Journal of Technical Science and Technologies; ISSN 2298-0032; Volume 3, Issue 1, 2014

Elaborating Admissibility Level of Implicit Dimensions in a Course via Singular Value Decomposition Based Principal Components

Mehtap ERGÜVEN *
Cabir ERGÜVEN **

Abstract

Teaching is the core of education and educational institutions focus to enhance the quality of teaching. There are various methods to control and enhance the quality. Our proposal is that, if teaching and assessment are multidimensional and integrated in a course; such educational process will increase the level and quality of education. In formal education individuals come with their abilities which are mostly based on pre-gained knowledge. Prerequisite courses are fundamental and they perform background of individuals and influence their future success as well. According to various studies there exists strong relation among new knowledge and pre-existing knowledge. In the research since examinations are used as indicator of gained knowledge, relation among previously implemented courses and actual courses is elaborated with respect to corresponding examinations. Using those relations, dimensionalities of an actual examination is determined for the quality interpretation of the actual course. Contribution of each of the prerequisite examinations is determined as impact of internal (implicit) dimensionality. On the strength of normally distributed grades of several congener courses, first pattern model is designated and it is compared with actual model. Research is realized based on different three cases. Purpose of the study is to compare actual model and expected pattern model, to describe the convenient acceptance level of the actual model for each of dimensions, and to determine the admissibility of the total impact of the internal dimensions of actual exam. In the article, quality of education process is interpreted with respect to the dimensionality of the course. Results of the analyses can be used as hints for the quality control in the education process.

Keywords: Multivariate data analysis, singular value decomposition, principal components, uncorrelated prerequisite courses, quality control, regression analysis

Introduction

(Mishra, 2007) implies that higher education institutions have three sub-systems, those are input, transformation and output. Mainly, inputs to the system are defined as students and teachers, a part of the transformation sub-system consists of the educational process and activities with respect to the curriculum, and employable graduates are outputs of the system. It is well known that, teaching is the core of the education process and educational institutions focus to enhance the quality of teaching. Efficient management of teaching-learning conditions increases the quality of education.

Effective teaching combines the principals of good classroom management, organization, effective planning, and the teacher's personal characteristics (Stronge, 2013). If a teacher follows mentioned essences, and has adequate knowledge and skills, he/ she would be able to provide required, accomplished teaching-learning in the education system. Beside those principals, proficient educator should follow the international standards on the assessment process too. Effective teacher gives various assessments on a regular basis. If a teacher is good at explaining the necessary materials and has a strong knowledge in a particular subject, based on this subject, examinations and assessments should reflect truly the performance of a students and teacher as well.

Knowledge is a continuously moving concept as once it is mastered it becomes a catalyst for further knowledge discovery (Hampson-Jones, 2011). As lecturers, we are responsible for providing high level teaching continuum by well-prepared curriculum plans, sufficient and convenient sources, and well-organized assessment methods to achieve accomplished program outcomes in support of students' overall educational aims. Because assessment is among vital components of effectiveness and adequateness of teaching and learning, institutions should check and support the quality of assessment.

Higher education institutions use student questionnaires to evaluate teaching quality during the education process. While students' opinions are important and should be included in the assessment of quality, it is obvious that, quality of teaching should be explained by different tools based on the meaningful reasons. If a group of students shows low performance in "well-organized" teaching and assess-

*PhD., Faculty of Computer Technologies and Engineering, International Black Sea University, Tbilisi, Georgia
E-mail: merguven@ibsu.edu.ge

**Assoc. Prof. Dr. Faculty of Computer Technologies and Engineering, International Black Sea University, Tbilisi, Georgia
E-mail: cabir.erguven@ibsu.edu.ge

ment process, their performance are expected to be similar in other assessments during the semester. If ability of them increased significantly in a new (actual) course, that situation should be questioned.

On the light of above mentioned bases, from the pedagogical point of view, there should be various reasons of that result. If an instruction was not well-organized, transferred knowledge given by instructor was not adequate and the questions in examination were easy, the adverted result can be considered as unavoidable. When we think about opposite situation it would be similar reasons to explain the situation. In both cases real ability level of pupils cannot be represented truly. Therefore, assessment tools, examinations should be controlled and enhanced to provide adequate, eligible and qualitative education.

Prior knowledge and background knowledge are themselves parent terms for many more specific knowledge dimensions and prior knowledge is associated with beneficial academic behaviors and higher academic performance (Strangman, N.; Hall, T., 2004). Idea is that an actual examination can be conceived as reflection and indicator of "a part of" prior knowledge. If actual examination's questions are solvable based on the pre-gained abilities and new abilities, this case should be the gauge of adequate transformation of new knowledge and satisfactory combination of pre-gained and ongoing acquired knowledge. Such kind of exam is multidimensional. One dimension is current course based actual examination. Other dimensions are prerequisite courses based prerequisite examinations. According to experts' opinions, a pattern model is simulated for the dimensionality and related quality analysis. Elaboration of the required pattern model, observed actual model and their comparison are done. Convenient acceptance level of the actual model for each of dimensions is described and admissibility level of the total impact of the internal dimensions of actual exam is determined.

Methodology

In the article, an empirical study is done based on the multivariate data analysis techniques and theory of conditional expectation. MatLab and Microsoft Excel Software are used for the required statistical analyses. Uncorrelated data generation is satisfied for 200 students. Generation of the grades of prerequisites is done by standard logistic distribution function within random number generator for higher accuracy to the normal distribution. Considering upper and lower groups, normally distributed grades are generated based on control parameters (F_{up}, g_{up}, F_{low}, g_{low}). Normality of those prerequisites is presented with their histograms. Singular value decomposition methodology is used to define principal component analysis' objects which can be defined as eigenvalues, transformation matrices and principal components. After that, principal components based multidimensional regression analysis is applied to describe coefficients of regression. Closeness and relations between actual model and pattern model are identified with respect to test of significance.

Uncorrelated Prerequisites

Research is based on previously implemented courses and their corresponding examinations which are named here as prerequisite examinations or only prerequisites. Relation between prerequisites and current course is elaborated. For that reason, pattern model is simulated with regard to experts' opinions.

Multicollinear, low correlated and uncorrelated types of data are generated to search in detail answers of the following questions:

- How can the acceptance level of an actual exam and corresponding course be determined on the strength of prerequisites with comparison of each of the estimated and desirable contribution of internal dimensions?
- What is the appropriate total impact of the contributions of internal dimensions between actual and pattern model?

Uncorrelated case is taken into consideration in the article. In the following table (see table 1), changes of the correlation coefficients and corresponding regression coefficients are represented on the strength of uncorrelated grades of students. In normal case, expected coefficient of correlation is 0.73 and 0.59 with respect to two prerequisite examinations. Beside this, corresponding regression coefficients are found as 0.44 and 0.53. If the grades significantly increase or decrease in the actual examination, changes among previous coefficients and new coefficients become obvious.

Table 1. Representation of changes in coefficients of correlation and regression with respect to three different cases.

	correlation coefficients	*regression coefficients*
normal case	0.73	0.44
	0.59	0.53
increased grades	0.10	0.06
	0.06	0.09
decreased grades	0.09	0.01
	0.01	0.05

Table 2a. Regression analysis results with respect to normal case.

normal case	
Regression Statistics	
Multiple R	0.96645
R Square	0.93402
Adjusted R Square	0.93335
Standard Error	0.90412
Observations	200

The ratio R2 (R-square) is called the coefficient of multiple determination, or more commonly the squared multiple correlation (Rencher, A. C., 2002). According to regression analysis results, R square is approximately 0.93. It can be interpreted that, independent predictor variables "prerequisite examinations" allow explaining 93% of all actual examination. However, if grades of actual examination increase or decrease explicitly, R-square changes significantly as well. In the table 2b, those results are represented. R-square decreases into 0.013 and multiple R (0.114) depicts low correlation among explanatory variables and dependent actual exam grades in the case of increment of grades (see table 2b).

Table 2b. Regression analysis results with respect to normal case.

Increased Grades			*Decreased Grades*	
Regression Statistics			*Regression Statistics*	
Multiple R	0.1142761		Multiple R	0.09472184
R Square	0.013059		R Square	0.00897223
Adjusted R Square	0.0030393		Adjusted R Square	-0.00108897
Standard Error	4.4811264		Standard Error	2.92879991
Observations	200		Observations	200

In contrast to previous case, if the grades decrease, Multiple R decreases to 0.095 and R-square to 0.00897. Only those base explanations are enough to think about the quality of a course when prerequisite examinations are taken into consideration. Obviously, there is significant abnormality in the coefficients of regression and coefficients of correlation in the mentioned experimental study.

On the light of given assumptions, a research is done through multivariate data analysis (Anderson, T., 1958) to compare adequateness of relation among prerequisites and currently implemented course to depict the dimensionality of that course and to give idea about quality of actual course with respect to that dimensionality. Thus, uncorrelated examinations are represented below with scatter plot.

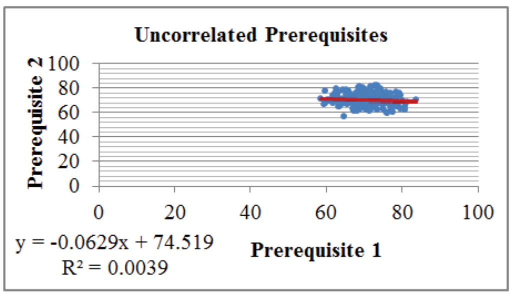

Figure 1. *Scatter plot of uncorrelated two prerequisites*

Correlation among those generated prerequisites is -0.06. It is clear that, correlation among two prerequisites is almost zero (see y=-0.0629x+74.519).

Average of the grades of the prerequisite 1 is 70.8 and average value of the grades of the prerequisite 2 is 70.1. Standard deviation of the grades of prerequisite 1 is 5.01 and standard deviation is 5.04 for the second prerequisite's grades. Normality of those prerequisites is represented below with their histograms:

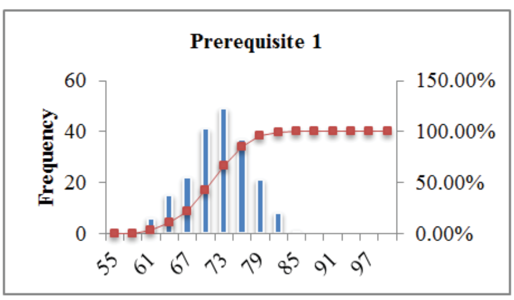

Figure 2. *Gaussian distribution of grades of prerequisite 1*

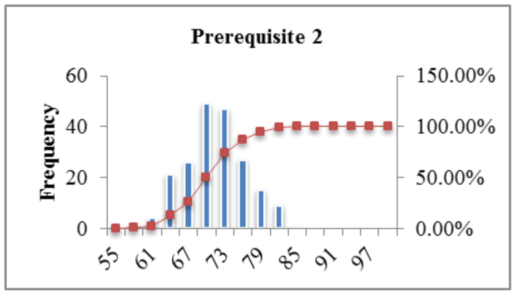

Figure 3. *Gaussian distribution of grades of prerequisite 2, by histogram*

PCA is the eigen-vector and eigenvalue based multivariate analysis and principal components (Smith, L., 2002) based regression is done in the research. Correlation among principal components is 0.0048, almost zero! Our purpose is detecting effectiveness and convenience of components based method in uncorrelated data.

The singular value decomposition (SVD) (Hardle, W.; Simar, L., 2003) is one of the most important tools in linear algebra (Ipsen, 2009). Using SVD methodology, matrices of right and left singular vectors are calculated and determined as V and U. The square roots of the eigenvalues of XX^T are singular values of X (Grodner & Grove, 2007). This connection helps to use PCA and SVD for the purpose of the study within convention. To obtain principal components of independent variables, prerequisites, dot product of U and matrix of singular values L is used (P=U*L). It is possible to use X*V to find same principal components. Uncorrelated principal components are represented below.

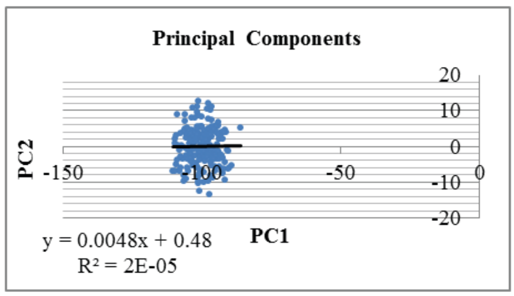

Figure 4. *Uncorrelated principal components*

Diagonal elements of matrix of singular values L are 1410.40296318200 and 73.1705854815163.

Using the fundamental assumptions of conditional expectation (Anderson, T., 1958), contributions of internal dimensions are determined. $\Sigma_{12}\,\Sigma_{22}^{-1}$ gives regression coefficients of principal components and Inner Product Matrix (IPM) of the system of vector of actual exam grades and principal components are used to find Σ_{12}.

$$\Sigma_{12} = [-633103.934024348 \;\; -160.717131424402]$$

Dot product of L and transpose of L gives Σ_{22} which is suggested below:

$$\Sigma_{22} = \begin{bmatrix} 1989236.51855258 & 0 \\ 0 & 5353.93457970789 \end{bmatrix}$$

Therefore, beta coefficients ($\Sigma_{12}\,\Sigma_{22}^{-1}$) are:

$$\beta = [-0.318264785569598 - 0.0300185086372816]$$

When right singular vectors matrix V is

$$V = \begin{bmatrix} -0.710786350358517 & 0.703407964231299 \\ 0.703407964231299 & -0.710786350358517 \end{bmatrix}$$

Estimated contributions of internal dimensions are found as:

$$\alpha = \begin{bmatrix} 0.205 \\ 0.245 \end{bmatrix}$$

Closeness comparison of estimated coefficients and desirable contributions of internal dimensions is done with MatLab among actual model and pattern model. It is represented that, difference between estimated coefficients in theoretical implementation and in MatLab code is approximately 0.005. Such a small difference can be considered negligible. Hence, usage of the code is offered for simplicity and saving from time.

To estimate coefficients of regression following part of MatLab code is implemented:

```
clear
X=xlsread('Examples.xlsx','TST1','C5:D204');% input of prerequisites
Y=xlsread('Examples.xlsx','TST1','E5:E204');% input of an exam under interest
ALFA=xlsread('Examples.xlsx','TST1','F5:F6');% input of expected alphas
[U, L,V]=svd(X);
S12=Y'*U*L;
S22=L'*L;
Beta=S12*inv(S22);
ALET=V*Beta';
```

Case I: Desirable Closeness of Actual and Pattern Model with respect to Uncorrelated Prerequisites.

Estimated coefficients are found as follows with MatLab code:

$$\alpha = \begin{bmatrix} 0.1998 \\ 0.2400 \end{bmatrix}$$

Desirable level of those contributions should be determined with respect to experts' opinions and quality assurance service. In the case study, desirable contributions of internal dimensions of pattern model are described as 0.15 and 0.20. We need to evaluate covariance matrix of the vectors of two prerequisites for the closeness comparison of pattern model coefficients and actual model coefficients which is given below:

$$\Sigma_X = \begin{bmatrix} 25.1884880816866 & -1.58379075719854 \\ -1.58379075719854 & 25.4497617211255 \end{bmatrix}$$

Using that covariance matrix, mean squared error is 0.1655, and both expected and estimated coefficients t-values of deviation of implicit contributions of pattern model from estimated ones are found as: 0.6145 and 0.4962. Those values are statistically not significant. This situation requires acceptance of the null hypothesis which is defined as "estimated contribution of inner dimensions of actual course is equal to the pattern model's desirable inner dimensions". Hence, contributions of both internal dimensions are acceptable.

Besides, total impact of estimated contributions of actual model is approximately defined as 44%. Total impact was expected as 35% in the pattern model. T-value of deviation of sum of pattern model's inner contributions from estimated one is -0.7856. That is the indicator of acceptable total impact of actual model. The actual examination has satisfactory relation with pattern model's prerequisites. It can be considered that, examination of current course is three dimensional. That should be result of convenient implementation of knowledge transformation process with regard to complementary usage of pre-learned knowledge and recently gained knowledge and abilities.

Based on the results of regression on principal components, corresponding p-values are 8.6491e-118 and 1.8672e-007 for the first and second regression coefficients of PCs. Both of p-values are less than 0.05. Thus, they are statistically significant. It means that, those coefficients are not zero and they are acceptable. Concurrently, t-values with regard to principal components are -52.5081 and -5.4043. These t-values show the same result like p-values.

Case II: In the Case of Significant Increment of Grades: Undesirable Model

In reality it is possible to encounter with various situations. Here new possibility is examined. In the pattern model, desirable coefficients of regression are determined as $\alpha_1^0 = 0.20$ and $\alpha_2^0 = 0.20$ respectively. Estimated regression coefficients of actual model are found as α_1 =0.0598 and α_2=0.0800. Correlation among prerequisite 1 and actual examination is 0.057187354 where correlation among prerequisite 2 and actual examination is 0.095166416. Both coefficients of correlation are very low. It is obvious that, while contribution of internal dimension decreases, correlation among prerequisites and actual examination decreases as well. In fact, those correlations were considered higher on the strength of desirable model. A part of MatLab code results are represented below with regard to highly increased grades:

```
Principal Components Coefficients
  -0.0988   -0.0148
Estimated Contributions of Inner Dimensions
   0.0598    0.0800
Pattern Contributions of Inner Dimensions
   0.2000    0.2000
t-values of deviation of pattern inner contribution from
estimated ones
  -4.2505   -3.6567
t-value of deviation of sum of pattern inner contribution
from estimated one
   5.5923
```

Comparison of estimated regression coefficients and pattern model coefficients suggested that, those regression coefficients are not close to each other since both t-values (-4.2505 and -3.6567) exceeds critical t-value. Deviation of sum of pattern model's internal contributions from sum of estimated ones allows finding t-value with regard to total impact. The t-value (5.5923) is the indicator of unacceptable total impact. In fact, total impact of the estimated contributions of implicit dimensions is very low (almost 14%). However, desirable total impact was considered as 40%. Undesirable contribution of both internal dimensions and unacceptable total impact are observed. Abnormality of situation in actual examination is obvious. Examination is mostly focused on one dimension actual examination itself. There is not found satisfactory relation among pre-courses and current course.

Case III: Undesirable Contribution of Both Internal Dimensions and Unacceptable Total Impact with regard to Significant Decrement of Grades

If the grades are very low, related MatLab code outputs are given as below.

Analysis of Results of Regression on Principal Components

Table_of_Reg_Coeff =

Coef	StdErr	tStat	pVal
0.74003	0.59041	1.2534	0.21154
-0.035122	0.0059201	-5.9327	1.3194e-008
-0.02154	0.0055604	-3.8739	0.00014575

Table_of_ANOVA

Source	df	SS	MS	F	P
Regr	2.00	8.34	4.17	25.21	0.00
Resid	197.00	32.61	0.17		
Total	199.00	40.95			

Principal Components Coefficients
-0.0351 -0.0215
Estimated Contributions of Inner Dimensions
0.0098 0.0400
Pattern Contributions of Inner Dimensions
0.2000 0.2000
t- values of deviation of pattern inner contribution from estimated ones
-5.7664 -4.8758
t-value of deviation of sum of pattern inner contribution from estimated one
7.5268
t-value of t-criterion for 0.95 confidence level
and for degrees of freedom197 is 1.6526.

Estimated contributions of internal dimensions are 0.0098 and 0.0400. Estimated regression coefficients and desirable coefficients are compared with regard to t-test. T-values are -5.7664 and -4.8758. They exceed t-critical. Hence, they represent unsatisfactory relation among pattern and actual model. Similar to previous case, both estimated coefficients are not acceptable in new implementation.

Correspondingly, total impact of estimated coefficients is almost 5% and based on the covariance matrix of prerequisites, mean squared error (0.17); t-value is found as 7.5268 which is not admissible. Last two examples are problematic. The situation and main reasons should be analyzed in detail. Otherwise, the real ability level of pupils cannot be represented truly.

If we compare with the first example, while the impact of the estimated contributions of internal dimensions decreases, consequently correlation among prerequisites and actual examination decreases as well. According to prerequisite 1 and prerequisite 2, the correlations are 0.09 and 0.01 respectively "very low".

Conclusion

Roschelle, J. (1995) implies that, prior knowledge has diverse and pervasive effects on the learning. A large body of findings shows that, learning proceeds primarily from prior knowledge, and secondarily from the presented materials. Neglect of prior knowledge can result in the audience learning something opposed to the educator's intentions.

According to Strangman, N., and Hall, T. (2004) prior knowledge promotes better learning and higher performance as a conclusion of various researches. Because new knowledge is constructed on the pre-gained abilities, providing adequate transformation of prior knowledge and connecting prior information and actual materials are very important for the higher level education and better quality. Therefore, relation among prerequisite courses and current course should be analyzed on the base of those purposes.

Based on the reasonable evidences, dimensions of a current course are used to identify and interpret quality of a course with respect to its previously implemented prerequisite courses. The purpose of the study was giving idea about how to identify dimensionality and correspondingly how to interpret quality of an examination with respect to that dimensionality. Relation between prerequisites and current (actual) course is depicted for uncorrelated data set.

Three different cases are examined. In the first case, desirable closeness of the actual and pattern model is provided with respect to the prerequisites, which can be indicated as the suitable utilization of broad perspective of the background knowledge. In the second and third implementations, undesirable actual model is faced. Those implementations are developing a methodology to determine the appropriate relation among prerequisites and parallel actual examination. Besides, they allow using prerequisites as estimator of dimensionality. Similar implementation is done in low correlated and multicollinear data as well. Eventually, approximate results are obtained in normally distributed three types of data. Results of the analyses can be used as hints for the quality determination in the education process.

References

Anderson, T. (1958). A*n Introduction to Multivariate Statistical Analysis*. Jhon Wiley & Sons ,Inc.

Hampson-Jones, N. (2011, December). Higher Education and Standardization: Knowledge Management Between Generations.

Hardle, W.; Simar, L. (2003). Elementary Properties of the Multinormal. In W. Hardle, & L. Simar, *Applied Multivariate Statistical Analysis* (p. 157). Germany: Springer-Verlag Berling Heidelberg.

Ipsen, I. C. (2009). Singular Value Decomposition. In I. C. Ipsen, Numerical Matrix Analysis: Linear Systems and Least Squares (p. 77). USA: Siam.

Mishra, S. (2007). Quality Assurance in Higher Education. In S. Mishra. Karnataka, India: The Director National Assessment and Accreditation Counsil (NAAC).

Rencher, A. C. (2002). Methods of Multivariate Analysis. In A. Rencher, Methods of Multivariate Analysis. USA: John Wiley & Sons Inc.

Roschelle, J. (1995). Learning in Interactive Environments: Prior Knowledge and New Experience. Dartmouth: Public Institutions for Personal Learning: Establishing a Research Agenda, American Association of Museums.

Smith, L. (2002, February 26). A tutorial on Principal Components Analysis Retrieved 10 1, 2013, from http://math.arizona.edu/~hzhang/waeso/pcrTutorial.pdf

Stronge, J. (2013, December). *Qualities of Effective Teachers*. Retrieved from http://www.mpsaz.org

Strangman, N., & Hall, T. (2004). *Background Knowledge Curriculum Enhancement Report*. National Center on Accessing the General Curriculum.