

PSI-2533 Modelagem em Processamento de Sinais

Exercício 3 do Módulo de Modelagem em Processamento de Voz

24 de agosto de 2016

1. Faça, pelo método da autocorrelação, a análise preditiva de ordem $p = 3$ do sinal

$$s = [1 \quad 0,64 \quad 0,81 \quad 0,20 \quad 0,12 \quad 0,05 \quad 0,08 \quad -0,02 \quad 0,25 \quad 0,32],$$

anotando todas as suas etapas intermediárias, incluindo as funções de transferência dos filtros de erro de predição (ou filtros inversos) progressivos e regressivos intermediários e finais e seus erros quadráticos de predição. Apresente também os coeficientes de correlação parcial. Compare os resultados das etapas intermediárias através de seus erros quadráticos e de seus coeficientes de correlação parcial.

Obs.: Use janela retangular.

2. Faça as conversões entre coeficientes de predição e coeficientes de correlação parcial (PARCOR) ou vice-versa conforme solicitado, identifique o filtro de síntese em questão e determine se ele é estável. Determine e apresente o diagrama de fluxo de sinal em treliça de Itakura-Saito dos filtros de síntese no caso de redução de ordem.

a) São dados os coeficientes PARCOR $k_1 = -3/8$, $k_2 = 1/2$ e $k_3 = 1/4$.

b) É dado o filtro inverso $A(z) = 1 - 9/20z^{-1} - 2/3z^{-2} + 1/2z^{-3}$.

c) É dado o filtro inverso $A(z) = 1 - 9/20z^{-1} - z^{-2} + 1/2z^{-3}$.

3. Use o mesmo sinal de voz do Exercício 1, que está disponível no stoa, onde também estão disponíveis as funções MATLAB `snrsege` e `disseg`, respectivamente para o cálculo da relação sinal-ruído segmentada (SNRSEG) e a distorção de Itakura-Saito (DIS), compactadas no arquivo `snrsegdisseg.zip`. Nosso objetivo nesta questão é a síntese dos primeiros 1,5 s deste sinal com o sintetizador misto de Itakura-Saito. Antes disso, porém, faremos a reconstrução perfeita do sinal usando o sinal residual da predição linear de ordem p . Recomenda-se ordem de predição $p = 40$ para frequência de amostragem $F_a = 44100$ Hz. Use uma janela de Hamming $w(n)$

com ou sem sobreposição para a análise e uma janela retangular (obviamente sem sobreposição) para a síntese. A suavidade da síntese é controlada pela identificação das condições iniciais do filtro de síntese no segmento seguinte com as condições finais dele no segmento corrente – também imponha estas identificações na filtragem inversa da análise para obter um sinal de erro $d(n)$ compatível. O comprimento L dos segmentos deve ser suficiente para resolver frequências fundamentais entre 50 Hz e 500 Hz. Recomenda-se a pré-ênfase prévia do sinal usando, para realce das altas frequências, o filtro de pré-ênfase

$$H_E(z) = 1 - \mu z^{-1}$$

com coeficiente $0 \leq \mu \leq 1$. Ajuste o valor que lhe pareça mais adequado e forneça-o em seu relatório. Sugere-se $\mu = 0,98$, mas outro valor pode ser mais adequado dependendo de sua percepção. Após a síntese, o sinal sintetizado deve ser devidamente desenfaturado pelo filtro inverso de $H_E(z)$.

- a) Faça e apresente o espectrograma dos primeiros 1,5 s do sinal dado, que são nosso sinal $s(n)$ de referência.
- b) Faça as análises preditivas dos segmentos de $s(n)$ e a filtragem inversa com os cuidados observados para obter o sinal residual $d(n)$. Com o sinal residual e seus filtros de síntese, obtenha a reconstrução quase-perfeita $\tilde{s}_1(n)$. Obtenha a SNRSEG desta reconstrução usando segmentos de 16 ms. Que valor esperava? Corresponde ao valor que obteve? Que erros influíram no processo?
- c) Obtenha os parâmetros da excitação do sintetizador misto de IS, que são os períodos fundamentais P , as sonoridades V e seus complementos U e os ganhos sonoros G_v e surdos G_u . Tendo todos os parâmetros necessários, faça a síntese para obter o sinal sintetizado $\tilde{s}_2(n)$. Recomenda-se que procure posicionar os pulsos de sua excitação sonora levando em conta a posição exata do último pulso do segmento anterior – esta pode ser representada por um número negativo para o segmento corrente. Ouça seu sinal sintetizado – o MATLAB possui o comando `soundsc` – e obtenha sua SNRSEG e sua DISSEG com segmentos de 16 ms. O critério de Papamichalis considera inaceitável o sinal sintetizado quando sua DIS supera 1,4. Observe seus parâmetros de excitação e procure editá-los para melhorar a qualidade do sinal sintetizado. Descreva os pontos significativos de sua edição e o incremento de qualidade que obteve. Obtenha também o espectrograma de $\tilde{s}_2(n)$ nas mesmas condições que você usou no item a) e apresente-o.
- d) Selecione um segmento interessante da sua síntese em c), que represente bem as duas componentes da excitação e apresente a determinação de seus parâmetros com todos os cálculos. Apresente também o gráfico da função de autocorrelação normalizada $\rho_{dd}(m)$, eventualmente suprimindo o ponto trivial com $m = 0$ do gráfico para ressaltar sua variação na região mais significativa.

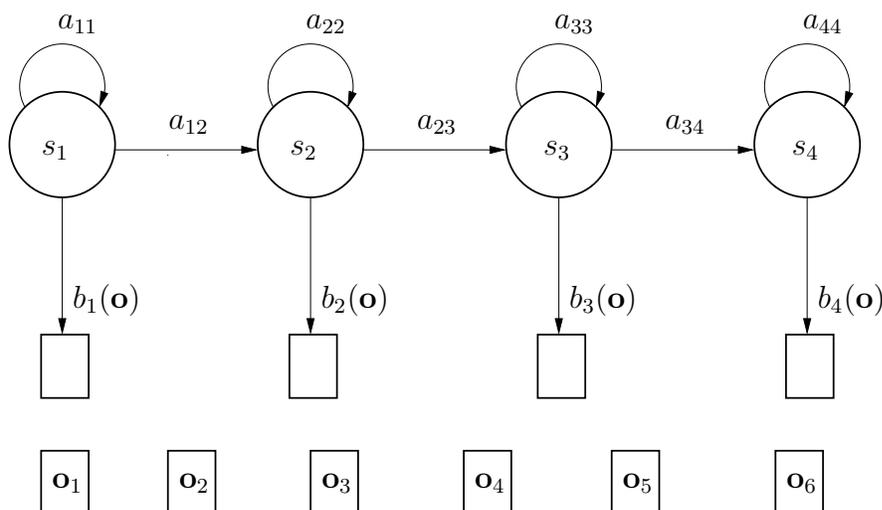
4. Será usada a máquina de Markov oculta M com massa de probabilidade do estado inicial $\pi = [1/3, 1/3, 1/3, 0]$ em $t = 0$, matriz de transição entre estados

$$A = \begin{bmatrix} 0,1 & 0,9 & 0 & 0 \\ 0 & 0,2 & 0,8 & 0 \\ 0 & 0 & 0,3 & 0,7 \\ 0 & 0 & 0 & 1,0 \end{bmatrix},$$

cujas linhas representam os estados de origem e as colunas os estados de destino na ordem s_1, s_2, s_3 e s_4 e matriz das funções massa de probabilidade dos códigos de observações

$$B = \begin{bmatrix} 0,3 & 0,2 & 0,6 & 0,2 \\ 0,4 & 0,3 & 0,3 & 0,7 \\ 0,3 & 0,5 & 0,1 & 0,1 \end{bmatrix},$$

cujas colunas estão ordenadas por estado na ordem s_1, s_2, s_3 e s_4 e as linhas representam os códigos de observação 0, 1 e 2 respectivamente.



Pedem-se os itens abaixo.

- Calcule a máxima verossimilhança de que a sequência de observações $O = [2, 0, 1, 1, 0]$ tenha sido gerada pela máquina M e determine também a sequência ótima de estados para tal.
- Calcule a probabilidade $P(O, Q|M)$ de que a máquina M emita a sequência de observações $O = [1, 0, 2]$ com a sequência de estados $Q = [s_1, s_1, s_2]$.