

# 用PSO方法搜索基于MLE的ARMA模型参数

罗 航, 黄建国, 龙 兵, 王厚军

(电子科技大学自动化工程学院 成都 610054)

**【摘要】**针对小样本条件下用矩估计(ME)方法获取ARMA模型参数粗略的缺点, 将粒子群优化算法(PSO)用于小样本ARMA模型参数的极大似然估计(MLE), 以获得概率上最优的数字解。在分析基于ARMA模型似然函数的基础上, 详细分析了PSO的思想、方法和评价指标。以实际例证显示了联合PSO优化方法估计AMAR模型参数的优良特性, 并从算法和似然函数角度分别阐释了形成利弊的原因。

**关键词** ARMA模型; 极大似然估计; 粒子群; 参数; 小子样

中图分类号 TP274

文献标识码 A

doi:10.3969/j.issn.1001-0548.2010.01.016

## Searching ARMA Model Parameters MLE-Based by Applying PSO Algorithm

LUO Hang, HUANG Jian-guo, LONG Bing, and WANG Hou-jun

(School of Automation Engineering, University of Electronic Science and Technology of China Chengdu 610054)

**Abstract** On the condition of small samples, it is difficult to achieve ARMA model parameters by only using moment estimation (ME). In this paper, a particle swarm optimization (PSO) algorithm is used to obtain optimal numeric solutions of small sample ARMA model in the sense of maximum likelihood estimation (MLE). The principle and evaluation index of PSO are discussed in detail, which is based on analyzing likelihood function of ARMA model. Actual example shows that the joint PSO optimization method used for estimating AMAR model parameters has better characteristics in comparison with other methods. At the same time, some advantages and disadvantages are expatiated and analyzed from perspectives of algorithm and likelihood function.

**Key words** ARMA model; maximum likelihood estimation; particle swarm optimization; parameter; small sample

ARMA模型参数的矩估计可以通过联合解Yule-Walker方程<sup>[1]</sup>和Newton-Raphson迭代<sup>[2]</sup>快速求得相应的参数。但是该方法存在明显的弊端, 因为它仅用到了 $p+q$ 个样本的自相关系数, 而失效数据的其他有用信息被忽略。对于小子样失效数据而言, 这种估计的精度常常很差。

将智能优化算法用于ARMA模型的研究<sup>[3-6]</sup>是一个有潜力的方向。本文的主要目的是用粒子群优化(PSO)算法<sup>[7]</sup>计算基于极大似然估计方法(MLE)的ARMA模型参数的数值解, 其原因在于: (1) 从估计层面上考虑, MLE有效利用了每个观察值所提供的信息, 它以发生的事件作为最可能出现的事件为出发点, 求其最大概率而得到参数的估计值。(2) 从求解角度考虑, ARMA模型下的似然方程是复杂的超越方程组, 无法用常规的代数方法求得解析解。而传统数值方法(牛顿法、信赖域算法<sup>[8]</sup>、梯度法等)

常使用目标函数的导数信息, 容易陷入局部最优。PSO算法的显著优势在于它不依赖于目标函数的解析性质, 是在似然方程的基础上确定适应度函数, 进而以并行方式搜索全局和局部极值求其数字解。而与此同时, 基于矩估计的模型参数正好方向性地提供了粒子群优化的初始解。

### 1 ARMA(p,q)模型参数的极大似然方程

在极大似然准则下, 小样本数据序列的对数似然函数为:

$$l(\tilde{\beta}, \tilde{x}) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma_\epsilon^2) - \frac{1}{2} \ln |\Omega| - \frac{1}{2\sigma_\epsilon^2} S(\tilde{\beta}) \quad (1)$$

式中  $\tilde{\beta} = [a_1, a_2, \dots, a_p, b_1, b_2, \dots, b_q]^T$ ;  $\tilde{x} = [x_1, x_2, \dots, x_n]^T$ ;  $\Omega = E(\mathbf{x}\mathbf{x}^T)/\sigma_\epsilon^2$ ;  $S(\tilde{\beta}) = \mathbf{x}^T \Omega^{-1} \mathbf{x}$ ;  $\sigma_\epsilon^2$  为噪声方差。分别对式(1)的未知参数  $\sigma_\epsilon^2$  和  $\tilde{\beta}$  求偏导数, 有:

收稿日期: 2008-08-09; 修回日期: 2008-10-23

基金项目: 部级预研基金

作者简介: 罗 航(1972-), 男, 博士生, 主要从事故障诊断和可靠性预测等方面的研究。

$$\begin{cases} \frac{\partial}{\partial \sigma_\varepsilon^2} l(\tilde{\beta}, \tilde{x}) = -\frac{n}{2\sigma_\varepsilon^2} + \frac{1}{2\sigma_\varepsilon^4} S(\tilde{\beta}) \\ \frac{\partial}{\partial \tilde{\beta}} l(\tilde{\beta}, \tilde{x}) = -\frac{1}{2} \frac{\partial \ln |\Omega|}{\partial \tilde{\beta}} - \frac{1}{2\sigma_\varepsilon^2} \frac{\partial S(\tilde{\beta})}{\partial \tilde{\beta}} \end{cases} \quad (2)$$

分别令式(2)中的各式为零, 并化简为:

$$\begin{cases} S(\tilde{\beta}) - n\sigma_\varepsilon^2 = 0 \\ \frac{\partial \ln |\Omega|}{\partial \tilde{\beta}} + \frac{1}{\sigma_\varepsilon^2} \frac{\partial S(\tilde{\beta})}{\partial \tilde{\beta}} = 0 \end{cases} \quad (3)$$

由于  $\Omega$  决定着  $S(\tilde{\beta})$  的值, 所以求解式(3)的关键在于寻找  $\Omega$  和  $\tilde{\beta}$  的关系。事实上,  $\Omega$  由  $\tilde{\beta}$  唯一确定:

$$\Omega = \begin{bmatrix} \sum_{i=0}^{\infty} G_i^2 & \sum_{i=0}^{\infty} G_i G_{i+1} & \cdots & \sum_{i=0}^{\infty} G_i G_{i+n-1} \\ \sum_{i=0}^{\infty} G_i G_{i+1} & \sum_{i=0}^{\infty} G_i^2 & \cdots & \sum_{i=0}^{\infty} G_i G_{i+n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=0}^{\infty} G_i G_{i+n-1} & \sum_{i=0}^{\infty} G_i G_{i+n-2} & \cdots & \sum_{i=0}^{\infty} G_i^2 \end{bmatrix} \quad (4)$$

式中  $G_i$  称为 ARMA( $p, q$ ) 模型的格林函数<sup>[9]</sup>, 由  $p+q$  个模型参数通过递推方式得到。式(3)是  $p+q+1$  个超越方程, 设其对应的解和目标函数分别为  $\mathbf{X} = [a_1, a_2, \dots, a_p, b_1, b_2, \dots, b_q, \sigma_\varepsilon^2]$  和  $F_i(\mathbf{X})$ ,  $i=1, 2, \dots, p+q+1$ 。为了使函数  $F_i(\mathbf{X})=0$ , 总是让数值寻优的  $|F_i(\mathbf{X})|$  最小, 故PSO优化采用的适应度函数可考虑为:

$$\begin{cases} \min_{\mathbf{X} \in R^{p+q+1}} \max_{1 \leq i \leq p+q+1} \{|F_i(\mathbf{X})|\} \\ \text{s.t. } \mathbf{X} \in [\mathbf{X}_0 - \delta, \mathbf{X}_0 + \delta] \end{cases} \quad (5)$$

式中  $\mathbf{X}_0$  是优化的初始值, 用矩估计方法得出;  $\delta$  是搜索的半径。

## 2 PSO优化算法

### 2.1 基本思想

在  $R^n$  空间中, 一个群体含有  $N$  个粒子, 而每个粒子的位置  $\mathbf{X}_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T$  ( $i=1, 2, \dots, N$ ) 是寻优的可能解。PSO最关键的操作在于: 每次迭代寻优过程中的所有粒子要通过跟踪两个“极值”更新自己。第一个“极值”是粒子本身目前所找到的最优解(局部极值pBest); 另一个“极值”是整个种群目前找到的最优解(全局极值gBest)。这两类极值构成粒子向下一代位置飞行的必要条件。粒子在计算其当前适应度的过程中, 不断地更新自己的位置, 最终逼近全局最优解。

### 2.2 粒子群的更新

设第  $k$  代粒子的位置为  $\mathbf{X}_i^{(k)} = [x_{i1}^{(k)}, x_{i2}^{(k)}, \dots,$

$x_{in}^{(k)}]^T$  ( $i=1, 2, \dots, N$ ), 各粒子的速度为  $\mathbf{V}_i^{(k)} = [v_{i1}^{(k)}, v_{i2}^{(k)}, \dots, v_{in}^{(k)}]^T$ , 各粒子的局部最优位置为  $\mathbf{P}_i^{(k)} = [p_{i1}^{(k)}, p_{i2}^{(k)}, \dots, p_{in}^{(k)}]^T$ , 各粒子的全局最优位置为  $\mathbf{P}_g^{(k)} = [p_{g1}^{(k)}, p_{g2}^{(k)}, \dots, p_{gn}^{(k)}]^T$ 。需要强调的是,  $\mathbf{P}_g^{(k)}$  是群体中每个粒子的全局最优位置, 因而对于群体本身而言, 每个粒子都是相同的。从第  $k$  代粒子到  $k+1$  代粒子的更新速度为<sup>[10]</sup>:

$$\mathbf{V}_i^{(k+1)} = \omega^{(k)} \mathbf{V}_i^{(k)} + c_1 r_{i1} (\mathbf{P}_i^{(k)} - \mathbf{X}_i^{(k)}) + c_2 r_{i2} (\mathbf{P}_g^{(k)} - \mathbf{X}_i^{(k)}) \quad i=1, 2, \dots, N \quad (6)$$

式中  $c_1$  和  $c_2$  为学习因子;  $r_{i1}$  和  $r_{i2}$  为  $[0, 1]$  上均匀分布的随机数;  $\omega$  为惯性权重系数。 $\omega$  的重要性在于: 用粒子的当前速度控制下代粒子的速度, 当  $\omega$  较大时, 粒子全局搜索能力强; 当  $\omega$  较小时, 粒子局部搜索能力强。为了快速搜索全局最优值,  $\omega$  常设置为随迭代次数增加而减小的函数<sup>[11]</sup>:

$$\omega^{(k)} = \omega_{\max} - k(\omega_{\max} - \omega_{\min}) / k_{\max} \quad (7)$$

第  $k$  代粒子到  $k+1$  代粒子的位置更新为<sup>[12]</sup>:

$$\mathbf{X}_i^{(k+1)} = \mathbf{X}_i^{(k)} + \alpha \mathbf{V}_i^{(k+1)} \quad (8)$$

式中  $\alpha$  是对更新速度进行加权的速度约束因子。

### 2.3 衡量粒子群搜索更新的两个指标

式(6)是粒子群更新的核心步骤。它蕴涵着群体向下一代“飞翔”的原因和方式, 其作用是从全局的角度控制粒子搜索的“度”。第一项直接源于粒子本身的速度, 是一种维持当前速度的记忆行为; 第二项是粒子自身“认知”对下一步速度的影响, 它用粒子当前位置和自己最好位置的距离“修正”继承的速度; 第三项用粒子对其他粒子经验的“社会”认知控制粒子飞向下一步的速度, 它用粒子当前位置与群体最佳位置的距离衡量影响的大小。速度的动态更新反映了粒子搜索的自适应调整过程。为了衡量这种动态行为对粒子收敛的影响, 常常考虑离散指标和速度指标<sup>[13]</sup>:

$$\bar{\delta} = (\sum_{i=1}^N \sum_{d=1}^n |x_{id} - p_{gd}|) / Nn \quad (9)$$

$$\bar{v} = (\sum_{i=1}^N \sum_{d=1}^n |v_{id}|) / Nn \quad (10)$$

式(9)是考察所有粒子如何围绕最佳粒子运动的指标, 该指标解释了粒子群搜索的覆盖区域。式(10)度量了粒子群在一定的迭代范围内移动快慢的程度。显然, 上述两个指标随迭代次数增加而同步(几乎同步)减小时, 则群体将很快收敛。联合优化 ARMA 模型参数的程序框图如图1所示。

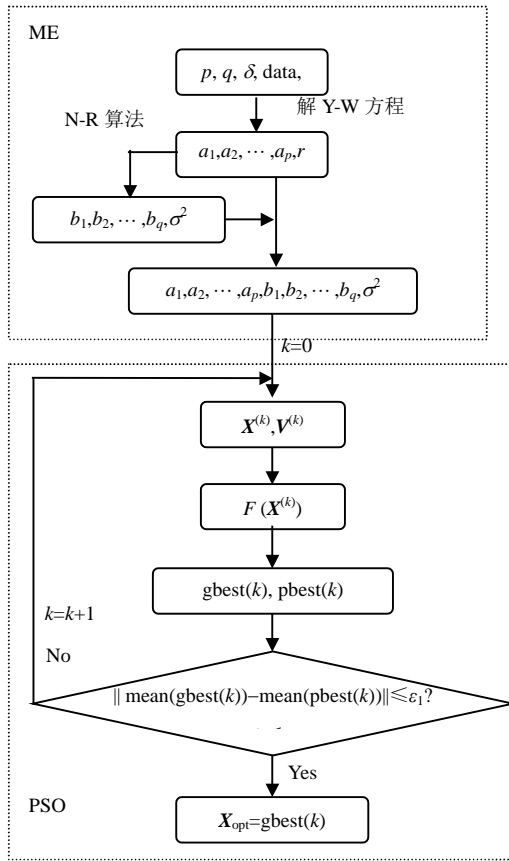


图1 联合优化程序框图

### 3 实例分析

#### 3.1 轴承失效数据的ARIMA(1,2,1)模型参数

表1 轴承失效数据的ARIMA(1,2,1)模型参数

	初始值	1	2	3	SAS分析
$a_1$	-0.601 4	-0.903 7	-0.903 7	-0.903 7	-0.910 7
$b_1$	-0.116 2	-0.016 2	-0.016 2	-0.016 2	-0.016 6
$\sigma^2$	126.680 0	76.399 0	76.390 0	76.356 0	83.747 0
$F$	536.680 0	0.0563 3	0.0563 5	0.0563 9	76.138 0
时间/s	0.598 30	48.269 9	46.373 2	70.966 5	0.212 4

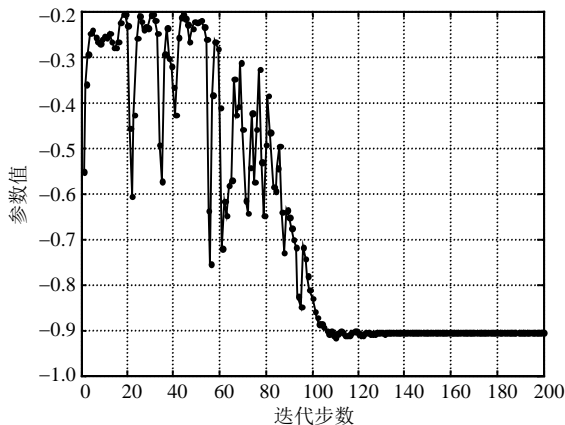


图2 AR参数 $a_1$ 的PSO优化过程

根据ARMA模型建立的条件和方法, 考虑轴承

失效数据<sup>[14]</sup>的ARIMA(1,2,1)模型。本质上, 估计参数是 $a_1$ 、 $b_1$ 和 $\sigma_\varepsilon^2$ 。以初始矩估计值为中心, 考虑搜索半径为 $\delta = [0.4 \ 0.1 \ 55]$ , 则PSO优化的结果如表1所示。图2~图5分别表示了第3次实验中3个参数的迭代过程及收敛指标。

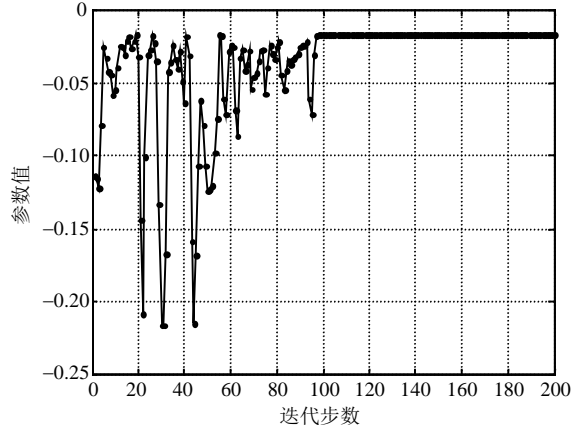


图3 MA参数 $b_1$ 的PSO优化过程

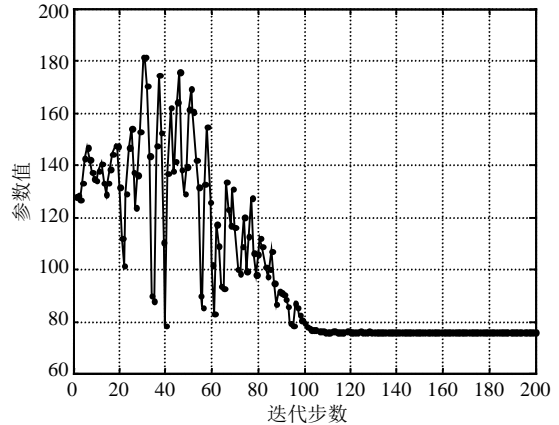


图4 误差方差 $\sigma^2$ 的PSO优化过程

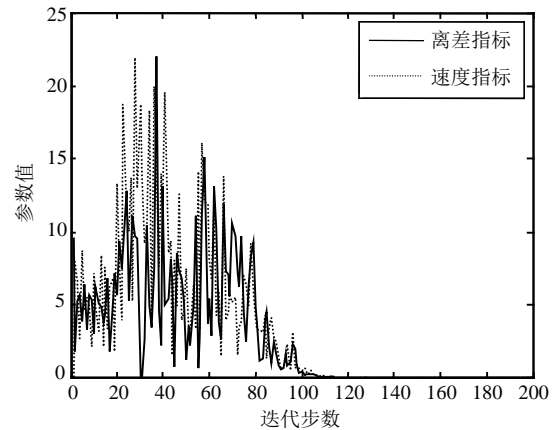


图5 两个指标的PSO收敛过程

#### 3.2 参数估计分析

首先, 从适应度角度考察, 联合PSO方法具有较高的搜索精度。究其原因: 矩估计由于数据信息利用率低导致了估计精度最差。基于SAS(statistical analysis system)软件的MLE估计精度虽然高于前者,

但又低于联合PSO方法。SAS软件的设置方式使其在本文实验中遇到鞍点而停止迭代,很明显,SAS分析是一个局部极值点。而联合PSO方法在解空间中以全局和局部并行的方式搜索,这种寻优是一种从全局强搜索逐渐过度到局部强搜索的此消彼长过程。基于这种先粗化后细化的特点,联合PSO方法在搜索空间中较为精确地得到了参数的数字解。

其次,联合PSO方法的搜索过程具有较强的稳健性,在一定的范围内,实验所得到的结果较为平稳,不因为重复计算而出现较大的差异。一方面,矩估计方法奠定了PSO搜索的初始位置,而这个初始位置方向性地给出了PSO的入口,这在较大程度上避免了PSO初始值选择的盲目性;另一方面,联合优化方法的核心是PSO,只要给定模型的似然方程,则具有全局和局部优化功能的PSO方法在给定范围内能够搜索到最优解。

然而,在本文实验中,联合PSO搜索的准确性和稳健性是以长时间开销为代价的。事实上,本文实验是多维搜索,增加了搜索空间的自由度;另一方面,适应度函数太过复杂,尽管 $\Omega$ 和 $S(\hat{\beta})$ 是AR和MA参数的函数,但它们不是显式表达式,必须通过格林函数来间接反映,因而似然函数参数的导数表述是用定义来反映的。此外,由于样本数据较少,式(3)保留了第二个方程的第一项(该方程的第一项在大样本条件下可以忽略)。基于上述原因,联合PSO方法用较长时间搜索到参数的数字解。如图2~图5所示,参数和指标的收敛消耗了较多的迭代步数。

## 4 结束语

本文分析了用PSO优化算法对ARMA模型参数进行极大似然估计的原因。在详细分析PSO方法的基础上,形成了基于矩估计和PSO的联合优化策略。实际例证表明,基于联合PSO参数估计方法较为合理准确。更为重要的是,传统MLE方法对似然函数作了省略的近似处理,而联合PSO方法则完全保留了似然函数的表述,使得小子样ARMA模型的参数估计更接近真实。

### 参 考 文 献

[1] FATTAH S A, ZHU W P, AHMAD M O. An identification technique for noisy ARMA systems in correlation domain[C]//IEEE International Symposium on Circuits and Systems, ISCAS 2007. New Orleans: IEEE, 2007: 349-350.  
[2] 张贤达. 现代信号处理[M]. 北京: 清华大学出版社, 2003: 111-112.  
ZHANG Xian-da. Modern signal processing[M]. Beijing:

Tsinghua University Press, 2003: 111-112.  
[3] CHON K H, SHENG Lu. A new algorithm for autoregression moving average model parameter estimation using group method of data handling[J]. Annals of Biomedical Engineering, 2001, 29: 92-98.  
[4] CHON K H, COHEN R J. Linear and nonlinear arma model parameter estimation using an artificial neural network[J]. IEEE Transactions on Biomedical Engineering, 1997, 44(3): 168-173.  
[5] 陈丹, 方康玲, 陈乔礼. 基于GA遗传算法的ARMA模型的辨识方法[J]. 武汉科技大学学报(自然科学版), 2007, 30(1): 87-89.  
CHEN Dan, FANG Kang-ling, CHEN Qiao-li. A system identification method based on genetic algorithm[J]. Journal of Wuhan University of Science and Technology (Natural Science Edition), 2007, 30(1): 87-89.  
[6] 彭喜元, 王军, 彭宇. 一种新型时间序列多分辨预测模型研究[J]. 电子学报, 2007, 35(11): 2148-2149.  
PENG Xi-yuan, WANG Jun, PENG Yu. A novel multi-scale predictor for complex time series[J]. Acta Electronica Sinica, 2007, 35(11): 2148-2149.  
[7] EBERHART R, KENNEDY J. A new optimizer using particle swarm theory[C]//Proc of the 6th Int'l Symposium on Micro Machine and Human Science. Piscataway NJ: IEEE Service Center, 1995: 39-43.  
[8] WANG Xiao-fei. Quasi-Newton algorithm with non-monotonic trust region methods for nonlinear equations[J]. Journal of Shanghai Normal University (Natural Science), 2003, 32(2): 16-22.  
[9] 王燕. 应用时间序列分析[M]. 北京: 中国人民大学出版社, 2005: 66-67.  
WANG Yan. Application time series analysis[M]. Beijing: China Renmin University Press, 2005: 66-67.  
[10] SHI Y, EBERHART R. A modified particle swarm optimizer[C]//IEEE International Conference on Evolutionary Computation. Piscataway NJ: IEEE Press, 1998: 69-73.  
[11] 史永丽, 侯朝桢, 苏海滨. 基于粒子群优化算法的自抗扰控制器设计[J]. 系统仿真学报, 2008, 20(2): 434-435.  
SHI Yong-li, HOU Chao-zhen, SU Hai-bing. Auto-disturbance-rejection controller design based on particle swarm optimization algorithm[J]. Journal of System Simulation, 2008, 20(2): 434-435.  
[12] CLERC M. The swarm and the queen: Towards a deterministic and adaptive particle swarm optimization [C]//Proc CEC 1999. [S.l.]: [s.n.], 1999: 1951-1957  
[13] AI T J, KACHITVICHYANUKUL V. Dispersion and velocity indices for observing dynamic behavior of particle swarm optimization[C]//IEEE Congress on Evolutionary Computation. Singapore: CEC 2007, 2007: 3264-3271.  
[14] 芮延年, 傅戈雁. 现代可靠性设计[M]. 北京: 国防工业出版社, 2007: 68-69.  
RUI Yan-nian, Fu Ge-yan. Design of modern reliability[M]. Beijing: National Defence Industry Press, 2007: 68-69.