\mathrm{d}t}h(t)$ and $Th(t) = th(t)$. (**Points 20**)

5. **Minimum-phase, maximum-phase, complex cepstrum**
Consider a minimum-phase all-zero model $D(z) = \sum_{k=0}^{Q} d_k z^{-k}$ with complex cepstrum $c[k]$. We create another all-zero model with coefficients $\tilde{d}_k = \alpha^k d_k$ and the complex cepstrum $\tilde{c}[k]$.

(a) if $0 < \alpha < 1$, find the relation between $\tilde{c}[k]$ and $c[k]$.

(b) Choose α so that the new model has a maximum phase. (**Points 20**)

6. **Maximum Mutual Information (MMI) - HMM**
In the conventional HMM, the association between the state sequence and observation sequence is not taken into account. In a modified formulation, we would like to add this component in the modeling. Let $H(\mathbf{O}|\mathbf{q})$ denote the conditional entropy of the

observation sequence $\mathbf{O} = \{\mathbf{o}_t\}_{t=1}^{T}$ given the state sequence $\mathbf{q} = \{q_t\}_{t=1}^{T}$. The conditional entropy is defined as,

$$H(\mathbf{O}|\mathbf{q}) = -\sum_{t}\sum_{q_t}\int_{\mathbf{O}_t} P(\mathbf{o}_t, q_t)\log\left(P(\mathbf{o}_t|q_t)\right)\partial\mathbf{o}_t$$

In the HMM formulation, we would like to minimize the conditional entropy (maximizing the mutual information). Thus, we modify the ML method in the following manner,

$$J = -(1-\epsilon)H(\mathbf{O}|\mathbf{q}) + \epsilon\log P(\mathbf{O}),$$

where ϵ is a positive constant whose maximum value is 1 and J is the function to be maximized to determine the parameters. We assume continuous density HMM with state emission probabilities as single Gaussian distribution ($b_j(\mathbf{o}_t) \sim \mathcal{N}(\mathbf{o}_t, \mu_j, \mathbf{\Sigma}_j)$). Derive the update equations for mean and covariance of the emission probabilties for the MMI- HMM. (**Points 15**)