第 25 卷 第 6 期 电子与信息学报 Vol.25 No.6
2003 年 6 月 JOURNAL OF ELECTRONICS AND INFORMATION TECHNOLOGY Jun. 2003

一种改进的分段平稳随机过程的参数估计方法 [1]

陈 颖 李在铭 *

(中国电子科技集团公司第十研究所 成都 610036)
*(电子科技大学通信信息学院 成都 610054)

摘 要 将非平稳随机信号划分为分段平稳随机信号进行处理，为非平稳随机信号的研究提供的一种新的分析方法．为最优地将非平稳随机信号划分为分段平稳随机信号， Djuric 等人用 Bayes 方法，通过递推多维条件分布概率来估计最优划分参数值，但计算相当复杂．本文在研究 AR 模型本身的一些特性的基础上，通过直接递推多维联合分布概率来估计最优划分参数，大大地减少了计算量．

关键词 非平稳随机信号，分段平稳随机信号， AR 模型，参数估计
中图号 TN911.23

1 引 言

一些非平稳随机信号可以划分为分段平稳随机信号来加以研究．这类信号的统计特性在一些未知时刻发生突变，而在这些时刻之间又保持平稳性．我们将这些未知时刻称为分界点，将这类信号称为分段平稳随机信号．

研究分段平稳随机信号需要解决的问题有： (1) 估计出分段的段数； (2) 估计出各段之间的分界点； (3) 估计出各平稳段的参数模型．文献 [1] 采用 Bayes 方法对此问题进行了研究，给出了一个关于分段数、各段 AR 模型阶数及各分界点的优化方程．但该方程求解非常复杂，计算量特别大．文献 [2, 3] 对该优化方程做了进一步的研究，给出了一个递推公式，从而减少了一定的计算量．但作者是通过递推数据的条件分布概率来得到结果，忽略了 AR 模型本身的一些特性，因此计算量仍然较大．本文通过研究 AR 模型本身的一些特性，给出一种新的递推算法，按数据的联合分布概率来递推，从而大大减小了计算量．

2 分段平稳随机过程的 Bayes 估计

设数据矢量 $Y=[y(1),y(2),\cdots,y(n)]$ 为分段平稳随机信号的观测值．假设 Y 的最大分段数为 Q 段，实际分为 q 段， $Q\geq q$ ，且将 Y 分段后的各段统计独立．各段用 $S_1,S_2,\cdots,S_q$ 表示，对应的信号模型为 $M^{s_1},M^{s_2},\cdots,M^{s_q}$，而 M^{s_m}，$m=1,2,\cdots,q$ 是阶数为 p_m 的 AR 模型，即

$$y(n)=a_1y(n-1)+a_2y(n-2)+\cdots+a_{p_m}y(n-p_m)+w(n) \tag{1}$$

式中 $w(n)$ 为零均值、方差为 σ_m 的正态白噪声．分段中的分界点为 $j_1,j_2,\cdots,j_{q-1}$，$j_1<j_2<\cdots<j_{q-1}$ ．假设最大可能的阶数为 P，$P\geq p_m$，$m=1,2,\cdots,q$ ．现要求解的是最佳分段数 q ，分界点 $J_{q-1}=[j_1,j_2,\cdots,j_{q-1}]^T$ 及模型阶数 $P_q=[p_1,p_2,\cdots,p_q]^T$ ．

设 H_k 为有 k 个分段的假设，则由 Bayes 定理有后验概率 $P(H_k,J_{k-1},p_k|Y)$ 为

$$P(H_k,J_{k-1},p_k|Y)=\frac{f(Y|H_k,J_{k-1},p_k)P(H_k,J_{k-1},p_k)}{f(Y)} \tag{2}$$

式中 $f(Y)$ 为数据的边缘概率密度， $P(H_k,J_{k-1},p_k)$ 为先验概率， $f(Y|H_k,J_{k-1},p_k)$ 为数据的

[1] 2001-11-29 收到， 2002-07-28 改回

条件概率密度.

在 H_k 的假设下，可认为 J_{q-1} 与 p_k 相互独立，则有

$$P(H_k, J_{k-1}, p_k) = P(J_{k-1}|H_k)P(p_k|H_k)P(H_k) \tag{3}$$

按文献 [2] 假定的先验概率有

$$P(H_k) = 1/Q, \qquad k = 1, 2, \cdots, Q \tag{4}$$

$$P(J_{k-1}|H_k) = \frac{(k-1)!}{\prod_{l=1}^{k-1}(N-k-l)} \tag{5}$$

$$P(p_k|H_k) = 1/(P+1)^k \tag{6}$$

由假设各段统计独立，有

$$f(Y|H_k, J_{k-1}, p_k) = \prod_{i=1}^{k} f(Y_{j_{i-1}, j_i - 1}|p_i) \tag{7}$$

式中 $Y_{j_{i-1}, j_i-1} = [y(j_{i-1}), y(j_{i-1}+1), \cdots, y(j_i - 1)]^T$.

根据最大后验概率估计的最优准则，综合以上各式，可得出关于求解分段数、各段 AR 模型阶数及各段之间的分界点估值 $(\hat{q}, \hat{P}_q, \hat{J}_{q-1})$ 的优化方程为

$$\begin{aligned}(\hat{q}, \hat{P}_q, \hat{J}_{q-1}) = \underset{k, J_{k-1}, p_k}{\arg\min} \Bigg[& -\sum_{i=1}^{k} \ln(f(Y_{j_{i-1}, j_1-1}|p_i)) \\ & + \ln\frac{\prod_{l=1}^{k-1}(N-k-l)}{(k-1)!} + k\ln(P+1)\Bigg]\end{aligned} \tag{8}$$

式中 N 表示整个数据段的长度. (8) 式的求解是非常复杂的，主要困难在于求解

$$\underset{k, J_{k-1}, p_k}{\arg\min}\left[-\sum_{i=1}^{k}\ln(f(Y_{j_{i-1}, j_1-1}|p_i))\right] \tag{9}$$

对于 (9) 式的求解，文献 [4] 给出了按数据条件分布概率计算的递推公式为

$$\begin{aligned} f(y(l)|Y_{j_{i-1},j}, p_i > 0) = & \frac{1}{\sqrt{2\pi}} \frac{|H_{p_k}^T(l-2)H_{p_k}(l-2)|^{1/2}}{|H_{p_k}^T(l-1)H_{p_k}(l-1)|^{1/2}} \\ \times & \frac{\Gamma\left(\dfrac{l-2p_k-j_{i-1}+1}{2}\right)}{\Gamma\left(\dfrac{l-2p_k-j_{i-1}}{2}\right)} \times \frac{(Y_{p_k+j_{i-1}, j_i-1}^T P_{p_k}(l-2) Y_{p_k+j_{i-1}, j_i-1})^{(l-2p_k-j_{i-1})/2}}{(Y_{p_k+j_{i-1}, l}^T P_{p_k}(l-1) Y_{p_k+j_{i-1}, l})^{(l-2p_k-j_{i-1}+1)/2}} \end{aligned} \tag{10}$$

其中

$$\Gamma(x) = \int_0^{\infty} t^{x-1}\exp(-t)\mathrm{d}t, \qquad t > 0$$

$$P_{p_k}(l) = I - H_{p_k}(l)(H_{p_k}^T(l)H_{p_k}(l))^{-1}H_{p_k}^T(l)$$

$$H_{p_k}(l) = \begin{bmatrix} y(p_k + j_{i-1} - 1) & y(p_k + j_{i-1} - 2) & \cdots & y(j_{i-1}) \\ y(p_k + j_{i-1}) & y(p_k + j_{i-1} - 1) & \cdots & y(j_{i-1} + 1) \\ \vdots & \vdots & \ddots & \vdots \\ y(l) & y(l-1) & \cdots & y(l - p_k + 1) \end{bmatrix}$$

文献 [2] 在 (10) 式的基础上，做了进一步的研究，给出了计算 (10) 式的一个递推公式，并在此基础上做了数值仿真工作．但 (10) 式的计算量仍然很大，主要原因是因为它是通过递推多维联合分布概率的条件概率来计算 (9) 式的．事实上，我们可以直接通过递推多维联合分布概率 $f(Y_{j_{i-1},j_i-1}|p_i)$ 来计算 (9) 式．

3 改进的计算方法

假设分段为 H_k，各段统计独立且可用 AR 模型表示，$w(n)$ 为零均值正态白噪声，则根据 AR 模型特性，第 i 个分段 S_i 的数据序列 $Y_{j_{i-1},j_i-1}=[y(j_{i-1},y(j_{i-1}+1),\cdots,y(j_i-1)]^T$ 服从多维正态分布，即

$$f(Y_{j_{i-1},j_i-1}|p_i)=(2\pi)^{-m/2}|\Gamma_m|^{-1/2}\exp\left(-\frac{1}{2}Y^T_{j_{i-1},j_i-1}\Gamma_m^{-1}Y_{j_{i-1},j_i-1}\right) \tag{11}$$

其中 m 为向量 Y_{j_{i-1},j_i-1} 中数据的个数，$m=j_i-1-j_{i-1}$，Γ_m 为 Y_{j_{i-1},j_i-1} 的自协方差阵．

分段 S_i 的最佳阶数 p_i 应由 AIC(Akaike Information Criterion) 准则确定．因此，当段 S_i 的数据序列 Y_{j_{i-1},j_i-1} 给定时，其最佳阶数 p_i 也就确定了，所以

$$f(Y_{j_{i-1},j_i-1}|p_i)=f(Y_{j_{i-1},j_i-1}) \tag{12}$$

根据 AIC 准则，最佳阶数 p_i 应使下式最小：

$$\mathrm{AIC}(p_i)=-2\ln f(Y_{j_{i-1},j_i-1}|p_i)+2p-m-m\ln(2\pi) \tag{13}$$

(11) 式的计算仍然较复杂．下面，我们考虑它的递推计算．

引理 1　设序列 $\{X(m),m=1,2,\cdots\}$ 是均值为零、自协方差为 $\gamma(i,j)$ 的随机序列．如果对一切 $m=1,2,\cdots$，协方差阵 $[\gamma(i,j)]_{i,j=1}^m$ 是非奇异的，则一步预测 $\hat{X}_{m+1}$ 和预测均方误差 v_m 可用下式计算：

$$\left.\begin{aligned}
&\hat{X}_{m+1}=\begin{cases}0, & m=0\\ \displaystyle\sum_{j=1}^{m}\theta_{m,j}(X_{m+1-j}-\hat{X}_{m+1-j}), & m\ge 1\end{cases}\\
&v_0=\gamma(1,1)\\
&\theta_{m,m-k}=v_k^{-1}\left(\gamma(m+1,k+1)-\sum_{j=0}^{k-1}\theta_{k,k-1}\theta_{m,m-j}v_j\right),\quad k=0,1,\cdots,m-1\\
&v_m=\gamma(m+1,m+1)-\sum_{j=0}^{m-1}\theta^2_{m,m-j}v_j
\end{aligned}\right\} \tag{14}$$

该引理的证明可见文献 [5]．

在引理 1 的基础上，我们可递推求出序列 Y_{j_{i-1},j_i-1} 的联合概率密度．为表示上的方便，我们用序列 $X(m)=[x_1,x_2,\cdots,x_m]^T$ 表示序列 Y_{j_{i-1},j_i-1}．令 $\theta_{i,0}=1$，$\theta_{i,j}=0$，当 $j<0$，$i=0,1,\cdots,m$．定义 $m\times m$ 下三角阵 C 及 $m\times m$ 对角阵 D 为

$$C=[\theta_{i,i-j}]_{i,j=0}^{m-1},\qquad D=\mathrm{diag}(v_0,v_1,\cdots,v_{m-1}) \tag{15}$$

则根据 (14) 式可将一步预测 $\hat{X}_{m+1}$ 改写为 $\hat{X}_m = (C-I)(X_m - \hat{X}_m)$，其中 I 为 $m \times m$ 单位阵．则有

$$X_m = X_m - \hat{X}_m + \hat{X}_m = C(X_m - \hat{X}_m) \tag{16}$$

显然，有对角阵 D 为 $(X_m - \hat{X}_m)$ 的协方差阵，因此有

$$\Gamma_m = E(X_m X_m^T) = CE((X_m - \hat{X}_m)(X_m - \hat{X}_m)^T)C^T = CDC^T \tag{17}$$

根据 (16) 和 (17) 式有

$$\begin{aligned} X_m^T \Gamma_m^{-1} X_m &= (X_m - \hat{X}_m)^T D^{-1} (X_m - \hat{X}_m) = \sum_{j=1}^{m} (X_j - \hat{X}_j)^2 / v_{j-1} \\ |\Gamma_m| &= |C||D||C^T| = v_0 v_1 \cdots v_m \end{aligned} \tag{18}$$

因此，多维正态概率密度 (11) 式可按下式递推：

$$f(x_1, x_2, \cdots, x_m) = (2\pi)^{-1/2} (v_0 v_1 \cdots v_m)^{-1/2} \exp\left(-\frac{1}{2}\sum_{j=1}^{m}(X_j - \hat{X}_j)^2 / v_{j-1}\right) \tag{19}$$

引理 1 的递推预测算法既适合非平稳随机过程，又适合平稳随机过程，但它必须预先知道数据序列的协方差阵 $[\gamma(i,j)]_{i,j=1}^m$ ．对于可用 AR 模型表示的平稳随机过程，我们可采用更简单的递推预测算法，如 Durbin-Levinson 算法．

引理 2 (Durbin-Levinson 算法) 设序列 $\{X(m), m=1,2,\cdots\}$ 是均值为零、自协方差为 $\gamma(\cdot)$ 的平稳随机过程，可用阶数为 p_m 的 AR 模型表示．则一步预测 $\hat{X}_{m+1}$ 和预测均方误差 v_m 可用下式递推计算：

$$\begin{aligned} \hat{X}_{m+1} &= \hat{a}_{p,1} X_m + \hat{a}_{p,2} X_{m-1} + \cdots + \hat{a}_{p,p} X_{m-p} \\ v_0 &= \gamma(0) \\ \hat{a}_{p,p} &= \left[\gamma(n) - \sum_{j=1}^{p-1} \hat{a}_{p-1,j}\gamma(p-j)\right] v_{p-1}^{-1} \\ \begin{bmatrix} \hat{a}_{p,1} \\ \vdots \\ \hat{a}_{p,p-1} \end{bmatrix} &= \begin{bmatrix} \hat{a}_{p-1,1} \\ \vdots \\ \hat{a}_{p-1,p-1} \end{bmatrix} - \hat{a}_{pp} \begin{bmatrix} \hat{a}_{p-1,p-1} \\ \vdots \\ \hat{a}_{p-1,1} \end{bmatrix} \\ v_p &= \begin{cases} v_{p-1}[1 - \hat{a}_{p,p}^2], & \text{当 } p \le p_m \\ v_{p_m}, & \text{当 } p > p_m \end{cases} \end{aligned} \tag{21}$$

在引理 2 的基础上，我们可以将 (20) 和 (21) 式计算出来的预测值 $\hat{X}_{m+1}$ 及预测均方误差 v_m 代入 (19) 式来递推计算多维联合正态分布概率密度 $f(x_1, x_2, \cdots, x_m)$ ．

综上所述，可以采用搜索法寻找 (8) 式的最优值．具体算法如下：

第 1 步 先令 $k=1$ ，即将整个数据序列作为一段，改变 AR 模型阶数，计算不同阶数情况下，(8) 式及 (13) 式的值，取使 AIC 值最小的阶数作为模型阶数，记下与此模型阶数对应的 (8) 式的值，以便最后比较．

第 2 步　其次令 $k=2$, 将整个数据序列分为 2 段，这时需改变前后两段 AR 模型的阶数 p_1,p_2 及分界点 j 的位置。让 $j_1=j_0+2,\ j_0+3,\cdots,N-2(j_0=1)$; $p_1=0,1,\cdots,P$; $p_2=0,1,\cdots,P$, 逐点搜索，按第 1 步类似的方法计算、取值、比较，得出最小值.

第 3 步　依此类推，令 $k=3,4,\cdots,Q$, 按上述方法进行计算、取值、比较，记下最小值.

第 4 步　完成上述搜索后，其中最小值所对应的分段数、模型阶数及分界点即为所求.

4 计算机仿真实验

我们采用与文献 [2] 相同的模型．它是由两段可用 AR 模型表示的随机过程合在一起形成的非平稳随机信号，其中第一段为下列 AR(2) 模型:

$$y(n)=1.37y(n-1)-0.56y(n-2)+w(n) \tag{22}$$

第二段为下列 AR(1) 模型:

$$y(n)=0.8y(n-1)+w(n) \tag{23}$$

(22) 和 (23) 式中 $w(n)$ 为零均值、单位方差的正态白噪声.

实验中，设 AR 模型最大阶数 $P=3$，最大可能的分段数为 Q=3．采样点 N=180，分界点 j=101, 即第一段 AR 模型产生 100 个采样点，第二段模型产生 80 个采样点．实验采用 MATLAB 语言编程，在奔腾 166 微机上运行，一次实验所需的时间约为 40s．重复实验 100 次，实验结果如表 1 所示.

表 1　100 次重复实验中分段数出现的次数

分段数 k	1	2	3
出现的次数	0	92	8

表 2　k=2 时，分界点在各区间出现的次数

采样点区间	2-85	86-95	96-100	101-105	106-115	116-180
出现次数	0	3	36	51	2	0

实验结果表明，算法的效果是令人满意的．图 1 和图 2 给出了一次实验所产生的随机数据及当 $k=2$，分界点 j_1 从 20 ~160 变化时，计算出的 $F=\sum_{i=1}^{k}\ln(f(Y_{j_{i-1},j_1-1}|p_i))$ 值的变化图．从图 2 可知，在分界点 j_1 =101 处，F 取得最大值为 −170．

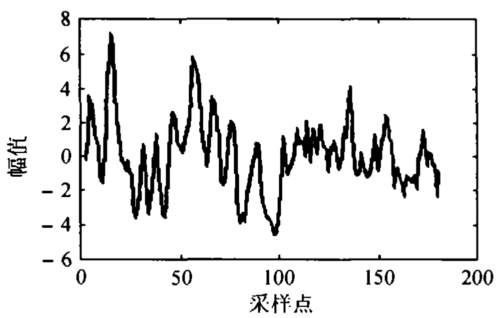

图 1　一次实验产生的随机数据图

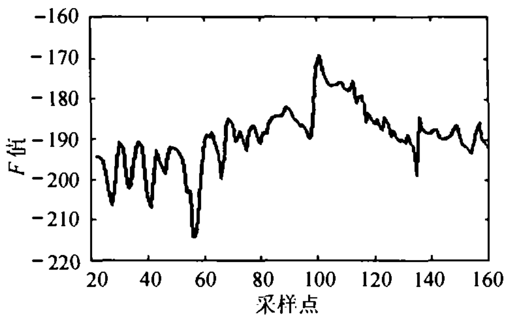

图 2　一次实验计算出的 F 值

5 结 论

针对满足一定条件的 AR 模型产生的数据服从多维联合正态分布这一事实，在最优估计分段平稳随机信号的相关参数下，采用直接递推多维联合正态分布概率密度来估计最优划分参数值比采用递推多维条件分布概率密度来估计要简单得多.

参 考 文 献

[1] P. M. Djuric, *et al.*, Segmentation of nonstationary signal, Proc. of IEEE ICASSP, San Franciso, 1992,5, 161-164.

[2] 王文华等，分段平稳随机过程的参数估计方法，电子科学学刊， 1997, 19(3), 311-317.

[3] 王宏禹，非平稳随机信号分析与处理，北京，国防工业出版社， 1999, 230-244.

[4] P. M. Djutic, *et al.*, Order selection of autoregressive models, IEEE Trans. on Signal Processing, 1992, 40(11), 2829-2833.

[5] 田铮，动态数据处理的理论与方法—时间序列分析，西安，西北工业大学出版社， 1995, 89-92.

AN ADVANCED METHOD TO ESTIMATE PARAMETERS OF PIECEWISE STATIONARY STOCHASTIC PROCESS

Chen Ying Li Zaiming*

(*No.10th Research Institute, CETC, Chengdu* 610036, *China*)

*(*Inst. of Communication and Information, UEST of China, Chengdu* 610054, *China*)

Abstract A new way to analysis nonstationary stochastic process is to divide it into piecewise stationary stochastic process. Djuric(1992) used Bayes method to estimate the parameters, which can optimally divide the nonstationary stochastic process into stationary stochastic process. Some authors estimated the optimum parameters through calculating recursively the multivariate conditional likelihood function, which made the computation very complex. Basing on some natural characteristics of AR mode, a new recursive method is provided, which can improve the computation efficiently, to estimate the optimum parameters.

Key words Nonstationary stochastic signal, Piecewise stationary stochastic signal, AR model, Parameters estimating

陈 颖： 男， 1973 年生，讲师，博士，研究方向包括通信信号处理，随机信号处理，图像信号分析等.

李在铭： 男， 1939 年生，教授，博士生导师，研究方向为多媒体通信网及信号处理技术.