Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Fitting Gompertz Curve Using Grey Method

Lifeng Wu
College of Mathematics and Information Science
Wenzhou University
Wenzhou, 325035, Zhejiang, China
wlf6666@126. com

Yinao Wang
College of Mathematics and Information Science
Wenzhou University
Wenzhou, 325035, Zhejiang, China
wyinao@sina.com

***Abstract*—Based on the grey system direct modeling method, we give an approach for fitting Gompertz curve under the criteria of minimized the mean absolute percentage error (MAPE) in this article, case analysis shows that the fitting accuracy in traditional modeling has been significantly improved by the approach we proposed.**

***Keywords*—grey system direct modeling method, fitting Gompertz curve, MAPE.**

I. Introduction

Grey system theory has already been widely used in many fields since 1982 [1-2], grey modeling method is an important component of the grey system theory, as one of the grey modeling methods, direct modeling method has been presented [3], the characteristics of the direct modeling method are studied [4-7].

In this paper, we consider the fitting problem for Gompertz function

$$x(t) = e^{a-be^{-ct}}, \quad b, c > 0, \quad a \in R \tag{1}$$

which is a solution of the so-called Gompertz growth rate model [8]

$$\frac{dx}{dt} = cx\ln\left(\frac{e^a}{x}\right), \quad c > 0, a \in R. \tag{2}$$

Gompertz curve occurs in various of applied research, for example, for fitting seedling emergence [9], continuous heterogeneous mortality [10], forestry product value forecasting model [11], forecasting Shanghai composite index [12], establish the model which reflect the relationship between the development expenses and the development time of a missle [13].

Although Gompertz function had been widely adopted, its fitting performance still could be improved. Up to the present, there are many scholars proposed new method to improve the fitting precision of the Gompertz function. Jukic et al gave necessary and sufficient conditions which guarantee the existence of least square estimate Gompertz curve parameter and suggest a choice of a good initial approximation [8]. Uniqueness theorems for estimation of the age-dependent parameter in the Gompertz survival model with a mortality deceleration rate had been proven [14]. The coordinates and differential coefficients of the Gompertz function inflection points and its parameters can be uniquely determined each other , according to this property, the initial value of its single non-linearized parameters can be searched during the model is fitting in the meaning of least square by means of GNL method which can fit the curve of implicit function [15]. In all the cases above, the adopted approach includes fitting the Gompertz curve to the data by means of a well-known estimation procedure, such as least squares, least squares method requires an initial approximation which is as good as possible, if we chose the initial approximation is not good, the fitting accuracy is not high after the iterations [8]. In this paper, we suggest a somewhat different approach, which consists of expressing the curve through its differential equation and searching for the coefficient a, b and c under the criteria of minimized the mean absolute percentage error, this approach dose not require an initial approximation, parameters optimization problem of Gompertz curve is formulated as the linear programming problem.

II. Direct Modeling Method of GM(1,1)

Let $X = \{x(1), x(2), \cdots, x(n)\}$ be a raw sequence, using these data to direct modeling GM(1,1), the steps are as follows [3]:

A. *Ascertain data matrixe* B, Δ

$$B = \begin{bmatrix} -\frac{1}{2}(x(1)+x(2)) & 1 \\ -\frac{1}{2}(x(2)+x(3)) & 1 \\ \vdots & \vdots \\ -\frac{1}{2}(x(n-1)+x(n)) & 1 \end{bmatrix}$$

$$\Delta = \begin{bmatrix} x(2)-x(1) \\ x(3)-x(2) \\ \vdots \\ x(n)-x(n-1) \end{bmatrix}$$

B. Calculation

$$\begin{bmatrix} \hat{a} \\ \hat{u} \end{bmatrix} = \left(B^T B\right)^{-1} B^T \Delta$$

we obtain the estimated value $\hat{a}, \hat{u}$ of the parameters a, u of the differenial equation model $\frac{dx}{dt} + ax = u$, solve differenial equation $\frac{dx}{dt} + \hat{a}x = \hat{u}$, we obtain

$$x(t) = ce^{-\hat{a}t} + b \tag{3}$$

where b, c is unknown parameters.

C. Using the least squares method to obtain the estimated value $\hat{b}, \hat{c}$ of parameters b, c,

$$\begin{bmatrix} \hat{b} \\ \hat{c} \end{bmatrix} = \left(D^T D\right)^{-1} D^T X$$

where

$$D = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{-\hat{a}} & e^{-2\hat{a}} & \cdots & e^{-n\hat{a}} \end{bmatrix}, \; X = \begin{bmatrix} x(1) \\ x(2) \\ \vdots \\ x(n) \end{bmatrix}$$

then
$$\hat{x}(t) = \hat{c}e^{-\hat{a}t} + \hat{b} \tag{4}$$

this method can be described as follows: firstly, estimating the development coefficient $(-a)$ by whitening the grey differenial equation

$$\frac{d \otimes x(t)}{dt} + a \otimes x(t) = u$$

then estimating the parameters b and c of equation (3) by least squares method once again.

However, in most references, estimating the parameters of the grey model by least squares method, evaluating the error by mean absolute percentage error, in other words, method of optimization goal function is not consistent with test method of model accuracy, so in order to solved this consistency problem. Liu and Zhang set the absolute of mean relative error as objective function, then obtaining parameters of GM(2,1, λ ρ) use the particle swarm optimization [16]. Zhou et al set MAPE as objective function and obtain its parameters of grey Bernoulli mode by the particle swarm optimization [17]. In the two cases above, the particle swarm optimization is stochastic algorithm. We set MAPE as objective function and adopt linear programming to fit Gompertz curve by using the grey direct modeling method.

III. FITTING GOMPERTZ CURVE USING GREY DIRECT MODELING METHOD

Let $X = \{x(1), x(2), \cdots, x(n)\}$ be a raw sequence, we have the following differential equation of Gompertz curve:

$$\frac{dx}{dt} = cx\ln\left(\frac{e^a}{x}\right), c > 0, a \in R.$$
$$= cx\left(\ln\left(e^a\right) - \ln x\right)$$
$$= acx - cx\ln x$$

the steps of fitting Gompertz curve using grey direct modeling method are as follows:

A. According to the grey direct modeling method, so we can obtain

$$x(k) - x(k-1) \approx acz(k) - cx\ln z(k),$$

where

$$z(k) = \frac{x(k) + x(k-1)}{2} \quad k = 2, \cdots n.$$

a and c is unknown parameters, therefore, the problem of parameter optimization can be formulated as the following optimization problem:

min MAPE

$$= \sum_{k=2}^{n} \left| \frac{acz(k) - cz(k)\ln z(k) - \left(x(k) - x(k-1)\right)}{x(k) - x(k-1)} \right|$$
$$= \sum_{k=2}^{n} \left| \frac{acz(k) - cz(k)\ln z(k)}{x(k) - x(k-1)} - 1 \right| \tag{5}$$

assume $ac = s$, then $\frac{sx(k) - cz(k)\ln(z(k))}{x(k) - x(k-1)} - 1$ is linear concerning s and c, the problem (5) can be formulated as linear programming problem [18].

$$\min \sum_{k=2}^{n} r_k$$

$$\text{s.t.} \begin{bmatrix} I_{n-1} & A & -A \\ I_{n-1} & -A & A \end{bmatrix} \begin{bmatrix} r \\ d_1 \\ d_2 \end{bmatrix} \ge \begin{bmatrix} y \\ -y \end{bmatrix}, \begin{bmatrix} r \\ d_1 \\ d_2 \end{bmatrix} \ge 0 \tag{6}$$

where $A=\begin{bmatrix} \frac{sz(2)}{x(2)-x(1)} & \frac{cz(2)\ln z(2)}{x(2)-x(1)} \\ \vdots & \vdots \\ \frac{sz(n)}{x(n)-x(n-1)} & \frac{cz(n)\ln z(n)}{x(n)-x(n-1)} \end{bmatrix}$, $r=\begin{bmatrix} r_2 \\ \vdots \\ r_n \end{bmatrix}$,

$r_k=\left|\frac{acz(k)-cz(k)\ln z(k)}{x(k)-x(k-1)}-1\right|$, $k=2,\cdots n$. y is the $(n-1)\times 1$ matrix $\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$, d_i is the 2×1 matrix, $i=1,2$.

I_{n-1} is a unit matrix of (n-1)-order.

We can obtain the global optimal solution of the problem (6) in mathematical software Lingo environment, then the estimated value $\hat{s}$ and $\hat{c}$ of parameters s and c is

$$\begin{bmatrix} \hat{s} \\ \hat{c} \end{bmatrix}=d_1-d_2$$

B. By substituting $\hat{c}$ into

$$x(t)=e^{a-be^{-ct}},\ t=1,2,\cdots n.$$

we obtain that

$$x(t)=e^{a-be^{-\hat{c}t}}$$

using the logarithmic transformation to obtain that

$$\ln x(t)=a-be^{-\hat{c}t},\ t=1,2,\cdots n.$$

a and b is unknown parameters, therefore, the problem of parameter optimization can be formulated as the following optimization problem under the criteria of minimized the mean percentage absolute error:

$$\min \quad \text{MAPE}$$

$$=\sum_{t=1}^{n}\left|\frac{\left(a-be^{-\hat{c}t}\right)-\ln x(t)}{\ln x(t)}\right|$$

$$=\sum_{t=1}^{n}\left|\frac{a-be^{-\hat{c}t}}{\ln x(t)}-1\right| \tag{7}$$

by $\frac{a-be^{-\hat{c}t}}{\ln x(t)}-1$ is linear concerning a and b, the problem (7) can be formulated as linear programming problem [18].

$$\min \sum_{t=1}^{n} l_t$$

$$\text{s.t.}\begin{bmatrix} I_n & G & -G \\ I_n & -G & G \end{bmatrix}\begin{bmatrix} l \\ q_1 \\ q_2 \end{bmatrix}\ge\begin{bmatrix} y \\ -y \end{bmatrix},\begin{bmatrix} l \\ q_1 \\ q_2 \end{bmatrix}\ge 0 \tag{8}$$

where $l=\begin{bmatrix} l_1 \\ \vdots \\ l_n \end{bmatrix}$, $l_t=\left|\frac{a-be^{-\hat{c}t}}{\ln x(t)}-1\right|$, $t=1,2,\cdots n$. I_n is a unit matrix of n-order, y is the $n\times 1$ matrix $\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$, q_i is the 2×1 matrix, $i=1,2$.

$$G=\begin{bmatrix} \frac{a}{\ln x(1)} & \frac{be^{-\hat{c}}}{\ln x(1)} \\ \vdots & \vdots \\ \frac{a}{\ln x(n)} & \frac{be^{-\hat{c}n}}{\ln x(n)} \end{bmatrix}$$

we can obtain the global optimal solution of the problem (6) in mathematical software Lingo environment, then the estimated value $\hat{b}$ and $\hat{a}$ of parameters b and a is

$$\begin{bmatrix} \hat{b} \\ \hat{a} \end{bmatrix}=q_1-q_2$$

C. By substituting $\hat{c}$, $\hat{b}$ and $\hat{a}$ into (1), we obtain that

$$\hat{x}(t)=e^{\hat{a}-\hat{b}e^{-\hat{c}t}},\ t=1,2,\cdots n.$$

IV. EXAMPLE

We use the example came from paper [15] to explain the validity of this kind of fitting method, using the method of paper [15] and the method we put forward to simulate, respectively, and then compare each result. we will simulate the datum values from 1 to 7 using this paper method, then we have

$$x(t)=e^{4.032443-8.572319e^{-0.4300680t}},\ t=1,2,\cdots n.$$

Refer to table I (raw series and simulation value) and table II (simulation precision), where R_t is raw data at time t, S_t is simulated data at time t, error $=R_t-S_t$.

TABLE I. RAW SERIES AND SIMULATION VALUE

number	R_t	S_t paper[15]	S_t this paper
1	0.16	1.224	0.214
2	1.5	3.137	1.500
3	3.61	6.627	5.329
4	14.24	12.007	12.156
5	22.59	19.256	20.785
6	26.47	28.03	29.463
7	37	37.777	36.970

TABLE II. RAW SERIES AND ERROR

number	R_t	error paper[15]	error this paper
1	0.16	-1.064	-0.054
2	1.5	-1.637	0.000
3	3.61	-3.017	-1.719
4	14.24	2.233	2.084
5	22.59	3.334	1.805
6	26.47	-1.560	-2.993
7	37	-0.777	0.030

V. CONCLUSIONS AND ANALYSIS

It is difficult to estimate the parameters of the Gompertz curve, to solve this problem, we give a method for estimating its parameters based on grey direct modeling, this method is very easy to implement in mathematical software Lingo environment, comparisons of the obtained results with the results of existing algorithm demonstrate that this method can improve fitting accuracy of Gompertz model.

ACKNOWLEDGMENT

This work is partially supported by the National Natural Science Foundation of China under Grant No. 60674071, the Research Project on Teaching Reform of Wenzhou University under Grant No.08JG06A, and by the Postgraduates Innovation Foundation of Wenzhou University under Grant No. YCX0810.

REFERENCES

[1] Deng Julong, Control Problems of Grey System, System and Control Letters 1(1), pp.288-294, 1989,.

[2] Sifeng Liu, Yi Lin. Grey Information Theory and Practical Applications. Springer-Verlag London Limited, 2006.

[3] Wang Yinao, Properties and Method of direct GM(1,1) Model, System Engineering Theory & Practice, vol. 1,1988, pp.27-31.

[4] Wang Yinao, Liu Guangzhen, Liu Kaidi, A step by step optimum direct modeling method of GM(1,1), System Engineering Theory & Practice, vol. 20,2000, pp.99-105.

[5] Wang Yinao, An extended step by step Optimum direct modeling method of GM(1,1), System Engineering Theory & Practice, vol. 23, 2003, pp.120-124.

[6] Yong. Mu, A direct modeling method of the unbiased GM(1,1), Systems Engineering and Electronics, vol. 25, 2003, pp.1094-1096.

[7] Shui Naixiang, Qin Yuchun, On Some Theoretical Problems of GM(1,1) Model of Grey Systems, System Engineering Theory & Practice, vol. 18(4), 1998,pp.59-63.

[8] Dragan Jukic,Gordana Kralik,and Rudolf Scitovski, Least-squares fitting Gompertz curve, Journal of Computational and Applied Mathematics, vol .(169),2004, pp. 359-375.

[9] M.R. Haj Seyed Hadi and J.L. Gonzalez-Andujar. Comparison of fitting weed seedling emergence models with nonlinear regression and genetic algorithm, Computers and Electronics in Agriculture, vol. (65), Janu 2009, pp19-25.

[10] W.J. Willemse, and R. Kaas. Rational reconstruction of frailty-based mortality models by a generalisation of Gompertz' law of mortality. Insurance: Mathematics and Economics, vol.40,May 2007,pp. 468-484.

[11] Yu Yong-qing, Fu Shun-hua, Application of Gompertz increase model to prediction of forestry product value, Journal of Zhejiang forestry College, vol. 18(1), 2008, pp.85-88.

[12] Tao Yong-hong, Positive analysis of forecasting Shanghai composite index based on Gompertz function, Mathematics in Practice and theory vol 35, Dec.,2005,pp.51-56.

[13] Jiang Cheng, Zhong Jinchun, and WangJie, Utilizing Gompertz Curve to establish the model between the development expenese and the development time of a missle, Systems Engineering and Electronics, vol 24, 4,2002,pp.23-25.

[14] M,pitchaimani,T.Eakin.Unique estimation of Gompertz paramters with mortality deceleration rate, Mathematical and Computer Modelling, vol..47,2008, pp.104-114.

[15] Zhu Min-ren, Fitting Gompertz Model and Logistic Model Mathematics in Practice and theory vol 32, sep.,2002, pp.705-70.

[16] Liuhong, Zhang Qishan, GM(2,1, λ , ρ) based on particle swarm optimization, System Engineering Theory & Practice, vol. 28(10), 2008, pp.96-101

[17] Jianzhong Zhou, Rengcun Fang, Yuhong Li, Yongchuan Zhang, Bing Peng, Parameter optimization of nonlinear grey Bernoulli Model using particle swarm optimization, Applied Mathematics and Computation, vol.207, Jan, 2009,pp..292-299.

[18] Chen Xiru, Least absolute deviate of linear regression, Application of Statistics and Management. vol..6, 1989,pp.48-56.