Journal of Theoretical and Applied Information Technology
15th October 2012. Vol. 44 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

# A NOVEL MODEL FOR THE TECHNICAL INNOVATION FACTORS EVOLUTION

[1] CHAO ZHENG, [2]YOUWEN YANG, [3]DEYU YUAN
[1]School of Management, Hefei University of Technology, Hefei, China
[2]Institute of Science & Technology, Hefei University of Technology, Hefei, China
[3] School of Management, Hefei University of Technology, Hefei, China
E-mail: [1]ahzhengchao@y163.com, [2]hfutyyw@163.com , [3]deyu1988@sina.com

**ABSTRACT**

In recent years, evaluating the technical innovation activity has gained a renewed interest in both growth economists and trade economists. An evolution model of the technical innovation activity is proposed by applying Kernel principal component analysis (KPCA) and the grey model in this paper. The KPCA is used to features selection and grey model is used to evolution. The proposed method is feasible and effective by the results, and it provides a better forecast and estimate tool for the technical innovation activity. It also provides a novel way for the evolution design of the other engineering.

**Keywords:** *Kernel Principal Component Analysis, Features Selection, The Grey Model, Evolution*

## 1. INTRODUCTION

The technical innovation activity is perceived as being at the core of observed divergences among countries in the areas of long-run growth. The development of demand for innovation is of interest as well to economists who seek to technological development. There are many traditional regression methods of evolution patent activity that could be used as linear regression model [1]. However, the training procedure of an artificial intelligence model is time consuming. Therefore, some other combined approaches must be proposed.

Grey theory, first derived by Prof. Deng Julong [2], is a novel mathematical approach, particularly designed for handling situations in which only limited data are available. This theory describes random variables as a changeable interval number that varies with time factors and uses color to represent the degree of uncertainty in a dynamic system. It implies that a grey environment is a system that consists of partially known and partially unknown information. It is believed that the uncertainties existing in the whitening process mainly come from the insufficiency of understandable information. Therefore, by way of increasing system information, the degree of uncertainty could be changed or diminished over time. The output from a grey environment must have a certain degree of implication for behavior in this system. The measured database with discrete features can be formulated as a typical solution based on a so-called pseudo-differential equation. With a solution from the proposed pseudo-differential equation, the next output from a grey environment can be predicted, even based on a few observed data. Since the grey model can characterize such an unknown system and can make predictions based on a few data, it has been effectively applied in many fields such as agricultural, socioeconomic and environmental evaluations where limited samples are required. so, Grey system theory utilizes accumulated generating data instead of original data to build forecasting model, which makes raw data stochastic weak, or reduces noise influence in a certain extent, therefore, intrinsic regularity of data can be searched easily, and model can be built with relatively little data.

The GM (1, N) model is a very important model of the grey theory. In the GM (1, N) modeling, all available indicators can be used as the inputs of GM (1, N), but irrelevant or correlated features could deteriorate the generalization performance of GM (1, N) due to the "curse of dimensionality" problem. Thus, it is very necessary to perform feature extraction in model. Principal component analysis (PCA) is a well-known method for feature extraction [3, 4]. By calculating the eigenvectors of the covariance matrix of the original inputs, PCA linearly transforms a high-dimensional input vector into a low-dimensional one whose components are uncorrelated [5]. Kernel principal component

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

analysis (KPCA) is one type of nonlinear PCA developed by generalizing the method into PCA [6]. KPCA firstly maps the original inputs into a high dimensional feature space using the kernel method and then calculates PCA in the high dimensional feature space [7, 8].

If combine grey system theory with KPCA, we can exploit sufficiently characteristic of grey system model requiring less data and feature of nonlinear map of KPCA, and develop both advantages, thus raise predicting precision much more. The experiment shows that this kind information manipulation and evolution method based on GM (1, N) with KPCA is of validity and feasibility.

## 2. KERNEL PRINCIPAL COMPONENT ANALYSIS

KPCA is one type of nonlinear PCA developed by generalizing the kernel method into PCA. The kernel method is demonstrated to be able to extract the complicated nonlinear structures embedded on the data set.

The ideal of KPCA is to firstly map the original input vectors $x_i$ into a high dimensional feature space F through a nonlinear function $\Phi$ $(x_i)$ and then to solves the Eigenvalues problem.

$$\lambda_i p_i = \tilde{C}\, p_i, i = 1, \cdots, N \tag{1}$$

Where

$$C = \frac{1}{N}\sum_{i=1}^{N}(\phi(x_i) - u)(\phi(x_i) - u)^T$$

$$u = \frac{1}{N}\sum_{1}^{N}\phi(x_i)$$

Where $C$ is the sample covariance matrix of $\Phi$ $(x_i)$, $\lambda_i$ is one of the non-zero Eigenvalues of C. $P_i$ is the corresponding eigenvector. Eq.1 can be transformed to the Eigenvalues problem.

$$\lambda_i \alpha_i = K\alpha_i, i = 1, \ldots, N \tag{2}$$

Where $\lambda_i = N\lambda_i$.

K is the N×N kernel matrix. The value of each element of K is equal to the inner product of two vectors $x_i$ and $x_j$ in the high dimensional feature space $\Phi$ $(x_i)$ and $\Phi$ $(x_j)$. $K$ $(x_i, x_j)$ is equal to the inner product of Φ $(x_i)$ and Φ $(x_j)$ of two vectors $x_i$ and $x_j$ in the high dimensional feature space. That is $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$.

$\lambda_i$ is one of the Eigenvalues of K. $\alpha_i$ is the corresponding Eigenvalues of K, satisfying:

$$p_i = \sum_{j=1}^{l}\alpha_i(j)\phi(x_j) \tag{3}$$

Furthermore, for assuring the eigenvectors of $\Phi$ $(x_i)$ is of unit length $P_i \cdot P_j$=1, each $\alpha_i$ must be normalized using the corresponding Eigenvalues by:

$$\tilde{\alpha}_i = \alpha_i \Big/ \sqrt{\tilde{\lambda}_i} \tag{4}$$

Finally, we can calculate the Kth nonlinear principle component of $x_i$ as the projections of $\Phi$ $(x_i)$ onto the eigenvector $P_k$.

$$s_k(x_i) = p_k{}^T\phi(x_i) = \sum_{j=1}^{N}\tilde{\alpha}_k(j)K(x_j, x_i), k = 1, \ldots, N \tag{5}$$

## 3. EVOLUTION WITH GM (1, N) MODEL

After feature extraction using KPCA, the evolution is GM(1, N). The GM (1, N) model is the grey compatibility model, and it is the one that describes the relationship between one main factor as output, and all the other N-1 factors as input in a system [6]. As well as their evaluation and logical, it is a model quite suitable for the dynamics relative analysis of all variables and forecasting the output by input.

In our research, the mathematics model is GM (1, N) model. Sequences $y_1{}^{(0)}$ $(k)$ is called the main factor in system, and $x_i{}^{(0)}$ $(k)$. Where i =1, 2, 3... N-1 are called the influencing factors in system. The analysis steps of GM (1, N) model:

Step1, Build up accumulated generating operation (AGO) sequences from original sequences according to the formula from grey system theory,

AGO

$$y^{(0)} = y^{(1)} = \left(\sum_{k=1}^{1}y^{(0)}(k), \sum_{k=1}^{2}y^{(0)}(k), \ldots, \ldots, \sum_{k=1}^{n}y^{(0)}(k)\right)$$

AGO

$$x^{(0)} = x^{(1)} = \left(\sum_{k=1}^{1}x^{(0)}(k), \sum_{k=1}^{2}x^{(0)}(k), \ldots, \ldots, \sum_{k=1}^{n}x^{(0)}(k)\right)$$

Step2, The GM (1, N) model is defined as

$$y_1{}^{(0)}(k) + az_1{}^{(1)}(k) = \sum_{i=1}^{N-1}b_i x_i{}^{(1)}(k) \tag{6}$$

Journal of Theoretical and Applied Information Technology
15th October 2012. Vol. 44 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Where $z_1^{(1)}(k)=0.5y_1^{(1)}(k)+0.5y_1^{(1)}(k-1) \quad k\geq 2$

Step2, Obtain the white equation of Eq. 7

$$\frac{dy_1^{(1)}}{dt}+ay_1^{(1)}=\sum_{i=1}^{N-1}b_i x_i^{(1)}(k) \tag{7}$$

Step3, Inverse and forecasting according to Eq. 7, 8, it gets Eq. 9

$$\hat{y}_1^{(1)}(k+1)=\left(y_1^{(1)}(0)-\frac{1}{a}\sum_{i=1}^{N-1}b_i x_i^{(1)}(k+1)\right)e^{-ak}+\frac{1}{a}\sum_{i=1}^{N-1}b_i x_i^{(1)}(k+1) \tag{8}$$

Where $y_1^{(1)}(0)$ is $y_1^{(1)}(1)$.

It will obtain the value of forecasting of GM (1, N) by inversing-accumulated generation, which is

$$\hat{y}_1^{(0)}(k+1)=\hat{y}_1^{(1)}(k+1)-\hat{y}_1^{(1)}(k) \tag{9}$$

## 4. THE TECHNICAL INNOVATION FACTORS

The factors mainly consist of the basic condition index and input index of the enterprise technology innovation, that is, the ratio of R&D funds expense in product sales income $x_1$, the ratio of enterprises with scientific and technological mechanism $x_2$, the ratio of enterprises with scientific and technological activities $x_3$, the number of employees at the end of the year $x_4$, the ratio of engineering and technological employees in the staff $x_5$, the ratio of scientific and technological activity employees in the staff $x_6$, the ratio of scientists and engineers in the scientific and technological activity employees $x_7$, the ratio of micro-electron controlling equipment in the equipment used in production and operation $x_8$, the ratio of new product developing funds in scientific and technological activity funds $x_9$, the ratio of scientific and technological activity funds in product sales income $x_{10}$, the ratio of R&D funds expense in product sales income $x_{11}$, average expense of every enterprise's scientific and technological activity $x_{12}$, average expense of every scientific and technological activity employee $x_{13}$, the ratio of government capital in fund collecting $x_{14}$, the ratio of enterprise capital in fund collecting $x_{15}$, the ratio of loan from financial organizations in fund collecting $x_{16}$, research and experiment develop personnel's full-time equivalent $x_{17}$, expense of technology reformation $x_{18}$, expense of technology introduction $x_{19}$, expense of digestion and absorption $x_{20}$, expense of buying domestic technology $x_{21}$.

## 5. RESULTS AND DISCUSSION

By using KPCA and GM(1, N), we can calculate the dynamic situation of the enterprise's technology innovation. It is shown in table 1.

*Table 1 The Results Of The GM(1,N) Model*

| | | |
|---|---|---|
| $a_1$=1.8533 | $b_2$=-0.1861 | $b_3$=-1.2793 |
| $b_4$=0.2206 | $b_5$=-0.0800 | $b_6$=-0.8365 |
| $b_7$=0.1982 | $b_8$=0.7110 | $b_9$=0.1553 |
| $b_{10}$=0.9367 | $b_{11}$=-0.4863 | $b_{12}$=0.1861 |
| $b_{13}$=1.1643 | $b_{14}$=-0.6312 | $b_{15}$=1.1705 |
| $b_{16}$=3.2401 | $b_{17}$=0.1632 | $b_{18}$=0.2626 |
| $b_{19}$=2.0208 | $b_{20}$=0.8341 | $b_{21}$=1.7881 |

The above-mentioned coordinating dynamic model reflects the influence degree of every factor to the technology innovation in recent years. The development modulus $a$>0 indicates that the efficiency of technology innovation in China is low and the independent innovation ability is not sufficient.

All these factors restrain the development of technology innovation of Chinese enterprises .The current ranking of the affecting factors of technology innovation is as follows:

$$x_{16}\succ x_{21}\succ x_{13}\succ x_{20}\succ x_{8}\succ x_{4}\succ x_{12}\succ x_{17}\succ x_{9}\succ x_{5}\succ x_{2}\succ x_{7}\succ x_{18}\succ x_{11}\succ x_{14}\succ x_{6}\ x_{10}\succ x_{15}\succ x_{3}\succ x_{19}$$

## 6. CONCLUSIONS

This paper focuses on the reality of the technical innovation activity in China and collects data in a period as long as 10 years. In this paper, the GM(1, N) with KPCA for nonlinear regression is presented. It is shown that the features selection approach based on grey model is effective and feasible. Also, the GM(1, N) and the KPCA model are integrated together to evolutes the technical innovation activity, which provides a novel way for the evolution design of the other engineering.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

**REFRENCES:**

[1] C. Bock. Forecast of trademark applications in Switzerland, *World Patent Information*. 26(4), 2004, pp.275-282.

[2] Julong Deng. A Course in Grey System Theory, *Wuchang: Huazhong University of Science and Technology Press*, china. 2002, pp.45-60.

[3] B. Scholkopf, A.Smola and K.R.Muller: *Neural Computation* 10, 1988, pp.1299-1319.

[4] I. T. Jolliffe. Principle Component Analysis, *Springer-Verlag*, New York 1986.

[5] L.J. Cao, W.K. Chong. Feature extraction in support vector machine: a comparison of PCA, *Proceedings of the 9th International Conference on Neural Information Processing*. 2(2002), pp.1001-1005.

[6] A. Kocsor, L. Tóth. Kernel-Based Feature Extraction with a Speech Technology Markov Models, *IEEE Transactions on Signal Processing*. 52(8), 2004, pp.2250-2263.

[7] R. Rosipal, M. Girolami, L.J. Trejo, A. Cichocki. Cichocki: Kernel PCA for feature extraction and de-noising in non-linear regression, *Neural Computing & Applications*. 10(3), 2001, pp.231-243

[8] I.Guyon, S.Gunn, M.Nikravesh, L.Zadeh. Nonlinear feature selection with the potential support vector machine, Foundations and *Applications in Feature extraction, Eds. Springer*, 2005.