

一种高分辨率的极大熵谱估计算法¹

黄 勇 彭应宁 王秀坛

(清华大学电子工程系 北京 100084)

摘 要 经典的基于 Yule-Walker 等式的极大熵方法——Levinson-Durbin 递推算法, 忽略了数据的自相关序列估计值代替理论值时的误差, 因而降低了算法性能。为了减小该误差的影响, 提出一种新的极大熵算法。该算法在计算过程中利用了一些现有算法的思路, 来提高谱估计的性能并减小计算量。计算机模拟结果证明, 这种新算法的估计性能远高于 Levinson-Durbin 算法。与同样具有高性能的 LUD(Lower and Upper triangular matrix Decomposition) 算法相比, 在低信噪比时新算法的分辨率更高, 高信噪比时, 二者性能则相似。

关键词 极大熵方法, Yule-Walker 等式, 谱分辨率

中图分类号 TN911.6

1 引 言

1967 年 Burg 提出的极大熵法 (Maximum Entropy Method, MEM) 由于把过程的自相关函数进行了无限的外推, 因此大大提高了谱分辨率。已经证明了一维极大熵谱与自回归 (AutoRegressive, AR) 谱是等效的^[1], P 阶的 AR 模型可以表示为

$$x(n) = - \sum_{k=1}^P a_k x(n-k) + w(n) \quad (1)$$

其中 $a_k (k = 1, \dots, P)$ 是 AR 系数, $w(n)$ 是均值为 0, 方差为 σ^2 的白噪声, $x(n) (n = 0, 1, 2, \dots, N-1)$ 是数据点。AR 功率谱密度可以表示为^[2]

$$S(f) = \frac{\sigma^2 \Delta t}{|1 + \sum_{k=1}^P a_k \exp(-j2\pi f k \Delta t)|^2} \quad (2)$$

其中 Δt 是取样间隔。

对于系数 a_k 的求解有许多方法^[2-7], 其中最经典的方法包括建立在 Yule-Walker 等式基础上的 Levinson-Durbin 递推算法^[2], 以及 Burg 算法和 Marple 算法^[3]等。Marple 算法由于很好地克服了 Burg 算法的谱线分裂以及谱峰偏移现象^[3], 因而具有很大的吸引力。

本文提出的新的极大熵算法, 基于经典的 Yule-Walker 等式, 但克服了其传统算法 (如 Levinson-Durbin 算法) 的不足。文中首先分析了数据自相关值的估计量代替理论值时对 Yule-Walker 等式的影响, 然后详细介绍了新的算法, 最后给出模拟计算结果和结论。仿真计算证明它具有比三种经典极大熵算法更高的谱分辨率。

2 经典 Levinson-Durbin 递推算法及其误差

对于 (1) 式具有以下经典的 Yule-Walker 等式^[2]:

¹ 1999-04-14 收到, 2000-02-24 定稿

$$\begin{bmatrix} r_{xx}(0) & r_{xx}(-1) & \cdots & r_{xx}(-P+1) \\ r_{xx}(1) & r_{xx}(0) & \cdots & r_{xx}(-P+2) \\ \vdots & \vdots & \ddots & \vdots \\ r_{xx}(P-1) & r_{xx}(P-2) & \cdots & r_{xx}(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_P \end{bmatrix} = - \begin{bmatrix} r_{xx}(1) \\ r_{xx}(2) \\ \vdots \\ r_{xx}(P) \end{bmatrix} \quad (3)$$

其中定义 $r_{xx}(k) = E[x(n)x^*(n-k)]$, $*$ 代表共轭, 则有 $r_{xx}(-k) = r_{xx}^*(k)$.

(3) 式写成矩阵形式为

$$\mathbf{r} = -\mathbf{R}\mathbf{a} \quad (4)$$

其中 $\mathbf{r} = [r_{xx}(1) \ r_{xx}(2) \ \cdots \ r_{xx}(P)]^T$, (T 表示转置), $\mathbf{a} = [a_1 \ a_2 \ \cdots \ a_P]^T$,

$$\mathbf{R} = \begin{bmatrix} r_{xx}(0) & r_{xx}(-1) & \cdots & r_{xx}(-P+1) \\ r_{xx}(1) & r_{xx}(0) & \cdots & r_{xx}(-P+2) \\ \vdots & \vdots & \ddots & \vdots \\ r_{xx}(P-1) & r_{xx}(P-2) & \cdots & r_{xx}(0) \end{bmatrix}$$

由于在实际计算中, $r_{xx}(k)$ 的理论值无法得到, 因此使用其无偏估计值来代替, 即

$$\hat{r}_{xx}(k) = \frac{1}{N-|k|} \sum_{n=|k|}^{N-1} x(n)x^*(n-k)$$

此时 (4) 式将不再满足, 考虑到 $r_{xx}(k)$ 的估计误差, (4) 式应改写为

$$\hat{\mathbf{r}} = -\hat{\mathbf{R}}\mathbf{a} + \mathbf{e} \quad (5)$$

式中 $\hat{\mathbf{r}}$ 和 $\hat{\mathbf{R}}$ 分别为 $\mathbf{r}$ 和 $\mathbf{R}$ 的估计值, $\mathbf{e}$ 为误差项。由于利用 $r_{xx}(k)$ 的无偏估计值来代替理论值, 因此误差项的均值为 0。

经典的 Levinson-Durbin 递推算法并没有考虑上述误差, 即认为 $\mathbf{e}$ 为零, 而在此基础上利用递推公式直接求解 (3) 式^[2], 显然这会大大降低谱估计的性能。

3 MEM 的新算法

由 (5) 式给出的方程组为

$$\hat{r}_{xx}(n) = - \sum_{k=1}^P a_k \hat{r}_{xx}(n-k) + e(n), \quad n = 1, 2, \cdots, P \quad (6)$$

上面的等式无论在形式上还是结构上均与 (1) 式很类似, 可以看作是“数据”序列 $\{\hat{r}_{xx}(-P+1), \hat{r}_{xx}(-P+2), \cdots, \hat{r}_{xx}(0), \hat{r}_{xx}(1), \cdots, \hat{r}_{xx}(P)\}$ 的线性预测模型, 这一点提示我们可以将最大熵谱估计方法用于 (6) 式。同时如果将 (3) 式改为超定方程 (即方程个数超过未知数个数), 则 (5) 式依然成立, 并且此时“数据”序列长度将增大, 即有

$$\hat{r}_{xx}(n) = - \sum_{k=1}^P a_k \hat{r}_{xx}(n-k) + e(n), \quad n = 1, 2, \cdots, H, \quad P \leq H \leq N-1$$

基于以上的认识, 可以推得此时“数据”序列前向线性预测误差为 (证明见附录)

$$e_f(m) = \hat{r}_{xx}(m) + \sum_{k=1}^P a_k \hat{r}_{xx}(m-k) = \sum_{k=0}^P a_k \hat{r}_{xx}(m-k), \quad m = P, \dots, L \quad (7)$$

式中取 $a_0 = 1$, $P \leq L \leq N-1$. 同时得后向线性预测误差为 (证明见附录)

$$e_b(m) = \sum_{k=0}^P a_k^* \hat{r}_{xx}(m-P+k), \quad m = P, \dots, L \quad (8)$$

这样, 前后向预测误差功率为

$$e_M = \sum_{m=P}^L |e_f(m)|^2 + \sum_{m=P}^L |e_b(m)|^2 \quad (9)$$

将 (7) 式和 (8) 式代入 (9) 式, 则为了使 e_M 极小化, 可以求 e_M 对 a_k 的微分, 并令它等于零, 得到 [8,9]

$$\frac{\partial e_M}{\partial a_i} = 2 \sum_{k=0}^P a_k R_M(i, k) = 0 \quad (10)$$

上式中 $i = 1, 2, \dots, P$, 注意 $a_0 = 1$. 同时

$$R_M(i, k) = \sum_{m=P}^L [\hat{r}_{xx}(m-k) \hat{r}_{xx}^*(m-i) + \hat{r}_{xx}^*(m-P+k) \hat{r}_{xx}(m-P+i)], \quad 0 \leq i, \quad k \leq P \quad (11)$$

最小预测误差功率为

$$e_{\min} = \sum_{k=0}^P a_k R_M(0, k) \quad (12)$$

(10) 式可写为矩阵形式

$$\begin{bmatrix} R_M(1,1) & R_M(1,2) & \cdots & R_M(1,P) \\ R_M(2,1) & R_M(2,2) & \cdots & R_M(2,P) \\ \vdots & \vdots & \ddots & \vdots \\ R_M(P,1) & R_M(P,2) & \cdots & R_M(P,P) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_P \end{bmatrix} = - \begin{bmatrix} R_M(1,0) \\ R_M(2,0) \\ \vdots \\ R_M(P,0) \end{bmatrix} \quad (13)$$

对 (13) 式求解, 就得到 AR 系数 a_k 的估计值.

(12) 式和 (13) 式可合写为矩阵形式

$$\mathbf{R}_M \mathbf{a}_M = \mathbf{E}_M \quad (14)$$

其中

$$\mathbf{R}_M = \begin{bmatrix} R_M(0,0) & R_M(0,1) & \cdots & R_M(0,P) \\ R_M(1,0) & R_M(1,1) & \cdots & R_M(1,P) \\ \vdots & \vdots & \ddots & \vdots \\ R_M(P,0) & R_M(P,1) & \cdots & R_M(P,P) \end{bmatrix}, \quad \mathbf{a}_M = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_P \end{bmatrix}, \quad \mathbf{E}_M = \begin{bmatrix} e_{\min} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

由 (11) 式知 R_M 不是 Toeplitz 矩阵, 因此不能利用 Levinson-Durbin 递推算法来计算 a_k 的值。但可将 R_M 变换为 Toeplitz 矩阵的组合, 即存在关系:

$$R_M = (T_M)^H T_M + (T_M')^H T_M' \quad (15)$$

H 表示共轭转置, T_M 是一个 $(L - P + 1) \times (P + 1)$ 的 Toeplitz 矩阵:

$$T_M = \begin{bmatrix} \hat{r}_{xx}(P) & \hat{r}_{xx}(P-1) & \cdots & \hat{r}_{xx}(0) \\ \hat{r}_{xx}(P+1) & \hat{r}_{xx}(P) & \cdots & \hat{r}_{xx}(1) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{r}_{xx}(L) & \hat{r}_{xx}(L-1) & \cdots & \hat{r}_{xx}(L-P) \end{bmatrix}$$

而 $T_M' = (T_M \cdot J)^*$, J 是 $(P + 1) \times (P + 1)$ 的交换矩阵

$$J = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \cdots & 0 & 0 \end{bmatrix}$$

由于存在 (15) 式的关系, 对 (13) 式的求解可利用 Marple 算法^[3], 或者根据 R_M 的一些特性 (如共轭对称性和中心对称性等), 利用更有效的快速算法——LUD 算法^[4] (Lower and Upper triangular matrix Decomposition, 下上三角矩阵分解法) 及其修正算法——MLUD 算法^[5] (Modified LUD, 修正 LUD)。一般来说, 后两种算法 (尤其 MLUD 算法) 更为简洁和快捷, 因此本文对 (13) 式的求解主要利用 MLUD 算法来实现。至于阶数 P 的选择, 则可以利用 (12) 式和各种判阶准则 (如 FPE 准则^[8]、AIC 准则^[8], 以及 MFPE 准则^[4] 等) 来进行阶数估计。

4 模拟结果 (算法性能的研究)

由于本文的算法是建立在 Yule-Walker 等式基础上的, 因此将与 Levinson-Durbin 递推算法进行比较。同时 Marple 算法各方面的性能均比 Burg 算法优越^[3], 所以也只需对比 Marple 算法与本文算法之间的性能。而在阶数相同时, LUD 算法又与 Marple 算法得到几乎一样的估计结果^[4], 这样本文将对新算法、Levinson-Durbin 算法和 LUD 算法之间的性能。之所以不考虑与 MLUD 算法进行对比, 是因为 MLUD 算法只是 LUD 算法的一种快速解法, 二者性能相似。

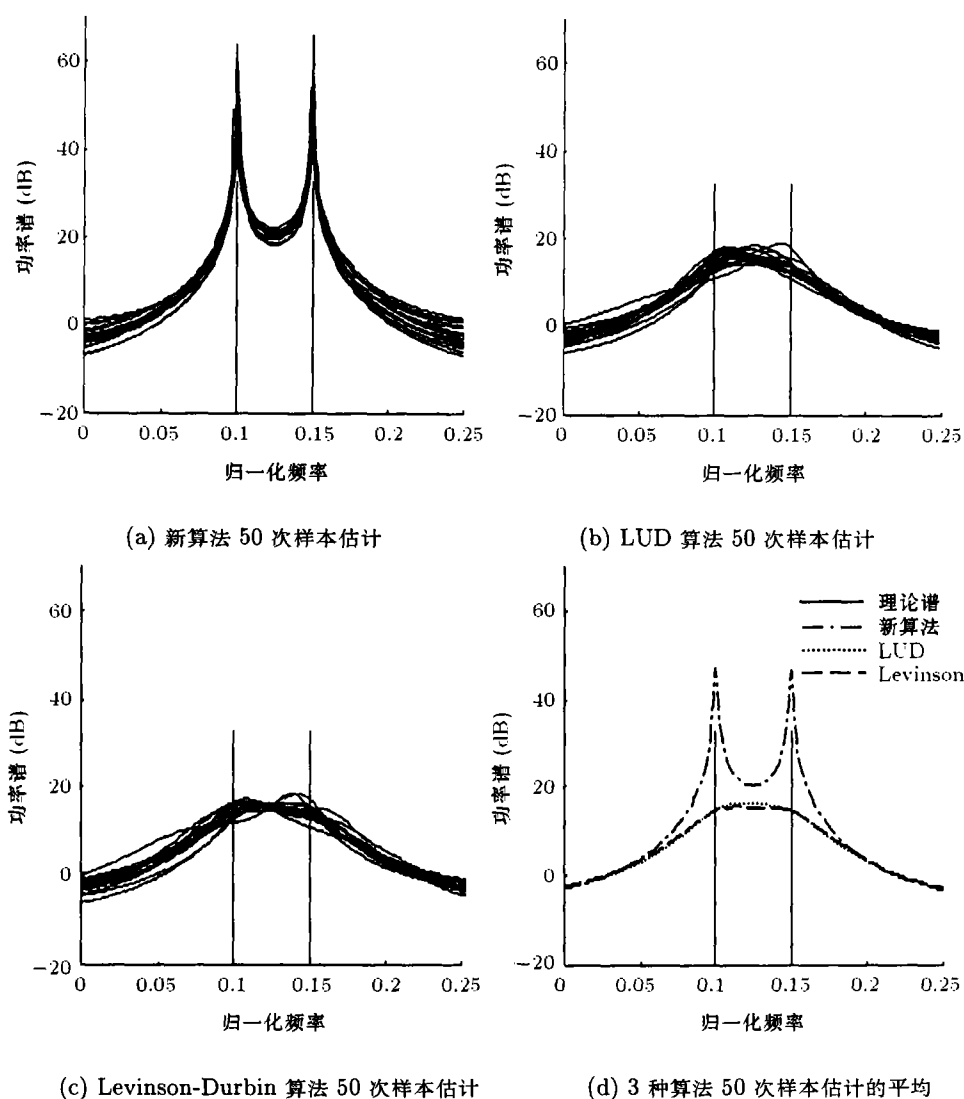
在评估谱估计算法的有效性时, 经典问题之一是检测加性白噪声中的复正弦信号^[10], 为此采用如下的数据产生模型

$$x(n) = A_1 \exp(2\pi f_1 n) + A_2 \exp(2\pi f_2 n) + w(n), \quad n = 0, 1, \cdots, N-1 \quad (16)$$

其中 $w(n)$ 是均值为 0, 方差为 1 的高斯白噪声。

需要指出的是, 对于谱估计这种典型的统计估计问题, 单次样本的仿真结果是不能说明问题的, 因此在对比各算法间性能时, 采用对 50 次样本的估计求平均的方法来进行比较和说明。同时为了尽可能在相同的情况下来评估各算法, 算法中阶数 P 是统一给定的。

实例 1 $A_1 = \sqrt{2}$, $f_1 = 0.10$; $A_2 = \sqrt{2}$, $f_2 = 0.15$, 即信号 1 的信噪比 (用 SNR1 表示) 为 0 dB, 信号 2 的信噪比 (用 SNR2 表示) 为 0 dB。数据长度 $N=64$, 阶数 $P=6$, $L=7P$ 。图 1 是估计结果。

图 1 3 种算法对白噪声中两个复正弦信号的分辨 ($\text{SNR}_1=\text{SNR}_2=0\text{dB}$)

实例 2 $f_1 = 0.10$, $f_2 = 0.15$, 即依然采用图 1 中信号的位置, 但是调整两个信号各自的信噪比。数据长度 $N=32$, 阶数 $P=4$, $L = 4P$ 。图 2 是估计结果, 限于篇幅, 图中仅给出 50 次样本估计的平均, 以下均如此。

图 1 的结果说明在低信噪比时, 新算法不仅具有远高于 LUD 算法和 Levinson-Durbin 递推算法的谱分辨率, 而且还具有较小的估计方差。当提高信噪比, 并减少数据点 N 和降低阶数 P 时, 图 2(a) 的结果显示, 新算法依然有良好的分辨率, 而 LUD 分辨率也有了提高, 但 Levinson-Durbin 算法还不能分辨出较弱的信号。进一步增大弱信号的信噪比后, 图 2(b) 说明, 此时三者分辨率相近。

实例 1 和实例 2 主要研究了不同信噪比情况下, 三种算法分辨率的比较, 为此采用的信号频率间隔较宽。下面研究频率间隔较窄时, 三者的性能比较。

实例 3 $f_1 = 0.10$, $f_2 = 0.12$, 信噪比参见图中说明。 $N=64$, $P=5$, $L = 3P$ 。图 3 是估计结果。

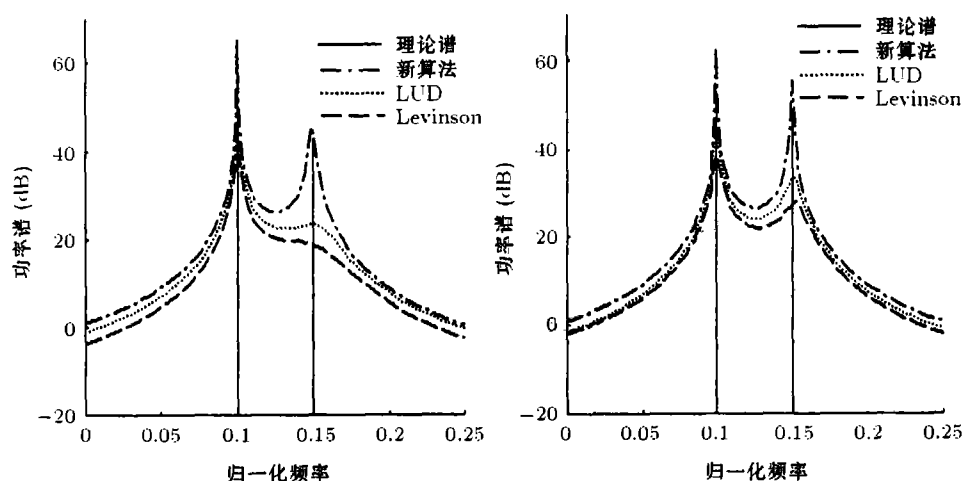
(a) $\text{SNR}_1=20\text{dB}$, $\text{SNR}_2=10\text{dB}$ (b) $\text{SNR}_1=20\text{dB}$, $\text{SNR}_2=15\text{dB}$

图 2 3 种算法分辨率比较

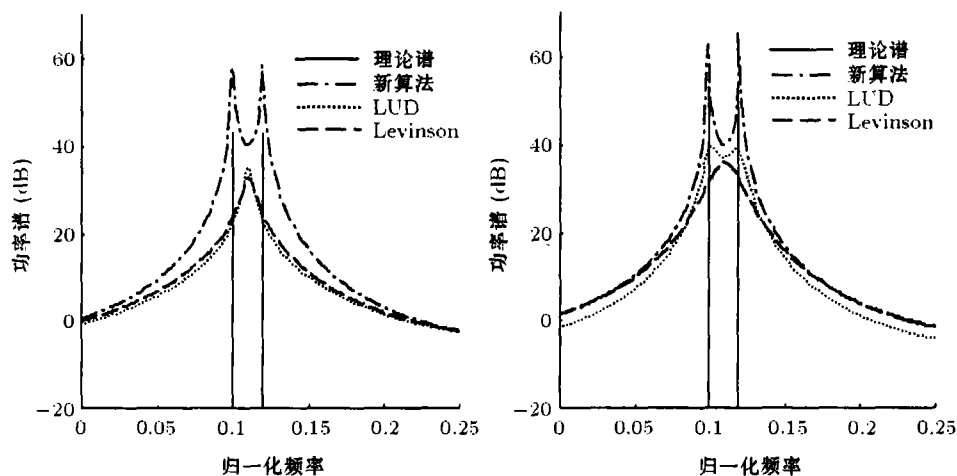
(a) $\text{SNR}_1=\text{SNR}_2=10\text{dB}$ (b) $\text{SNR}_1=\text{SNR}_2=20\text{dB}$

图 3 3 种算法分辨率比较

图 3(a) 的模拟结果说明在窄频率间隔、低信噪比时, 新算法的分辨率仍要高于 LUD 算法和 Levinson-Durbin 算法。在增大信噪比后, 尽管前两种算法已能明显分辨出信号 (新算法的谱峰更尖锐), 但是 Levinson-Durbin 算法仍无法分辨。

3 个仿真实例的结果表明, Levinson-Durbin 算法只有在信号间隔较大且有高信噪比时, 才具有与新算法相同的分辨率, 显然它的估计性能要明显低于后者。实验中还发现, 如果在 N 和 P 不变的情况下, 增大 L 的取值, 一般来说新算法会有更好的谱估计性能。

5 结 论

本文介绍的最大熵谱估计的新算法, 从经典的 Yule-Walker 等式出发, 但其估计性能要远高于传统的 Levinson-Durbin 递推算法。与高性能的 LUD 算法 (或 Marple 算法) 相比, 在

低信噪比时新算法的分辨率更高; 高信噪比时, 二者性能则相似。由于新算法首先需要计算数据的自相关值, 并在此基础上利用 MLUD 算法来计算 AR 系数, 因此它的缺点在于计算量较大 (近似与 NL 成正比), 但这可以通过采用自相关值的快速算法来给予一定的克服, 如文献 [11] 中的算法。

附 录

对文中 (7) 式和 (8) 式的成立进行证明。

对于 (1) 式这样一个 AR 过程, 前向线性预测误差为 [3]

$$e_{f1}(m) = x(m) + \sum_{k=1}^P a_k x(m-k) = \sum_{k=0}^P a_k x(m-k), \quad m = P, \dots, N-1 \quad (\text{A1})$$

式中取 $a_0 = 1$ 。后向线性预测误差为 [3]

$$e_{b1}(m) = \sum_{k=0}^P a_k^* x(m-P+k), \quad m = P, \dots, N-1 \quad (\text{A2})$$

(A1) 式两边分别乘以 $x^*(0)$, 并求数学期望, 得

$$e_{f2}(m) = \sum_{k=0}^P a_k r_{xx}(m-k), \quad m = P, \dots, N-1 \quad (\text{A3})$$

当 (A3) 式中的 $r_{xx}(k)$ 用其无偏估计值 $\hat{r}_{xx}(k)$ 来代替时, 有

$$e_f(m) = \sum_{k=0}^P a_k \hat{r}_{xx}(m-k), \quad m = P, \dots, N-1 \quad (\text{A4})$$

在计算中为了减小运算量, (A4) 式中的 m 的定义域可做适当调整, 即

$$e_f(m) = \sum_{k=0}^P a_k \hat{r}_{xx}(m-k), \quad m = P, \dots, L \quad (\text{A5})$$

其中 $P \leq L \leq N-1$ 。(A5) 式即为文中的 (7) 式。

证毕

同理利用 (A2) 式进行相似的推导, 即可得到文中 (8) 式。

参 考 文 献

- [1] A. Van Den Bos, Alternative interpretation of maximum entropy spectral analysis, IEEE Trans. on Information Theory, 1971, IT-17(4), 493-494.
- [2] S. M. Kay, S. L. Marple, Spectrum analysis—A modern perspective, Proc. IEEE, 1981, 68(11), 1380-1419.
- [3] S. L. Marple, A new autoregressive spectrum analysis algorithm, IEEE Trans. on Acoustics, Speech, and Signal Processing, 1980, ASSP-28(4), 441-454.
- [4] Lin Daimao, Mao Yuhai, A fast algorithm of maximum entropy method for spectral estimation suitable for short data sequence, Scientia Sinica, Series A, 1984, XXVII(2), 196-212.
- [5] 李武皋, 茅于海, 一种新的 MEM 快速算法, 电子学报, 1989, 17(4), 19-24.

- [6] S. M. Kay, Noise compensation for autoregressive spectral estimation, IEEE Trans. on Acoustics, Speech, and Signal Processing, 1980, ASSP-28(3), 292-303.
- [7] W. Campbell, D. N. Swingler, Frequency estimation performance of several weighted Burg algorithms, IEEE Trans. on Signal Processing, 1993, 41(3), 1237-1247.
- [8] S. M. Kay, Modern Spectral Estimation: Theory and Application, Englewood Cliffs, NJ., Prentice-Hall, 1988, Chapter 7.
- [9] 陆明泉, 肖先赐, 李乐民, 从 GAR 模型参数提取特征的数字调制识别新方法, 电子科学学刊, 1999, 21(2), 145-151.
- [10] J. A. Cadzow, High performance spectral estimation — a new ARMA method, IEEE Trans. on Acoustics, Speech, and Signal Processing, 1980, ASSP-28(5), 524-529.
- [11] C. M. Rader, An improved algorithm for high speed autocorrelation with applications to spectral estimation, IEEE Trans. on Audio and Electroacoustics, 1970, AU-18(4), 439-441.

A KIND OF MAXIMUM ENTROPY SPECTRAL ESTIMATION ALGORITHM WITH HIGH RESOLUTION

Huang Yong Peng Yingning Wang Xiutan

(Department of Electronic Engineering, Tsinghua University, Beijing 100084, China)

Abstract Based on the Yule-Walker equations, a traditional Maximum Entropy Method (MEM) is the Levinson-Durbin recursion algorithm. The algorithm omits errors that are produced when estimated values of autocorrelation sequence take the place of their ideal values. A new algorithm is presented in order to decrease the errors as much as possible. The algorithm improves its performance of spectral estimation and reduces computational complexity by means of the clues of some existing algorithms. The results of simulation demonstrate that the new algorithm has much more excellent performance of spectral estimation than that of the Levinson-Durbin algorithm. Compared with the LUD (Lower and Upper triangular matrix Decomposition) algorithm which also has high estimation performance, the new algorithm has higher resolution under low Signal-to-Noise Ratio(SNR) condition, or same resolution under high SNR.

Key words Maximum Entropy Method (MEM), Yule-Walker equations, Spectral resolution

黄 勇: 男, 1972 年生, 博士生, 主要研究方向包括雷达信号处理、谱估计、神经网络、无线通信等。

彭应宁: 男, 1939 年生, 教授, 博士生导师, 目前的研究方向为: 谱估计、自适应滤波、自适应阵列信号处理、并行信号处理、雷达信号处理、神经网络与子波变换等。

王秀坛: 女, 1947 年生, 副教授, 主要研究方向为: 高速实时信号处理、微光夜视图像处理、雷达信号处理等。