

論文 / 著書情報
Article / Book Information

論題(和文)	極零モデルによる音声スペクトルの最尤推定
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. 321-322
Citation(English)	, , , pp. 321-322
発行日 / Pub. date	1976, 5

3-2-6 極零モデルによる音声スペクトルの最尤推定*

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

音声スペクトルを極零(AR-MA)モデルにより推定することは、声道伝達関数や音源の極と零点の推定に関連して、最近国内外で活発に研究が進められている。本報告は、統計的推定論の方向から、最も良い性質を持つ推定量—最尤推定量について議論する。

2 AR-MA 過程モデル

$\{\varepsilon_t\}$ (t は整数) を、平均 0、分散 σ^2 の白色過程とするとき、($\alpha_0 = \beta_0 = 1$)

$$\sum_{\mu=0}^p \alpha_{\mu} x_{t-\mu} = \sum_{\nu=0}^q \beta_{\nu} \varepsilon_{t-\nu} \quad \text{----- (4)}$$

により生成される定常確率過程 $\{x_t\}$ を、次数 (p, q) の AR-MA (自己回帰移動平均) 過程と呼ぶ。そのパワースペクトルは、

$$f(\lambda) = \frac{\sigma^2}{2\pi} \left| \sum_{\nu=0}^q \beta_{\nu} e^{-j\nu\lambda} \right|^2 / \left| \sum_{\mu=0}^p \alpha_{\mu} e^{-j\mu\lambda} \right|^2 \quad \text{--- (2)}$$

で与えられる。(1)において、 $\{\varepsilon_t\}$ を入力、 $\{x_t\}$ を出力とする線形フィルタを考えると、その伝達関数は、

$$G(z) = \frac{Q(z)}{P(z)} ; \begin{cases} P(z) = \sum_{\mu=0}^p \alpha_{\mu} z^{\mu} \\ Q(z) = \sum_{\nu=0}^q \beta_{\nu} z^{\nu} \end{cases} \quad \text{--- (3)}$$

となる。

3 AR-MA パラメータの最尤推定

式(1)の AR-MA 過程 $\{x_t\}$ から有限個のサンプル値が得られたとき、パラメータ $\{\alpha; \beta; \sigma^2\}$ を推定することを考える。推定量には種々のものがあるが、最尤推定量は、一貫性、漸近的な不偏性と有効性などを持つ、統計的に最もすぐれた推定

量であるので、ここでは、パラメータの推定に、最尤推定法を採用する。

$\{x_t\}$ が正規過程であれば、 N 個のサンプル $\{x_1, x_2, \dots, x_N\}$ が与えられたときの対数尤度は、 N が大きいとき、次式で近似できる⁽⁴⁾

$$\tilde{L}_N(\alpha; \beta; \sigma^2) = -\frac{N}{2} \left\{ \log \sigma^2 + \frac{1}{\sigma^2} \int_{-\pi}^{\pi} \frac{I_N(\lambda)}{|G(e^{j\lambda})|^2} d\lambda \right\} \quad \text{--- (4)}$$

$$\text{但し、} I_N(\lambda) \triangleq \frac{1}{2\pi} \left| \sum_{t=1}^N x_t e^{j\lambda t} \right|^2 \quad (\text{ヒルボグラム})$$

これを、 σ^2 に関して最大化することにより、式(4)の最大化(最尤推定)は、結局、 $\{\alpha_1, \dots, \alpha_p; \beta_1, \dots, \beta_q\}$ の関数、

$$F(\alpha; \beta) = \int_{-\pi}^{\pi} \frac{I_N(\lambda)}{|G(e^{j\lambda})|^2} d\lambda \quad \text{--- (5)}$$

の最小化の問題に帰着する。

4 最尤推定の物理的意味

式(5)は、パワースペクトルが $I_N(\lambda)$ である信号を、フィルタ $G(z)$ (式(3)の逆フィルタ)に通したときの出力パワを表わす。従って、最尤推定は、「逆フィルタ出力パワ最小化」の定式化に同等であり、Markel 流の LPC の、極零モデルへの拡張となっている。

また、このことは、 x_t の予測を、

$$\hat{x}_t = - \sum_{\mu=1}^p \alpha_{\mu} x_{t-\mu} + \sum_{\nu=1}^q \beta_{\nu} e_{t-\nu} \quad \text{--- (6)}$$

但し、 $e_t = x_t - \hat{x}_t$ (予測誤差列)

によって行ない、予測誤差 $\{e_t\}$ の分散を最小にする、Atal 流の LPC の極零モデルへの拡張にもなっている。

* Maximum Likelihood Estimation of Speech Spectrum by Pole-Zero Modeling, by Shigeki SAGAYAMA, Sadaoki FURUI, and Shuzo SAITO, Musashino Electrical Communication Laboratory, N.T.T.

5 最小位相条件

式(3)のフィルタ $G(z)$, およびその逆フィルタが安定であるためには, $P(z)$ および $Q(z)$ のすべての根が, z 平面上の単位円内になければならない (これを最小位相条件と呼ぶことにする)。

図1は, $F(\alpha; \beta)$ の形の例である。 $\{\alpha\}$ に関しては単純な形であるが, $\{\beta\}$ に関しては複雑で, 最小位相領域 (図中の三角形内) の外では, 最小解が存在しない様子がわかる。

6 最尤推定アルゴリズム

式(5)の $F(\alpha; \beta)$ の最小化問題は, 非線形最小化問題であり, 数値的に解くには, 繰り返し (iteration) 算法によらざるを得ないが, この問題では Newton-Raphson 法が有利である。そのアルゴリズムは, $\theta = (\alpha_1 \dots \alpha_p \beta_1 \dots \beta_q)'$ とし, θ の第 m 近似を $\theta^{(m)}$ と書くと,

$$\theta^{(m+1)} = \theta^{(m)} - H(\theta^{(m)})^{-1} \nabla F(\theta^{(m)}) \quad \dots (7)$$

但し, $\nabla F(\theta) = \text{grad} F = \left(\frac{\partial F}{\partial \theta_i} \right) (\text{gradient})$

$$H(\theta) = \left(\frac{\partial^2 F}{\partial \theta_i \partial \theta_j} \right) (\text{Hesse行列})$$

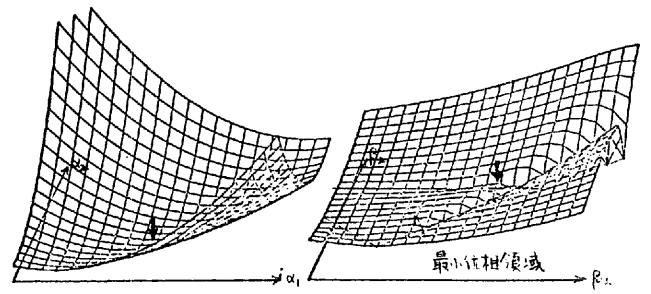
により与えられる。しかし, この方法では, 必ずしも (最小位相領域内での) global な最小解が得られる保証はないため, ここでは, 式(7)の代わりにその変形:

$$\theta^{(m+1)} = \theta^{(m)} - \frac{\nabla F(\theta^{(m)}) \nabla F(\theta^{(m)})'}{\nabla F(\theta^{(m)})' H(\theta^{(m)}) \nabla F(\theta^{(m)})} \nabla F(\theta^{(m)}) \quad (8)$$

(必ず下降方向へ進む) を用いる。また初期値となるべくよいものを与える必要がある。

7 実験結果例

図2に実験結果例を示す。音声サンプルは, 「銀世界の高原」の /n/ (男声) の部分である。図中の $\hat{I}_N(\lambda)$ は, ピリオ



(a) $(p, q) = (2, 0)$ の場合の $F(\alpha_1, \alpha_2)$ (b) $(p, q) = (0, 2)$ の場合の $[F(\beta_1, \beta_2)]^{-1}$

図1. $F(\alpha; \beta)$ の形の立体表示例 (矢印は真値)

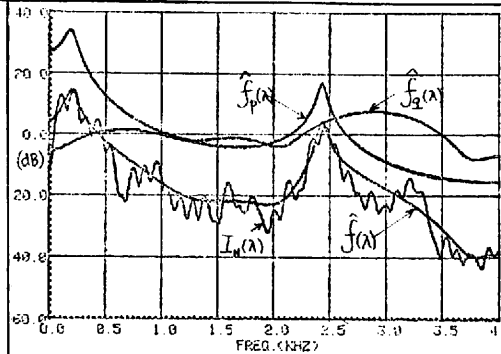


図2. 本方法による実験結果例 (最尤推定 $p=6, q=8$)

ドグラム (標本スペクトル) $I_N(\lambda)$ を, Log window 法⁽²⁾により少し平滑化したものである。スペクトル反転法⁽²⁾による一致推定を初期値として, 式(8)の iteration により解を求め, 推定スペクトル $\hat{f}(\lambda)$ を得た。(図中, $\hat{f}_p(\lambda), \hat{f}_q(\lambda)$ は, そのうちの, 極のみ, 零点のみのスペクトルを表わす。)

8 おわりに

音声スペクトルの極零モデルによる最尤推定法を試み, その物理的意味とアルゴリズムについて考察し, 実験を行なった。本方法は, 今後の問題であるモデルの次数 (p, q) の決定法に結びつけることができよう。

謝辞 日頃 御指導・御協力を蒙り, 野田整理部長, 橋本第1研究室長, 板倉調役 ほか, 斎藤持明研究室, 第4研究室の諸氏に感謝する。

文献 (1) 山崎山古井・斎藤「極零モデルによる音声スペクトルの推定」信学技報 PRL75-81

(2) 東倉 板倉 橋本「スペクトル平滑化による PARCOR 合成音の品質向上」本論文集