基于子空间迭代的 Prony 模型参数估计

谢维波[1], 王永初[2]
(1.华侨大学计算机系,福建泉州 362021;2.华侨大学机电学院,福建泉州 362021)

摘 要: 在Prony模型满足充分必要条件的非线性最小二乘最优解的基础上,给出了Prony模型基于子空间迭代的非线性最小二乘参数估计(ISNLSE).基于对最优解几何结构的认识,导出了一个合理的"收敛控制条件",并构造了一个充分有效的算法;既加深了对问题求解的认识,又大大地改进了算法的收敛性和有效性.最后,用一个简单的例子阐明这一迭代过程,其结果和"扩充的ESPRIT算法"的结果作了比较.

关键词: 非线性最小二乘估计; 子空间迭代; Prony模型
中图分类号: TN957 文献标识码: A 文章编号: 0372-2112(2008)06-1206-04

Iterative Subspace Method for Parameter Estimation of Prony Model

XIE Wei-bo[1], WANG Yong-chu[2]
(1. Department of Computer Science, Huaqiao University, Quanzhou, Fujian 362021, China;
2. Colleges of Mechanics and Automation, Huaqiao University, Quanzhou, Fujian 362021, China)

Abstract: An iterative subspace method for parameters estimation of Prony model (superimposed exponential signals in noise) in the sense of nonlinear least-squares is presented. With condition meeting in the optimization is both necessary and sufficient, such that the solution is sole and globally optimal. The optimal condition is interpreted in the geometric language of abstract vector spaces. In the foundation, a fully effective iterative algorithm is acquired. Thus, both the understanding for the solution is deepened and the convergence property and effectiveness is greatly improved. Finally, the procedure is illustrated with a simple example, and the result compared with one's of Pro-ESPRIT method.

Key words: nonlinear least-squares estimation; iterative subspace; Prony model

1 引言

早在1795年,Prony提出了用"复指数函数"的线性组合来描述等间隔采样数据的数学模型,常称为"Prony模型".Prony模型作为Fourier级数的一种拓展,在理论和应用上都有十分重大的意义;作为一种线谱估计方法,广泛适用于雷达、水声、声音和脉冲响应分析等领域的信号处理[1~3].如果能把该模型成功地应用于雷达信号处理,将从根本上提高雷达目标的分辨率.基于该模型的谱估计具有短时数据的特点,这在某些应用场合是特别重要的[2].然而,Prony模型的严格最优求解是一个困难的非线性问题、是一个经典的信号处理问题.其模型表达如下:

$$f(k)=\sum_{n=0}^{N-1}A_nZ_n^k+\varepsilon_k \tag{1}$$

其中$f(k)$为均匀采样的观测数据($k=0,\cdots,M$),ε_k为高斯白噪声,N为模型阶数;未知的复数A_n和Z_n分别为信号的第n个留数和极点.式(1)精确的非线性最小二乘辨识,通过最小化误差函数Q来实现.本文仅考虑$f(k)$、ε_k为实数的情况,相应地A_n和Z_n将分别成共轭复数对出现.差函数Q表达如下:

$$Q=\sum_{k=0}^{M}\varepsilon_k^2=\sum_{k=0}^{M}\left(\sum_{n=0}^{N-1}A_nZ_n^k-f(k)\right)^2 \tag{2}$$

式(2)的最小化求解涉及高度复杂的非线性.长期以来对于式(1)的研究是换个角度来进行的.人们从"线性化处理"及"依据信号的统计量"两方面,关于式(1)提出了许多求解算法;代表性的算法分别是:扩充的Prony方法、扩充的ESPRIT方法;存在的问题分别是:线性化处理使估计性能变差、难于获取信号的真实统计量.事实上,Prony模型在数值上是病态的,其性能严重地依赖于信噪比;这一点极大地限制了Prony模型的应用.尽管ESPRIT方法是精确优化的且十分完美,信号统计量的估计误差影响了它的精度;为此提出了类似ESPRIT方法的"直接矩阵束方法",在工程上得到广泛应用;"直接矩阵束方法"不需要估计信号的相关函数或协方差[4].但是,"直接矩阵束方法"存在这样的缺陷:①具有

收稿日期:2006-12-25;修回日期:2008-03-31
基金项目:福建省自然科学基金(No. A0540005),国务院侨务办公室科学研究基金(No. 03QZR06)

“线性化处理”的特点, ②缺少合理的优化准则.

从非线性处理的角度, 在“仅仅满足最优化必要条件”的前提下: 文献[5]依据“最大似然准则”给出了式(1)模型的一维搜索迭代求解方案, 存在算法收敛的不确定性及解的唯一性问题; 文献[6]在研究了最优解应满足的一些性质之后构造了一种迭代方案, 由于没能给出“最优化的充分性条件”, 同样存在算法收敛的不确定性及解的唯一性问题. 文献[7]给出了式(1)模型满足“充分必要条件”的非线性最小二乘解的代数表达及迭代算法, 为研究解的几何结构提供了可能; 文献[8]给出了文献[7]中代数表达的几何结构及更新的迭代算法, 为定量描述该结构的特征奠定了基础. 本文依据对最优解几何结构的认识, 构造了一个充分有效的基于子空间迭代的数值算法. 算法的性能是: 在干扰噪声高达 10dB 或 0dB 的情况下, 计算精度达到 0.5‰, 1000 次独立的数值试验中有 1~2 次不能收敛. Prony 模型的应用已成为现实.

2 Prony 模型的最优解[7,8]

2.1 最优解的代数表达[7]

定义

$$\sum_{m=0}^{N} b_m Z^m = \prod_{n=0}^{N-1} (Z - Z_n) \tag{3}$$

$$\sum_{m=0}^{2N} B_m Z^m = \prod_{n=0}^{N-1} (Z - Z_n)^2 \tag{4}$$

维数为 $(M+1)\times(M+1-N)$ 的矩阵 $\boldsymbol{b}$

$$\boldsymbol{b} = \begin{pmatrix} b_0 & b_1 & \cdot & \cdot & b_N & & & & \\ & b_0 & b_1 & \cdot & \cdot & b_N & & 0 & \\ & & \cdot & \cdot & & & & \cdot & \\ & 0 & & \cdot & \cdot & & & & . \\ & & & & b_0 & b_1 & \cdot & \cdot & b_N \end{pmatrix} \tag{5}$$

维数为 $(M+1)\times(M+1-2N)$ 的矩阵 $\boldsymbol{B}$

$$\boldsymbol{B} = \begin{pmatrix} B_0 & B_1 & \cdot & \cdot & B_{2N} & & & & \\ & B_0 & B_1 & \cdot & \cdot & B_{2N} & & 0 & \\ & & \cdot & \cdot & & & & \cdot & \\ & 0 & & \cdot & \cdot & & & & . \\ & & & & B_0 & B_1 & \cdot & \cdot & B_{2N} \end{pmatrix} \tag{6}$$

维数为 $(M+1-N)\times(M+1-2N)$ 的矩阵 $\boldsymbol{C}$

$$\boldsymbol{C} = \begin{pmatrix} b_0 & b_1 & \cdot & \cdot & b_N & & & & \\ & b_0 & b_1 & \cdot & \cdot & b_N & & 0 & \\ & & \cdot & \cdot & & & & \cdot & \\ & 0 & & \cdot & \cdot & & & & . \\ & & & & b_0 & b_1 & \cdot & \cdot & b_N \end{pmatrix} \tag{7}$$

$\boldsymbol{B}$ 和 $\boldsymbol{b}$ 的关系为

$$\boldsymbol{B} = \boldsymbol{bC} \tag{8}$$

维数为 $(M+1-N)\times(N+1)$ 的矩阵 $\boldsymbol{F}_b$

$$\boldsymbol{F}_b = \begin{pmatrix} f(0) & f(1) & \cdot\ \cdot & f(N) \\ f(1) & f(2) & \cdot\ \cdot & f(N+1) \\ \vdots & \vdots & & \vdots \\ f(M-N) & f(M-N+1) & \cdot\ \cdot & f(M) \end{pmatrix} \tag{9}$$

记矩阵 $\boldsymbol{b}$ 的 SVD(奇异值分解)为

$$\boldsymbol{b} = \boldsymbol{U}_b \boldsymbol{S}_b \boldsymbol{V}_b^{\mathrm{T}} \tag{10}$$

$$\boldsymbol{U}_b = [\boldsymbol{U}_{b1}, \boldsymbol{U}_{b2}] \tag{11}$$

其中 $\boldsymbol{U}_{b1}$ 由 $\boldsymbol{U}_b$ 的第 1 列到第 $(M+1-N)$ 列组成. 记矩阵 $\boldsymbol{B}$ 的 SVD 为

$$\boldsymbol{B} = \boldsymbol{U}_B \boldsymbol{S}_B \boldsymbol{V}_B^{\mathrm{T}} \tag{12}$$

$$\boldsymbol{U}_B = [\boldsymbol{U}_{B1}, \boldsymbol{U}_{B2}] \tag{13}$$

其中 $\boldsymbol{U}_{B1}$ 由 $\boldsymbol{U}_B$ 的第 1 列到第 $(M+1-2N)$ 列组成.

记 $\boldsymbol{f} = [f(0), f(1), \ldots, f(M)]^{\mathrm{T}}$ 为 $(M+1)\times 1$ 的数据向量, $\boldsymbol{f}^*$ 为信号向量的估值. 由文献[7]可知: $\boldsymbol{U}_{B1}$ 构成了 $(M+1-2N)$ 维的噪音子空间, 而 $\boldsymbol{U}_{B2}$ 组成了 $2N$ 维的信号子空间. 则

$$\boldsymbol{f}^* = \boldsymbol{U}_{B2} \boldsymbol{U}_{B2}^{\mathrm{T}} \boldsymbol{f} \tag{14}$$

依照式(9), 用 $\boldsymbol{f}^*$ 替代 $\boldsymbol{f}$ 构成矩阵 $\boldsymbol{F}_b^*$. 得到

$$\boldsymbol{F}_b^* [b_0, b_1, \ldots, b_N]^{\mathrm{T}} \tag{15}$$

式(1)模型在非线性最小二乘估计意义上的最优解, 由式(14)和(15)唯一地确定.

2.2 最优解的几何结构[8]

由式(14)和(15)联立方程, 容易构造一个求解参数向量 $[b_0, b_1, \ldots, b_N]^{\mathrm{T}}$ 的迭代算法. 尽管式(14)和(15)唯一地确定了最优解, 迭代过程仍然需要非线性计算, 如式(14)所示. 如何确保迭代算法的效率, 尤其是收敛性对于最优解几何结构的认识, 有助于构建一个有效的迭代算法. 迭代过程收敛时, 数据向量 $\boldsymbol{f}$ 应落在 $\boldsymbol{U}_{B1}$ 和 $\boldsymbol{U}_{b2}$ 合并的 $(M+1-N)$ 维子空间内[8]. 那么, 收敛时 $\boldsymbol{f}$ 在 $\boldsymbol{U}_{B1}$ 和 $\boldsymbol{U}_{b2}$ 合并的 $(M+1-N)$ 维子空间的正交投影应该等于零.

文献[8]迭代算法的收敛控制条件涉及两个控制精度 δ_1 和 δ_2. 依据对最优解几何结构的认识, 本文的迭代算法构造了一个更为合理的收敛控制条件, 仅涉及一个控制精度 δ: 记 $\boldsymbol{u}$ 为数据向量 $\boldsymbol{f}$ 在 $\boldsymbol{U}_{B1}$ 和 $\boldsymbol{U}_{b2}$ 合并的 $(M+1\text{-}N)$ 维子空间的正交投影

$$\boldsymbol{u} = (\boldsymbol{I}\boldsymbol{U}_{b2}\boldsymbol{U}_{b2}^{\mathrm{T}} - \boldsymbol{U}_{B1}\boldsymbol{U}_{B1}^{\mathrm{T}})\boldsymbol{f} = (\boldsymbol{I}\boldsymbol{U}_{b1}\boldsymbol{U}_{b1}^{\mathrm{T}} - \boldsymbol{U}_{B1}\boldsymbol{U}_{B1}^{\mathrm{T}})\boldsymbol{f} \tag{16}$$

本文算法的收敛控制条件为

$$Q_n = \boldsymbol{u}^{\mathrm{T}}\boldsymbol{u} < \delta \tag{17}$$

3 迭代方案

文献[8]迭代算法的效率不能令人满意, 尤其是收

敛率. 考虑对于收敛性十分重要的若干因素: (1) 参数向量 $[b_0, b_1, \ldots, b_N]^T$ 初值的选择; (2) 简化计算; (3) 减少构造子空间 U_{B1} 和 U_{B2} 时的计算误差; (4) 迭代过程中, $[b_0, b_1, \ldots, b_N]^T$ 的更新; (5) 收敛控制条件. 得到了一个改进的迭代算法, 称之为 ISNLSE(迭代子空间非线性最小二乘估计), 以下是 ISNLSE 算法的具体步骤:

(a) 初始化: 用较高阶数的方程 $F_b[b_0, b_1, \ldots, b_{Ne}]^T = 0$, 基于"低秩逼近"的方法[9]来计算 $[b_0, b_1, \ldots, b_N]^T$ 的初值, 其中 $Ne > N$; 由方程 $Q_u = f^T f$ 计算 Q_u 的初值.

(b) 检查收敛控制条件: $Q_u < \delta$ (δ 为给定的控制精度). 满足, 转向 (h); 不满足, 转向 (c).

(c) 用 QR 分解替代上述推导中使用的 SVD 分解, 以简化计算; 同时避免通过式 (12) 计算 U_B ($B = bC$ 为两个矩阵的乘积, B 阵的 SVD 分解在数值上是不稳定的[9]). 具体叙述如下:

(1) 计算 b 阵压缩形式的 QR 分解 $[Q_b, R_b] = qr(b, 0)$, 其中 Q_b 的维数为 $(M+1)\times(M+1-N)$, R_b 的维数为 $(M+1-N)\times(M+1-N)$. Q_b 的列向量和 U_{b1} 的列向量构成相同的子空间.

(2) 比较式 (5) 和式 (7) 可知: C 阵的 QR 分解 $[Q_C, R_C] = qr(C, 0)$, 其 $Q_C = Q_b(1{:}M+1-N, 1{:}M+1-2N)$.

(3) 计算 R_b 和 Q_C 乘积的 QR 分解 $[Q, R] = qr(R_b Q_C, 0)$, 由 $B = bC = Q_b R_b Q_C R_C = Q_b Q_R R_C$ 得到 B 阵的 QR 分解, 即 $Q_B = Q_b Q$. Q_B 的维数为 $(M+1)\times(M+1-2N)$, Q_B 的列向量和 U_{B1} 的列向量构成相同的子空间.

(d) 用 Q_B 替代 U_{B1}, 由式 (14) 计算 $f^* = U_{B2} U_{B2}^T f = (I - U_{B1} U_{B1}^T) f = (I - Q_B Q_B^T) f$.

(e) 用 Q_b 替代 U_{b1}、Q_B 替代 U_{B1}, 由式 (16)、(17) 计算 Q_u 的更新值.

(f) 考虑到参数向量 $[B_0, B_1, \ldots, B_{2N}]^T$ 的维数为 $(2N+1)$, 由式 (15), 用阶数为 $(2N+1)$ 的方程 $F_{b^*}[b_0, b_1, \ldots, b_{2N}]^T = 0$, 基于"低秩逼近"的方法来计算 $[b_0, b_1, \ldots, b_N]^T$ 的更新值.

(g) 转向 (b).

(h) 由式 (3) 计算式 (1) 模型中信号参数 (Z_n, $n = 0, 1, \ldots, N-1$) 的估计.

4 仿真结果

为了便于和"扩充的 ESPRIT 算法"相比较, 选择了一个和文献[9]介绍的相同的例子用于仿真. 在关于"ISNLSE 算法"的 Monte Carlo 试验中, 依据下列模型生成 128 个数据点, 构成数据向量 f

$$f(k) = \sqrt{20}\cos(2\pi f_1 k) + \sqrt{20}\cos(2\pi f_2 k + \pi/3) + \sqrt{2}\cos(2\pi f_3 k + \pi/4) + w(k)$$

其中, $f_1 = 0.1$, $f_2 = 0.12$, $f_3 = 0.2$; $w(k)$ 是伪随机高斯白噪声序列, $w(k)$ 的均值为 0、方差为 1. 这样, f_1 或 f_2 频率成分的信噪比是 10dB, f_3 频率成分的信噪比是 0dB. 表 1 给出了"ISNLSE 算法"和"扩充的 ESPRIT 算法"频率估计的均值和离差, 在 1000 次独立的试验中仅有 1 次不能收敛(迭代次数超过 100), 迭代的平均收敛速率为 7.016 步, 精度达到 0.5‰. 显然, "ISNLSE 算法"在均值和离差两方面都比"扩充的 ESPRIT 算法"具有更好的估计性能.

表 1 "**ISNLSE** 算法"和"扩充的 **ESPRIT** 算法"频率估计的对比; $Ne = 20$, $N = 6$, $M = 127$, $\delta = 1.0e-8$

	ISNLSE 算法		扩充的 ESPRIT 算法	
	Average	Square-Root-of MSE	Average	Square-Root-of MSE
f_1	0.1000	0.0001328	0.09961	0.00059
f_2	0.1200	0.0001294	0.12031	0.00064
f_3	0.2000	0.0003893	0.20008	0.00092

5 结论

"ISNLSE 算法"相对于文献[8]算法的主要改进在于"收敛控制条件"的确定. "ISNLSE 算法"的"收敛控制条件"仅涉及一个控制精度 δ, 是依据对最优解几何结构认识的基础上导出的, 更为合理、更具可操作性. 文献[8]算法的收敛条件涉及两个控制精度 δ_1 和 δ_2.

"ISNLSE 算法"相对于文献[8]算法的另一重要改进在于: 用 QR 分解替代 SVD 分解, 既简化了计算, 又避免了矩阵乘积的 SVD 分解在数值上的不稳定.

然而"ISNLSE 算法"还存在不收敛的情形(迭代次数超过 100)或称其为难于收敛的情形. 究其原因在于: 一方面, 计算过程的累积误差影响了算法的收敛性; 另一方面, 迭代过程中参数向量 $[b_0, b_1, \ldots, b_N]^T$ 的更新缺乏足够合理的方案. 问题的彻底解决在于"对最优解几何结构的代数解析".

"ISNLSE 算法"是全局优化的, 尽管它的收敛性仍需要进一步的理论研究和实践; 从理论推导和仿真结果已经显示, 相比于以往的各种方法, "ISNLSE 算法"提供更好的收敛性和估计精度. "ISNLSE 算法"不需要估计信号的协方差函数, 因而适用于短时数据的情形, 这在某些应用场合是尤其重要的.

本文的第一作者感谢他的妻子在撰写本文的过程中给予他的鼓励和支持, 感谢论文的审阅者提出的有价值的评论和意见.

参考文献:

[1] Carriere R, Moses RL. High resolution radar target modeling using a modified Prony estimator [J]. IEEE Trans on AP, 1992,

40(1): 13- 18.

[2] Marple L, Brotherton T. Detection and classification of short duration underwater acoustic signals by Prony's method [A]. Proceedings of International Conference on Acoustics, Speech, and Signal Processing [C]. San Diego, Canada 1991, 2: 1309-1312.

[3] Paraskevas M, Tsoukalas D. A Prony method for high-quality audio coding [A]. Proceedings of International Conference on Acoustics, Speech, and Signal Processing [C]. Minneapolis, USA 1993, 1: 205- 208.

[4] Hua Y, Sarkar T K. Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise [J]. IEEE Trans Acoustics Speech Signal Processing, 1990, ASSP-38: 814- 824.

[5] Bresler Y, Macovski A. Exact maximum likelihood parameter estimation of superimposed exponential signals in noise [J]. IEEE Trans. Acoustics, Speech, Signal Processing, 1986, ASSP-34: 1081- 1089.

[6] McDonough R. N, Huggins W. H. Best least squares representation of signals by exponentials [J]. IEEE Trans on Automatic Control, 1968, AC-13(4): 408- 412.

[7] 谢维波. 求解复极点模型最优解的一种新方法[J]. 电子与信息学报, 2001, 23(12): 1311- 1315.

Weibo Xie. A new solution for complex pole model in optimization [J]. Journal of Electronics & Information Technology, 2001, 23(12): 1311- 1315. (in chinese)

[8] 谢维波, 林劲松. 复指数信号模型非线性最小二乘解的几何结构及迭代算法[J]. 电子学报, 2002, 30(5): 757- 759.

Weibo Xie, Jinsong Lin. The geometric structure of nonlinear least squares solution for signal by complex exponents and alternate algorithm [J]. ACTA ELECTRONICA SINICA, 2002, 30(5): 757- 759. (in chinese)

[9] 张贤达. 现代信号处理[M]. 北京: 清华大学出版社, 1995. 115- 142.

Zhang XD. Modern Signal Processing [M]. Beijing: Tsinghua University Press, 1995. 115- 142. (in chinese)

作者简介:

谢维波　男, 1964 年 10 月生于福建, 1988 年哈尔滨工业大学信息工程专业毕业, 主要研究方向为信号处理、模式识别、智能信息处理.
E-mail: xwblxf@hqu.edu.cn

王永初　男, 1937 年 12 月生于福建, 1962 年浙江大学自动化专业毕业, 长期从事过程控制、自动化仪表与装置、检控技术等领域的科研与教学.