

論文 / 著書情報
Article / Book Information

論題(和文)	極零モデルによるスペクトル最尤推定法の改良
Title(English)	
著者(和文)	嵯峨山 茂樹, 古井 貞熙, 斉藤 収三
Authors(English)	SADAOKI FURUI
出典(和文)	日本音響学会1976年講演論文集, , , pp. 545-546
Citation(English)	, , , pp. 545-546
発行日 / Pub. date	1976, 10

3-4-3 極零モデルによるスペクトル最尤推定法の改良*

嵯峨山 茂樹 古井 貞熙 斎藤 収三
(日本電信電話公社 武蔵野電気通信研究所)

1 はじめに

音声を極零モデル (AR-MA モデル) からのサンプル値列であるとして、そのスペクトルパラメータを最尤推定する方法については、既に述べた⁽¹⁾。この方法は統計的推定論の立場から、最も優れたスペクトル推定ができるが、計算が多くなる欠点があった。ここで述べる方法によれば、計算量をかなり減らすことができ、この欠点を補える。

2 スペクトルパラメータの最尤推定

$\{\varepsilon_t\}$ (t は整数) を平均 0、分散 σ^2 の白色過程とすると、

$$\sum_{k=0}^p \alpha_k x_{t-k} = \sum_{l=0}^q \beta_l \varepsilon_{t-l} \quad (1)$$

により生成される $\{x_t\}$ ((p, q) 次 AR-MA 過程) を音声のモデルとすると、パワースペクトルは、次式のようになる。

$$f(\omega) = \frac{\sigma^2}{2\pi} |G(e^{j\omega})|^2; \quad G(z) = \frac{Q(z)}{P(z)}; \quad (2)$$

$$P(z) = \sum_{k=0}^p \alpha_k z^{-k}, \quad Q(z) = \sum_{l=0}^q \beta_l z^{-l}$$

N 個の音声サンプル列 $\{x_1, \dots, x_N\}$ が得られたとき、パラメータ $\{\alpha; \beta; \sigma^2\}$ に関する対数尤度は、近似的に (N : 大)

$$L_N(\alpha; \beta; \sigma^2) = -\frac{N}{2} \left\{ \log \sigma^2 + \frac{1}{\sigma^2} \int_{-\pi}^{\pi} \frac{I_N(\lambda)}{|G(e^{j\lambda})|^2} d\lambda \right\} \quad (3)$$

$$\text{但し } I_N(\lambda) = \frac{1}{2\pi} \left| \sum_{t=1}^N x_t e^{j\lambda t} \right|^2 \quad (\text{ヒルバーク法})$$

で与えられる。最尤推定は、この式を最大化する $\{\alpha_1, \dots, \alpha_p; \beta_1, \dots, \beta_q; \sigma^2\}$ を求めることによりなされる。

3 定式化の改良

式(3)を、まず σ^2 について最大化すると、

$$\max_{\sigma^2} L_N = -\frac{N}{2} \left\{ 1 + \log \int_{-\pi}^{\pi} \frac{I_N(\lambda)}{|G(e^{j\lambda})|^2} d\lambda \right\} \quad (4)$$

となる⁽¹⁾。これを $\{\alpha; \beta\}$ に関して最大化する $(p+q)$ 次元の最大化問題は、計算量が多くなる欠点があった。

そこで、式(4)をさらに $\{\alpha\}$ について最大化して、 $\{\beta\}$ のみの関数とすれば、

$$\max_{\sigma^2, \{\beta\}} L_N = -\frac{N}{2} \left\{ 1 + \log (D_{pm} / D_p) \right\} \quad (5a)$$

$$\text{但し } D_n = \det \Phi_n,$$

$$\Phi_n = \begin{bmatrix} \phi_{n-j} \end{bmatrix}, \quad \phi_c = \int_{-\pi}^{\pi} \frac{I_N(\lambda)}{|Q(e^{j\lambda})|^2} e^{j\tau\lambda} d\lambda \quad (5b)$$

のようになる。結局、最尤推定問題は、

$$J[\beta] = \frac{D_{pm}}{D_p} \quad \text{あるいは} \quad K[\beta] = \log \frac{D_{pm}}{D_p} \quad (6)$$

を $\{\beta_1, \dots, \beta_q\}$ について最小化する、 q 次元の非線形最小化問題に帰着する。

4 物理的意味

式(6)の $J[\beta]$ は、実は、スペクトルが

$$g(\lambda) = I_N(\lambda) / |Q(e^{j\lambda})|^2 \quad (7)$$

であるような信号を PARCOR 分析した場合の残差パワになっている⁽²⁾。また、式(7)は、スペクトル $I_N(\lambda)$ を、線形フィルタ $Q(z)$ の逆フィルタに通したときのスペクトルであるから、式(6)の $J[\beta]$ は、Fig. 1 に示すような系の出力パワーを表わすことになる。

この事実は、計算にも役立つ。即ち、与えられた $\{\beta_1, \dots, \beta_q\}$ から、パワー伝達

* "An Improvement in Maximum Likelihood Estimation of Speech Spectrum by Pole-Zero Modeling" by Shigeki SAGAYAMA, Sadaoki FURUI, and Shuzo SAITO, Musashino ECL, NTT.