Expert Systems with Applications 42 (2015) 1806–1815

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Regularized multivariable grey model for stable grey coefficients estimation

Zhi He [a,b,*], Yi Shen [b], Junbao Li [c], Yan Wang [b]

[a] School of Geography and Planning, Sun Yat-Sen University (SYSU), Guangzhou 510275, China
[b] Department of Control Science and Engineering, Harbin Institute of Technology (HIT), Harbin 150001, China
[c] Department of Automatic Test and Control, HIT, Harbin 150001, China

ARTICLE INFO

Article history:
Available online 13 October 2014

Keywords:
Multivariable grey model
Ill-posed problem
Regularization
Industrial indicators

ABSTRACT

Recently, the convolution integral-based multivariable grey model (GMC(1,N)) has attracted considerable interest due to its significant performance in time series forecasting. However, this promising technique may occasionally confront ill-posed problem, which is a plague ignored by most researchers. In this paper, a regularized GMC(1,N) framework (R-GMC(1,N)) is proposed to estimate the grey coefficients in case there exists potential ill-posed problem. More specifically, we adopt two state-of-the-art regularization methods, i.e. the Tikhonov regularization (TR) and truncated singular value decomposition (TSVD), together with two regularization parameters detection methods, i.e. L-curve (LC) and generalized cross-validation (GCV), to identify the stable solutions. Numerical simulations on industrial indicators of China demonstrate that our methods yield more accurate forecast results than the existing GMC(1,N).

1. Introduction

Among all forecast techniques being developed over the last few decades, the grey model and its alternatives have emerged as powerful tools in various application domains, such as economy (Evans, 2014; Huang & Jane, 2009; Ma, Zhu, & Wang, 2013), industry (Benítez, Paredes, Lodewijks, & Nabais, 2013; Hsu, 2011; Hsu, Liou, & Chuang, 2013), society (Jin, Zhou, Zhang, & Tentzeris, 2012; Pao, Fu, & Tseng, 2012; Wei, Zhou, Wang, & Wu, 2014) and engineering (Chen & Wang, 2012; He, Liu, & Chen, 2012; Ye, Lu, & Liu, 2013). Compared with conventional statistical models, the grey model has three superiorities: (1) requiring few sample size; (2) providing convenient calculation; (3) achieving accurate prediction accuracy. In fact, the grey model is a special differential equation, according to which the output data can be significantly predicted. The general form of grey model can be expressed as GM(φ, ϕ), where φ is the order of differential equation and ϕ denotes the number of variables. In the literature, two widespread grey models are the first-order one-variable GM(1,1) ($\varphi = 1$, $\phi = 1$) and the multivariable one, i.e. GM(1,N) ($\varphi = 1, \phi = N$).

With regard to GM(1,1), much work has been carried out to improve the grey prediction performance since it was pioneered by Deng (1982). Two remarkable aspects can be highlighted: theory and application. Theoretically, several improved versions have been proposed, including the hybrid one (Xia, Chen, Zhang, & Wang, 2008), discrete one (Xie & Liu, 2009), intelligent algorithm-based ones (Bahrami, Hooshmand, & Parastegari, 2014; Hsu, 2010), least-squares-based one (Xu, Tan, Tu, & Qi, 2011), smart adaptive one (Truong & Ahn, 2012) as well as the fractional order-based one (Xiao, Guo, & Mao, 2014). On the other hand, the applications of GM(1,1) are also enormous, such as accurate reliability prediction (Li, Masuda, Yamaguchi, & Nagai, 2010), fashion color forecasting (Yu, Hui, & Choi, 2012), hyperspectral feature extraction (Yin, Gao, & Jia, 2013), and end effects mitigation in our previous works (He, Shen, & Wang, 2012; He et al., 2012).

As for GM(1,N), it has also drawn great attention recently due to its convenient calculation and accurate results in multi-factor forecasting. Intensive modified models have been developed since it was introduced by Deng et al. (1988) in the 1980s. One of the widespread models is GMC(1,N) (Wu & Chen, 2005; Tien, 2012), which is motivated by the utilization of convolution integral. In greater detail, note that it is inaccurate to assume the sum of the first-order accumulated generating operation (1-AGO) data (i.e. $\sum_{i=1}^{N-1} b_i x_{i+1}^{(1)}(t)$ in Eq. (1)) as a constant in GM(1,N), the GMC(1,N) takes $\sum_{i=1}^{N-1} b_i x_{i+1}^{(1)}(t)$ as a variable and detects the convolution integral by trapezoidal rule. Much work has been carried out in

* Corresponding author at: School of Geography and Planning, Sun Yat-Sen University (SYSU), Guangzhou 510275, China. Tel.: +86 451 86413411; fax: +86 451 86418378.

E-mail addresses: hzhdhz@126.com (Z. He), shen@hit.edu.cn (Y. Shen), lijunbao@hit.edu.cn (J. Li), wyabc@hit.edu.cn (Y. Wang).

http://dx.doi.org/10.1016/j.eswa.2014.09.058
0957-4174/

the literature to enhance the prediction accuracy of GMC(1,N) since its appearance. For instance, Tien (2008, 2009, 2011) is dedicated to the refinement of GMC(1,N) and proposes plenty of improved models, which are mainly associated with the calculation of grey derivative and selection of initial condition. Abdulshahed, Longstaff, Fletcher, and Myers (2013) build a thermal model by integrating artificial neural networks (ANNs) and GMC(1,N). The thermal model can improve the self-learn and self-adapt ability of GMC(1,N) and be able to extract realistic governing laws of the system with limited data pairs. Wang and Pei (2014) introduce *N* interpolation coefficients into the background series calculation and determine the optimized coefficients by particle swarm optimization algorithm (PSO). This method can improve the modelling accuracy by providing more flexible background series of grey derivation. Moreover, Wang (2014) proposes nonlinear GMC(1,N) by adding power exponent to the 1-AGO of every relative data series. Compare to traditional GMC(1,N), the nonlinear GMC(1,N) can reflect better relationship between cause and effect. On the other hand, considerable applications has sprung up in recent years. For instance, Hsu (2009) forecasts the integrated circuit output inspired by GM(1,N). Tien (2008, 2009, 2011, 2012) utilizes the improved GMC(1,N) for indirect measurement of tensile strength. Zhang and Hu (2013) build product quality prediction model for multi-varieties and small-batch production based on GMC(1,N). In our previous work (He, Wang, Shen, & Wang, 2013), the GMC(1,3) ($N = 3$) is adopted to mitigate the boundary effects of bi-dimensional empirical mode decomposition (BEMD).

Although many efforts have been devoted to improve the prediction performance of GMC(1,N), till now, there is no research which explicitly investigates its ill-posed problem. In general, ill-posed problem is an essential drawback of GMC(1,N), which may occasionally appear in detecting the grey coefficients. In case there exists ill-posed problem, the traditional least-squares solution of grey coefficients will be inaccurate and the forecast error of GMC(1,N) will increase sharply. As a consequence, it is very difficult to gain satisfactory forecast results, which would seriously impede the wide application of GMC(1,N). According to above analysis, the ill-posed problem of GMC(1,N) is a challenging issue that needs to be further investigated. In this paper, we propose a regularization framework (i.e. R-GMC(1,N)) to tackle this deficiency. That means, regularization methods (Hansen, 2007), which can solve ill-posed problem by investigating a tradeoff between the constructed new problem one can solve reliably and the solution that is close to the desired one, are adopted to estimate stable grey coefficients in ill-posed scenario. It is notable that two state-of-the-art regularization methods are the Tikhonov regularization (TR) (Tikhonov, Arsenin, & John, 1977) and truncated singular value decomposition (TSVD) (Hansen, 1987), whose regularization parameters can be evaluated by L-curve (LC) (Hansen & O'Leary, 1993) or generalized cross-validation (GCV) (Hansen, 1998). Integrating the regularization methods (i.e. TR and TSVD) with parameters detection strategies (i.e. LC and GCV), four hybrid methods can be revealed: LC-TR, GCV-TR, LC-TSVD and GCV-TSVD, by which different R-GMC(1,N) methods can be formed.

Compared to the existing literature, the main novelties and contributions of this paper lie in the following two aspects.

- This paper proposes R-GMC(1,N) to estimate stable grey coefficients (see Fig. 2). Unlike existing GMC(1,N), which calculates the grey coefficients by least-squares method all the time, the proposed R-GMC(1,N) utilizes regularization methods to identify the grey coefficients in ill-posed scenario.
- This paper proposes four versions of R-GMC(1,N) based on various regularization methods. In greater detail, $R_{LC\text{-}TR}$-GMC(1,N), $R_{GCV\text{-}TR}$-GMC(1,N), $R_{LC\text{-}TSVD}$-GMC(1,N) and $R_{GCV\text{-}TSVD}$-GMC(1,N), which are developed from the aforementioned hybrid regularization methods (i.e. LC-TR, GCV-TR, LC-TSVD and GCV-TSVD), are proposed in this paper. Moreover, the effectiveness of the proposed methods is assessed by two numerical tests on industrial indicators of China (see Section 4).

The outline of this paper is organized as follows. The GMC(1,N) and its ill-posed problem are illustrated in Section 2. This is followed by a detailed description of our proposed regularization framework in Section 3. Numerical results and discussions are stated in Section 4, and the conclusions are drawn in Section 5.

2. GMC(1,N) and its ill-posed problem

In this section, we illustrate the main steps of GMC(1,N) and discuss the causes of ill-posed problem. Since the details of GMC(1,N) can be found in Wu and Chen (2005) and Tien (2012), we only sketch the core steps of this method briefly.

Step 1: Obtain original series. Assume the original characteristic data series and relative data series are $\boldsymbol{X}_1^{(0)} = \{x_1^{(0)}(1+r), x_1^{(0)}(2+r), \ldots, x_1^{(0)}(n+r)\}$ and $\boldsymbol{X}_i^{(0)} = \{x_i^{(0)}(1), x_i^{(0)}(2), \ldots, x_i^{(0)}(n), \ldots, x_i^{(0)}(n+m)\}, i = 2,3,\ldots,N$, respectively, where r is the period of delay, n gives the length of original characteristic data series and m denotes the number of entries to be predicted.

Step 2: Execute 1-AGO. Generate the 1-AGO sequence $\boldsymbol{X}_1^{(1)} = \{x_1^{(1)}(1+r), x_1^{(1)}(2+r), \ldots, x_1^{(1)}(n+r)\}$ and $\boldsymbol{X}_i^{(1)} = \{x_i^{(1)}(1), x_i^{(1)}(2), \ldots, x_i^{(1)}(n), \ldots, x_i^{(1)}(n+m)\}, i = 2,3,\ldots,N$ based on $\boldsymbol{X}_i^{(0)}, i = 1,2,\ldots,N$, where $x_1^{(1)}(k+r) = \sum_{l=1}^{k} x_1^{(0)}(l+r), k = 1,2,\ldots,n$ and $x_i^{(1)}(k) = \sum_{l=1}^{k} x_i^{(0)}(l), k = 1,2,\ldots,n+m, i = 2,3,\ldots,N$. Additionally, the GMC(1,N) can be modeled by

$$\frac{dx_1^{(1)}(t+r)}{dt} + a x_1^{(1)}(t+r) = \sum_{i=1}^{N-1} b_i x_{i+1}^{(1)}(t) + b_N, \quad t = 1,2,\ldots,n+m \tag{1}$$

where a and $b_i, i = 1,2,\ldots,N$ are the grey developmental coefficient and the associated coefficients, respectively.

Step 3: Calculate grey coefficients. Notably that Eq. (1) can be approximately rewritten as

$$\boldsymbol{B\kappa} = \boldsymbol{Y} \tag{2}$$

where $\boldsymbol{\kappa} = [a \quad b_1 \quad b_2 \quad \ldots \quad b_N]^{\mathrm{T}}, \boldsymbol{Y} = \left[x_1^{(0)}(2+r) \quad x_1^{(0)}(3+r) \quad \cdots \quad x_1^{(0)}(n+r)\right]^{\mathrm{T}}$ and

$$\boldsymbol{B} = \begin{bmatrix} -\frac{1}{2}\left(x_1^{(1)}(1+r) + x_1^{(1)}(2+r)\right) & \frac{1}{2}\left(x_2^{(1)}(1) + x_2^{(1)}(2)\right) & \cdots & \frac{1}{2}\left(x_N^{(1)}(1) + x_N^{(1)}(2)\right) & 1 \\ -\frac{1}{2}\left(x_1^{(1)}(2+r) + x_1^{(1)}(3+r)\right) & \frac{1}{2}\left(x_2^{(1)}(2) + x_2^{(1)}(3)\right) & \cdots & \frac{1}{2}\left(x_N^{(1)}(2) + x_N^{(1)}(3)\right) & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -\frac{1}{2}\left(x_1^{(1)}(n-1+r) + x_1^{(1)}(n+r)\right) & \frac{1}{2}\left(x_2^{(1)}(n-1) + x_2^{(1)}(n)\right) & \cdots & \frac{1}{2}\left(x_N^{(1)}(n-1) + x_N^{(1)}(n)\right) & 1 \end{bmatrix}, \tag{3}$$

we can evaluate the grey coefficient $\boldsymbol{\kappa}$ by least-squares method

$$\boldsymbol{\kappa} = \left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{Y} \tag{4}$$

Step 4: Detect predicted value of $\boldsymbol{X}_1^{(1)}$. Unlike GM(1,N), which treats the $\sum_{i=1}^{N-1} b_i x_{i+1}^{(1)}(t)$ of Eq. (1) as a constant (actually that is inaccurate), GMC(1,N) regards that term as a variable in terms of defining $f(t) = \sum_{i=1}^{N-1} b_i x_{i+1}^{(1)}(t) + b_N$. Subsequently, the convolution integral $\int_1^t e^{-a(t-\tau)} f(\tau) d\tau$ can be exploited by trapezoidal rule, and the predicted value of $x_1^{(1)}(t+r), t = 2,3,\ldots,n+m$ yields

$$\hat{x}_1^{(1)}(t+r) = x_1^{(0)}(1+r)e^{-a(t-1)} + \frac{1}{2}e^{-a(t-1)}f(1) + \sum_{\tau=2}^{t-1}\left[e^{-a(t-\tau)}f(\tau)\right] + \frac{1}{2}f(t) \tag{5}$$