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

谢维波[1],林劲松[2]
(1. 华侨大学信息科学与工程学院,福建泉州 362011;2. 厦门大学自动化系,福建厦门 361005)

摘 要: 本文给出了复指数信号模型非线性最小二乘解的几何结构.从分析迭代算法的收敛性态入手得到解的几何结构,将有助于构造十分有效的迭代算法.另外,本文在低信噪比(10dB)及较小频率差(0.02Hz)的情况下,对迭代求解的收敛控制条件进行了研究.

关键词: Prony方法; 非线性最小二乘; 子空间分解
中图分类号: TN957 **文献标识码**: A **文章编号**: 0372-2112 (2002) 05-0757-03

The Geometric Structure of Nonlinear Least Square Solution for Signal by Complex Exponents and Alternate Algorithm

XIE Wei-bo[1], LIN Jin-song[2]
(1. *College of Information Science & Engineering, Huaqiao University, Quanzhou, Fujian* 362011, *China*;
2. *Dept. of Automation, Xiamen University, Xiamen, Fujian* 361005, *China*)

Abstract: The geometric structure of nonlinear least squares solution for signals by complex exponents is offered in this paper. Beginning with analysis for the convergent state of alternative algorithm solving two equations together contented by the model's nonlinear least squares solution, the recognition for geometric structure of nonlinear least squares solution is acquired. It would help to construct a fully effective algorithm and understanding for the solution's structure is deepened. The alternative algorithm presented by this paper is fully effective in higher SNR or when the difference of frequency in model is slightly increased. Nevertheless, the invalid convergence (large error) appears in the condition with lower SNR(10dB) and smaller difference of frequency (0.02 Hz), if the convergent control condition of alternative algorithm is only in accordance with varying quantity of least squares error.

Key words: prony method; nonlinear least squares; subspace decomposition

1 引言

早在1795年,PRONY就提出了用复指数函数的一线性组合来描述等间隔采样数据的数学模型,又常称为复指数模型.在白噪声背景下,该模型的最优求解是一个困难的非线性最小二乘问题.在复指数为复正弦的情况下,人们从依据信号的统计量及线性化处理两方面,提出了许多高分辨率的求解算法[1~3].文献[4]在最大似然准则下给出了非线性处理的一维搜索求解方案,存在算法收敛的不确定性及解的唯一性问题.文献[5]在文献[4、6]的基础上,给出了复指数模型满足充要条件的非线性最小二乘解的表达,为构造有效的迭代算法打下了基础.本文研究该表达蕴涵的几何结构,并在低信噪比(10dB)的情况下改进了文献[5]迭代算法的收敛控制条件.

2 若干表达式的定义

下面先定义若干式子,以便于问题的表达.

2.1 复指数信号模型

$$f(k) = \sum_{n=0}^{N-1} A_n Z_n^k + \varepsilon_k,\ k = 0,1,\cdots,M \tag{1}$$

其中 $f(k)$ 为观测数据,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.

2.2 最小二乘辨识的误差函数

$$Q = \sum_{k=0}^{M} \varepsilon_k^2 = \sum_{k=0}^{M} [f(k) - \sum_{n=0}^{N-1} A_n Z_n^k]^2 \tag{2}$$

2.3 由$\{Z_m\}_{m=0}^{N-1}$构造的基本对称函数

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

记向量 $b = [b_0, b_1, \cdots, b_N]^T$

2.4 以$\{Z_m\}_{m=0}^{N-1}$为二重根的多项式系数$\{B_m\}_{m=0}^{2N}$

$$\sum_{m=0}^{2N} B_m Z^m = \prod_{m=0}^{N-1} (Z - Z_m)^2 = [\sum_{m=0}^{N} b_m Z^m]^2 \tag{4}$$

记向量 $B = [B_0, B_1, \cdots, B_{2N}]^T$

2.5 M+1行,M+1-N列的矩阵b(列满秩)

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

收稿日期:2001-04-07;修回日期:2001-12-04

2.6 M+1 行,M+1-2N 列的矩阵 B(列满秩)

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

2.7 矩阵 b 和矩阵 B 的关系

$$B = bC \tag{7}$$

其中 C 为 $M+1-N$ 行,$M+1-2N$ 列的矩阵. 注:C 阵与 b 阵形式上相同,维数不同.

$$C=\begin{pmatrix} b_0 & b_1 & \cdot & \cdot & b_N & & & \\ & b_0 & b_1 & \cdot & \cdot & b_N & & 0 \\ & & \cdot & \cdot & & & \cdot & \\ & 0 & & \cdot & \cdot & & & \cdot \\ & & & & b_0 & b_1 & \cdot \cdot & b_N \end{pmatrix}^T$$

2.8 矩阵 b 的奇异值分解

$$b = U_b S_b V_b^T \tag{8}$$

记 $U_b=[U_{b1}, U_{b2}]$,U_{b1} 为 U_b 的第 1 列到第 $M+1-N$ 列的子矩阵

2.9 矩阵 B 的奇异值分解

$$B = U_B S_B V_B^T \tag{9}$$

记 $U_B=[U_{B1}, U_{B2}]$,U_{B1} 为 U_B 的第 1 列到第 $M+1-2N$ 列的子矩阵

2.10 M+1-N 行,N+1 列的数据矩阵 F_b

$$F_b=\begin{pmatrix} f(0) & f(1) & \cdot\cdot & f(N) \\ f(1) & f(2) & \cdot\cdot & f(N+1) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ f(M-N) & f(M-N+1) & \cdot\cdot & f(M) \end{pmatrix} \tag{10}$$

2.11 M+1-2N 行,2N+1 列的数据矩阵 F_B

$$F_B=\begin{pmatrix} f(0) & f(1) & \cdot\cdot & f(2N) \\ f(1) & f(2) & \cdot\cdot & f(2N+1) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ f(M-2N) & f(M-2N+1) & \cdot\cdot & f(M) \end{pmatrix} \tag{11}$$

3 非线性最小二乘解的几何结构

文献[5]给出了式(1)复指数模型非线性最小二乘解的表达,该表达可归纳为:解应同时满足如下两个方程.

$$f^* = U_{B2} U_{B2}^T f \tag{12}$$

$$F_b^* b = 0 \tag{13}$$

其中 $f=[f(0), f(1), \cdots, f(M)]^T$ 为已知的数据向量,$b=[b_0, b_1, \cdots, b_N]^T$ 为待求的参数向量;信号向量的估计 f^* 是数据向量 f 在 U_{B2} 子空间的投影,故 U_{B2} 构成了 $2N$ 维的信号子空间;F_b^* 是以 f^* 代替 f 按照式(10)构成的矩阵;U_{B1} 是 B 阵的列所张成的正交子空间,构成了 $M+1-2N$ 维的噪音子空间,故式(2)定义的 Q 可表达为:

$$Q = f^T U_{B1} U_{B1}^T f \tag{14}$$

下面从分析式(12)、(13)构成迭代的收敛性态入手,以获得对于式(1)非线性最小二乘解几何结构的认识.

3.1 式(13)矛盾方程组 $F_b^* b=0$ 的收敛性态

式(12)、(13)构成的迭代收敛时,矛盾方程组 $F_b^* b=0$ 应不再"矛盾",即 $F_b^* b=0$ 的剩余误差应为零. 由于 $F_b^* b = b^T f^*$ 是一个恒等式(其中 b 为式(5)定义的矩阵),则式(13)等价为 $b^T f^*=0$,由式(8)有.

$$U_{b1}^T f^* = 0 \tag{15}$$

式(15)即为收敛时信号向量的估计 f^* 应满足的条件. 综上所述,迭代收敛时有以下两方面的结论:①由式(12)知 f^* 为数据向量 f 在 U_{B2} 子空间的投影;②由式(15)知 f^* 垂直于 U_{b1} 子空间,故 f^* 为数据向量 f 在 U_{b2} 子空间的投影. 如何保证二者同时成立呢?

3.2 U_{b1} 和 U_{B1},U_{b2} 和 U_{B2} 子空间的关系

由式(7)知 B 阵的列向量是 b 阵列向量的线性组合,故 $M+1-2N$ 维空间 U_{B1} 是 $M+1-N$ 维空间 U_{b1} 的子空间,而 N 维空间 U_{b2} 是 $2N$ 维空间 U_{B2} 的子空间. 容易证明:U_{B1} 子空间既垂直于 U_{B2} 子空间,又垂直于 U_{b2} 子空间. 若数据向量 f 落在 U_{B1} 和 U_{b2} 合并的 $M+1-N$ 维子空间内,那么数据向量 f 在 U_{B2} 子空间的投影就等于数据向量 f 在 U_{b2} 子空间的投影,即上述两方面的结论①和②得以同时成立;因此收敛时,数据向量 f 应落在 U_{B1} 和 U_{b2} 合并的子空间内. 现在的问题是:如何由已知的数据向量 f 去求得满足这种性质的子空间 U_{B1} 和 U_{b2}(注意:这种性质的子空间 U_{B1} 和 U_{b2} 完全地由待求的参数向量 $b=[b_0, b_1, \cdots, b_N]^T$ 所决定,而 $b=[b_0, b_1, \cdots, b_N]^T$ 由数据向量 f 所决定).

4 迭代算法及收敛控制条件的改进

文献[5]依据式(12)、(13)构造了一个迭代算法,描述如下:

①由方程组 $F_b b=0$ 得 $b=[b_0, b_1, \cdots, b_N]^T$ 的初值;以 $Q=f^T f$ 作为噪声能量的初始估值.

②由 $b=[b_0, b_1, \cdots, b_N]^T$ 依据式(4)、(6)和(9)构造子空间 U_{B1} 和 U_{B2}.

③由式(12)得 f^*,以 f^* 构成矩阵 F_b^*;由方程组 $F_b^* b=0$ 得 $b=[b_0, b_1, \cdots, b_N]^T$ 的迭代更新.

④由式(14)得 Q 的迭代更新;转②循环迭代,直至 Q 的改变量趋于零.

数值实验采用的数据模型为:

$$f(k) = A1 * \cos(2\pi * f_1 k + pha1) + A2 * \cos(2\pi * f_2 k + pha2) + var * w(k)$$

$w(k)$ 为高斯白噪声($\sigma_w^2=1$);$A1=1.0$;$A2=1.5$;$pha1=0$,$pha2=\pi/3$;$f_1=0.1$,$f_2=0.12$;$var=0.1$. 数值实验的结果为:10 次计算有 1-2 次不收敛;收敛时频率的分辨精度很高;降底信噪比、缩小频率差值或减少数据点,会使不收敛的次数增多.

上述实验用的信噪比高达 22dB,没有反映出该算法存在的缺陷. 当信噪比为 10dB 时(实验参数为 $A1=\sqrt{20}$,$A2=$

$\sqrt{20}$; $pha1 = 0$, $pha2 = \pi/3$; $f_1 = 0.1$, $f_2 = 0.12$; $var = 1.0$) ,出现了“无效收敛”(收敛结果精度很差) 的情况. 分析上述算法可以发现:相邻两次迭代的两组不同的 $b = [b_0, b_1, \cdots, b_N]^T$ 可以对应相同的 Q 值. 在算法中以 Q 和 $b = [b_0, b_1, \cdots, b_N]^T$ 的改变量同时作为收敛的控制条件,解决了“无效收敛”的问题. 数值实验的统计结果在表 1 中列出. 共进行了 100 次计算,每次计算的迭代次数超过 100 被认为发散;每次计算的平均迭代次数为 26.1707;100 次计算有 18 次发散.

表 1　数值实验的统计结果($M = 63$, $N = 4$, $e - 3 = 0.001$)

统计结果	f_1	$-f_1$	f_2	$-f_2$	f_1 模	$-f_1$ 模	f_2 模	$-f_2$ 模
准确值	0.1	-0.1	0.12	-0.12	1.0	1.0	1.0	1.0
均值	0.0999	-0.0999	0.1202	-0.1202	0.9998	0.9998	0.9988	0.9988
标准差	0.4744×10^{-3}	0.4744×10^{-3}	0.4714×10^{-3}	0.4714×10^{-3}	0.0036	0.0036	0.0033	0.0033

5　结束语

由文献[5]可知式(12)、(13)给出的解满足的最优化条件是充要的,本文又给出了解的几何结构的认识,但求解算法的构造尚待深入研究. 已有的条件是:①同时满足式(12)、(13)的解是唯一的;②对于解的几何结构的认识有助于构造求解算法,或许可以由一系列的空间变换(刚体的)来达到这种结构. 感谢审稿者提出“采用低信噪比”的意见,这使文献[5]的迭代算法存在“无效收敛”的缺陷得以显现;本文修正了这一缺陷,同时加深了对解的几何结构的认识. 模型中频差的大小对算法收敛的影响当属意料之中,适当扩大频差(信噪比 10dB),本文算法的收敛性显著提高.

参考文献:

[1] 马淑芬,吴嗣亮. 有色噪声中谐波频率的频域非线性预滤波估计方法[J]. 电子学报,2000,28(6):48-50.

[2] 金梁,殷勤业,汪仪林. 广义谱相关子空间拟合 DOA 估计原理[J]. 电子学报,2000,28(1):60-63.

[3] 喻胜,陈光[illegible]. 一种检测噪声中正弦信号的 SVD 方法[J]. 电子学报,2000,28(6):108-110.

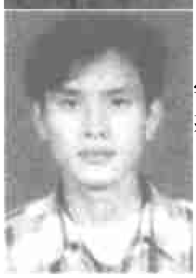
[4] Bresler Y, A Macovski. Exact maximum likelihood parameter estimation of superimposed exponential signals in noise [J]. IEEE Trans. Acoust, Speech, Signal Processing, 1986, ASSP-34:1081-1089.

[5] 谢维波. 复极点模型的非线性求解[A]. 第七届联合国际计算机会议论文集[C]. 汕头:汕头大学出版社,2000.65-68.

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

作者简介:

谢维波　男,1964 年 10 月生于福建,福建泉州华侨大学计算机系副教授,1988 年毕业于哈尔滨工业大学无线电系,从事信号处理和计算机硬件的教学和研究. E-mail: wxblxf@hqu.edu.cn



林劲松　男,1975 年 11 月生于福建,厦门大学自动化系讲师,1999 年毕业于厦门大学自动化系,从事系统工程和计算机软件的教学和研究.