Research on Maintenance and Repair Cost of Equipment Based on Improved PCA

GONG Yongqi*,
Air Force Logistics College, Xuzhou 221000, China

DENG Jian, LIU Shenyang
Air Force Logistics College, Xuzhou 221000, China

***Abstract*—In order to avoid information loss problem when maintenance and repair cost of equipment was forecasted by traditional PCA (principal component analysis), improved PCA was put forward in the paper. At first, processing the index data that affecting maintenance and repair cost of equipment by traditional PCA (principal component analysis), then improved PCA was used. Compared to traditional PCA, the result of improved PCA showed that the principal component analysis reduces the number of principal components, and improves the first principal component cumulative contribution rate and assess reasonably.**

***Index terms*—maintenance and repair cost, PCA, equalization, standardizations**

I. INTRODUCTION

Equipment maintenance is an essential activity that keeps equipment in good and function well, and maintenance cost for equipment is also important for plant budget. By assessing maintenance cost, many large-scale factory can reasonably distribute funds and make proper development plan. Nowadays assessment model of maintenance cost have principal component analysis, analytic hierarchy process, gray connect degree synthetic evaluation, fuzzy comprehensive evaluation model[1], and so on. In order to forecast rationally maintenance of equipment expense, the first thing is to find key indicators affected maintenance cost for equipment.

Through the investigation and research, the paper find that many correlative charges consist of maintenance and repair cost of equipment, which are cost of repaired personnel, spare part cost, maintenance material cost, training expenses of maintenance, shipping and handling fees, purchasing cost of equipment, support cost of equipment, technical information cost, management fee, maintenance power cost, repair test fee and technical renovation costs [2]. Due to these charges have a certain correlation, information reflected by collected data have an overlap in some degree. Therefore, this article aims to reduce many indicator to few targets, which can reflect the information of original data as much as possible.

Acquiring the covariance matrix's eigenvalues and eigenvectors of primary data is the key of PCA. But because different dimension of indicators have a great effect on PCA, it's imperative for original data to normalization processing. Traditional PCA is processing data by standardizing, whose effect is not optimal[3]. This paper main improve this aspect.

II. TRADITIONAL PCA

PCA also can be called principal component analysis or matrix data analysis. It can find the major features of original data by the way of changing variables. By getting rid of redundant data it can maintain the major features at the same time. For the purpose of reducing original dimension, cutting off relevant interference and finally we can find the dominant factors. By the interpretation of samples , we can find a more exact estimate[4].

PCA is an analytical method that put n index vectors of data matrix X as linear combination F_i .By Var(F_1)=max Var ($\sum_i^n c_i X_i$), we can know that F_1 is the maximum variance of all linear combination. The figures of variances show that how much information be extracted by every principal component. And original index correlative matrix corresponding eigenvalue λ_i equals contribution of each principle component variance[5-6]. The bigger the variance's contribution rate is, the stronger the ability of the relevant principal component to reflect the general information is.

This paper chooses annual maintenance and repair cost of equipment to analyze by PCA, the modeling structure is shown in Fig.1.

Corresponding author: Gong Yongqi*:
Email: 1256392494@qq.com.

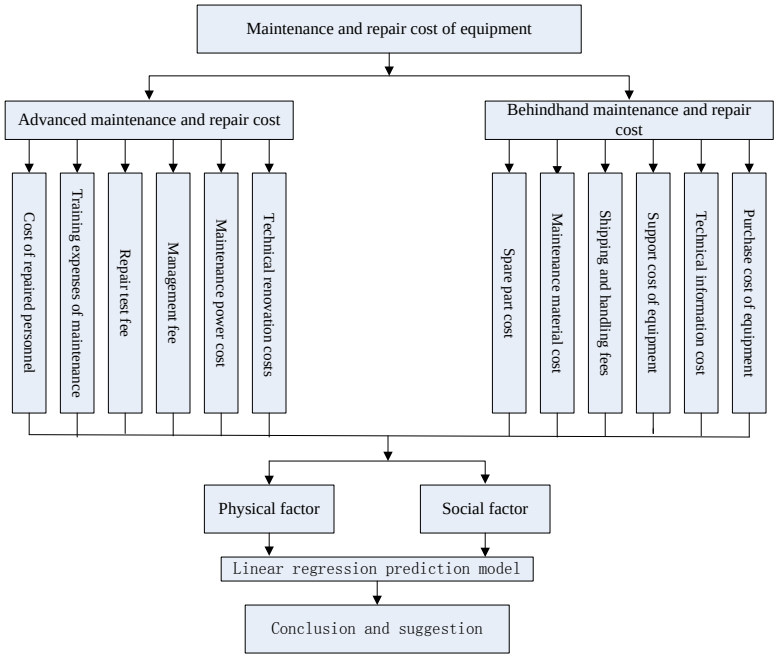

Fig.1 the modeling structure of maintenance and repair cost of equipment

Principle component analysis steps:

a) Sample standardization: the impact of normalization is to eliminate original indices' dimension and to ensure comparability of each index.

$$X_{ij} = Y_{ij} - \overline{Y_j} / S_j \quad (1)$$

$$\overline{Y_j} = \sum_{i=1}^{m} Y_{ij} \quad (2)$$

$$S_j^2 = \sum_{i-1}^{m} (Y_{ij} - \overline{Y_j})^2 / (m-1) \quad (3)$$

b) Estimating σ by standardized sample: through σ =E(XXT)=cov(XXT), we can estimate the covariance matrix by samples.

$$\sigma_{ij} = \sum_{k=1}^{m} X_{ki} X_{kj} / \sqrt{\sum_{k=1}^{m} X_{ki}^2 \sum_{k=1}^{m} X_{kj}^2} \quad (4)$$

c) Calculating each principle component: according to the covariance matrix σ calculated in front we can get n non negative characteristic roots $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ and thus get unit characteristic vectors and constitute orthogonal matrix finally.

d) Calculating variance contribution rate α and accumulated variance contribution rate γ of each principle component: If serve λ_i as its covariance matrix, it's variance contribution rate and cumulative contribution rate is:

$$S_{ij} = \lambda_i / \sum_{i=1}^{n} \lambda_i \quad (5)$$

$$\gamma_i = \sum_{i=1}^{r} \lambda_i / \sum_{i=1}^{n} \lambda_i \quad (6)$$

e) Extract principal components: Make sure principal components by selecting the first principal components on the principle of cumulative contribution rate which is not less than a certain threshold (95%) [7].

III. IMPROVED PCA

Initial data contain not only different information of variability about each index, but also contain the information which indicates how each index is interacted. Traditional principal component analysis using the standardized processing of raw data makes each variance of index 1. By this way, it can eliminate the variation of each index. And principal component extracted by standardized data or calculated by correlative matrix actually contains merely the information interacted by each index[8]. All the information contained by initial datum is not reflected fully.

Because of the above reasons, the effect on digging information of data through standardization is not very well. In order to reflect the different information and eliminate the influences of index dimension and the order of magnitude simultaneously, the improved way is to process the data by equalization.

a) Processing initial data by equalization: equalization means that each index divided by the relevant original data. Using (7) and (8) translates initial data matrix X_{ij} into equalized data matrix Y_{ij}.

$$y_{ij} = x_{ij} / \bar{x}_j \quad (7)$$

$$\bar{x}_j = \sum_{k=1}^{n} x_{kj} / n \quad (8)$$

b) Calculating covariance matrix U_{ij} through equalization matrix: because mean value of each vector among Y_{ij} is 1, we can request covariance matrix U_{ij} via (5) and (9).

$$u_{ij} = s_{ij} / (\bar{x}_i \bar{x}_j) \quad (9)$$

$$r_{ij}^* = u_{ij} / \bar{u}_{ij}^2 = s_{ij} / (\overline{s_{ii}} \overline{s_{jj}}) = r_{ij} \quad (10)$$

c) r_{ij}^* represents indices' correlation coefficient of equalization data, and r_{ij} from (10) represents indices' correlation coefficient (10) show equalization of primary data doesn't change correlation coefficient of index. Therefore covariance matrix by equalization can reflect

all the relevant information of different indices in the effect degree.

d) By using covariance matrix we can get and (4), (5) and (6), various principal component and variance contribution rate can be acquired.

e) When variance contribution rate of the first principal component F_1 is not very high, it's necessary to select the first r principal components on the principle of cumulative contribution rate which is not less than a certain threshold(95%). And then putting variance contribution rate of each principal component F_i as weight α_i, calculating the composite model of principal component (11) express this composite model that various principal component multiply by corresponding indices' coefficient[9].

$$H = \alpha_1 F_1 + \alpha_2 F_2 + \cdots + \alpha_r F_r \quad (11)$$

IV. CASE ANALYSIS

A. *Indices data of Maintenance and repair cost of equipment*

This paper makes use of a certain maintenance and repair cost of equipment as data sample. Based on all-rounded principle, we choose twelve Indices expenses as maintenance cost. Indices system of maintenance and repair cost of equipment are shown as indicated in Table.1. In order to fully grasp maintenance and repair cost of this machine, we collect indices data between 2004 to 2013 as listed in Table. 2.

TABLE I. INDICES SYSTEM OF MAINTENANCE AND REPAIR COST OF EQUIPMENT

Indices	Conventional letter	Interdependency	Unit
Cost of repaired personnel	X_1	+	Ten-thousand RMB
Spare part cost	X_2	+	hundred-thousand RMB
Maintenance material cost	X_3	+	Ten-thousand RMB
Training expenses of maintenance	X_4	+	Ten-thousand RMB
Shipping and handling fees	X_5	+	thousand RMB
Purchase cost of equipment	X_6	+	Ten-thousand RMB
Support cost of equipment	X_7	+	Ten-thousand RMB
Technical information cost	X_8	+	Ten-thousand RMB
Management fee	X_9	+	Ten-thousand RMB
maintenance power cost	X_{10}	+	Ten-thousand RMB
repair test fee	X_{11}	+	thousand RMB
technical renovation costs	X_{12}	+	Ten-thousand RMB

TABLE II. DATA OF MAINYENANCE AND REPAIR COST

Years	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
2004	5	42.5	2	25	0.75	62.5	110	45	45	2.2	1	4
2005	5.7	102.5	3.55	22.5	1.2	55	112	40	50	3	0.8	5.2
2006	4.7	100	3.15	18	1	75	81.5	40.25	48.75	1.2	0.6	3.2
2007	3.7	72.625	2.5	16.6	0.65	86.65	48	37.6	44.6	4.1	0.5	2.6
2008	3.3	45.626	2.15	20	0.45	80	25	32.5	40	1.1	0.45	4.27
2009	3.4	34	2.25	27	0.35	60.9	14.6	29.25	36.65	1.5	0.62	5.6
2010	3.55	43.25	2.5	35	0.45	42.5	12.5	32.5	35	2.7	0.7	6.5
2011	3.6	68	2.5	39.6	0.08	37	12.45	42.5	36	2.6	0.55	7.2
2012	3.35	97.5	2.5	37.5	0.75	42.5	12.5	52.5	37.5	1.8	0.51	5.5
2013	4.1	110	2.75	25	1.2	40	17.5	42.5	40	3.1	0.46	7.1

B. *THE COMPARISON BETWEEN MODIFIED PCA AND TRADITIONAL PCA*

Through traditional PCA, processing collected data in standardized methods acquires various principal component and variance contribution rate as listed in Table.3. Modified PCA adopts equalized method to deal with primary data, and acquired result was listed in Table.4.

TABLE III. TRADITIONAL PCA.

Indices	Initial eigenvalue			Extraction of principal components		
	Summation	Proportion	Cumulative	Summation	Proportion	Cumulative

		(%)	(%)		(%)	(%)
X_7	5.0994	42.4954	42.4954	5.0994	42.4954	42.4954
X_2	2.9013	24.1773	66.6727	2.9013	24.1773	66.6727
X_6	1.7212	14.3436	81.0163	1.7212	14.3436	81.0163
X_4	0.9743	8.1192	89.1355	0.9743	8.1192	89.1355
X_8	0.8413	7.0112	96.1467	0.8413	7.0112	96.1467
X_9	0.3475	2.8956	99.043			
X_{12}	0.1076	0.8963	99.9393			
X_{10}	0.0070	0.0583	99.9976			
X_1	0.0004	0.0032	99.9999			
X_3	0	0	99.9999			
X_5	0	0	99.9999			
X_{11}	0	0	99.9999			

TABLE IV. IMPROVED PCA.

Indices	Initial eigenvalue			Extraction of principal components		
	Summation	Proportion (%)	Cumulative (%)	Summation	Proportion (%)	Cumulative (%)
X_7	1.1528	58.8361	58.8361	1.1528	58.8361	58.8361
X_2	0.3586	18.3020	77.1381	0.3586	18.3020	77.1381
X_6	0.2015	10.2863	87.4244	0.2015	10.2863	87.4244
X_4	0.1499	7.6528	95.0772	0.1499	7.6528	95.0772
X_8	0.0550	2.8078	97.885			
X_9	0.0316	1.6105	99.4955			
X_{12}	0.0090	0.4576	99.9531			
X_{10}	0.0009	0.0454	99.9985			
X_1	0.0000	0.0016	100			
X_3	0	0	100			
X_5	0	0	100			
X_{11}	0	0	100			

Traditional PCA selects the first principal components on the principle of cumulative contribution rate which is not less than certain threshold (95%), and reducing dimensions to five indices from twelve indices as indicated in Fig.2. This five indices was ordered from large to small by contribution rate, which are support cost of equipment, spare part cost, purchasing cost of equipment, training expenses of maintenance, technical information cost, and have important influence on maintenance and repair cost of this equipment.

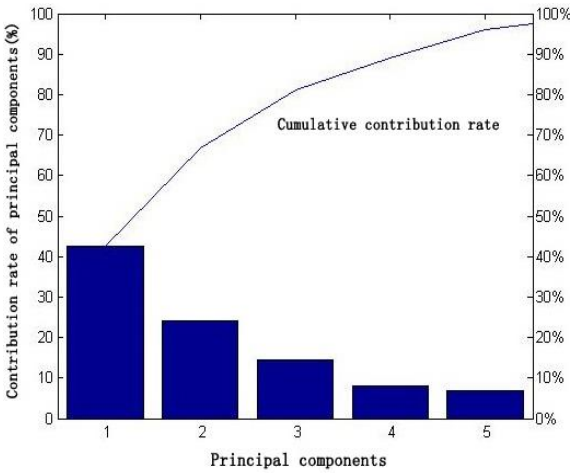

Fig.1. Extraction of principal components by traditional PCA

Modified PCA reduces dimensions to four indices from twelve indices as indicated in Fig.3. This four indices are support cost of equipment, spare part cost, purchasing cost of equipment and training expenses of

maintenance, which were ordered from large to small by contribution rate, and cumulative contribution rate of this five indices reach certain threshold (95%), which contains almost all the information on maintenance and repair cost of this equipment.

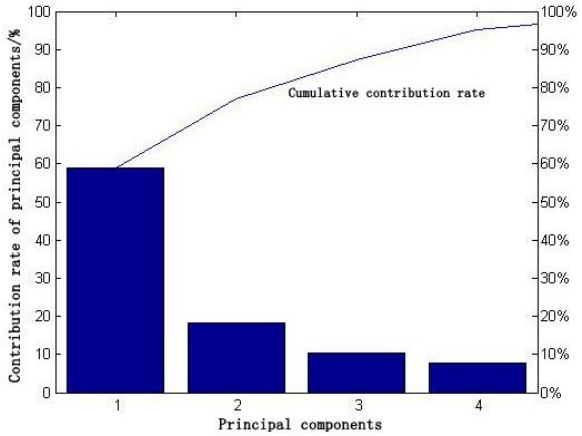

Fig.2. Extraction of principal components by modified PCA

V. CONCLUSION

Compared between traditional PCA and modified PCA, processing initial data by equalized method not only can eliminate indices dimension and order of magnitudes, but also get rid of information loss by standardized methods. Principal components extracted by modified PCA can contain more comprehensive information than principal components extracted by modified PCA. Contrast is known between Table 2 and Table 3, variance contribution rate of the first principal components in modified PCA is 58.8361%, which improves by 16.3407% in traditional PCA. By means of the comparison between Fig.2 and Fig.3, Fig.2 shows that traditional PCA extracted five principal components contained mostly all information, but Fig.3 indicates that modified PCA only chose four principal components reducing research's workload. At the same time, cumulative contribution rate of four principal components exceeds 95%, which is up to the mustard.

Making sure principal elements affected maintenance and repair cost of equipment contributes to predict upkeep cost of machine via practical data. Through analyzing principal components, the factory especially large-scale factory can identify development plan and allocate rationally funds.

REFERENCES

[1] Wang Yingluo, *Systems Engineering.* Beijing, MA: China Machine Press, 2008.

[2] M. E. Tipping, " Probabilistic principal component analysis," J. Journal of the Royal Statistical Society., vol. 61, no. 3, pp. 611-622, Mar. 1999.

[3] Ni Yuhui and Dan Bingjuan, " Upkeep costs of control equipment," J. Non-government Technological., vol. 83, no. 2, pp. 95-108, Mar.2007.

[4] Liu Muxiao, " Application of improved grey theory in the cost forecasting of vessel maintenance," J. Ship Electronic Engineering., vol. 12, pp. 23-27, Apr. 2010.

[5] Ji RongFang, " Data pre-processing method in improved PCA," J. Journal of Shangdong Universty of Science and Technology., vol. 43, no. 4, pp. 195-208, Apr.2007.

[6] S. Haykin, " The management of maintenance and repair cost," J. Equipment Maintenance and Management., vol. 13, no. 2, pp. 19-22, May.2010.

[7] Mao Zhaoyong, " Optimal policy research of preventive maintenance period to series systems," J. Fire Control&Command Control., vol. 34, no. 4, pp. 63-66, Apr.2009.

[8] Xu Yajing, " The improvement of the application method of principle component analysis," J. Mathematics in Practice and Theory., vol. 24, no. 3, pp. 21-25, May.2006.

[9] Richard A Johnson, " Applied multivariate statistical analysis," J. Mathematics in Practice and Theory., vol. 14, no. 3, pp. 123-127, Apr.2003.