

基于进化算法的模糊抗差理论在 输电线路故障定位中应用

王绍部, 龚庆武, 舒乃秋, 陈玉林, 陈 晟
(武汉大学 电气工程学院, 湖北 武汉 430072)

摘要: 从模糊理论的角度, 介绍了一种基于进化算法的数据处理方法。该方法通过建立模糊抗差回归模型确定数据的可能性回归区间, 然后剔除回归区间外的数据; 对保留的数据作最小二乘处理, 最终达到提高计算精度和鲁棒性的目的。仿真结果证明了算法的有效性。

关键词: 模糊抗差回归分析; 数据处理; 故障定位; 进化算法

中图分类号: TM 726

文献标识码: A

文章编号: 1006-6047(2005)03-0044-03

采用多端同步采样信号的微分方程算法进行故障定位时, 微分方程经过离散化处理后, 用相邻的三个采样点信息就能算出一个定位结果, 然后对这些结果进行处理得出最终的计算结果。对于严格服从正态分布的数据, 最小二乘法^[1]具有方差最小且无偏的统计特性, 其数学模型简洁, 计算方法简单。然而, 亦存在缺点, 即不具备处理粗差的能力。在实际装置中由于采样误差和算法本身的原因, 计算结果中可能有一部分是粗差。数据中的粗差严重歪曲最小二乘法的估计结果。实践证明, 严格服从某一特定分布的观测数据是不存在的。因此, 如何恰当地考虑粗差的影响, 使定位精度进一步提高是一件重要的工作。本文提出了一种基于进化算法的模糊抗差回归分析方法。该方法不考虑数据的分布特性, 而对观测数据进行模糊线性回归分析。仿真结果表明该方法是一种鲁棒性很强的数据处理方法。

1 模糊抗差回归分析

1.1 模糊回归分析

粗差数据和精确数据之间没有明确的分界线, 是一个相对的、模糊的概念。为了表征数据的这种特性, 引入模糊回归分析的概念。

与通常的回归模型不同, 模糊回归模型^[2]把模型和数据之间的误差, 看作系统本身的模糊性。在此, 认为表示系统的系数存在模糊性。模糊回归分析的方程式如下:

$$Y_j = K_0 + K_1 x_{1j} + K_2 x_{2j} + \cdots + K_n x_{nj} = \mathbf{K} \mathbf{x}_j \quad (1)$$

式中 回归系数 K_i 是一个模糊数, 可以表示为 $K_i = (a_i, c_i)$, a_i 是模糊数的中心, c_i 是其宽度。

式(1)也可简记为

$$\mathbf{Y} = \mathbf{K} \mathbf{X}, \quad \mathbf{Y} = [Y_1, Y_2, \cdots, Y_m]$$

$$\mathbf{K} = [K_0, K_1, \cdots, K_n], \quad \mathbf{X} = \begin{bmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nm} \end{bmatrix}$$

根据模糊数的计算规则有 $\mathbf{K} \mathbf{x}_j = (a, c) \mathbf{x}_j = (a \mathbf{x}_j, c |\mathbf{x}_j|)$ 。根据式(1), 输出结果 $\mathbf{Y}$ 也是模糊数。即 $\mathbf{Y} = \mathbf{K} \mathbf{X} = (a(\mathbf{x}), c |\mathbf{x}|)$ 。所以对于任意一 Y_j , 从理论上讲应该具有中心 $a \mathbf{x}_j$ 和宽度 $c |\mathbf{x}_j|$, 而对于一个具有中心 y 和宽度 d 的数据样本, 蕴涵在模型和数据之间的关系应该是

$$\begin{aligned} ax - c|x| &\leq y - d \\ ax + c|x| &\geq y + d \end{aligned} \quad (2)$$

即普通模糊回归分析模型在满足式(2)的条件下至少要包含所有的测量数据(对于 $Y_j = K_0 + K_1 x_{1j}$ (一维情况), 如图 1 所示)。

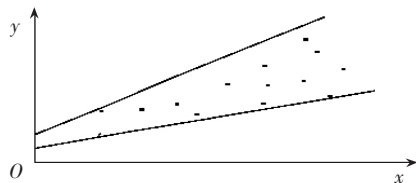


图 1 模糊回归模型

Fig.1 The fuzzy regression model

至此, 可以给出普通模糊回归分析的数学意义: 就是在满足式(2)的条件下, 求目标函数 $J = \sum_{j=1}^m c |\mathbf{x}_j|$ 的极小值问题。因为它包含了所有的数据点, 普通模糊回归模型没有抗差能力。

1.2 模糊抗差回归分析^[3]

为了恰当考虑粗差的影响, 引入了抗差模糊回归模型。其基本概念可以简述为: 设模型如图 2(a) 所示, 具有上界 $\bar{Y}$ 和下界 $\underline{Y}$, 则 y_j 相对于模型有三种分布情况, 如图 2(b) 所示。

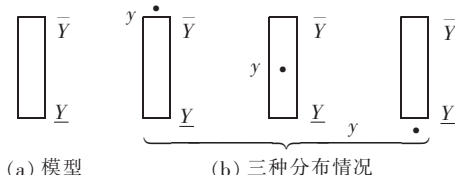


图 2 模型及其与数据点之间的关系

Fig.2 The relations between model and data

定义数据点与模型之间的距离为

$$r_j = \begin{cases} y_j - \bar{Y}_j & \bar{Y}_j < y_j \\ 0 & \bar{Y}_j \leq y_j \leq \underline{Y}_j \\ \underline{Y}_j - y_j & y_j < \underline{Y}_j \end{cases}$$

根据距离的定义,定义目标函数

$$J = (1-R) \sum_{j=1}^m |r_j| + R \sum_{i=1}^n c_i \quad 0 < R < 1 \quad (3)$$

式中 R 是一个参数,它决定了要求的模型特性,当 R 取较小的数时,目标函数侧重于将数据与模型之间的距离最小化;当 R 取一个较大的数值时,目标函数侧重于将模型的宽度最小化。

式(3)中的两项是矛盾的, $\sum_{j=1}^m |r_j|$ 增大, $\sum_{i=1}^n c_i$ 必然减小。求目标函数式(3)的最小值是一个最优化问题,本文采用了进化算法 EA (Evolutionary Algorithms) 求目标函数极值^[4]。

2 EA 概述

EA^[4]是一种迭代算法,即从一定的初值出发通过算子的运算不断地迭代,逐步改进收敛到最优解,每一次迭代称为一代。搜索过程是从原问题的一组解进化到一组更好的解,然后从这组解出发作进一步的改进,从而达到向最优解逼近的目的。每一组解为群体,每一个解为个体。

群体搜索策略和群体中的个体信息交换是 EA 的两大特点,适用于求解复杂优化问题(特别是大规模、非线性、多极值、目标函数无确切表达式的优化问题)。

3 模糊抗差回归分析在故障定位中应用

忽略输电线路电容,可建立一种 $R-L$ 近似模型。线路的两端分别标记为 m 和 n ,线路全长为 l 。假设三相线路均匀换位,其等值网络如图 3 所示(图中 R_0, L_0 是零序电阻和零序电感)。

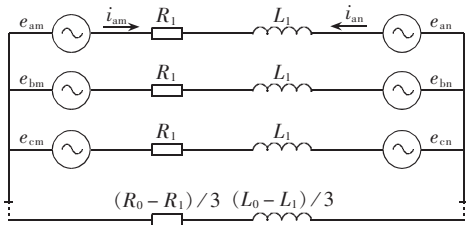


图 3 输电线路的 $R-L$ 模型

Fig.3 $R-L$ model of transmission line

对线路进行 Karenbauer^[5]变换后,由于线模分量对应的线路参数在数值上等于正序参数,因此线模分量回路可简化为图 4,图中 p 为短路点 f 到 m 侧的距离占线路全长的百分比。

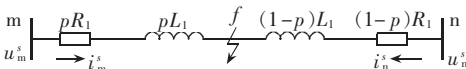


图 4 输电线路模示意图

Fig.4 The diagram of line mode for transmission line

由图 4 可得:

$$u_{mn}^s = p \left(R_1 i_m^s + L_1 \frac{di_m^s}{dt} + R_1 i_n^s + L_1 \frac{di_n^s}{dt} \right) - \left(R_1 i_n^s + L_1 \frac{di_n^s}{dt} \right) \quad (4)$$

式中 $i_m^s, i_n^s (s = \alpha, \beta)$ 分别为第 s 线模量两侧的电流; R_1, L_1 为正序电阻和电感; $u_{mn}^s (s = \alpha, \beta)$ 为第 s 线模量两侧的电压差。

由式(4)得:

$$p = \frac{u_{mn}^s + R_1 i_n^s + L_1 \frac{di_n^s}{dt}}{R_1 i_m^s + L_1 \frac{di_m^s}{dt} + R_1 i_n^s + L_1 \frac{di_n^s}{dt}} \quad (5)$$

离散化后可得:

$$p(k) = \frac{u_{mn}^s(k) + R_1 i_n^s(k) + L_1 \frac{i_n^s(k+1) - i_n^s(k-1)}{2\Delta t}}{R_1 i_m^s(k) + L_1 \frac{i_m^s(k+1) - i_m^s(k-1)}{2\Delta t} + R_1 i_n^s(k) + L_1 \frac{i_n^s(k+1) - i_n^s(k-1)}{2\Delta t}} \rightarrow$$

式中 k 是采样点的序号; Δt 是两个相邻采样点之间的时间差。

令

$$y(k) = u_{mn}^s(k) + R_1 i_n^s(k) + L_1 \frac{i_n^s(k+1) - i_n^s(k-1)}{2\Delta t}$$

$$x(k) = R_1 i_m^s(k) + L_1 \frac{i_m^s(k+1) - i_m^s(k-1)}{2\Delta t} + R_1 i_n^s(k) + L_1 \frac{i_n^s(k+1) - i_n^s(k-1)}{2\Delta t}$$

则有

$$p(k) = y(k) / x(k) \quad (6)$$

在距 m 侧 10% 处发生 B 相接地短路故障时, y 和 x 对应关系的分布如图 5 所示。

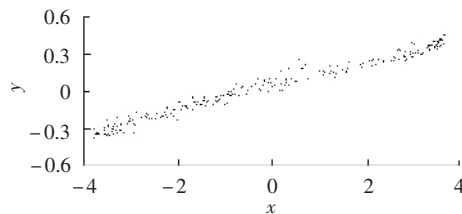


图 5 y 和 x 对应关系分布图

Fig.5 The distribution of y by x

令式(1)中的 $K_1 = P, K_0 = K_2 = \dots = K_n = 0$ 进行模糊抗差回归分析。即选择恰当的 R 用分配区间型进化算法^[6]确定直线方程 $y_1 = k_1 x$ 和 $y_2 = k_2 x$, 使得式(3)最小(式(3)中的 $c_1 = k_2 - k_1$ (设 $k_2 > k_1$))。其过程大致可归结为六步。

a. 产生初始群体 $(K_1, K_2, \dots, K_n), K_i = [k_{i1}, k_{i2}]$; 代数序号 $\text{gen} = 1$, 有两种方法: 在给定区间 $[a, b]$ 内随机生成(按均匀分布), 或者用正交试验优化选取。这里采用了前一种方法。

b. 变异产生 N 个个体 $(K_{n+1}, K_{n+2}, \dots, K_{2n}), K_{n+1} =$

$[k_{(n+1)1}, k_{(n+1)2}]$, $k_{(n+1)1} = k_1 + \Delta k_1$, $k_{(n+1)2} = k_2 + \Delta k_2$; Δk 是按柯西分布的随机数。如果 $k_{(n+1)}$ 的值超出 $[a, b]$, 则采用反射的方法使其重新落入 $[a, b]$ 内。

c. 对 $2N$ 个个体按适应度高低排(即由 $K_i = [k_{i1}, k_{i2}]$ 确定的直线 $y = k_{i1}x$ 和 $y = k_{i2}$ 使得式(3)的值较小的个体适应度高)。

d. 用距离函数使个体分配区间(即当 $|K_i - K_j| \leq L$ (L 为预先设定的阈值)则淘汰两者中适应度较小的个体, 以防止近亲繁殖)。

e. 选择适应度较高的个体(取前 N 个个体);

f. $gen = gen + 1$, 判断收敛与否, 若不收敛返回到 b, 直到收敛。本文采用的收敛准则: 超过预先设定的最大进化代数或者全局最优解(适应度最高的个体)在连续 30 代的进化过程中变化很小(小于预先设定的阈值), 可认为收敛。从而得到最优的 k_1, k_2 , 由进化算法得到的回归模型如图 6 所示。

于是 $P = K_1 = (a_1, c_1)$, 其中 k_1, k_2, K_1 之间的关系为 $K_1 = (a_1, c_1)$; a_1 是模糊数 K_1 中心, c_1 是其宽度, $c_1 = k_2 - k_1$ (设 $k_2 > k_1$)。然后保留 $p(k) \in P$ (即保留 $p(k_1 <$

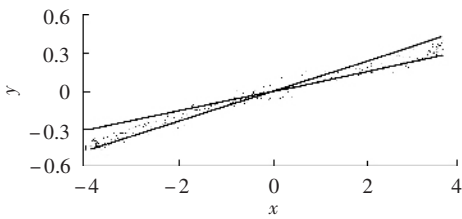


图 6 由进化算法得到的回归模型
Fig.6 The regression model from EA

$p < k_2$)), 对保留的 $p(k)$ 进行最小二乘^[1]处理得出最终定位结果。

4 RTDS 仿真结果

本文以宁夏石嘴山供电局 220 kV 新平线为对象建立模型。采样率 4 kHz, 线路参数 $z_1 = 0.053\ 4 + j0.316\ 3(\Omega/\text{km})$, $z_0 = 0.304\ 5 + j1.190\ 3(\Omega/\text{km})$, 线路全长 59.24 km, 式(3)中的 R 取 0.95, 故障类型为 B 相接地。计算结果见表 1。

由表 1 计算结果可以看出, 采用本算法所得计算结果精度较高且较为稳定。

表 1 计算结果
Tab.1 Computed results

p	经典最小二乘				模糊抗差最小二乘			
	$R_g=0\ \Omega$		$R_g=50\ \Omega$		$R_g=0\ \Omega$		$R_g=50\ \Omega$	
	计算结果	误差	计算结果	误差	计算结果	误差	计算结果	误差
5	6.374 6	1.374 6	5.463 0	0.463 0	4.824 8	-0.175 2	5.279 3	0.279 3
10	9.987 1	-0.012 9	10.706 7	0.706 7	9.878 0	-0.122 0	10.575 3	0.575 3
20	19.914 4	-0.085 6	20.610 7	0.610 7	19.936 0	-0.064 0	20.548 0	0.548 0
30	30.006 7	0.006 7	30.315 8	0.315 8	29.978 6	-0.021 4	30.238 7	0.238 7
40	40.008 2	0.008 2	40.493 3	0.493 3	39.996 6	-0.003 4	40.411 1	0.411 1
50	49.811 6	-0.188 4	50.348 8	0.334 8	50.028 6	0.028 6	50.320 2	0.320 2
60	59.576 4	-0.423 6	60.397 4	0.397 4	60.078 7	0.078 7	60.302 0	0.302 0
70	70.276 5	0.276 5	70.307 1	0.307 1	70.265 9	0.265 9	70.273 8	0.273 8
80	79.795 8	-0.204 2	80.282 8	0.282 8	80.247 2	0.247 2	80.260 3	0.260 3
90	88.792 9	-1.207 1	90.161 9	0.161 9	90.229 9	0.229 9	90.133 7	0.133 7
95	93.471 9	-1.528 1	95.092 7	0.092 7	95.272 5	0.272 5	95.202 1	0.202 1
平均误差	—	0.461 9	—	0.379 0	—	0.137 1	—	0.322 2
最大误差	—	1.528 1	—	0.706 7	—	0.272 5	—	0.575 3

5 结语

本文采用基于进化算法的模糊抗差思想使得定位结果的精度在很大程度上克服采样数据分布特性的影响, 提高了定位精度和鲁棒性。

参考文献:

[1] CHEN Yun-ping, GONG Qing-wu, FU Feng-nian, et al. An accurate fault location and its data processing by least square method[A]. **Proceedings PowerCon 2002 International Conference**[C]. Beijing: Power System Technology, 2002.13 – 17.

[2] [日]寺野寿郎, 浅居喜代治, 菅野道夫. 模糊系统理论及应用[M]. 刘维仲, 徐平, 姜延译. 天津: 天津大学出版

社, 1991.

[3] WATADA J, YABUUCHI Y. IEEE world congress on computational “Fuzzy robust regress analysis” fuzzy systems, intelligence [A]. **Proceedings of the Third IEEE Conference**[C]. Orland, FL USA: IEEE, 1994. 1370 – 1376.

[4] 云庆夏. 进化算法[M]. 北京: 冶金工业出版社, 2000.

[5] 解广润. 电力系统过电压[M]. 北京: 水利电力出版社, 1989.

[6] 胡云昌, 骆寒冰, 徐慧, 等. 解决多峰性函数全局最优解的进化算法及其应用研究[J]. 中国造船, 1998, (3): 84 – 95.

HU Yun-chang, LUO Han-bing, XU Hui, et al. Study on evolutionary algorithms to search global optimizations of multimodel function and its application[J]. **China Shipbuilding**, 1998, (3): 84 – 95.

(责任编辑: 李玲)

作者简介：

王绍部(1976－),男,山东临沂人,硕士研究生,主要从事电力系统继电保护和自动化方面的研究(E-mail:yinghuabu@sohu.com);

龚庆武(1967－),男,湖南怀化人,副教授,主要从事电力

系统继电保护和自动化方面的研究;

舒乃秋(1954－),男,湖北武汉人,教授,主要从事电力系统在线监测方面的研究;

陈玉林(1980－),男,湖北武汉人,博士研究生,主要从事电力系统继电保护和自动化方面的研究;

陈 晟(1979－),男,江西丰城人,硕士研究生,主要从事电力系统稳定方面的研究。

Application of evolutionary algorithm based fuzzy robust theory in transmission line fault locating

WANG Shao-bu,GONG Qing-wu,SHU Nai-qi,CHEN Yu-lin,CHEN Sheng
(Wuhan University,Wuhan 430072,China)

Abstract: A data processing method based on evolutionary algorithms is presented. The fuzzy robust regression model is established to determine the possible regression area of data. The out-area data are discarded and the reserved data are processed by least square method,which improves the calculation precision and robusticity. The simulative results verify its effectiveness.

Key words: fuzzy robust regression analysis; data processing; fault locating; evolutionary algorithm