Operations Research and Fuzziology 运筹与模糊学, 2019, 9(4), 292-298
Published Online November 2019 in Hans. http://www.hanspub.org/journal/orf
https://doi.org/10.12677/orf.2019.94034

Hans汉斯

An Improved Exponential Curve Model and its Application in the Total Water Resources in Tibet Autonomous Region

Ping Zhang[1*], Feng Luo[2]

[1]The Engineering & Technical College of Chengdu University of Technology, Department of Fundamental Education, Sichuan Leshan
[2]Leshan Vocational & Technical College, Department of Mechanical and Electrical Engineering, Sichuan Leshan
Email: *zhpingls@163.com

Received: Oct. 17th, 2019; accepted: Oct. 31st, 2019; published: Nov. 7th, 2019

Abstract

Based on the exponential curve model and the modified exponential curve model, an improved exponential curve model is proposed by introducing the time power term in this paper. According to the idea of non-linear least square method and trust region algorithm, the parameters of the new exponential curve model are estimated. On this basis, the model is applied to the total water resources of Tibet Autonomous Region. Firstly, the specific values of system parameters are solved by using the Matlab software, and then the specific expression of the new model is given. Furthermore, the results of the new model are compared with the classical exponential curve and the modified exponential curve model. The calculation results show that the new modified model has higher prediction accuracy on the total water resources in Tibet Autonomous Region.

Keywords

The Exponential Curve, Time Power Term, Non-Linear Least Squares Method, the Total Water Resources

改进的指数曲线模型及其在西藏自治区水资源总量中的应用

张　萍[1*]，罗　缝[2]

[1]成都理工大学工程技术学院基础教学部，四川 乐山
[2]乐山职业技术学院机电工程系，四川 乐山
Email: *zhpingls@163.com

*通讯作者。

文章引用: 张萍, 罗缝. 改进的指数曲线模型及其在西藏自治区水资源总量中的应用[J]. 运筹与模糊学, 2019, 9(4): 292-298. DOI: 10.12677/orf.2019.94034

收稿日期：2019年10月17日；录用日期：2019年10月31日；发布日期：2019年11月7日

摘　要

本文在指数曲线模型和修正的指数曲线模型的基础上引入时间幂次项，提出了一种改进的新型指数曲线模型。利用非线性最小二乘法的思想，结合信赖域算法对新提出的指数曲线模型的参数进行了估计。在此基础上，将模型应用于西藏自治区水资源总量中。首先，借助Matlab软件求解出系统参数的具体取值，并给出模型的具体表达式。进一步，将建模的结果与经典的指数曲线、修正的指数曲线进行比较，计算结果表明改进的指数曲线模型在西藏自治区水资源总量问题上有更高的预测精度。

关键词

指数曲线，时间幂次项，非线性最小二乘法，水资源总量

1. 引言

现实世界中，社会经济现象总是在不断的变化与发展的。当社会经济变化的长期趋势呈现出稳定的线性趋势时，可以使用移动平移法、指数平滑法、直线趋势方程拟合法对其进行预测研究。当社会经济变化的长期趋势不是线性时，但又呈现一定规律，就需要配合适当的非线性趋势曲线。比如抛物曲线型、指数曲线型、修正曲线指数型、龚珀兹曲线型等。对于长期趋势，若按照大体相同的增长速度递增或递减时，可以采用指数曲线模型或修正指数曲线模型来进行研究。

唐五湘[1]利用两步最小二乘法研究了指数曲线模型的参数估计，并以我国长途电话去话张数为例进行了数值计算。白雪梅和赵松山[2]分析了传统指数曲线模型参数估计存在的问题、适用条件并提出了改进方法和非线性近似估计法。杨佳元[3]根据指数曲线模型的特点，对指数曲线模型的假设、参数估计、拟合误差以及无偏性等问题进行了讨论。毛艺萍和王斌会[4]利用时间序列相邻两项之差建立回归方程，并用最小二乘法对模型参数进行估计。秦尚林等人[5]利用指数曲线模型并结合武广铁路客运专线部分实测资料讨论了路基沉降问题。张军等人[6]等人提出了修正指数曲线模型参数估计的一种新方法，并经过数据实验证明了参数估计新方法的有效性和实用性。欧阳明等人[7]在现有模型的基础上构造了一个新的修正的指数曲线模型。通过对不同类型的单桩静载荷试验数据进行拟合，验证了提出的新模型能够很好的对单桩 P-S 曲线进行拟合。陈希鸣等人[8]针对灰色 GM(1, 1)模型和指数曲线模型在高速公路沉降预测中的不足，采用不等权系数将二者形成组合模型预测沉降，从而达到提高沉降预测精度的目的。

本文在上述研究的基础上，提出了时间幂次项的指数曲线以及它的参数估计方法，并将其应用于西藏自治区水资源总量中。选择西藏自治区水资源总量的研究，主要是因为西藏自治区水资源总量是我国最丰富的地方，人均占水量和亩均占有量均居全国首位。对西藏自治区水资源总量的研究就显得格外重要。本文涉及到的数据是从中国统计年鉴上获得的，其中 2004 年至 2015 年的实际数据作为建模数据，2016 年至 2017 年的数据为模型的外推预测检验数据。对比分析指数曲线、修正的指数曲线以及本文提出的新型指数曲线，计算结果显示本文的指数曲线模型预测精度高于其它两个模型。

2. 指数曲线和修正的指数曲线

2.1. 指数曲线

由文献[9]，经典的指数曲线为

$$Y_t = ab^t \text{。} \tag{1}$$

为了估计参数 a,b，一般将(1)式两端取对数，得到

$$\ln Y_t = \ln a + t\ln b \text{。} \tag{2}$$

然后运用最小二乘法和方程(2)，得到如下方程

$$\begin{cases} \sum \ln Y = n\ln a + \left(\sum t\right)\ln b \\ \sum t\ln Y = \left(\sum t\right)\ln a + \left(\sum t^2\right)\ln b \end{cases}, \tag{3}$$

估计出参数 $\ln a$ 和 $\ln b$，再取反对数，即可得到参数 a,b 的估计值。

2.2. 修正指数曲线

在经典指数曲线的基础上增加一个常数 K，即得到修正指数曲线模型

$$Y_t = ab^t + K, \tag{4}$$

其中 K、a、b 为未知参数，$K \in (0,\infty)$，$a \neq 0$，$b \in (0,1) \cup (1,\infty)$。

参数 K、a、b 估计的基本思想是三和法：把整个时间序列分成相等的三个数组，每个组有 m 项，根据趋势值 Y_t 的三个局部总和分别等于原数列观察值 Y 的三个局部总和来确定三个参数。具体为：设观察值的三个局部总和分别为 S_1、S_2、S_3，得

$$S_1 = \sum_{t=0}^{m-1} Y_t,\quad S_2 = \sum_{t=m}^{2m-1} Y_t,\quad S_2 = \sum_{t=2m}^{3m-1} Y_t \text{。} \tag{5}$$

由三和法得到如下方程

$$\begin{cases} S_1 = mK + a + ab + ab^2 + \cdots + ab^{m-1} \\ S_2 = mK + ab^m + ab^{m+1} + \cdots + ab^{2m-1} \\ S_3 = mK + ab^{2m} + ab^{2m+1} + \cdots + ab^{3m-1} \end{cases} \text{。} \tag{6}$$

通过方程(6)解得

$$\begin{cases} b = \left(\dfrac{S_3 - S_2}{S_2 - S_1}\right)^{\frac{1}{m}} \\ a = \left(S_2 - S_1\right)\dfrac{b-1}{\left(b^m - 1\right)^2} \\ K = \dfrac{1}{m}\left(S_1 - \dfrac{a\left(b^m - 1\right)}{b-1}\right) \end{cases} \text{。} \tag{7}$$

3. 改进的幂次项指数曲线模型

在上面指数模型的基础上，本文提出幂次项指数曲线，其一般方程为

$$Y_t = ab^t + ct^\alpha + d, \tag{8}$$

其中参数为a,b,c,d,α。相比于经典的指数曲线模型和修正的指数曲线模型，最大的一点是引入$ct^{\alpha}+d$部分以描述序列的变化趋势。但是，改进的指数曲线带有5个未知参数，且方程(8)本身为非线性函数。所以，在对参数的估计时采用的是非线性最小二乘法。

令$X=(a,b,c,d,\alpha)$，$Y=(Y_1,Y_2,\cdots,Y_n)^{\mathrm{T}}$为观测值，则通过方程(8)可得到如下的非线性方程组

$$\hat{Y}=f(X)=\begin{bmatrix} ab+c+d \\ ab^2+c2^{\alpha}+d \\ \vdots \\ ab^n+cn^{\alpha}+d \end{bmatrix}。\tag{9}$$

求非线性方程组(9)的最小二乘估计量$\hat{X}=(\hat{a},\hat{b},\hat{c},\hat{d},\hat{\alpha})$，构造如下的残差平方和最小：

$$\min\quad I=\sum_{t=1}^{n}\Delta E_t^2=\left\|f(\hat{X})-Y\right\|^2=\left(f(\hat{X})-Y\right)^{\mathrm{T}}\left(f(\hat{X})-Y\right),\tag{10}$$

其中$\Delta E_t=\hat{a}\hat{b}^t+\hat{c}t^{\hat{\alpha}}+\hat{d}-Y_t$，这是一个非线性无约束最优化问题。

对于上述问题，本文采用信赖域算法来求解(10)式。信赖域算法最早可以追溯到Levenberg-Marquardt方法，它不是直接求解目标函数，而是寻找一个与之近似的问题，称为信赖域子问题，然后求得该问题的极小值进而获得目标函数的最优解。其基本思想是：先给出一个目标函数的初始估计解，然后求解目标函数的近似模型，得到的解称为试探步；通过试探步可以调节下一个迭代点，并调节信赖域半径；反复求解近似模型就可以不断更新迭代点和调节信赖域半径，直到求得目标函数的最优解为止。信赖域算法在每步的迭代过程中，都是在求解信赖域子问题，而子问题的解又总是限定在一个可信赖的广义球内。关于信赖域子问题中目标函数的构造本文选取二次模型。

对于式(10)的非线性无约束优化问题，利用二次逼近，信赖域模型的目标函数一般形式可以表示为：

$$Q_k(d_k)=I(\hat{X}_k)g_k^{\mathrm{T}}d_k+\frac{1}{2}d_k^{\mathrm{T}}H_kd_k,\tag{11}$$

于是信赖域子问题构造如下：

$$\begin{aligned} q_k(d_k)&=I(\hat{X}_k)g_k^{\mathrm{T}}d_k+\frac{1}{2}d_k^{\mathrm{T}}H_kd_k, \\ \text{s.t.}\quad &\|d_k\|\le\Delta_k \end{aligned}\tag{12}$$

其中$g_k=\nabla I(x_k)$为目标函数的梯度，有

$$g_k=\left(g_1^{(k)},g_2^{(k)},g_3^{(k)},g_4^{(k)},g_5^{(k)}\right)=\left(\frac{\partial I}{\partial a},\frac{\partial I}{\partial b},\frac{\partial I}{\partial c},\frac{\partial I}{\partial d},\frac{\partial I}{\partial \alpha}\right)\Bigg|_{X=X^{(k)}},k=0,1,\cdots,\tag{13}$$

x_k为第k次迭代点，$d_k=x_{k+1}-x_k$为试探步，$H_k=\nabla^2I(x_k)$为Hess矩阵

$$H_k=\begin{bmatrix} \frac{\partial^2 I}{\partial a^2} & \frac{\partial^2 I}{\partial ab} & \frac{\partial^2 I}{\partial ac} & \frac{\partial^2 I}{\partial ad} & \frac{\partial^2 I}{\partial a\alpha} \\ \frac{\partial^2 I}{\partial ba} & \frac{\partial^2 I}{\partial b^2} & \frac{\partial^2 I}{\partial bc} & \frac{\partial^2 I}{\partial bd} & \frac{\partial^2 I}{\partial b\alpha} \\ \frac{\partial^2 I}{\partial ca} & \frac{\partial^2 I}{\partial cb} & \frac{\partial^2 I}{\partial c^2} & \frac{\partial^2 I}{\partial cd} & \frac{\partial^2 I}{\partial c\alpha} \\ \frac{\partial^2 I}{\partial da} & \frac{\partial^2 I}{\partial db} & \frac{\partial^2 I}{\partial dc} & \frac{\partial^2 I}{\partial d^2} & \frac{\partial^2 I}{\partial d\alpha} \\ \frac{\partial^2 I}{\partial \alpha a} & \frac{\partial^2 I}{\partial \alpha b} & \frac{\partial^2 I}{\partial \alpha c} & \frac{\partial^2 I}{\partial \alpha d} & \frac{\partial^2 I}{\partial \alpha^2} \end{bmatrix}_{X=X^{(k)}},k=0,1,\cdots\tag{14}$$

DOI: 10.12677/orf.2019.94034

$\Delta_k > 0$ 为信赖域半径，其要在信赖域范围内，$\|.\|$ 为 Euclid 范数。

求解式(10)得到校正量 d_k 后，进一步求解

$$Ared_k = I\left(\hat{X}_k\right) - I\left(\hat{X}_k + d_k\right), \tag{15}$$

$$Pred_k = Q_k\left(0\right) - Q_k\left(d_k\right) = -\left(g_k^{\mathrm{T}} d_k + \frac{1}{2} d_k^{\mathrm{T}} H_k d_k\right), \tag{16}$$

$$r_k = \frac{Ared_k}{Pred_k}, \tag{17}$$

其中 $Ared_k$ 为实际下降量，$Pred_k$ 为预测下降量，下降比率 r_k 为下一次的迭代半径。根据比值 r_k 的大小来判断近似程度，r_k 越接近 1 时，收敛越好，可增大 r_k；r_k 小于 0 时，不收敛，需要缩小半径。

4. 具体应用

为了对新型的指数曲线的预测精度进行检验，本文选用西藏自治区水资源总量实际数据来验证分析，并将计算结果与已有的指数曲线、修正的指数曲线模型的计算结果进行对比分析。本文选取西藏自治区水资源总量 2004 年至 2017 的统计数据，见表 1。

Table 1. The raw data of the total water resources of tibet autonomous region (hundred million cubic meters)
表 1. 西藏自治区水资源总量的统计数据(亿立方米)

年份	2004	2005	2006	2007	2008	2009	2010
实际值	4665.16	4451.07	4157.14	4321.38	4560.20	4029.16	4593.00
年份	2011	2012	2013	2014	2015	2016	2017
实际值	4402.71	4196.35	4415.74	4416.30	3853.00	4642.20	4749.90

为了检验模型的预测精度，根据预测值与实际值确定绝对百分误差(APE)和平均绝对百分误差(MAPE)如下。

$$\text{APE} = \left|\frac{Y_t - \hat{Y}_t}{Y_t}\right| \times 100\%, t = 1, 2, \cdots, n, \tag{18}$$

$$\text{MAPE} = \frac{1}{v - l + 1} \sum_{t=l}^{v} \left|\frac{Y_t - \hat{Y}_t}{Y_t}\right| \times 100\%, v \le n\text{。} \tag{19}$$

从表达式(19)可知，当 $l = 2, v = m$，MAPE 为拟合误差，记为 MAPE_{fit}；当 $l = m + 1, v = n$，MAPE 为预测误差，记为 $\text{MAPE}_{\text{fore}}$；当 $l = 2, v = n$，MAPE 为总误差，记为 $\text{MAPE}_{\text{total}}$。

通过 Matlab 编写数值计算程序，利用 2004 年至 2015 年的数据建立三种指数曲线模型，其表达式分别为

$$Y_t = 4521.5564 \times 0.9936^t, \tag{20}$$

$$Y_t = -0.0012 \times 2.9199^t + 4398.7213, \tag{21}$$

$$Y_t = -4336.3911 \times \left(-0.1743\right)^t + 14.3774 \times t^{1.6839} + 3583.8882\text{。} \tag{22}$$

通过表达式(20)~(22)，可以计算得到相应的拟合值和预测值，具体见表 2。从计算结果看出改进的指数曲线对西藏自治区水资源总量的预测更接近。指数曲线的拟合误差、预测误差和总的误差分别为

4.3780%，11.6675%和 5.4995%；修正的指数曲线的拟合误差、预测误差和总的误差分别为 3.4466%，63.9725%和 12.7583%；幂次项的指数曲线的拟合误差、预测误差和总的误差分别为 9.9810%，0.8380%和 8.5744%。

修正的指数曲线在拟合上是最优的，其值为 3.4466%，比幂次项的指数曲线小了将近 3 倍。但是，幂次项的指数曲线预测误差是最低的，为 0.8380%，远远小于指数曲线和修正的指数曲线的值。这说明，幂次项的指数曲线在预测上比前两种模型有更高的精度，相应的结果的可靠性也更为可靠。同时，也说明了本文提出的新型幂次项指数曲线在一定程度上是能够提高模型的预测精度的。

Table 2. The calculation results of the total water resources in Tibet autonomous region by using the exponential curve, modified exponential curve and improved exponential curve
表 2. 指数曲线、修正的指数曲线和新型指数曲线对西藏自治区水资源总量的计算结果

年份	实际值	指数曲线	APE (%)	修正指数曲线	APE (%)	新型指数曲线	APE (%)
2004	4665.16	4492.7207	3.6963	4398.7176	5.7113	4354.4133	6.6610
2005	4451.07	4464.0690	0.2920	4398.7107	1.1763	3498.2293	21.4070
2006	4157.14	4435.6000	6.6984	4398.6904	5.8105	3698.3086	11.0372
2007	4321.38	4407.3125	1.9885	4398.6312	1.7877	3728.2893	13.7246
2008	4560.20	4379.2054	3.9690	4398.4584	3.5468	3800.6831	16.6553
2009	4029.16	4351.2776	7.9947	4397.9538	9.1531	3877.5154	3.7637
2010	4593.00	4323.5279	5.8670	4396.4802	4.2787	3964.7169	13.6791
2011	4402.71	4295.9551	2.4248	4392.1776	0.2392	4060.7058	7.7680
2012	4196.35	4268.5582	1.7207	4379.6141	4.3672	4165.3080	0.7397
2013	4415.74	4241.3360	3.9496	4342.9294	1.6489	4278.1742	3.1154
2014	4416.30	4214.2875	4.5742	4235.8119	4.0869	4399.0389	0.3908
2015	3853.00	4187.4114	8.6792	3923.0346	1.8177	4527.6633	17.5101
2016	4642.20	4160.7067	10.3721	3009.7415	35.1656	4663.8366	0.4661
2017	4749.90	4134.1723	12.9630	342.9741	92.7793	4807.3699	1.2099
$MAPE_{fit}$ (%)			4.3780		3.4466		9.9810
$MAPE_{fore}$ (%)			11.6675		63.9725		0.08380
$MAPE_{total}$ (%)			5.4995		12.7583		8.7544

5. 结论

本文在传统的指数曲线基础上，引入时间幂次项来描述现实世界的序列。通过非线性最小二乘法构建最优目标函数来求解模型的各个参数的取值，并进一步将其应用在西藏水资源总量的拟合预测中。从计算结果可以看出新型的指数曲线模型预测精度高于指数曲线和修正指数曲线模型。本文提出的模型在一定程度上能够提高模型的预测精度，在今后的研究中可将这种推广模型引入到如龚珀茨模型中。

参考文献

[1] 唐五湘. 指数曲线模型参数估计的两步最小二乘法[J]. 预测, 1994(6): 68-69.

[2] 白雪梅, 赵松山. 对指数曲线模型参数估计方法的探讨[J]. 数量经济技术经济研究, 1997(10): 50-51.

[3] 杨桂元. 指数曲线预测模型的参数估计与误差分析[J]. 运筹与管理, 2003, 2(4): 55-58.

[4] 毛艺萍, 王斌会. 增长型曲线模型参数估计的新思路[J]. 统计与决策, 2006(8): 14-16.

[5] 秦尚林, 陈善雄, 许锡昌. 路基沉降预测的拓展指数曲线模型[J]. 铁道标准设计, 2010(2): 28-30.

[6] 张军, 吕雄, 姚贵平. 修正指数曲线模型参数估计的一种新方法[J]. 内蒙古农业大学学报(自然科学版), 2012(1): 285-290.

[7] 欧阳明, 丁伯阳, 石吉森. 单桩荷载-沉降曲线的修正指数曲线模型拟合研究[J]. 水运工程, 2013(1): 31-38.

[8] 陈希鸣, 黄张裕, 王睿, 秦洁, 刘仁志. 基于灰色 GM(1,1)与指数曲线组合模型的高速公路沉降预测[J]. 勘察科学技术, 2019(1): 10-12.

[9] 贾俊平, 何晓群, 金勇进. 统计学(第 7 版)[M]. 北京: 中国人民大学出版社, 2019.