Available online www.jocpr.com

Journal of Chemical and Pharmaceutical Research, 2014, 6(7):970-976

Research Article

ISSN : 0975-7384
CODEN(USA) : JCPRC5

Principal component analysis-based sports dance development influence factors research

Xiaoping Xie

Public Sports Department, Jingchu University of Technology, Jingmen, Hubei, China

---

ABSTRACT

Nowadays material civilization and spiritual civilization of human beings have been rapidly developing; sports dance has become a sports event that is favored by people. Due to China covers a vast geographic area and has a large population, sports dance development suffers many factors influence as economic level is limited, people acceptance levels are not enough and so on .When research on sports dance development relative problems, excessive such factors will lead to inconvenience in researching. The paper takes sports dance development as research objects, takes teachers, referees, athletes' three types of population questionnaire survey results as evidence, targeted at economic level, competition system, scientific research level and other thirteen influence factors to make principal component analysis. Analysis result indicates that two main components can replace the thirteen influence factors. The two principal components are three influence factors' linear combination, the differences between the two is that every factor weight is different.

Key words: sports dance, influence factors, principal component analysis, linear combination, weight

---

INTRODUCTION

In recent years, sports dance has been developed in lots of universities, and become one of well-received sports events among numerous university students. To this day, sports dance development is still not balanced. The problem leads to people to think about sports dance development constraint factors.

In 2010, Cheng Wei-Hai in the article "Hebei province sports dance development status investigation and development countermeasures research", analyzed Hebei province sports dance development status, result showed that in the aspect of sports dance, Hebei province men athletes were fewer, coaches and referees cultural levels needed to improve. Competition organizations were disordered, management was poor, and these factors serious restricted Hebei province sports dance development [1-3]. In 2004, Zheng Chuan-Feng and others in the article "Research on sports dance development and countermeasures in universities", they made comprehensive analysis of present sports dance development status in universities. Result showed, each university sports dance development levels were different, students' interests and positivity in sports dance event were higher, but overall they presented as backward of theoretical knowledge and scientific researches [4-7]. With respect to this, authors provided countermeasures. In order to let sports dance to be able to better develop, author pointed out, it should propel to sports dance optional course and club-oriented [8-10]. In 2012, Bi Fei in the article "University sports dance event education research", applied teaching experiment method and others multiple research methods, analyzed universities sports dance teaching transforming towards sports dance education issues, result showed that sports dance education functions were not only letting students to master sports dance basic motions, but also can let students' attainments and quality to be improved [10-12]. Sports dance event teaching conformed to the trends of times that were worth promoting to each university. In 2003, Zhao Li in the article "Chinese sports dance organization status and counter measure research", applied multiple research methods, researched on sports dance development influence factors, research result showed Chinese sports dance lacked of self textbook system,

scientific research levels were lower, referees grade evaluation system was not normalized, referees education degrees were generally lower [13].

The paper takes Chinese university sports dance development status as research object, analyzes teachers, referees and athletes each kind of situations, and further gets conclusion.

MODEL ESTABLISHMENT

The model researches on sports dance development influence factors, takes each factor teachers occupied percentage, referee occupied percentage and athlete occupied percentage as evidence, and reduces multiple influence factors into fewer influence factors so as to easy for the kind of problems' late researching.

Table 1 data is from "Chinese sports dance development status investigation and countermeasure research".

Table 1: Original data table

Influence factor	Teacher	Percentage%	Referee	Percentage %	Athlete	Percentage %	Rank
Economic level constraints	132	80	22	43	630	61	8
Competition and referee system	140	84	37	73	932	90	3
Teachers level	154	93	51	100	1000	97	1
Funding issue	93	56	43	84	403	39	11
Related to non-Olympic Games events	103	62	41	80	726	70	4
Fewer international exchange	149	90	47	92	955	92	2
Public concept	57	34	30	59	745	72	6
Scientific research level	73	44	25	49	352	34	12
Sports level	62	37	32	63	601	58	9
Mass media influence	88	53	30	59	722	70	5
Disordered organizational management	77	46	24	47	561	54	10
Field facilities	96	58	17	33	677	65	7
Others	26	16	6	12	135	13	13

Main thought of principal component analysis is variable's dimension reduction. It is a statistical analysis method that transforms multiple variables into fewer main variables. It generally is used to data compression, system evaluation, regression analysis and weighted analysis so on.

Principal component analysis method

Main way of principal component analysis is reducing dimension of variables, which is recombining original many variables with correlation into a group of uncorrelated variables to replace original variables. Therefore, we can pay attention to every time observation's variables that have maximum variation, to every time observation's small changed variables that can be used as constant to process and get rid of them, so that it reduces variables number in problem that needs to be considered.

Assume that there is m pieces of original indicators to do principal component analysis, which are recorded as $x_1, x_2, \cdots, x_m$, now it has n pieces of samples, corresponding observation value is $x_{ik}(i=1,2,\cdots,n)$, and $k=1,2,\cdots,m$ takes standardization transformation, and then transform x_k into x_k^*, that:

$$x_k^* = \frac{x_k - \overline{x_k}}{s_k}, \quad k=1,2,\cdots,m \tag{1}$$

Among them, $\overline{x_k}$ and s_k are respectively x_k average number and standard deviation, x_k^* average number is 0, standard deviation is 1.

According to each sample original indicator observation value x_{ik} or after standardization observation value x_{ik}^*, it solves coefficient b_{kj}, establish indicator x_k^* that is transformed through standardization to express comprehensive indicator z_j equation $z_j = \sum_k b_{kj} x_k^*$, which can also establish equation that uses original indicator x_k to express comprehensive indicator z_j :

$$z_j = \sum_k \tilde{b}_{kj} x_k^* + a_j \tag{2}$$

There are two requirements on defining b_{kj} :
(1) Comprehensive indicators are mutual independent from each other or uncorrelated.
(2) Every comprehensive indicator reflected each sample gross information content is equal to corresponding feature vector(comprehensive indicator coefficient)feature values. In general, it is required that selected comprehensive indicator feature vales contribution ratios sum to be above 80%.

Principal component analysis general steps

(1) According to observed data, calculate $\overline{x_k}$ and $s_k\left(k, j = 1,2,\cdots,m\right)$.

(2) By correlation coefficient matrix R, it can get feature value $\lambda_j \left(j = 1,2,\cdots,m\right)$ and each principal component variance contribution、 contribution ratio and accumulative contribution ratio, and define principal component reserved number p with accumulative contribution ratio as evidence.

(3) m pieces of basic equations are as following:

$$\begin{cases} r_{11}x_1^{(j)} + r_{12}x_2^{(j)} + \cdots + r_{1m}x_m^{(j)} = \lambda_j x_1^{(j)} \\ r_{21}x_1^{(j)} + r_{22}x_2^{(j)} + \cdots + r_{2m}x_m^{(j)} = \lambda_j x_2^{(j)} \\ \cdots \\ r_{m1}x_1^{(j)} + r_{m2}x_2^{(j)} + \cdots + r_{mm}x_m^{(j)} = \lambda_j x_m^{(j)} \end{cases} \tag{3}$$

Among them, $j = 1,2,\cdots,m$.

Proceed with Schmidt orthogonalization, for every λ_i, solve its basic equations solution $x_1^{(j)}$, $x_2^{(j)}$, …, $x_m^{(j)}$ $\left(j = 1,2,\cdots,m\right)$, and then let:

$$b_{kj} = \frac{x_k^{(j)}}{\sqrt{\sum_k \left(x_k^{(j)}\right)^2}} \tag{4}$$

It can get expressed by x_1^*, x_2^*, …, x_m^* principal component $z_j = \sum_k b_{kj} x_k^*$, or input $x_k^* = \frac{x_k - \overline{x_k}}{s_k}$ and then get expressed by x_1, x_2, …, x_m principal component $z_j = \sum_k \tilde{b}_{kj} x_k^* + a_j$.

(4)Input x_1, x_2, …, x_m observed values into principal component expressions, calculate each component value.
(5) Calculate original indicator and principal component correlation coefficient that is also factor loading that use it to explain principal component significances.

Table 2 represents every variable communalities result. Table 2's left side represents every variable explainable variance from all factors, while the right side represents variable communalities. From table data, we can see that variable communalities are 1 that are very high, which shows most information in variables can be extracted by factors; it shows the analysis is valid.

Table 2: Variables communalities table

	Initial	Extract
Economic level constraints	1.000	1.000
Competition and referee system	1.000	1.000
Teachers level	1.000	1.000
Funding	1.000	1.000
With non-Olympic Games events	1.000	1.000
International exchange	1.000	1.000
Public concept	1.000	1.000
Scientific research level	1.000	1.000
Sports level	1.000	1.000
Mass media influence	1.000	1.000
Disordered organizational management	1.000	1.000
Field facilities	1.000	1.000
Others	1.000	1.000
Extract method: principal component analysis.		

Table 3: Factor contribution ratio table

Component	Initial feature value			Extract squares sum and load in			Rotate squares sum and load in		
	Total	Variance %	Accumulation %	Total	Variance %	Accumulation %	Total	Variance %	Accumulation %
1	7.568	58.212	58.21	7.568	58.212	58.212	7.411	57.010	57.01
2	5.432	41.788	100.00	5.432	41.788	100.00	5.589	42.990	100.00
3	5.919E-16	4.553E-15	100.00						
4	2.462E-16	1.894E-15	100.00						
5	2.419E-16	1.861E-15	100.00						
6	1.329E-16	1.022E-15	100.00						
7	9.404E-17	7.234E-16	100.00						
8	-1.660E-17	-1.277E-16	100.00						
9	-1.103E-16	-8.484E-16	100.00						
10	-1.479E-16	-1.138E-15	100.00						
11	-1.711E-16	-1.316E-15	100.00						
12	-2.965E-16	-2.281E-15	100.00						
13	-3.367E-16	-2.590E-15	100.00						
Extract method: principal component analysis.									

In Table 3, accumulation items' data indicates percentage of total feature values. From table data, it can easily see that factor 1 and factor 2 feature values are above 1, and the two factors' feature values sum are 100% of total feature values. Therefore, we use factor 1 and factor 2 as main factors.

Table 4: Rotational factor loading table

	Component	
	1	2
x1	-.939	.345
x2	-.318	.948
x3	.960	-.282
x4	.292	-.956
x5	.908	-.418
x6	.988	.155
x7	.878	.479
x8	-.034	-.999
x9	1.000	-.026
x10	.662	.750
x11	.464	.886
x12	-.453	.892
x13	-.996	.087
Extract method: principal component. *Rotational method: Orthogonal rotation method with Kaiser standardization.*		
a. Convergent after three times' iterating of rotation.		

Data in Table 4 indicates factor loading value after using Kaiser standard orthogonal rotation. By such rotating, every factor's significance is relative clear. From the table, it can see that two main factors are extracted. From Figure 1, it can more intuitive indicate.

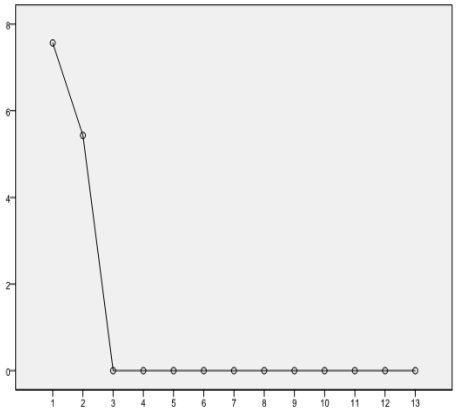

Fig.1: Scree plot

Figure 1 is feature values' scree plot. In general, the figure shows big factor steep slope and surplus factor gentle tail has obvious interruption. Generally selected main factors are in the very steeply slope, and factors lie in gentle slope have insignificant effects on total. From Figure 1, it is clear that the former two factors are in the relative steeply slope, and starts from the third factor, the slope turns to be gentle, while starts from the third factor, the slope is nearly zero, therefore select two factors as comprehensive factors.

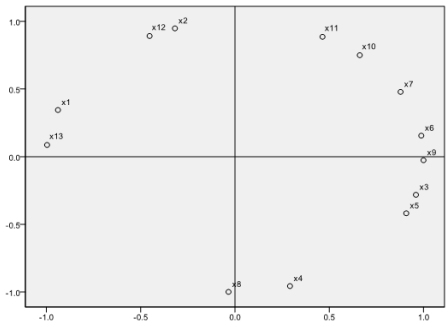

Fig.2: The view of the rotating components of the space

From Figure 2, it is clear that principal component analysis totally extracts two main factors this time, x_8, x_9 that get closer to coordinate axis have big factor loading and explanatory ability is relative strong.

After defining main factors amount, it should calculate feature vectors, feature vectors amount is the same as main factors amount. Feature vector matrix is as Table 5 show.

Table 5: Feature vector matrix

F1	F2
-1.58831	-0.42853
-1.01902	0.98213
1.03123	-0.32397
0.42783	-1.62071
1.01548	-0.5518
0.91186	0.3439
0.65869	0.78239
0.03037	-1.82201
0.99056	0.0755
0.28878	1.09916
-0.00898	1.21932
-1.16855	0.84169
-1.56994	-0.59705

By Table 5 feature vectors, it can get principal component computational formula:

$$z_1 = -1.58831x_1 - 1.01902x_2 + 1.03123x_3 + 0.42783x_4 + 1.01548x_5 + 0.91186x_6 + 0.65869x_7 + 0.03037x_8 + 0.99056x_9 + 0.28878x_{10} - 0.00898x_{11} - 1.16855x_{12} - 1.56994x_{13} \quad (5)$$

$$z_2 = -0.42853x_1 + 0.98213x_2 - 0.32397x_3 - 1.62071x_4 - 0.5518x_5 + 0.3439x_6 + 0.78239x_7 - 1.82201x_8 + 0.0755x_9 + 1.09916x_{10} + 1.21932x_{11} + 0.84169x_{12} - 0.59705x_{13} \quad (6)$$

Respectively input original data into formula (5) (6), it can get data as Table 6.

Table 6: Main components variables

z_1	z_2
0.97268	0.268955
0.853204	0.2337
0.24219	1.49121

Table 6 is two main components variables after factor analysis. z_1, z_2 are economic restriction level and other thirteen factors' linear combinations. That is principal component analysis reducing original thirteen factors into two factors so that is easy to research, but physical significances after factor dimensions reducing is hard to define.

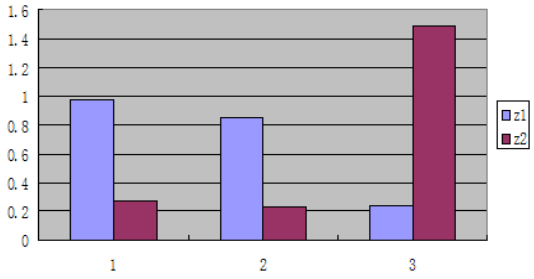

Fig.4: Principal component comparison chart

Figure 4 is principal component comparison chart. From Figure 4, it is clear that to teachers and referees, principal component 1 occupied weight is bigger than that of principal component 2, to athletes; principal component 1 occupied weight is smaller than that of principal component 2.

CONCLUSION

Utilize principal component analysis thought to use fewer variables to replace original multiple variables, these fewer variables can reflect original data most information. In addition, the model more focuses on information comprehensive evaluation. The method also has certain drawbacks, such as, when principal component factor loading positive and negative symbols are simultaneously existing, evaluation function significances will not be clear, naming clarity will be low, only involve a group of variables' correlations. The model's principal component is composed of original factors linear combinations, so principal components actual significances are hard to define, just functions as dimension reduction. Principal component analysis application field is very widely, such as "regional water resources carrying capacity problem", "Town land evaluation problem", "3G network comprehensive performance evaluation problem" and other aspects' analysis problems. The paper applies principal component analysis into sports dance development restriction factors, succeeds in reducing three influence factors into two principal components, and is convenient for later such kinds of problems researching.

Acknowledgment
The paper belongs to one of Hubei province science of education "the 12th five years" project issue in 2013-"Regular institution of higher learning sports classroom interactive teaching research" (No.2013B200) research attainments.

REFERENCES

[1] Wang Xiaoguang. *Research On Development*, **2007**, (5), 57-60.
[2] CHEN Nan - yue. *Journal of Yunnan Finance and Trade Institute*, **2005**, 21(6), 81-85.

[3] LIU Chang-ming, GUAN Bin. *Journal of Tianjin University(Social Sciences)*, **2009**, 11(5), 453-457.
[4] Fan Liwei, Liang Jiyao. *Value Engineering*, **2006**, 25(5), 29-31.
[5] SUN Yi-kai. *Journal of Anhui University(Philosophy & Social Sciences)*, **2006**, 30(6).
[6] LU Jian-hua. *Journal of Anhui University(Philosophy & Social Sciences)*, **2006**, 30(6), 5-6.
[7] XU Xiao-yue. *Journal of Anhui University(Philosophy & Social Sciences)*, **2006**, 30(6), 6-8.
[8]SU Bao-mei, LIU Zong-xian, *Liu Chang-ming. Journal of Jinan University*, **2002**, 12(5), 1-10.
[9] Zhang B.; Zhang S.; Lu G.. *Journal of Chemical and Pharmaceutical Research*, **2013**, 5(9), 256-262.
[10] Zhang B.; *International Journal of Applied Mathematics and Statistics*, **2013**, 44(14), 422-430.
[11] Zhang B.; Yue H.. *International Journal of Applied Mathematics and Statistics*, **2013**, 40(10), 469-476.
[12] Zhang B.; Feng Y.. *International Journal of Applied Mathematics and Statistics*, **2013**, 40(10), 136-143.
[13] Bing Zhang. *Journal of Chemical and Pharmaceutical Research*, **2014**, 5(2), 649-659.