PSO算法在非线性回归模型参数估计中的应用

陆克中,方康年
(池州学院 计算机科学系,安徽 池州 247100)

摘 要:非线性回归模型的参数估计是较为困难的寻优问题,经典方法常会陷入局部极值。由于粒子群算法是一种有效的解决优化问题的群集智能算法,它的突出特点是操作简便、容易实现且全局搜索功能较强,故将粒子群优化算法用于非线性系统模型参数估计,并通过对6种非线性回归模型的参数估计进行了验证。实验结果表明:粒子群优化算法是一种有效的参数估计方法。

关键词:粒子群优化;非线性系统;参数估计

中图分类号:TP301.6　　**文献标识码**:A　　**文章编号**:1673-629X(2008)12-0134-03

Application of PSO Algorithm in Parameter Estimation of Nonlinear Regression Models

LU Ke-zhong, FANG Kang-nian
(Department of Computer Science, Chizhou College, Chizhou 247100, China)

Abstract: Estimation of nonlinear regression model parameters is a tough searching problem. Unfortunately, the traditional approaches easily get stuck in a local minimum. Considering that the particle swarm optimization (PSO) algorithm is quite simple and easy to implement, it was used to estimate the nonlinear regression model parameters in this paper. Here six different models of nonlinear regression system were estimated by PSO algorithm and simulations demonstrated that PSO algorithm is an effective way for nonlinear system parameter estimation with global optimal.

Key words: particle swarm optimization; nonlinear system; parameter estimation

0 引 言

人们在进行科学和应用研究时,经常要进行实验数据的处理,以根据所掌握的数据来分析和推断所观察现象的内部规律。而回归分析就是实验数据处理中普遍使用的一种数学方法。但是,由于各学科的实验所涉及到的模型大都是非线性的,而非线性回归模型一般比较复杂,不像线性模型那样容易获得其参数估计,从而很大程度上限制了回归分析的应用和发展。目前常见的非线性回归模型参数估计的方法有最小二乘法、极大似然法、单纯形法[1]、模拟退火算法[2]、神经网络法[3]、遗传算法[4]等。最小二乘法和极大似然法都是基于估计目标函数对优化变量的梯度信息进行优化的,要求优化模型具有连续、可导等特性。而对非线性系统参数优化问题,由于系统具有非线性特性,使用梯度信息的局部搜索方法就不容易找到全局最优解。利用神经网络进行系统参数辨识虽然具有以任意精度逼近非线性函数的能力,但是在实际应用中,只有选择了合适的网络结构,才能获得好的结果,而选择合适的网络结构往往是非常困难的。利用遗传算法对模型参数进行估计时,复制、交叉和变异功能以及群体寻优的方式可避免陷入局部最优解,但要涉及繁琐的编码、解码过程,并且算法结构比较复杂,影响其执行效率。

粒子群优化算法(particle swarm optimization, PSO)是Kennedy和Eberhart于1995年提出的一种全局优化算法[5,6]。PSO算法同遗传算法、蚁群算法等大多数进化计算方法一样,也是一类基于群智能的随机优化算法。但与其它进化计算方法相比,PSO算法具有搜索能力强、收敛速度快、设置参数少、程序实现异常简洁、具有深刻的智能背景等特点,既适合科学研究,又特别适合工程应用[7]。文中使用PSO算法进行非线性回归模型参数估计,并与已知结果进行了比较,仿真结果表明该优化算法优化效率高、参数估计精度优,是一种有效的非线性系统模型参数估计方法,且其

收稿日期:2008-04-09
基金项目:安徽省高校青年教师科研资助项目(2006jql244)
作者简介:陆克中(1976-),男,硕士,讲师,研究方向为进化计算、生物信息学。

只利用函数值进行搜索，不需要模型连续、可导、单峰等条件，因而有着更广泛的使用范围。

1　问题的提出描述

非线性回归模型的一般形式为：

$$y_i = g(\boldsymbol{x}_i, \boldsymbol{\theta}) + e, e \sim N(0, \delta) \tag{1}$$

其中 $\boldsymbol{x}_i$ 为系统输入向量，y_i 为系统输出变量，$\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_k]^T$ 为待定参数向量，e 是均值为0、方差 δ^2 为的白噪声。模型结构 g 的形式已知。现已知观测数据对 $(\boldsymbol{x}_i, y_i)$，$i = 1, \cdots, n$。非线性参数的估计问题就是要求根据已知的数据对估计出向量 $\boldsymbol{\theta}$ 的值，即求解偏差平方和

$$f(\boldsymbol{\theta}) = \sum_{i=1}^{n} (y_i - g(\boldsymbol{x}_i, \boldsymbol{\theta}))^2 \tag{2}$$

为最小的 $\boldsymbol{\theta}$ 值。可以看出非线性回归模型参数优化问题实质就是非线性函数优化问题。由于PSO算法可以用于解决非线性、不可微和多峰值的复杂函数优化问题，所以其为解决复杂非线性系统模型的参数估计问题，提供了一条可行的途径。

2　粒子群优化算法

美国的Kennedy和Eberhart受鸟群觅食行为的启发，提出了PSO算法。PSO算法求解优化问题时，问题的解就是搜索空间中的一只鸟的位置，称这些鸟为“粒子”。所有的粒子都有一个由被优化函数决定的适应值和一个决定它们飞翔方向与距离的速度。在优化过程中，每个粒子记忆、追随当前的最优粒子，在解空间中进行搜索。PSO算法初始化为一群随机粒子，然后通过迭代找到最优解。在每一次迭代过程中，粒子通过追逐两个极值来更新自己的位置。一个是粒子自身所找到的当前最优解，这个解称为个体极值pbest；另一个是整个群体当前找到的最优解，这个解称为全局极值gbest。PSO算法数学表示如下[7]：

D 维搜索空间中，有 M 个粒子，其中第 i 个粒子的位置是 $X_i = (x_{i1}, x_{i2}, \cdots, x_{iD})$，其速度为 $V_i = (v_{i1}, v_{i2}, \cdots, v_{iD})$，$i = 1, 2, \cdots, M$。记第 i 个粒子搜索到的最优位置为 $P_i = (p_{i1}, p_{i2}, \cdots, p_{iD})$，也称为pbest，整个粒子群搜索到的最优位置为 $P_g = (p_{g1}, p_{g2}, \cdots, p_{gD})$，也称为gbest。粒子状态更新操作如下：

$$v_{id}(t+1) = w.v_{id}(t) + c_1 r_1 (p_{id} - x_{id}(t)) + c_2 r_2 (p_{gd} - x_{id}(t)) \tag{3}$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \tag{4}$$

其中，$i = 1, 2, \cdots, M$，$d = 1, 2, \cdots, D$；w 是非负常数，称为惯性因子；学习因子 c_1 和 c_2 也是非负常数；r_1 和 r_2 是介于[0,1]之间的随机数；$v_{id} \in [-v_{max}, v_{max}]$，$v_{max}$ 是之前设定的最大速率(边界值)，t 为当前迭代次数。

用于参数估计的PSO算法具体实现过程如下：

a) 设置初始参数，如种群规模 M，惯性因子 w，学习因子 c_1、c_2，估计参数向量 $\boldsymbol{\theta}$ 的大致论域范围等，并置迭代次数 $t = 1$。

b) 产生初始种群，即在一定范围内随机生成 M 个初始个体及相应的初始速度。

c) 根据方程式(2)计算每个粒子的适应值 f。

d) 将每个粒子的当前适应值与其自身的个体极值pbest和群体的历史全局极值gbest进行比较，如果某个粒子的适应值优于其个体极值pbest，则设置pbest等于此粒子的当前适应值；如果其当前适应值还优于gbest，则重设gbest等于此粒子的当前适应值。

e) 根据方程式(3)和(4)，更新每个粒子的速度与新的当前位置，并把它们限制在一定范围内。

f) $t = t + 1$，返回到步骤c)，直到获得一个预期的适应值或 t 达到设定的最大迭代次数。

3　仿真实验

为了验证利用PSO算法进行非线性回归模型参数估计的有效性以及便于比较的需要，这里采用了文献[8]中的两类6个非线性回归模型，具体如下：

1) 渐近回归模型

$$y = \alpha - \beta\gamma^x \tag{5}$$

2) S形生长模型

a) Gompertz

$$y = \alpha\exp(-\exp(\beta - \gamma x)) \tag{6}$$

b) Logistic

$$y = \alpha/(1 + \exp(\beta - \gamma x)) \tag{7}$$

c) Richards

$$y = \alpha/(1 + \exp(\beta - \gamma x))^{1/\delta} \tag{8}$$

d) MMF

$$y = (\beta\gamma + \beta x^\delta)/(\gamma + x^\delta) \tag{9}$$

e) Weibull type

$$y = \alpha - \beta\exp(-\gamma x^\delta) \tag{10}$$

在仿真实验中，选取种群规模 $M = 20$；$w = 0.9 \sim 0.4$，表示惯性因子从0.9逐渐线性衰减到0.4；学习因子取为 $c_1 = c_2 = 2$；粒子维数 D 分别与要估计的参数个数一致；最大迭代次数为2000；适应度函数 f 采用偏差平方和形式(见公式2)。仿真结果如表1、表2所示。表中的 σ^2 为残差方差估计，用于评价算法的好坏，定义为：

$$\sigma^2=\sum_{i=1}^{n}(y_i,y'_i)^2/(n-d) \tag{11}$$

其中 y_i 为实际值，y_i' 为估计值，d 为模型中所含参数的个数。

第一个计算实例是估计渐近回归模型的参数 $\boldsymbol{\theta}=(\alpha,\beta,\gamma)$，其中 $0<\boldsymbol{\theta}<1$。对文献[8]附录 5 中提供的 7 个数据集采用 PSO 算法分别进行了参数估计，其结果与文献[8]中的最小二乘结果一并列于表 1 中。PSO 算法结果是 10 次随机计算所得到的平均结果。

第二个计算实例是选取文献[8]中附录 4.A 中的数据集 1a，对 S－生长模型式(6)～(10)分别进行了参数估计，结果列于表 2 中。

表 1　渐近回归模型参数估计结果

数据集	参数	PSO结果	文献[8]结果
1	α	2.66663299	2.66663
	β	0.97253550	0.972536
	γ	0.87349897	0.873499
	σ^2	0.0077816316	0.0077816335
2	α	0.33502051	0.335021
	β	0.295512081	0.2955112
	γ	0.98382971	0.98383
	σ^2	2.01878445e－4	2.01878447e－4
3	α	41.70790280	41.7079
	β	15.83881005	15.8388
	γ	0.95628398	0.956284
	σ^2	1.454473944	1.454473945
4	α	33.80226778	33.8023
	β	26.697982028	26.698
	γ	0.752993688	0.752995
	σ^2	0.03491542967685	0.034915430
5	α	72.43262795	72.4326
	β	28.25187460	28.2519
	γ	0.59678992	0.59679
	σ^2	1.784402398	1.784402403
6	α	539.07787174	539.078
	β	307.54197123	307.542
	γ	0.53745977	0.53746
	σ^2	28.02649064	28.062491
7	α	22.48704364	22.487
	β	1.95860387	1.9586
	γ	0.70553953	0.705539
	σ^2	0.0058764270	0.005876431

表 1 和表 2 的仿真结果表明，对于给定的多组数据和多种非线性回归模型，PSO 算法对参数的估计结果与文献[8]的结果相一致，并且要好于文献[8]的结果，这说明将 PSO 算法用于非线性回归模型参数估计是有效的、可行的。

4　结束语

利用 PSO 算法对 6 个非线性回归模型进行了参数估计。在仿真过程中，PSO 算法充分体现了它具有多点寻优、算法简单、需要调整的参数较少、易于编程实现、计算量小等优点，并克服了传统算法对模型的限制条件较强、通用性较差的缺点。仿真结果证明了基于 PSO 算法的非线性回归模型参数估计是有效的、可行的，从而为非线性回归模型参数估计问题提供了一种新方法。

表 2　S－生长模型参数估计结果

数据集	参数	PSO结果	文献[8]结果
Gompertz	α	82.83102953	82.83
	β	1.22372860	1.224
	γ	0.03707588	0.037
	σ^2	3.63233170	3.6415
Logistic	α	72.46223736	72.46
	β	2.61807685	2.618
	γ	0.06735920	0.067
	σ^2	1.34275382	1.3879
Richards	α	69.53573476	69.62
	β	4.31591279	4.255
	γ	0.09018607	0.089
	δ	1.75266097	1.724
	σ^2	1.21032663	1.23195
MMF	α	80.95923807	80.96
	β	8.89450926	8.895
	γ	49577.23284308	49577
	δ	2.82799310	2.828
	σ^2	2.71143630	2.71144
Weibull type	α	69.95539860	69.96
	β	61.68181508	61.68
	γ	0.00010015	0.000100
	δ	2.37778949	2.378
	σ^2	1.67517672	1.675234

参考文献：

[1] 王　凌，李令莱，郑大钟．非线性系统参数估计的一类有效搜索策略[J]．自动化学报，2003，29(6)：953－958.

[2] 王新生，姜友华，李仁东，等．模拟退火算法及其在非线性地学模型参数估计中的应用[J]．华中科技大学学报，2001，35(1)：103－106.

[3] 蔡煜东，陈常庆，周　斌，等．用人工神经网络法辨识发酵动力学模型参数[J]．生物数学学报，1994，9(4)：103－107.

[4] 蔡煜东，陈常庆．用遗传算法辨识发酵动力学模型参数[J]．化工学报，1995，46(3)：338－342.

[5] Eberhart R C, Kennedy J. A new optimizer using particle swarm theory[C]//in: Proc. the sixth international symposium on Micro Machine and Human Science. Nagoya, Japan: [s.n.], 1995: 39－43.

[6] Kennedy J, Eberhart R C, Shi Y. Swarm intelligence[M]. San Francisco: Morgan Kaufmann Publishers, 2001.

[7] Eberhart R C, Shi Y. Particle swarm optimization: developments, applications and resources[C]//in: Proc. Congress on Evolutionary Computation. Piscataway: [s.n.], 2001: 81－86.

[8] Ratkowsky D A. Nonlinear Regression Modeling— a unified practical approach[M]. New York: Marcel Dekker Inc, 1983.