

PŘEDMĚT B2M31DSP/PŘ. 2

PS

Přednáška 2: Lineární predikce a její využití



- 1 ÚVOD
- 2 MOTIVACE A HISTORIE
- 3 ZÁKLADNÍ POJMY LP ANALÝZY
- 4 AUTOKORELAČNÍ METODA
- 5 ANALÝZA A SYNTÉZA
- 6 INTERPRETACE LP ANALÝZY VE SPEKTRÁLNÍ OBLASTI
- 7 SHRUTÍ
- 8 PŘÍKLAD AR SPEKTRA
- 9 POUŽITÍ LP PRO ZTRÁTOVOU KOMPRESI

ÚVOD

Cíle přednášky:

- základní pojmy
- použití lineární predikce (LP) pro parametrizaci signálu a pro spektrální analýzu
- použití lineární predikce pro ztrátovou kompresi řeči

MOTIVACE A HISTORIE

Hledání

- alternativních metod spektrální analýzy¹
- efektivních metod parametrizace a komprese signálů²

Použití lineárního predikčního kódování (LPC):

- spektrální analýza
- analýza/syntéza a komprese řeči i audio signálů

Historie:

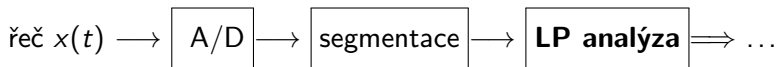
- Gauss - metoda nejmenších čtverců (≈ 1800),
- Wiener - modifikace pro stacionární procesy (≈ 1930)
- Yule, Walker - analýza ekonomických dat (≈ 1930)
- Saito, Itakura, Burg - kódování řeči & spektrální analýza - (≈ 1970)

¹ Zejména pro krátké signály (radar, sonar) nebo nestacionární signály (seismické) neposkytující dostatečné rozlišení ve spektru

² Např. řeč, audio signály

Příklad

Vokodér^a využívající LPC = základ algoritmů kódování^b a komprese řeči v mobilních telefonech:



... přenos parametrů ...



^aVoice-coder: US ARMY \approx 1970

^bBudeme se zabývat zvýrazněnými bloky analýza & syntéza

ZÁKLADNÍ POJMY LP ANALÝZY

- **predikce**

$$\hat{x}[n] = - \sum_{k=1}^M a_k x[n-k], \quad M \text{ je řád prediktoru a } a_k \text{ jeho koeficienty}$$

- **chyba predikce**

$$e[n] = x[n] - \hat{x}[n] = x[n] + \sum_{k=1}^M a_k x[n-k], \quad n = 0, 1, \dots, M + N - 1$$

= přeurčená soustava rovnic

- řešení \leftrightarrow účelová funkce

- $\min(\sum e^2[n])$ = metoda nejmenších čtverců (Gauss)
- $\min(E[e^2[n]])$ = metoda min. střední kvadratické chyby (Wiener)

- výsledek³ minimalizace = soustava lineárních rovnic obsahujících korelace (normální rovnice)

³ Podle typu účelové funkce a dalších podmínek získáme různé metody. Např. autokorelační, kovarianční, Burgova, ...

AUTOKORELAČNÍ METODA

Odvození autokorelační metody pro $M = 2$:

- máme chybu predikce⁴ $e[n]$ získanou FIR filtrem (= bělicím nebo chybovým nebo dekorelačním filtrem)

- odvození rovnic pro určení $a_k \leftrightarrow \min(E[e^2[n]]$:

$$\begin{aligned} \frac{\partial(E[e^2[n]])}{\partial a_1} &= 2E[e[n] \frac{\partial(e[n])}{\partial a_1}] = 2E[e[n]x[n-1]] \\ &= 2E[x[n]x[n-1]] + a_1 2E[x[n-1]x[n-1]] + a_2 2E[x[n-2]x[n-1]] = 0 \\ &\rightarrow 2R_x[1] + a_1 R_x[0] + a_2 R_x[-1] = 0 \rightarrow a_1 R_x[0] + a_2 R_x[1] = -R_x[1] \\ &\text{podobně pro } a_2 \end{aligned}$$

- normální rovnice pro autokorelační metodu = **Yuleovy-Walkerovy rovnice**:

$$a_1 R_x[0] + a_2 R_x[1] = -R_x[1]$$

$$a_1 R_x[1] + a_2 R_x[0] = -R_x[2]$$

⁴ viz předchozí strana

Obecný tvar Yuleových-Walkerových rovnic⁵ :

$$\begin{bmatrix} R_x[0] & R_x[1] & \dots & R_x[M-1] \\ R_x[1] & R_x[0] & \dots & R_x[M-2] \\ \vdots & \vdots & \ddots & \vdots \\ R_x[M-1] & R_x[M-2] & \dots & R_x[0] \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ \vdots \\ \vdots \\ a_M \end{bmatrix} = - \begin{bmatrix} R_x[1] \\ \vdots \\ \vdots \\ R_x[M] \end{bmatrix}$$

a v maticovém tvaru:

$$\mathbf{R}_x \mathbf{a} = -\mathbf{r}_P$$

matice soustavy \mathbf{R}_x je:

- symetrická \leftrightarrow autokorelační funkce je sudá
 - ekvidiagonální = prvky na hlavní a vedlejších diagonálách jsou stejné
- pro inverzi matice soustavy existují efektivní algoritmy

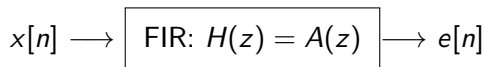
⁵Vysvětlit vlastnosti a konstrukci matice soustavy a vztah matice a vektoru na pravé straně - posun korelačních koeficientů v matici i mezi maticí a vektorem

ANALÝZA AR PROCESU POMOCÍ FILTRACE FIR FILTREM

Analýza signálu $x[n]$ = filtrace chybovým FIR⁶ filtrem:

$$e[n] = x[n] + \sum_{k=1}^M a_k x[n-k], \quad n = 0, 1, \dots$$

Graficky lze vyjádřit náčrtnem



Tvrzení bez důkazu:

Chybový signál $e[n]$ je, v případě splnění podmínek

- $x[n]$ je AR signál
- koeficienty FIR filtru jsou správně určeny,

bílým šumem

⁶ Jeho koeficienty získány např. autokorelační metodou

PŘESNÁ REKONSTRUKCE PROCESU

Rekonstrukce signálu = použití filtru inverzního k FIR - ten získáme vyjádřením $x[n]$ z předchozí rovnice - je to diferenční rovnice pro filtr se zpětnými vazbami = „all-pole“ filtr:

$$x[n] = e[n] - \sum_{k=1}^M a_k x[n-k], \quad n = 0, 1, \dots$$

Problém: chybový signál $e[n]$ pro syntézu v mnoha případech nemáme k dispozici (zvláště při vysoké kompresi signálu pro komunikační účely) - v tomto případě použijeme libovolnou realizaci bílého šumu $u[n]$ a nikoliv chybový signál $e[n]$.

Musí být splněno, že PSD šumů $e[n]$ a $u[n]$ jsou obě konstantní

Pozn.: Náhradou chyby $e[n]$ šumem $u[n]$ ovšem neprovádíme přesnou rekonstrukci signálu

SYNTÉZA AR PROCESU

Syntéza⁷ AR signálu pomocí „all-pole“ filtru při nepoužití chyby $e[n]$ má tvar

$$\tilde{x}[n] = u[n] - \sum_{k=1}^M a_k x[n-k], \quad n = 0, 1, \dots$$

Graficky lze vyjádřit náčrtkem

$$u[n] \longrightarrow \boxed{\text{all-pole filtr: } H(z) = 1/A(z)} \longrightarrow \tilde{x}[n]$$

All-pole filtr má přenos inverzní k chybovému FIR filtru

Pozn.: v případě, že analyzovaný signál nebyl AR, pak syntetizovaný signál, který je AR, se od původního signálu může spektrálně lišit

⁷ V případě syntézy řeči ve vokodéru nahrazujeme $e[n]$ bílým šumem $u[n]$ pro neznělé hlásky a jednotkovými impulsy pro znělé hlásky - řeč zní přirozeněji než při buzení pouze šumem

APROXIMACE SPEKTRA SIGNÁLU AR SPEKTRUM

- analýza = filtrace chybovým (FIR) filtrem:

$$e[n] = a_n * x[n]$$
- z předchozí rovnice lze pomocí autokorelace signálu $e[n]$ odvodit vztah mezi spektrálními hustotami: $S_e(e^{j\Theta}) = |A(e^{j\Theta})|^2 S_x(e^{j\Theta})$
- je-li chyba predikce bílý šum $\rightarrow S_e(e^{j\Theta}) = \sigma_e^2 \rightarrow$
- vyjádření PSD signálu $x[n]$ pomocí frekvenční charakteristiky chybového filtru má tvar:
$$S_x(e^{j\Theta}) = \frac{\sigma_e^2}{|A(e^{j\Theta})|^2} \quad (\text{R1})$$
- vztah⁸ $\frac{\sigma_e^2}{|A(e^{j\Theta})|^2}$ označujeme jako **LP spektrum (autoregresní spektrum, AR spektrum)**

⁸Výraz $1/|A(e^{j\Theta})|^2$ roven kvadrátu modulu frekvenční charakteristiky all-pole filtru

SHRNUTÍ POUŽITÍ LP PRO SPEKTRÁLNÍ ANALÝZU

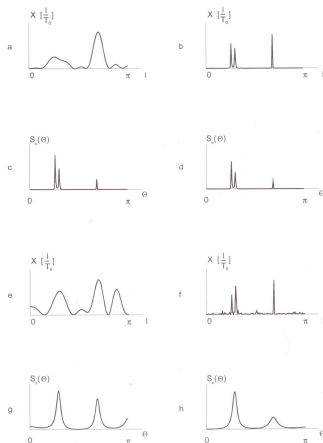
- LP analýza poskytuje aproximaci spektrální hustoty signálu $x[n]$ pomocí rozptylu chyby predikce σ_e^2 a kvadrátu modulu frekvenční charakteristiky chybového filtru $|A(e^{j\Theta})|^2$ ve tvaru (R1), přičemž **frekvenční rozlišení (AR spektra⁹) téměř nezávisí na délce signálu**
- chybový filtr řádu M získaný LP analýzou zcela dekoreluje¹⁰ signál $x[n]$ pouze tehdy, je-li $x[n]$ generován průchodem bílého šumu filtrem řádu M s přenosem $H(z) = 1/A(z)$ (= autoregresní model řádu M)
- pro signály, které nejsou autoregresními, chybový filtr dekoreluje signál $x[n]$ pouze částečně $\rightarrow e[n]$ není bílým šumem a vztah (R1) platí přibližně¹¹: $S_x(e^{j\Theta}) \approx \frac{\sigma_e^2}{|A(e^{j\Theta})|^2}$

⁹ LP spektra

¹⁰ Požadavek min. výkonu chyby predikce \rightarrow vzorky chyby predikce $e[n]$ jsou pro autoregresní signál $x[n]$ nekorelované (= chyba predikce je bílý šum). Pouze v tomto případě vztah (R1) platí přesně

¹¹ V tomto případě se jedná o aproximaci skutečné spektrální hustoty S_x racionální lomenou funkcí $\sigma_e^2/|A(e^{j\Theta})|^2$

PŘÍKLAD AR SPEKTRA



Analýza směsi 3 sinusovek se stejnými amplitudami pomocí DFT (viz označení X) a pomocí AR spektra (označení S)

Analýza směsi 3 sinusovek v šumu (SNR=0 dB) pomocí DFT a pomocí AR spektra

levý sloupec: délka signálu neumožňující rozlišení frekvencí pomocí DFT

pravý sloupec: délka signálu umožňující rozlišení frekvencí pomocí DFT

POUŽITÍ LP PRO ZTRÁTOVOU KOMPRESI ŘEČI

Postup ztrátové komprese řeči pomocí LP - zjednodušený princip vokodéru¹²

Analýza

- vstupem je vzorkovaný a kvantovaný akustický signál - např. pro digitální telefonii kvantování (nelineární¹³) 8 bitů, vzorkovací frekvence 8 kHz
- segmentace řeči na úseky přibližně 30 ms (240 vzorků)
- analýza jednoho segmentu poskytne: AR koeficienty $a[k]$ (typicky $k = 8 - 16$), typ buzení: znělé/neznělé: Z/N, v případě znělého buzení jeho periodu L_0 a energii signálu E
- přenos získaných parametrů: místo 64 kbit/s se přenesou pouze 2-6 kbit/s (závisí na počtu parametrů a požadované kvalitě)

¹²Detailnější popis ztrátové komprese řeči je náplní kursů věnovaných zpracování řeči a komunikacím

¹³Nelineární převod z 12 bitů na 8 (A-law, μ -law) zajistí vyrovnaní poměru výkonu signálu k výkonu kvantizačního šumu

POUŽITÍ LP PRO ZTRÁTOVOU KOMPRESI ŘEČI

Syntéza

- pro každý segment se provede syntéza:
 - vstupem je posloupnost impulsů vzdálených o L_0 pro znělé úseky nebo pseudonáhodný šum pro neznělé úseky
 - vstup je upraven pomocí zisku $\sqrt{(E)}$
 - upravený signál vstupuje do all-pole filtru s koeficienty $a[k]$
 - výstupem filtru je řečový signál, který se spektrálně shoduje s původním řečovým signálem
- následuje skládání úseků generovaného signálu a jeho rekonstrukce poskytující signál spojitý v čase