

基于主成分分析的 NBA 球员综合能力评价

满 帅¹, 龙 华¹, 熊 新¹, 李一民¹, 刘霖璇²

(1. 昆明理工大学 信息与自动化学院, 云南 昆明 650500;
2. 加利福尼亚大学欧文分校 数量经济学院, 美国 加利福尼亚 92697)

摘 要: 目前对球员综合能力的评价方法有 TOPSIS 法、灰色关联分析法等, 但当评价指标量多时, 存在计算效率低等缺陷, 同时有些评价指标并不是值越大越好, 且会对最优排序和最劣排序造成影响。为求取最适合的算法、作出准确的 NBA 球员综合能力评价, 采用主成分分析法, 以现役 NBA 联盟中 538 名运动员的得分、助攻、三分命中率等 13 项指标为实例, 进行评价方法研究。分析球员多方面能力, 得出球员在各项成分中得分排名并计算出各项能力最强的前十名球员, 与体育界分析结果进行比较, 分析球员的强项和弱项。实验结果表明, 主成分分析能够高效地将数据降维, 表现各项数据之间关联性, 并且分析结果正确, 适用于 NBA 球员综合能力评价。

关键词: 主成分分析; NBA; 球员综合能力评价

DOI: 10.11907/rjdk.172838

中图分类号: TP391

文献标识码: A

文章编号: 1672-7800(2018)006-0185-05

Evaluation of NBA Players' Comprehensive Ability Based on Principal Component Analysis

MAN Shuai¹, LONG Hua¹, XIONG Xin¹, LI Yi-min¹, LIU Lin-xuan²

(1. Department of Information and Automation, Kunming University of Science and Technology, Kunming 650500, China;

2. Department of Number Economics, University of California, Irvine, California 92697, United States)

Abstract: At present, the evaluation methods of the player's comprehensive ability mainly include Topsis method, gray correlation analysis and so on. However, the calculation efficiency is low when the evaluation index is large; indicator magnitude has an impact on optimal scheduling and worst ranking while not all large evaluation indicators equal better performance. Aiming to find the most suitable algorithm to make accurate evaluation of NBA players' comprehensive ability, we use the principal component analysis to research the evaluation method with 13 indicators involving 538 NBA league athletes' scores, assists, three point rates and other performance as examples. We analyze the players' ability in various aspects, calculate the final scores, and select the top ten players to compare with the official analysis of NBA; and then we analyze the strengths and weaknesses of the players. The experimental results show that the principal component analysis can reduce the data dimensionality, present the relationship between the performance of the data and the results are correct. Therefore, it is applicable in NBA players' comprehensive ability evaluation.

Key Words: principal component analysis; NBA; evaluation of players' comprehensive ability

0 引言

篮球作为一种主流运动, 受到广大球迷(特别是大、中学生)的关注。NBA 是世界男子篮球竞技水平最高的联

赛, 随着人民生活水平的提高, 以及通信技术的飞速发展, 中国对 NBA 联赛转播力度加大, 我国人民对 NBA 的关注越来越多, 在王治郅、姚明、易建联、周琦加入 NBA 后, 由于其巨大的商业价值而受到了社会各界的关注。^[1]

篮球运动员的综合能力评价一直是联盟、球队管理高

收稿日期: 2017-11-09

基金项目: 云南省科技厅资助项目(2014RA051)

作者简介: 满帅(1993-), 男, 昆明理工大学信息工程与自动化学院硕士研究生, 研究方向为无线通信; 龙华(1963-), 女, 博士, 昆明理工大学信息工程与自动化学院教授, 研究方向为无线网络; 熊新(1976-), 男, 博士, 昆明理工大学信息工程与自动化学院副教授, 研究方向为信息处理; 李一民(1963-), 男, 博士, 昆明理工大学信息工程与自动化学院教授, 研究方向为智能信号处理; 刘霖璇(1997-), 女, 加利福尼亚大学欧文分校, 数量经济学院学生, 研究方向为数据分析。

层和球迷非常关心的事情。联盟每年都会颁发各种奖项,比如常规赛 MVP、最佳新秀、进步最快球员、最佳防守球员等^[2],这也是对球员素质的评价。最终这些奖项归谁所有,仅仅根据球员一个赛季的上场表现,由美国资深体育评论员、体育记者、NBA 技术官员和球迷投票选举产生是远远不够的^[3],因为这种评价中人为因素影响巨大,缺乏一套科学而严谨的评选体系,应该根据球员们的各项技术统计进行分析。现在 NBA 进入了大数据时代,NBA 赛场采用 SportVU 技术^[4],通过使用 6 只摄像头追踪捕捉场上球员动作,每秒可记录 25 个动作,一场比赛可以收集到 72 000 个数据点,采集得到的原始数据积累成为海量数据库。利用这些海量原始数据,通常采用 TOPSIS 法、灰色关联分析法进行评价分析,但当评价指标量多时,计算效率低^[5],并且有些评价指标并不是值越大越好,对最优排序和最劣排序造成影响^[6],本文利用多元统计中的主成分分析法进行球员的综合能力分析,很好地解决了这个问题。

1 研究对象

本文选取 NBA 联盟中 30 只球队的所有运动员在 2016—2017 赛季的数据作为样本。选取球员上场时间、投篮命中率、三分命中率、罚球命中率、前场篮板、后场篮板、总篮板、助攻、抢断、盖帽、失误、犯规、得分,作为原始指标进行统计分析。

2 主成分分析法在 NBA 球员能力评价中的研究

2.1 TOPSIS 法

TOPSIS 法^[7]即逼近理想解排序法,意为与理想方案相似性的顺序选优技术,是系统工程中有限方案多目标决策分析的常用方法。它是基于归一化后的原始数据矩阵,找出有限方案中最优方案和最劣方案(分别用最优向量和最劣向量表示),然后分别计算评价对象与最优方案和最劣方案的距离,获得各评价对象与最优方案的相对接近程度,以此作为评价优劣的依据。

但是利用 TOPSIS 法作 NBA 球员综合评价时,对高优指标和低优指标都很高的球员排名影响大,不能准确说明球员综合能力的高低^[8]。

2.2 灰色关联分析法

灰色关联分析法^[9]是一种通过灰色关联度分析和确定系统因素间的影响程度或因素对系统主行为贡献的测度方法。灰色关联分析的基本思想是根据序列曲线几何形状的相似程度判断其联系是否紧密。

但是利用灰色关联分析法作 NBA 球员综合评价时,存在分辨系数取值问题,并且球员的某些最优指标并不是越大越好,比如失误和犯规,利用该方法对球员的综合能力作评价存在误差^[10]。

2.3 主成分分析法基本概念及原理

主成分分析法(Principal Component Analysis)^[11]是

一种通过“降维”的思想将多个指标的数学统计量转化为少数有代表性的指标量的数学统计方法,采用降维的思想解决高维系统的寻优问题^[12]。其基本原理是:通过一个正交变换把问题中本来相互关联的随机变量转换为一些不相关的新变量。数学上来说,这种方法是将变量的协方差矩阵转化为相应的对角矩阵,将原来的变量系统通过正交变换,构成一个新的正交系统^[13],使原本相关的样本点散列成正交方向,成为不相关的正交变量。

2.4 算法步骤

(1)进行数据处理,假设所统计的球员有 q 个, $q=(1, 2 \cdots d \cdots n)$,每名球员的技术统计指标有 13 个,即 $m=(1, 2 \cdots p \cdots 13)$,则样本统计矩阵见公式(1)。

$$X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1m} \\ X_{21} & X_{22} & \cdots & X_{2m} \\ \cdots & \cdots & X_{dp} & \cdots \\ X_{n1} & X_{n2} & \cdots & X_{nm} \end{bmatrix} \quad (1)$$

X_{dp} 表示第 d 个球员的第 p 个技术统计参数,其中 $p=(1, 2 \cdots 13)$,分别代表球员的场均上场时间、投篮命中率、三分球命中率、罚球命中率、场均前场篮板个数、场均后场篮板个数、场均总篮板个数、场均助攻数、场均抢断个数、场均盖帽数、场均失误数、场均犯规个数以及场均得分。

在评价过程中,因为 X_{dp} 是有量纲的变量,容易对评价结果产生不必要影响,因此,根据主成分分析法对样本进行评价,需要对样本按照公式(4)进行无量纲化处理^[14]。其中,定义 $x_{d'p'}$ 是无量纲化后的样本值, X_{dp} 是样本值, \bar{X}_p 为样本均值,见公式(2):

$$\bar{X}_p = \frac{\sum_{d=1}^n X_{dp}}{n} \quad p = (1, 2 \cdots 13) \quad (2)$$

S_p 为样本标准差,见公式(3):

$$S_p = \sqrt{\frac{\sum_{d=1}^n (X_{dp} - \bar{X}_p)^2}{n}} \quad p = (1, 2 \cdots 13) \quad (3)$$

$$x_{d'p'} = \frac{X_{dp} - \bar{X}_p}{S_p} \quad d = (1, 2 \cdots n), p = (1, 2 \cdots 13) \quad (4)$$

从而得到 $x_{d'p'}$, 标准化后的样本值会出现负值。

首先,对原始样本数据作标准化计算以消除量纲。再通过原始样本矩阵让样本矩阵中心化,即每一维度减去该维度的均值^[15],即 $X' = X_{dp} - \bar{X}_p$, $d = (1, 2 \cdots n)$, $p = (1, 2 \cdots m)$, $m = 13$ 。

$$X' = \begin{bmatrix} X_{11} - \bar{X}_1 & X_{12} - \bar{X}_2 & \cdots & X_{1m} - \bar{X}_m \\ X_{21} - \bar{X}_1 & X_{22} - \bar{X}_2 & \cdots & X_{2m} - \bar{X}_m \\ \cdots & \cdots & X_{dp} - \bar{X}_p & \cdots \\ X_{n1} - \bar{X}_1 & X_{n2} - \bar{X}_2 & \cdots & X_{nm} - \bar{X}_m \end{bmatrix}$$

然后直接用新得到的样本矩阵 X' 的转置乘以它本身,再除以 $(n-1)$ 即可,从而计算出其相关系数矩阵 R 。

$$R = \frac{X'^T X'}{n-1} \quad (5)$$

(2)根据样本的相关系数矩阵 R ,计算出相关系数矩阵的特征方程,并计算出相应的特征值 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$,

$m=13$ 。根据方差特征值计算出方差贡献率则有:

$$\alpha_i = \frac{\lambda_i}{m} \tag{6}$$

其中, α_i 为第 i 个主成分的方差贡献率, $i=(1,2\cdots e\cdots$

13)。又有 $\frac{\sum_{i=1}^e \lambda_i}{m}$ 为前 e 个主成分的累积方差贡献率。

(3) 计算出主成分 $Z_j = c_j^0 \star x_{d'p'}$, $j=(1,2\ldots e)$, c_j^0 是单位特征向量。

(4) 选取方差贡献率超过 85% 以上的主成分构成综合评价函数。

2.5 算法流程图

主成分分析法的算法流程图, 如图 1 所示。

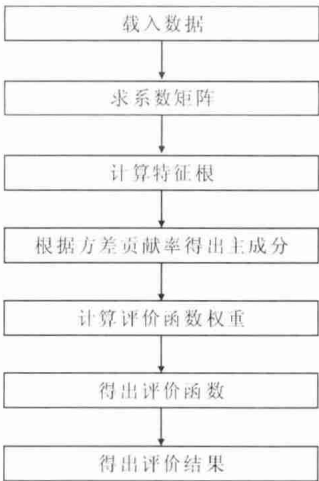


图 1 主成分分析法算法步骤

3 主成分分析法在球员综合能力评价的应用

3.1 数据源

从 <http://nba.sports.sina.com.cn/teams.php?dpc=1> 网站上摘取 NBA 球员的 13 项统计数据。

$R = \begin{bmatrix} 1 & 0.237 & 0.309 & 0.323 & 0.389 & 0.712 & 0.652 & 0.634 & 0.748 & 0.360 & 0.742 & 0.708 & 0.869 \\ 0.237 & 1 & -0.043 & 0.010 & 0.466 & 0.356 & 0.411 & 0.025 & 0.129 & 0.381 & 0.157 & 0.347 & 0.234 \\ 0.309 & -0.043 & 1 & 0.334 & -0.286 & 0.006 & -0.086 & 0.270 & 0.215 & -0.171 & 0.201 & 0.027 & 0.311 \\ 0.323 & 0.010 & 0.334 & 1 & -0.103 & 0.115 & 0.053 & 0.256 & 0.243 & -0.096 & 0.220 & 0.175 & 0.330 \\ 0.389 & 0.466 & -0.286 & -0.103 & 1 & 0.745 & 0.869 & 0.011 & 0.242 & 0.644 & 0.291 & 0.589 & 0.324 \\ 0.712 & 0.356 & 0.006 & 0.115 & 0.745 & 1 & 0.977 & 0.347 & 0.528 & 0.614 & 0.609 & 0.687 & 0.664 \\ 0.652 & 0.411 & -0.086 & 0.053 & 0.869 & 0.977 & 1 & 0.261 & 0.469 & 0.658 & 0.543 & 0.696 & 0.596 \\ 0.634 & 0.025 & 0.270 & 0.256 & 0.011 & 0.347 & 0.261 & 1 & 0.640 & 0.044 & 0.813 & 0.343 & 0.661 \\ 0.748 & 0.129 & 0.215 & 0.243 & 0.242 & 0.528 & 0.469 & 0.640 & 1 & 0.194 & 0.629 & 0.547 & 0.645 \\ 0.360 & 0.381 & -0.171 & -0.096 & 0.644 & 0.614 & 0.658 & 0.044 & 0.194 & 1 & 0.305 & 0.533 & 0.312 \\ 0.742 & 0.157 & 0.201 & 0.220 & 0.291 & 0.609 & 0.543 & 0.813 & 0.629 & 0.305 & 1 & 0.571 & 0.820 \\ 0.708 & 0.347 & 0.027 & 0.175 & 0.589 & 0.687 & 0.696 & 0.343 & 0.547 & 0.533 & 0.571 & 1 & 0.558 \\ 0.869 & 0.234 & 0.311 & 0.330 & 0.324 & 0.664 & 0.596 & 0.661 & 0.645 & 0.312 & 0.820 & 0.558 & 1 \end{bmatrix}$

特征根 λ_i 与相对应的特征向量 c_j^0 , 如表 1 所示。

表 1 特征根与特征向量

特征根	特征值	特征向量
λ_1	0.000 2	$c_1^0 = [0.001\ 0\ 0.570\ 6\ 0.536\ 1\ 0.105\ 3\ 0.424\ 0\ 0.062\ 7\ -0.037\ 0\ 0.166\ 0\ 0.079\ 4\ 0.021\ 3\ -0.024\ 1\ 0.169\ 2\ 0.022\ 0]$
λ_2	0.064 9	$c_2^0 = [-0.000\ 0\ 0.030\ 7\ -0.001\ 6\ -0.116\ 3\ 0.002\ 0\ -0.040\ 9\ 0.043\ 2\ 0.047\ 7\ -0.411\ 5\ -0.714\ 7\ -0.461\ 4\ -0.245\ 1\ 0.160\ 1]$
λ_3	0.131 2	$c_3^0 = [-0.000\ 3\ -0.019\ 6\ -0.058\ 8\ 0.081\ 8\ -0.282\ 2\ 0.023\ 4\ 0.027\ 4\ 0.688\ 9\ -0.246\ 9\ -0.456\ 8\ 0.400\ 4\ 0.057\ 2]$
λ_4	0.148 9	$c_4^0 = [0.000\ 2\ 0.030\ 2\ -0.028\ 7\ 0.031\ 9\ -0.111\ 1\ -0.120\ 2\ -0.035\ 5\ -0.186\ 2\ -0.430\ 4\ 0.535\ 2\ -0.598\ 0\ 0.309\ 3\ 0.098\ 2]$
λ_5	0.262 7	$c_5^0 = [0.246\ 5\ 0.035\ 5\ -0.012\ 1\ 0.719\ 5\ -0.282\ 4\ 0.066\ 5\ 0.285\ 9\ 0.056\ 0\ 0.015\ 0\ 0.109\ 1\ -0.038\ 5\ -0.412\ 7\ 0.259\ 1]$
λ_6	0.322 5	$c_6^0 = [0.577\ 6\ -0.062\ 4\ 0.052\ 2\ -0.581\ 5\ -0.051\ 9\ -0.090\ 9\ 0.320\ 0\ -0.044\ 6\ 0.160\ 7\ 0.162\ 6\ -0.011\ 4\ -0.166\ 2\ 0.355\ 4]$
λ_7	0.442 4	$c_7^0 = [-0.778\ 2\ -0.038\ 7\ 0.040\ 7\ -0.202\ 9\ -0.127\ 2\ -0.046\ 6\ 0.328\ 7\ -0.015\ 5\ 0.123\ 5\ 0.155\ 3\ -0.020\ 9\ -0.253\ 4\ 0.345\ 6]$
λ_8	0.482 9	$c_8^0 = [0.003\ 0\ -0.283\ 7\ 0.506\ 2\ 0.049\ 8\ -0.452\ 9\ -0.047\ 6\ -0.019\ 3\ -0.229\ 3\ -0.228\ 1\ -0.208\ 3\ 0.336\ 5\ 0.364\ 7\ 0.242\ 5]$
λ_9	0.640 5	$c_9^0 = [-0.001\ 2\ -0.035\ 3\ -0.301\ 2\ 0.045\ 6\ -0.063\ 6\ -0.555\ 0\ -0.103\ 6\ 0.642\ 7\ -0.099\ 2\ -0.004\ 5\ 0.171\ 2\ 0.215\ 8\ 0.290\ 6]$
λ_{10}	0.717 8	$c_{10}^0 = [-0.000\ 3\ -0.011\ 7\ 0.033\ 4\ 0.049\ 4\ 0.023\ 4\ -0.427\ 8\ -0.651\ 3\ -0.404\ 3\ 0.224\ 3\ 0.017\ 1\ -0.039\ 8\ -0.345\ 1\ 0.232\ 1]$
λ_{11}	0.972 3	$c_{11}^0 = [-0.004\ 0\ 0.484\ 2\ -0.566\ 7\ -0.034\ 5\ -0.134\ 4\ 0.219\ 4\ -0.010\ 2\ -0.365\ 1\ -0.095\ 6\ -0.124\ 2\ 0.264\ 1\ 0.209\ 2\ 0.324\ 6]$
λ_{12}	2.529 9	$c_{12}^0 = [-0.000\ 1\ -0.241\ 2\ -0.036\ 0\ -0.115\ 2\ -0.114\ 0\ 0.648\ 5\ -0.491\ 2\ 0.345\ 8\ -0.032\ 1\ 0.123\ 3\ -0.065\ 2\ -0.091\ 6\ 0.332\ 5]$
λ_{13}	6.283 7	$c_{13}^0 = [0.001\ 3\ -0.539\ 0\ -0.182\ 2\ 0.222\ 0\ 0.623\ 2\ 0.051\ 8\ 0.160\ 8\ -0.217\ 8\ 0.030\ 2\ -0.067\ 6\ 0.005\ 2\ 0.205\ 8\ 0.338\ 8]$

根据公式(7)得到计算表达式(8)~式(12)。

$$Z_j = c_j^0 * x_{d'p'}^T \quad j = (1, 2, 3, 4, 5) \quad (7)$$

根据不同的特征根算出不同特征根的方差累积贡献值,如表2所示。

表2 方差贡献率

成分	初始化特征根		
	特征根	方差贡献(%)	累积方差贡献(%)
1	6.284	48.336	48.336
2	2.530	19.460	67.797
3	0.972	7.479	75.276
4	0.718	5.522	80.789
5	0.640	4.927	85.724

$$Z_1 = 0.0013X_1 - 0.5390X_2 - 0.1822X_3 + 0.2220X_4 + 0.6232X_5 + 0.0518X_6 + 0.1608X_7 - 0.2178X_8 + 0.0320X_9 - 0.0676X_{10} + 0.0052X_{11} + 0.2058X_{12} + 0.3388X_{13} \quad (8)$$

$$Z_2 = -0.0001X_1 - 0.2412X_2 - 0.0360X_3 + 0.1152X_4 - 0.1140X_5 + 0.6485X_6 - 0.4912X_7 + 0.3458X_8 - 0.0321X_9 + 0.1233X_{10} - 0.0652X_{11} - 0.0916X_{12} + 0.3225X_{13} \quad (9)$$

$$Z_3 = -0.0040X_1 + 0.4842X_2 - 0.5667X_3 - 0.0345X_4 - 0.1344X_5 + 0.2194X_6 - 0.0102X_7 - 0.3651X_8 - 0.0956X_9 - 0.1242X_{10} + 0.2641X_{11} + 0.2092X_{12} + 0.3246X_{13} \quad (10)$$

$$Z_4 = -0.0003X_1 - 0.0117X_2 + 0.0334X_3 + 0.0494X_4 + 0.0234X_5 - 0.4278X_6 - 0.6513X_7 - 0.4043X_8 + 0.2234X_9 + 0.0171X_{10} - 0.0398X_{11} - 0.3451X_{12} + 0.2321X_{13} \quad (11)$$

$$Z_5 = -0.0012X_1 - 0.0353X_2 - 0.3012X_3 + 0.0456X_4 - 0.0636X_5 - 0.5550X_6 - 0.1036X_7 + 0.6427X_8 - 0.0992X_9 - 0.0045X_{10} + 0.1712X_{11} + 0.2158X_{12} + 0.2906X_{13} \quad (12)$$

5种评价指标的方差贡献率达到85.724%,所以使用前五位主成分综合得出评价函数权重 $Y_j = U * x_{d'p'}^T$, $U =$

$$\begin{bmatrix} u_{11} & \cdots & u_{1m} \\ \vdots & \ddots & \vdots \\ u_{n1} & \cdots & u_{nm} \end{bmatrix}, U = c_j^0 X_j^T, j = (1, 2, \dots, 5)。得出最终的球$$

员综合能力评价函数公式(13)~公式(17)。

$$Y_1 = 0.473X_1 + 0.321X_2 - 0.152X_3 - 0.044X_4 + 0.881X_5 + 0.846X_6 + 0.906X_7 - 0.070X_8 + 0.249X_9 + 0.793X_{10} + 0.278X_{11} + 0.672X_{12} + 0.373X_{13} \quad (13)$$

$$Y_2 = 0.744X_1 + 0.048X_2 + 0.211X_3 + 0.195X_4 + 0.054X_5 + 0.419X_6 + 0.328X_7 + 0.935X_8 + 0.775X_9 + 0.049X_{10} + 0.881X_{11} + 0.440X_{12} + 0.776X_{13} \quad (14)$$

$$Y_3 = 0.202X_1 + 0.011X_2 + 0.150X_3 + 0.955X_4 - 0.049X_5 + 0.066X_6 + 0.034X_7 + 0.033X_8 + 0.130X_9 - 0.138X_{10} + 0.023X_{11} + 0.171X_{12} + 0.170X_{13} \quad (15)$$

$$Y_4 = 0.241X_1 - 0.001X_2 + 0.943X_3 + 0.147X_4 - 0.202X_5 + 0.039X_6 - 0.035X_7 + 0.033X_8 + 0.063X_9 - 0.023X_{10} + 0.047X_{11} + 0.000X_{12} + 0.235X_{13} \quad (16)$$

$$Y_5 = 0.046X_1 + 0.943X_2 - 0.003X_3 + 0.009X_4 + 0.200X_5 + 0.051X_6 + 0.101X_7 + 0.012X_8 - 0.009X_9 + 0.151X_{10} + 0.038X_{11} + 0.124X_{12} + 0.084X_{13} \quad (17)$$

3.3 结果分析

由5个函数关系式可以看出: Y_1 中 X_5 前场篮板、 X_6 后场篮板、 X_7 总篮板、 X_{10} 盖帽、 X_{12} 犯规的方差贡献率高。可以确定 Y_1 是一个体现球员争抢篮板能力与防守篮筐能力的主成分,所以推断出该主成分计算出结果得分较高的球员,大多是中锋、大前锋位置球员。

Y_2 中 X_1 球员上场时间、 X_8 助攻、 X_9 抢断、 X_{11} 失误、 X_{13} 得分的方差贡献率高。所以推断出该主成分计算出结果得分较高的球员,大多是控球后卫、得分后卫、全能小前锋位置球员。

Y_3 中明显只有 X_4 罚球命中率的方差贡献率高,所以推断出该主成分计算出结果得分较高的球员是各个球队中罚球稳定的球员。这些球员在关键时刻、关键罚球得分时派上场,以确保稳定的罚球得分。

Y_4 中明显只有 X_3 三分球命中率的方差贡献率高,所以推断出该主成分计算出结果得分较高的球员,是各个球队中的稳定三分球投手。这些球员在球队比赛比分胶着时,起到迅速拉大比分作用,在球队落后较多时,起到稳定追分作用,在现代篮球中占有非常重要的地位。

Y_5 中明显的只有 X_2 投篮命中率的方差贡献率高,所以推断出该主成分计算出结果得分较高的球员是各个球队中的稳定得分手,多为各支球队的核心球员。这些球员是衡量球队水平的一大重要因素。

根据公式(13)~公式(17),计算出得分排序,如表3所示。

表3 各主成分得分排序结果

球员姓名	得分	球员姓名	得分	球员姓名	得分	球员姓名	得分	球员姓名	得分
安东尼-戴维斯	52.4217	威斯布鲁克	74.6469	威斯布鲁克	15.1123	威斯布鲁克	16.5335	唐斯	7.27879
唐斯	52.3083	詹姆斯-哈登	73.446	詹姆斯-哈登	14.9326	詹姆斯-哈登	16.5298	威斯布鲁克	7.46605
德马库斯-考辛斯	51.8747	勒布朗-詹姆斯	68.2916	勒布朗-詹姆斯	14.3006	勒布朗-詹姆斯	16.0173	安东尼-戴维斯	7.17504
哈桑-怀特塞德	51.0436	德马库斯-考辛斯	65.1776	德马库斯-考辛斯	14.1681	达米安-利拉德	15.7186	德马库斯-考辛斯	7.1611
威斯布鲁克	50.04	约翰-沃尔	64.1314	安东尼-戴维斯	14.1501	以赛亚-托马斯	15.6725	詹姆斯-哈登	6.9448
鲁迪-戈贝尔	48.6458	安东尼-戴维斯	62.9664	唐斯	13.696	德玛尔-德罗赞	15.4311	哈桑-怀特塞德	6.6962
D-乔丹	47.7962	唐斯	61.8427	达米安-利拉德	13.85	安东尼-戴维斯	15.3857	勒布朗-詹姆斯	6.84424
德拉蒙德	47.2045	达米安-利拉德	60.7761	德玛尔-德罗赞	13.6645	德马库斯-考辛斯	15.2034	鲁迪-戈贝尔	6.6276
詹姆斯-哈登	45.8255	安特托孔波	60.707	以赛亚-托马斯	13.5963	凯里-欧文	15.1135	安特托孔波	6.46612

结合上表,可以得出以下结论:

(1) Y_1 得分主要与前场篮板、后场篮板、总篮板、盖帽和犯规有强正相关,得分的高低表示球员的拼抢篮板能力强防守效率。根据上表中的数据可以看出:安东尼·戴维斯、卡尔·安东尼·唐斯、德马库斯·考辛斯、哈桑·怀特塞德、威斯布鲁克、戈贝尔、小乔丹、德拉蒙德和詹姆斯·哈登的得分在全联盟538名球员中排名前十,说明这些球员在拼抢篮板能力强、防守效率较高,这也与实际能力情况相符。安东尼·戴维斯、卡尔·安东尼·唐斯、德马库斯·考辛斯是当今NBA联盟中公认的内线基石球员,他们在场上的位置都为内线球员,可以扛起球队的后场篮板球保护、前场篮板球拼抢以及守护篮筐,不让对手轻易得分。

(2) Y_2 得分主要与助攻、抢断、失误以及得分贡献率有关,得分高低表示球员的带动球队攻防能力和个人得分能力。根据上表中的数据可以看出:威斯布鲁克、詹姆斯·哈登、勒布朗·詹姆斯、德马库斯·考辛斯、约翰·沃尔、安东尼·戴维斯、卡尔·安东尼·唐斯、达米安·利拉德和安托托孔波的得分在全联盟538名球员中排名前十,说明这些球员在带队攻防能力和个人得分能力出众,与实际情况相符。威斯布鲁克、詹姆斯·哈登、勒布朗·詹姆斯是当今NBA联盟中带队能力最好的三位球员,他们在场上的位置为后卫和小前锋,均为外线球员,能够在场上创造机会,精准把球传给队友,关键时刻还可以依靠超强的个人能力得分。这10个人中除了德马库斯·考辛斯、安东尼·戴维斯和卡尔·安东尼·唐斯是内线球员,其他均为外线球员,证明了 Y_2 得分较高的球员,大多是控球后卫、得分后卫、全能小前锋位置球员的推断。内线球员能上该排名的前十,说明他们不仅有强壮的身体、细腻的手感,还有后卫般控球能力和传球意识。这也是德马库斯·考辛斯、安东尼·戴维斯和卡尔·安东尼·唐斯成为NBA联盟中最好的内线基石球员的有利证据。

(3) Y_3 得分主要与罚球命中率的贡献率有关,得分高低表示球员的罚球稳定性和持球进攻杀伤的能力。根据上表中的数据可以看出:威斯布鲁克、詹姆斯·哈登、勒布朗·詹姆斯、德马库斯·考辛斯、安东尼·戴维斯、卡尔·安东尼·唐斯、达米安·利拉德、德马尔·德罗赞、以赛亚·托马斯和保罗·乔治的得分在全联盟538名球员中排名前十,说明这些球员的罚球稳定性和持球进攻造杀伤能力很强,与实际情况相符。这些球员都是球队中核心球员,比赛中是执行关键罚球和关键时刻持球个人进攻的最好人选。

(4) Y_4 得分主要与三分命中率的贡献率有关,得分高低表示球员的三分稳定性和迅速得分能力。根据表3中的数据可以看出:威斯布鲁克、詹姆斯·哈登、勒布朗·詹姆斯、达米安·利拉德、以赛亚·托马斯、德马尔·德罗赞、安东尼·戴维斯、德马库斯·考辛斯、凯里·欧文和约翰·沃尔的得分在全联盟538名球员中排名前十,说明这些球员三分命中率很稳定,除了德马尔·德罗赞(三分命中率26.6%)和安东尼·戴维斯(三分命中率29.9%)受

得分因素影响较大、排名进入前十以外,其他人均于实际情况相符。

(5) Y_5 得分主要与投篮命中率的贡献率有关,得分高低表示球员的投篮稳定性。根据表3中的数据可以看出:卡尔·安东尼·唐斯、威斯布鲁克、安东尼·戴维斯、德马库斯·考辛斯、詹姆斯·哈登、哈桑·怀特塞德、勒布朗·詹姆斯、鲁迪·戈贝尔、安托托孔波和的德安德烈·乔丹的得分在全联盟538名球员中排名前十,说明这些球员的投篮命中率稳定,是球队中的稳定得分手,与实际情况符合。而且这些球员除了威斯布鲁克、詹姆斯·哈登、勒布朗·詹姆斯,其他均为内线球员,实际比赛中内线球员的得分范围较接近篮筐,所以命中率较高,验证计算准确。

基于主成分分析结果,可以得出基于5个主成分的得分靠后的运动员应该对自己薄弱的环节进行补强。以达米安·利拉德和安托托孔波为例,达米安·利拉德从数据上和实际能力上都已经达到了全明星水准,可是这几年却频频入选不了全明星,这不排除球迷的主观喜爱程度和对其能力的评价标准,但是基于本文中的数据分析评价,达米安·利拉德在 Y_1 的得分排名是第29名,在 Y_2 的得分排名是第8名,在 Y_3 的得分排名是第7名,在 Y_4 的得分排名是第4名,在 Y_5 的得分排名是22名。很显然他需要提高的是 Y_1 和 Y_5 上体现出的能力,即提升篮板球拼抢能力和投篮命中率。但是实际情况是,达米安·利拉德是后卫位置球员,身材相对矮小,很难提高篮板球拼抢能力,而且将后卫和内线高大球员在抢篮板球的上作比较也是不合理的,所以他切实需要加强的是在场上减少犯规数和通过训练提高投篮命中率,这样才能有目的地训练自己,提升为全明星球员。安托托孔波拥有极强的身体天赋,身材高大但不缺灵活性,四肢奇长,有出色的运动能力,被业界评论是可以超越杜兰特的球员,但是要真正超越杜兰特,安托托孔波还需要提高一些方面的能力。安托托孔波在 Y_1 的得分排名是第11名,在 Y_2 的得分排名是第9名,在 Y_3 的得分排名是第15名,在 Y_4 上的排名是第22名,在 Y_5 的得分排名是第9名。可以明显得出安托托孔波需要提高在 Y_3 和 Y_4 上的能力,即三分球命中率和罚球命中率。在实际情况中,安托托孔波的罚球命中率为77%,三分球命中率为27.2%,预期比较的杜兰特罚球命中率为87.5%,三分球命中率为37.5%,两者差距较大。所以安托托孔波需要努力提升罚球命中率和三分球命中率,才能成为像杜兰特一样的超级得分手。

4 结语

在利用主成分分析法进行NBA球员综合能力评价的过程中,算法在应用中表现出一些优势:首先,可以在保证变量数据信息损失很小的情况下,实现从高维空间到低维空间的转化。在有538组原始数据的情况下,通过主成分分析法降维,使变量之间的分析更为容易。其次,主成分

(下转第197页)

3.3 安全管理体系

通过电子政务外网管理平台^[14]运作,对网络实时监控、实时故障报警。通过对链路状态、流量、网络及安全设备运行状态、日志信息、服务器 CPU 和内存存储空间的使用率判断整体网络的健康度。通过关联分析技术,在网络发生故障时可通知管理人员快速发现问题,并对故障区进行定位。电子政务外网管理平台很多策略通过部署组件的方式实现。

(1)设备认证管理组件。网络设备是政务外网建设的核心,也是信息安全的重要层面,为保证网络设备安全,需要验证登录到网络设备的用户身份,明确管理用户可访问的设备、可执行的命令和可完成的操作。

(2)用户接入管理组件。网络管理是对网络基础资源的管理,包括路由器、交换机、服务器等,而用户接入管理是对用户管理包括身份认证、用户信息的组织管理。

4 结语

本文以电子政务外网安全为研究对象,分析了电子政务外网存在的安全威胁,在电子政务组网安全的基础上,从技术安全和管理安全两个方面进行设计。在技术层面,利用现代化的网络安全技术构建一个安全的网络环境;在管理层面,运用最新的管理技术进行管理,做到技术和管理有效结合,提高电子政务外网的安全等级。

电子政务外网安全体系建设应注重两个方面:①不断进行技术创新,运用最新技术抵御新兴病毒;②使用 IPv6,统一 IP 地址规划,建立网络信任体系^[15]。电子政务网络安全体系建设不会一蹴而就,不仅在技术、管理层面,还要在人员素质培养、专业水平提高等方面不断探索,使

电子政务外网持续健康发展。

参考文献:

- [1] 朱梦琳. 电子政务外网安全问题研究[J]. 电子世界, 2017(4): 49-50.
- [2] 周民.“互联网+”时代国家电子政务外网的安全发展之路[J]. 信息安全研究, 2015, 1(2): 98-104.
- [3] 王建敏, 管征峰, 李挺, 等. 台州市电子政务外网安全和应用现状及对策[J]. 信息化建设, 2009(5): 33-35.
- [4] 关翔宇. OSPF 路由协议算法的研究与仿真[D]. 武汉: 华中科技大学, 2011.
- [5] 杨继良. 浅谈政务外网的安全机制和管理[J]. 网络与信息, 2007, 21(2): 80-81.
- [6] 盖慧玲. 电子政务外网安全建设中的问题与对策分析[J]. 网络安全技术与应用, 2016(9): 101-102.
- [7] 阎彩英. 山西省电子政务外网安全等级保护工作研究[J]. 中国信息界, 2012(5): 48-50.
- [8] 郭伟. 省级电网互联网信息安全关键技术研究与应用[J]. 电力信息与通信技术, 2012, 10(6): 82-86.
- [9] 储庄, 张全海, 李建华. 上海市电子政务外网安全保障体系研究与设计[J]. 信息网络安全, 2012(4): 1-3.
- [10] ZHANG Z. Research and design of E-government security architecture[J]. Network Security Technology & Application, 2010(11): 25-27.
- [11] BO D U, CUI Y, ZHANG H, et al. The basic technical architecture of network security of E-government[J]. Computer Security, 2013(9): 22-27.
- [12] 黎水林. 基于安全域的政务外网安全防护体系研究[J]. 信息网络安全, 2012(7): 3-5.
- [13] 蔡昌许, 蔡昌曙. 国家电子政务外网入侵防御系统应用实践研究[J]. 曲靖师范学院学报, 2014, 33(6): 32-35.
- [14] 郭红, 王勇, 吴亚非. 安全管理平台(SOC)在国家电子政务外网中的应用[J]. 电子政务, 2008(8): 78-85.
- [15] WANG Z, HAN W B, LIN Y, et al. Research on security assurance architecture of E-government[J]. Journal of Computer Applications, 2008(S1): 63-66.

(责任编辑: 杜能钢)

(上接第 189 页)

分析法选取正交化的综合评价指标作为主成分, 规避了球员综合能力评价指标间的不确定关系, 因为提取的综合变量之间相互正交, 所以消除了原始变量之间的多重相关性。综上所述, 主成分分析法在篮球运动员的综合能力评价中具有良好的适用性。

参考文献:

- [1] 吴翔峰. 14—16 年 NBA 优秀控球后卫与得分后卫攻防能力的对比分析[D]. 北京: 北京体育大学, 2017.
- [2] 王晓明. 2015—2016 赛季 NBA 季后赛各队攻、防能力比较研究[D]. 济南: 山东师范大学, 2017.
- [3] 刘国璧, 孙群, 赵姝. 基于 LS-SVM 的 NBA 前锋/中锋球员综合能力评价[J]. 西安文理学院学报: 自然科学版, 2012, 15(4): 10-14.
- [4] 郭兆辉. NBA 智能系统的启示[J]. 软件和信息化服务, 2012(7): 45.
- [5] 王彦霞, 齐秀英. 运用 TOPSIS 法和综合指数法综合评价医院医疗工作质量[J]. 中国病案, 2015, 16(10): 21-23.
- [6] 王平, 王云峰. 综合权重的灰色关联分析法在河流水质评价中的应用[J]. 水资源保护, 2013, 29(5): 52-54+64.
- [7] 何思长, 赵大仁, 孙渤星, 刘志会, 张瑞华. 基于 TOPSIS 法和 RSR

法模糊联合的医疗联合体实效评价[J]. 中国医院管理, 2016, 36(5): 7-9.

- [8] 顾洲一, 邱瑾. 基于熵权 TOPSIS 评价法的 NBA 教练执教水平评价模型研究[J]. 南京体育学院学报: 自然科学版, 2015, 14(5): 152-157.
- [9] 黄囊, 王玮. 基于改进灰色关联分析法的工业园区循环经济评价研究[J]. 软科学, 2016, 30(6): 130-133+144.
- [10] 董伦红. 我国男篮优秀大前锋若干特征分析与评价[J]. 广州体育学院学报, 2005(4): 76-78.
- [11] 苏为华. 我国多指标综合评价技术与应用研究的回顾与认识[J]. 统计研究, 2012, 29(8): 98-107.
- [12] 王湘玉, 董丽沙. 基于主成分分析和 Q 型聚类分析的 NBA 球队综合研究[J]. 九江学院学报: 自然科学版, 2017, 32(1): 120-122.
- [13] 刘静, 钱淦宇, 王广硕. 基于主成分分析的大学教练排名——以美国篮球教练的模糊综合评价为例[J]. 赤子(上中旬), 2015(16): 35-37.
- [14] 曹伟. 基于改进型 PSO-BP 神经网络算法的水环境质量评价[D]. 昆明: 昆明理工大学, 2016.
- [15] 陈刚强, 宋光辉, 李蕊. 顺序主成份分析——多元数据排序问题探讨[J]. 统计与决策, 2005(24): 27-29.

(责任编辑: 江 艳)