

Practical Teaching Research Based on the Evaluation of the Principal Component Analysis

Haima Yang^{1,*}, Yan Jin¹, Mingyu Ma¹, Jun Li², Dawei Zhang¹, Jin Liu³

¹School of Optics-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

²Sino-German College, University of Shanghai for Science and Technology, Shanghai 200093, China

³School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China

*Corresponding author. Email: snowyhm@sina.com

ABSTRACT

In this paper, the Principal Component Analysis method is introduced into the curriculum evaluation system of teachers for students. Taking Proteus measurement and control circuit simulation application course as an example, we use PCA method to describe the calculation method of multi index weight in curriculum evaluation, so as to extract the principal component of multi index. The results obtained improve the objectivity, simplicity and comprehensiveness of curriculum evaluation. The method described in this paper can be used as a reference for teachers of similar practical courses in methods and models, so as to obtain better evaluation effect of practical courses.

Keywords: Principal Component Analysis, Course evaluation, comprehensive evaluation, Proteus practice class.

基于主成分分析方法评价的实践教学研究

杨海马^{1,*}, 金焱¹, 马明宇¹, 李筠², 张大伟¹, 刘瑾³

¹上海理工大学光电学院, 上海, 中国

²上海理工大学中德学院, 上海, 中国

³上海工程技术大学电子电气工程学院, 上海, 中国

*通讯作者. 邮箱: snowyhm@sina.com

中文摘要

本文将主成分分析法引入教师对学生的课程评价体系中. 以Proteus测控电路仿真应用课程为例, 我们利用PCA的方法简述了在课程评价中多指标的配重计算方法, 从而提取多维指标的主成分. 由此获得的结果提高了课程评价的客观性, 凝练性与综合性. 本文所述的方法以供相似实践类课程任课教师在方法上和模型上进行参考, 获得更优的实践课评价效果.

关键词: 主成分分析法, 课程评价, 综合评价, Proteus 实践课

1. 引言

在传统的教学过程中, 任课教师通常凭借课堂提问、留书面作业、期末笔试等方式和方法来对学生的知识掌握情况做出评定. 传统的期末成绩评定方式以文字作答为主. 这样的方式容易引起学生心态上的焦虑从而影响到考卷上的呈现发挥. 对大部分的学生来讲, 往往难以从书面成绩中获得客观、全面的成绩评定^[1-2]. 同时, 由于信息技术的发展, 学生对于电子产品的熟悉度、依赖度逐渐提高, 若老师仍旧采用传统的教学方式, 生搬硬套地讲解理论知识, 学生的学习热情与求知欲将大大减低^[3]. 因此, 传统的教学过程需要创新、改革、多元化. 在互联网和大数据技术的快速发展背景下, 需要以现代信息技术推动教育创新、激活更高水平的教育发展^[4].

针对一门实践课程, 课程的内容通常包括了多学科的融合, 例如测控电路、单片机原理、嵌入式系统等. 学生们通常是以分组的形式, 合作完成实践课程的实验案例任务. 实践课程不仅检验了学生对于课堂知识的掌握程度, 而且锻炼了实际操作与编程设计的整体思路. 能够达到加深学生的知识体验、激励学生的学习兴趣等目的. 这一过程离不开教师综合多指标评价过程, 从而给予学生成绩上的认可. 培养学生进行小组形式的答辩评估, 重视评价学习的过程, 注重科学态度的培养、科学方法的训练, 养成善于发现问题、分析问题的习惯, 充分调动学生的学习积极性, 能够让每一位学生参与到成绩评估的全流程中去^[2,5].

本文主要研究的就是如何运用数学模型, 更加客观合理地评价学生的课程学习效果, 使得实践教学进入一个师生互动、效果良好的循环过程.

2. 教师和学生协同对于课程评价指标的分析

本文对于 Proteus 测控电路仿真应用这门实践课程, 罗列了可供参考的综合评定指标. 教师对学生的学习成绩评估可以有以下两个方面的评估指标:

(1) 首先, 是教师对于学生小组完成课程作业的主观评价. 其指标可以是: 测控电路的背景意义、研究现状是否全面了解与掌握以及对设计测量任务的分析是否恰当、对典型的电路模块功能是否清楚、对典型的电路设计的误差分析是否准确、PPT 内容是否详细、报告流畅、讲述条理清晰等一些其他的参考指标, 来作为实践课程的评估指标去评价学生的学习效果.

(2) 其次, 除了有老师的主观评价外, 还可以引入学生之间的评价. 其自评指标可以包括: 小组内工作是否平均分配、该小组成员是否配合交流讨论、是否具有团队精神、完成任务的及时程度等. 互评指标与教师的评价大致相似. 从学生的角度出发, 引入学生的视角, 可以给教师提供更准确、更客观的评价.

3. 主成分分析法

3.1. 方法原理

在实际的课程实践教学之中, 由于学生个体性格、特质的多样性, 对其学习效果的评价不应只从单一维度进行考量. 所以, 需要从多维度、多角度进行学生实践课课程的综合评判. 面对多维的数据, 教师需要评定不同的指标在综合成绩中的权重. 为了在降维数据的同时, 也获得更加合理的成绩结果, 引入了主成分分析法. 它是一种识别数据中模式的方法, 并以突出其相似性和不同性的方式表达数据^[6].

主成分分析法 (Principal Component Analysis), 简称 PCA, 它是一种多变量统计分析中非常高校的方法. 主成分分析方法得到的主成分可以在被观测对象为实现提供最佳的表示空间功能^[7]. 它的目的是要从 p 个指标出发, 综合样本数据的信息, 得到 m 个综合指标 ($m < p$), 然后利用这 m 个综合指标加权求和, 得到综合评价值^[8]. 主成分分析法可以完成数据降维, 不仅消除了各个指标之间相对严重的相关关系, 而且尽量保证了原始数据的完整度. 这种系统简化的方法体现了抓住事物主要矛盾的思维.

3.2. PCA 的数学原理

PCA 方法主要运用了方差、协方差、协方差矩阵等概念来选取“主成分”. 方差表示的是数据离散的程度, 即随机变量与均值之间的差值. 方差越大就表示数据间的离散程度越大, 即包含的信息越多. PCA 的目的就是要找到能够包含样本指标间差异信息最多的指标, 并把最多的指标作为“第一主成分”, 其次把方差较小的作为“第二主成分”, 以此类推, 在第二个主成分里面将不再有第一个主成分的信息. 协方差用于衡量两个变量的总体误差, 可以衡量数据的协变趋势. 当处理多维问题时, 要使用协方差矩阵来解决. PCA 算法求解的大致流程^[9]如图 1 所示.

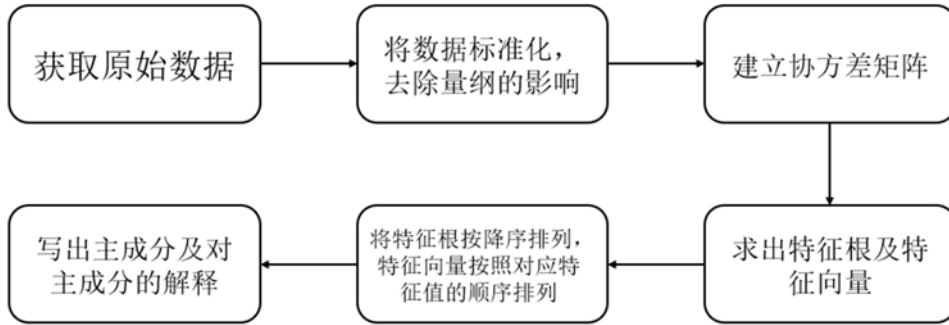


图 1 PCA 方法大致流程

PCA 算法的具体实施步骤^[10]如下:

第一, 我们需要获取一组原始数据. 其中, 样本集 X 是一个 $m \times n$ 的矩阵, m 表示样例个数, n 表示特征维度个数.

$$X = \begin{bmatrix} x_1^{(1)} & x_2^{(1)} & \dots & x_n^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \dots & x_n^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{(m)} & x_2^{(m)} & \dots & x_n^{(m)} \end{bmatrix}_{m \times n} = [X_1, X_2, \dots, X_n] \quad (1)$$

在消除数据量纲对实验结果的影响的过程中, 我们需要对数据进行标准化处理. 最终达到矩阵的每一列的均值为 0, 方差为 1 的效果. 事先需要明确的是, 如果两个数据的相关系数的绝对值大于或等于 0.8, 可以认为是两者高度相关. 如果变量之间不存在高度相关性, 是不能进行降维处理的^[11].

第二, 需要对变量建立协方差矩阵 C .

$$C = \begin{bmatrix} \text{var}(x_1) & \text{cov}(x_1, x_2) & \dots & \text{cov}(x_1, x_n) \\ \text{cov}(x_1, x_2) & \text{var}(x_2) & \dots & \text{cov}(x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(x_1, x_p) & \text{cov}(x_2, x_n) & \dots & \text{var}(x_p) \end{bmatrix} \quad (2)$$

3.3. PCA 的优点

(1) 减少各个评价指标之间的相关性影响

主成分分析法是通过原指标的变量进行正交变换后, 从而形成了相互独立的主成分, 而且通过实践发现如果各个指标之间相关程度越高, 主成分分析效果越好.

(2) 可减少指标选择的工作量

主成分分析完成的目标就是在数据降维的同时使丢失的信息量最小. PCA 本质上是把方差最大的维度作为最主要特征, 并且在各个正交方向上消除数据的相关性, 也就是让它们在不同正交方向上不相关. 所以说, 当评价指标较多时, 还可以在保留绝大部分信息的情况下, 用少数几个综合指标代替原指标进行分析^[12].

(3) 规范性与简洁性

由公式(2)得知, 根据协方差矩阵 C , 求出该矩阵的特征根并将其降序排列, 会得到 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$. 第三, 求出每个特征根所对应的特征向量, 并将其写为特征向量矩阵 U

$$U = [a_1, a_2, \dots, a_n] \quad a_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}, a_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix}, \dots, a_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{pmatrix} \quad (3)$$

第四, 我们需要选择提取特征向量矩阵的前 k 列. 在解决实际问题时, 一般不是取 n 个主成分, 而是根据累计贡献率的大小取前 k 个, 且特征值的大小应大于 1^[12].

第五, 就可以得出主成分 F_i :

$$F_i = a_{1i}X_1 + a_{2i}X_2 + \dots + a_{ki}X_k, i = 1, 2, \dots, k \quad (4)$$

PCA 方法的计算比较规范, 便于在计算机上实现, 可以在 MATLAB 上运用代码实现 PCA 分析. 还可利用专门的软件, 如 SPSS, 进行 PCA 方法的实现.

4. PCA 在实践课程当中的实际应用

通过如第二章中所说的 PCA 的数学原理, 我们使用 MATLAB 实现了 PCA 在课程评价中的实体应用. 我们整理了 20 位同学的打分信息. 其中, X_1 - X_{10} 是从老师和同学互评的两个角度制定的指标. X_1 为测控电路的背景意义、研究现状是否全面了解及掌握. X_2 为对设计测量任务的分析是否恰当. X_3 为对典型的电路模块功能是否清楚. X_4 为对典型的电路设计的误差分析是否准确. X_5 为 PPT 内容是否详细. X_6 为报告过程是否流畅、讲述是否条理清晰. X_7 为小组内工作是否平均分配. X_8 为该小组成员是否配合交流讨论. X_9 为该小组成员是否具有团队精神. X_{10} 为该小组成员是否及时完成了任务. 其中, X_1 - X_6 是从老师的角度设计的评价指标, X_7 - X_{10} 是从学生互评的角度设计的评价指标. 同学互评分数已做了组内平均处理.

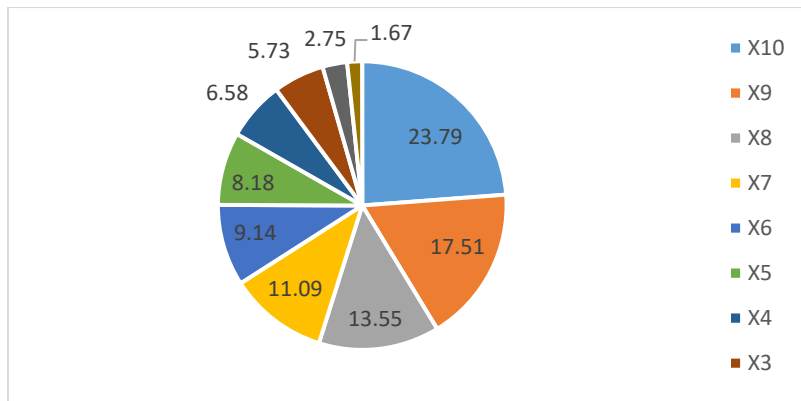


图 2 特征值及其贡献率、累计贡献率

计算得到这 10 个指标各自的特征值、贡献率(%)和累计贡献率(%)如图 2 所示. 我们取主成分保留率 $T=0.85$, 结果得到 X4, X5, X6, X7, X8, X9, X10 为排名前

七的指标, 累计贡献率为 89.85%. 将各指标主成分 $F_i (i=1, 2, \dots, 7)$ 的特征向量及贡献率整理, 得到表 1.

表 1 各指标主成分的特征向量及贡献率

主成分		F1	F2	F3	F4	F5	F6	F7
特征值		2.3793	1.7514	1.3551	1.1086	0.9144	0.8178	0.6580
贡献率%		23.79	17.51	13.55	11.09	9.14	8.18	6.58
累计贡献率%		23.79	41.31	54.86	65.94	75.09	83.27	89.85
特征向量	X1	0.4298	0.2461	0.2420	0.2447	-0.1613	0.2975	-0.076
	X2	0.0792	-0.2135	-0.6218	-0.3975	-0.3079	0.1223	-0.3279
	X3	0.1492	0.4465	-0.3016	-0.2485	0.3927	0.1375	0.5352
	X4	0.2796	0.3716	-0.4874	0.1552	0.1861	-0.2256	-0.2020
	X5	-0.0339	0.5215	0.0922	-0.0889	-0.4397	0.4750	-0.2223
	X6	-0.3584	0.3438	-0.0625	-0.0701	-0.3817	-0.5749	-0.0171
	X7	0.4068	0.1060	0.2677	-0.2989	-0.3553	-0.3972	0.3307
	X8	-0.4575	0.3819	0.0880	0.1817	0.2319	-0.057	-0.2121
	X9	0.4507	0.0226	0.0601	0.2635	0.1550	-0.3291	-0.4372
	X10	0.0317	0.0927	0.3618	-0.7012	0.3839	-0.0361	-0.4054

PCA 的分析方法将 10 个罗列的评价指标转换为 7 个独立综合指标, 这 7 个指标能够代表 10 个评价学生课程作业信息量的 89.85%. 其中, 特征值大于 1 的是 X10, X9, X8 和 X7, 这表明老师在对学生进行实践课程小组作业评价时, 需要更多地注意同学之间互评的结果. 这些指标也更能反映出, 以分组的形式完成实践课程作业时, 学生的合作精神、积极性更为重要. 使用 PCA 的分析方法结果也符合我们的预期结果, 鼓励了学生自身对于课题的探讨与表达能力, 促进了课程的正向循环.

这七个主成分分析如下: 第一主成分关键指标为 X1, X8 和 X9, 这三个指标表明了对学生的理解能力、合作精神至关重要, 评价贡献率为 23.79%; 第二主成分关键指标为 X3 和 X5, 这两个指标显示了学生表达能力的水平, 评价贡献率为 17.51%; 第三主成分关键指标是 X2 和 X4, 这两个指标反映了学生对任务认识及分析能力, 评价贡献率为 13.55%; 第四主成分的关键指标为 X10, 该指标指向学生完成任务的积极性与准时性, 评价贡献率为 11.09%. 由此可以推断不同主成分关键评价指标的作用效果, 对建立多维度的实践课程评价体系有指导意义.

但同时还是可以看出来, 仅单独 PCA 的方法选取得来的 7 个主成分中仍有一些重复、冗余之处, 可后续通过与其他方法的交叉结合使用, 如层次聚类分析法、熵权法等, 提高分析能力. 除此之外, 由 PCA 方法得来的七个指标, 对其含义的解释是人为的、主观的, 其含义比起原指标来说是有一些含糊不清的, 没有原来的指标解释那么清晰, 这是降维过程中不得不付出的代价.

5. 结束语

2000-2020 年期间, 我国高校教学表现出了对教学成绩评定改革的持续关注, 发文量呈现出逐年快速上升的态势; 在期刊载文分布上, 我国成绩评定教学研究文献主要集中分布在教育理论与教育管理和高等教育类的 20 余种核心期刊群; 在教学课程的领域上, 大部分集中在高等教育的话题上, 而对自动化技术的专科学研究仅占约 1%.

本文是对实践教学课程评定方法的研究, 本文通过阐述 PCA 的方法原理、数学原理及其独特的优势, 构建了以 Proteus 测控电路仿真应用为例的实践课程

评价中,师生共同参与的模型结构和评判过程.结果表明,基于主成分分析法评价的实践教学,提高了实践课程评价的客观性以及学生参与实践课堂的积极性.为有类似需求的实践类课程的教学提供了一种可供参考的评判组织模型和评价总结.

致谢

基金资助:上海理工大学教师发展项目(CFTD203016);上海理工大学虚拟仿真实验教学培育项目(2020GD06)

REFERENCES

- [1] Li X H, Huang G, Wang R R, Discussion on practical teaching reform of measurement and control circuit, in: J. Electronic Test, vol.22, 2018, pp. 130-131. DOI: 10.16520/j.cnki.1000-8519.2018.22.059
- [2] Xiong S R, Wu X D, Xiao X M, Zhang Q L, Study on the assessment system of scientific inquiry learning, in: J. Curriculum, Teaching Material and Method, vol.03,2006, pp. 82-86. DOI: 10.19877/j.cnki.kcjcf.2006.03.022
- [3] Zhou H Q, Zheng F L, Introduction to Computer under the New Situation Reform and Practice of Teaching, in: J. Practical Electronics, vol.02, 2020, pp. 56-57. DOI: 10.16589/j.cnki.cn11-3571/tn.2020.02.021
- [4] Yuan J H, Research Hotspots and Dynamics of Virtual Teaching in Colleges and Universities in China, in: J. Forum on Contemporary Education, vol.06,2019, pp. 80-88. DOI: 10.13694/j.cnki.ddjylt.2019.06.003w
- [5] Chen C F, Shen Z L, Research on the evaluation of education and teaching in the implementation of new curriculum, in: J. Forum on Contemporary Education, vol.06, 2005, pp. 13-16. DOI: 10.13694/j.cnki.ddjylt.2005
- [6] F. Yousefi, Z. Amoozandeh. A new model to predict the densities of nano fluids using statistical mechanics and artificial intelligent plus principal component analysis[J]. Chinese Journal of Chemical Engineering, vol.25(9), 2017, pp. 1273-1281. DOI: <https://doi.org/10.1016/j.cjche.2016.10.016>
- [7] Yonathan AFLALO, Ron KIMMEL. Regularized Principal Component Analysis, in: J. Chinese Annals of Mathematics, Series B, vol. 38(1), 2017, pp. 1-12. DOI: 10.1007/s11401-016-1061-6
- [8] Zhou S L, Tong Y F, Application in Evaluation on Student Score of Industrial Engineering Based on Principal Component Analysis, in: J. Theory and Practice of Contemporary Education, vol.10(03), 2018, pp. 79-84. DOI : 10.13582/j.cnki.1674-5884.2018.03.016
- [9] Lin H M, Du Z F, Some Problems in Comprehensive Evaluation in the Principal Component Analysis, in: J. Statistical Research, vol.30(08), 2013, pp. 25-31. DOI: 10.19343/j.cnki.11-1302/c.2013.08.004
- [10] Li Z L, Zhang Z Q, A DOI: 10.13582/j.cnki.1674-5884.2018.03.016, in: J. Science & Technology Vision, vol.18, 2018, pp. 107-109. DOI : 10.13582/j.cnki.1674-5884.2018.03.016
- [11] Wang X M, Problems in the application of principal component analysis and factor analysis, in: J. Statistics & Decision, vol.08, 2009, pp. 140-141. DOI: 10.3969/j.issn.1002-6487.2007.11.071
- [12] Wu D T, Wu D, The problems should be cared when using principal component analysis to make multi index comprehensive evaluation, in: J. Mathematics in Practice and Theory, vol.45(20), 2015, pp. 143-150.