Evaluation of Teaching Effectiveness Based on Gray Markov Chain in the Context of MOOC

ZHANG Li

(Department of Mathematical and Physics, Mobile College of Chongqing University of Posts and Telecommunications, Chongqing 401520, China)

Author's brief introduction: ZHANG Li (1982-), master's degree, lecturer, Mainly engaged in mathematics education and functional analysis research.

Abstract: Markov chain is an analytical method to study the law of change between different states of an event in a random process. In this paper, the GM (1, 1) model and Markov model are combined under the background of MOOC to construct the gray Markov model. The Grey Markov model evaluates the using effect of MOOC in the teaching process.

1. Introduction
In recent years, MOOC has become an important means of teaching reform in colleges and universities. With flexible learning time, MOOC can offer students more space to think and reflect. The MOOC platform, with resource richness and flexibility, can improve teachers' teaching level. It provides learners with a new way of learning and educators with a new teaching mode and teaching method. MOOC, as a mode of teaching method, has cultivated college students' ability of independent learning and lifelong learning. With rich resources, MOOC enables a large number of students and teachers to learn all kinds of high quality courses of good teachers from famous universities around the world for free, so as to stimulate students' self-learning ability and cultivate students' ability to solve practical problems.

The traditional Grey model is mainly applicable to the system objects with short test time, little data and little fluctuation. It has poor fitting for the data series with large random fluctuation and low prediction accuracy. In Markov chain theory, the transition probability $P_{ij}$ can reflect the influence of random factors. The transition probability $P_{ij}$ makes up the limitation of the Grey prediction. The transition probability $P_{ij}$ is suitable to predict the dynamic process with large random fluctuation. But Markov chain prediction object must have the characteristics of Markov chain stationary process. However, most of the prediction problems in the life are non-stationary processes with fuzzy change trend with time.

The organic combination of GM (1,1) model and Markov model in this paper can not only be applied to the prediction of non-stationary processes, but also improve the accuracy of GM (1,1) model.

2. Application of Grey Markov model in teaching effect evaluation

2.1 The gray GM (1,1) model was constructed
The GM (1,1) model was constructed based on the average score of students' test scores:
(1) The average score of the students in the test
\[ X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(10)\} = \{68.7, 74.6, 79.4, 85.2, 91.8, 98.5, 105.3, 113.7, 118.3, 124.6\} \]

(2) Do a sum \( X^{(0)} \) to \( X^{(i)} \)
\[ X^{(i)} = \{x^{(i)}(1), x^{(i)}(2), \ldots, x^{(i)}(10)\} = \{68.7, 143.3, 222.7, 307.9, 399.7, 498.2, 603.5, 717.2, 835.5, 960.1\} \]

Among them \( x^{(i)}(n) = \sum_{i=1}^{n} x^{(i)}(i) \quad n = 1, 2, \ldots, 10 \)

(3) Construct matrices \( B \) and vectors \( Y_n \), and solve coefficients
\[
B = \begin{bmatrix}
-\frac{1}{2} [x^{(i)}(1) + x^{(i)}(2)] & 1 \\
-\frac{1}{2} [x^{(i)}(2) + x^{(i)}(3)] & 1 \\
-\frac{1}{2} [x^{(i)}(3) + x^{(i)}(4)] & 1 \\
-\frac{1}{2} [x^{(i)}(4) + x^{(i)}(5)] & 1 \\
-\frac{1}{2} [x^{(i)}(5) + x^{(i)}(6)] & 1 \\
-\frac{1}{2} [x^{(i)}(6) + x^{(i)}(7)] & 1 \\
-\frac{1}{2} [x^{(i)}(7) + x^{(i)}(8)] & 1 \\
-\frac{1}{2} [x^{(i)}(8) + x^{(i)}(9)] & 1 \\
-\frac{1}{2} [x^{(i)}(9) + x^{(i)}(10)] & 1 \\
\end{bmatrix}
\]
\[ Y_n = \begin{bmatrix}
-106 & 1 \\
-183 & 1 \\
-265.3 & 1 \\
-353.8 & 1 \\
-448.95 & 1 \\
-550.85 & 1 \\
-660.35 & 1 \\
-776.35 & 1 \\
-897.8 & 1 \\
\end{bmatrix} \]

(4) By solving the equation \( \hat{a} = (B^T B)^{-1} B^T Y_n \), getting \( \hat{a} = \begin{bmatrix} -0.064964 \\ 68.421953 \end{bmatrix} \)

(5) Return the obtained parameters to the original equation, then
\[ \hat{x}^{(i)}(k+1) = [x^{(0)}(1) - \frac{\hat{u}}{\hat{a}} e^{-\frac{\hat{a}}{\hat{a}}} + \frac{\hat{u}}{\hat{a}}] \quad k = 1, 2 \cdots \]
\[ = 1121.9288 e^{0.064964 \cdot \cdot \cdot} - 1053.2288 \]

(6) Reduction yields the predicted value
\[ \hat{x}^{(0)}(k+1) = \hat{x}^{(i)}(k+1) - \hat{x}^{(i)}(k) \quad k = 1, 2 \cdots \]

2.2 Construct gray Markov model
(1) The status is divided according to the relative value (the ratio of the average score of test results to the average score of predicted results). This is shown in table 1.
The state transition probability matrix is used to compile the prediction table of students' performance. We select four test scores closest to the predicted score to compile the table.

In the transition matrix corresponding to the number of transition steps, the row vectors corresponding to the initial state are selected to form a new matrix. The column vectors of the new matrix are summed up to the largest, which is the turn of the system in the next time.

(3) The state transition probability matrix is used to compile the prediction table of students' performance.

We select four test scores closest to the predicted score to compile the table.

In the transition matrix corresponding to the number of transition steps, the row vectors corresponding to the initial state are selected to form a new matrix. The column vectors of the new matrix are summed up to the largest, which is the turn of the system in the next time.

Tab.2 The prediction of the average score of the last test

| ordinal | initial state | transfer steps | state |
|---------|--------------|----------------|-------|
|         |              |                | 1     | 2     | 3     | 4     |
| 9       | 2            | 1              | 0.0000 | 0.5000 | 0.5000 | 0.0000 |
| 8       | 2            | 2              | 0.0000 | 0.3750 | 0.5000 | 0.1250 |
| 7       | 1            | 3              | 0.1250 | 0.4375 | 0.3750 | 0.0625 |
| 6       | 1            | 4              | 0.0625 | 0.3750 | 0.4688 | 0.0938 |
| summation |              |                | 0.1875 | 1.8875 | 1.8438 | 0.2813 |

\[
\max(P_{ij}) = 1.8875. \text{ Therefore, the last test result will turn to the status 2.}
\]

\[
\hat{Y} = \frac{F_{21} + F_{22}}{2} = 123.44.
\]

The average score on the last test was 124.6. The predicted value of GM(1,1) model is 126.61. The predicted value of the gray Markov model is 123.44.

(4) The gray GM(1,1) model and the gray Markov model were used to predict the average score of the test, as shown in table 3.
| ordinal | test scores | GM model | residual | Gray Markov model | residual |
|---------|-------------|----------|----------|------------------|----------|
| 1       | 68.7        | 68.7     | 0.0      | 68.7             | 0.0      |
| 2       | 74.6        | 75.29    | -0.69    | 74.92            | -0.32    |
| 3       | 79.4        | 80.36    | -0.96    | 79.82            | -0.42    |
| 4       | 85.2        | 85.75    | -0.55    | 85.30            | -0.10    |
| 5       | 91.8        | 90.89    | 0.91     | 91.35            | 0.45     |
| 6       | 98.5        | 97.52    | 0.98     | 98.30            | 0.20     |
| 7       | 105.3       | 105.98   | -0.68    | 105.51           | -0.21    |
| 8       | 113.7       | 108.29   | 5.41     | 111.40           | 2.30     |
| 9       | 118.3       | 118.65   | -0.35    | 118.40           | -0.10    |
| 10      | 124.6       | 126.61   | -2.01    | 125.80           | -1.20    |

**mean relative error%**

1.21 0.51

We can see from table 3: The simulated value of GM(1,1) model is basically the same as the change trend of test score, but there was still significant residual error, while the simulated value of gray Markov model was closer to test scores, and the residual error was significantly reduced.

### 3. Accuracy test

Posterior difference test is a statistical concept, which is tested according to the probability distribution of residuals. Let the variance of the initial data sequence $X^{(0)}$ and the residual sequence $e$ be $S_1^2$ and $S_2^2$ respectively, so

$$S_1^2 = \frac{1}{n} \sum_{i=1}^{n} (X^{(0)}(i) - \bar{x}^{(0)})^2, \quad S_2^2 = \frac{1}{n} \sum_{i=1}^{n} (e(i) - \bar{e})^2$$

Among them

$$\bar{x}^{(0)} = \frac{1}{n} \sum_{i=1}^{n} x^{(0)}(i), \quad e(i) = x^{(0)}(i) - \hat{x}^{(0)}(i), \quad \bar{e} = \frac{1}{n} \sum_{i=1}^{n} e(i)$$

So the posterior difference ratio is calculated

$$C = \frac{S_2}{S_1}$$

Small error probability

$$P = P\{|e(i) - \bar{e}| < 0.6745S_1\}$$

GM (1, 1) model:

$$S_1 = 18.264953 \quad S_2 = 1.186467 \quad C = \frac{S_2}{S_1} = 0.064959 \quad P = 1$$

Gray Markov model:

$$S_1 = 18.264953 \quad S_2 = 0.853194 \quad C = \frac{S_2}{S_1} = 0.046712 \quad P = 1$$

The precision of the model is characterized by $C$ and $P$. If the $P$ values of the two models are the same, the value of $C$ is examined. The smaller the value of $C$ is, the higher the model accuracy will be. According to the size of $P$ and $C$, the model accuracy can be divided into four categories: good, qualified, barely qualified and unqualified. The values of $P$ and $C$ are shown in table 4.
Tab. 4 Table of model’s precision

| Model accuracy level | $P$       | $C$         |
|----------------------|-----------|-------------|
| Level 1 (good)       | $P \geq 0.95$ | $C \leq 0.35$ |
| Level 2 (qualified)  | $0.80 \leq P < 0.95$ | $0.35 < C \leq 0.5$ |
| Level 3 (barely qualified) | $0.70 \leq P < 0.80$ | $0.5 < C \leq 0.65$ |
| Level 4 (unqualified) | $P < 0.70$  | $C > 0.65$  |

It can be seen from the table 4 that both GM(1,1) model and gray Markov model have a good prediction of test performance.

4. Conclusions
MOOC cannot completely replace our traditional classrooms, but they promote the innovation and development of teaching models. We believe that we can give priority to the physical classroom teaching. MOOC, as an auxiliary teaching means, continues to take advantage of their online resources. MOOC further promotes the deep integration of classroom teaching and online teaching, so as to comprehensively improve the effect of college mathematics classroom teaching.

The classical GM(1,1) model and Markov model are combined to make full use of the advantages of the two models and solve the problems of data instability, volatility and randomness. The method given in this paper is effective and practical. By considering the average simulation error of the two models, the GM(1,1) model is 1.21%, and the gray Markov model is 0.51%. The gray Markov model is more suitable for test result prediction than GM(1,1) model.

Acknowledgement
Project: Research project of higher education teaching reform in Mobile college of Chongqing university of posts and telecommunications (YTJG201811)

References
[1] LIU Sifeng, DANG Yaoguo, FANG Zhigeng. Grey System Theory and its Application[M].3rd edition. Beijing: Science Press, 2004.
[2] CAI Anning, Lai Si. The Grey Dynamic Model is used to Forecast Power Demand in China [J]. Statistics and Decision-Making, 2009(23): 47-49.
[3] XIE Naiming, LIU Sifeng. Modeling Mechanism of Discrete GM(1,1) Model and Gray Prediction Model [J]. Systems Engineering Theory and Practice, 2005, 25(1): 94-99.
[4] HE Chenggang. Research and Application of Markov Model Prediction Method [D]. Hefei: Anhui University, 2011.
[5] ZHI Zewen, ZHANG Wen, QIU Shufang. Improvement of Grey-Markov Model and its Parameter Calculation Method [J]. Mathematical Practice and Cognition, 2009, 39(1): 125-131.
[6] CHEN Lihui, CUO Tsueiyang. The Application of a Grey Markov Model to Forecasting Annual Maximum Water Levels at Hydrological Stations [J]. QUALITY & QUANTITY, 2011, 45: 911-922.
[7] Li Xiang, WANG Xinyuan, SHAO Wei. Forecast of Flood in Chaohu Lake Basin of China Based on Grey-Markov Theory [J]. CHINESE GEOGRAPHICAL SCIENCE, 2007, 17(1): 64-68.