Teaching Factor Analysis in Terms of Variable Space and Subject Space Using Multimedia Visualization

Ho Yu Chong, Andrews Sandra, Winograd David, Jannasch-Pennell Angel & A. DiGangi Samuel

Published online: 01 Dec 2017.

Submit your article to this journal ☑

Article views: 166

View related articles ☑

Teaching Factor Analysis in Terms of Variable Space and Subject Space Using Multimedia Visualization

Chong Ho Yu
Arizona State University

Sandra Andrews
Arizona State University

David Winograd
State University of Potsdam

Angel Jannasch-Pennell
Arizona State University

Samuel A. DiGangi
Arizona State University

Key Words: Biplot; Eigenvector; Hypermedia; Vector space.

Abstract

There are many common misconceptions regarding factor analysis. For example, students do not know that vectors representing latent factors rotate in subject space, rather than in variable space. Consequently, eigenvectors are misunderstood as regression lines, and data points representing variables are misperceived as data points depicting observations. The topic of subject space is omitted by many statistics textbooks, and indeed it is a very difficult concept to illustrate. An animated tutorial was developed in an attempt to alleviate this problem. Since the target audience is intermediate statistics students who are familiar with regression, regression in variable space is used as an analogy to lead learners into factor analysis in the subject space. At the end we apply the Gabriel biplot to combine the two spaces. Findings from a textbook review, a survey, and a "think aloud" protocol were taken into account during the program development and are discussed here.

1. Introduction

Teaching and learning factor analysis is challenging. As Pedhazur and Schelkin (1991) point out,

> The literature of factor analysis (FA) is vast and generally complex. Perusing even small segments of this literature in an effort to understand what FA is, how it is applied, and how the results are interpreted is bound to bewilder and frustrate most readers. This is due to a wide variety of contrasting and contradictory views on almost every aspect of FA, serious misconceptions about it, and lack of uniformity in terminology and notation. (p. 590)

To trace the sources of misconceptions, a review of textbooks and Web sites dedicated to teaching factor analysis was conducted. Textbooks and Web sites were identified through a review of books in print, Web-based search engines (Google, Alta Vista, Infoseek, and Yahoo), and discussions with faculty teaching quantitative methods courses as well as members of the Educational Statistics Listserv group (EDSTAT-L).

An assessment examining concepts of factor analysis was constructed and administered to students at differing levels of statistical literacy. The results of these investigations formed the basis for construction and implementation of a computer-based multimedia instructional program, centering on the perspective of "subject space." The impact of this instructional program was evaluated using "think aloud" protocol (Someren, Barnard, and Sandberg 1994), a method of recording subjects' mental processes by having them verbalize their thinking as they navigate through the program. Based upon the findings, a multimedia program was developed to counteract those misceptions. The target audience for this hypermedia program is graduate students in all disciplines who have learned the basic concepts of regression. Students will be led from regression in variable space to factor analysis in subject space.

2. Review of Approaches of Teaching Factor Analysis

Strategies of teaching factor analysis can be classified into the conceptual, mathematical, and geometric approaches. These three approaches are used in varying combinations by the following authors.

2.1. Conceptual Approach

Examples of the conceptual approach can be found in Ingram (1998), Thapalia (1998), and Wulder (1998). Although the purpose and the application of factor analysis are emphasized in this approach, questions regarding the underlying dimensions of the data and their relationship to the Cronbach alpha coefficient are not mentioned. Researchers commonly claim that they have extracted several subscales by factor analysis and that all subscales are strongly correlated to the total scale. This claim is invalid when all subscales are correlated with each other, and in fact there is only one dimension of the data. Factor analysis is usually placed in textbooks under multivariate analysis; it is assumed that students understand how multivariate techniques are used to address the multi-dimensionality of the data. This assumption may not be applicable to all students, as Huberty (1994) saw; he therefore laid emphasis upon the fact that multivariate analysis techniques are analyses of data vectors for each individual observation under study that consists of two or more scores.

In the conceptual approach, practical uses of factor analysis are emphasized. Technical terms such as "eigenvalue" and "orthogonality" are omitted, or mentioned only in passing. Serious misconceptions may arise when explanations of these terms are omitted. For example, Tabachnick and Fidell (1996), Thapalia (1998), and Wulder (1998) state that the researcher can "rotate" factors to gain a better

interpretation of the data, possibly leading students to impose their own non-technical understanding on what appear to be words with everyday meanings. For example, vectors representing factors are rotated in subject space. Learners with no understanding of vectors and subject space may assume that this rotation implies spinning a plot to get a better perspective, or using different variables at different times as if they were tires to be rotated.

Even common terms such as "weight," "model," and "factor" that instructors assume will be understood by students may cause confusion. Students often confuse the meaning of the term "factor" in factorial analysis with that in factor analysis. In the former a factor is an observed variable with clearly identified levels while in the latter a factor is a latent and abstract mathematical construct. This major difference was not emphasized in the texts reviewed, which merely define a factor as a latent variable or a hypothetical construct (see, for example, Harman 1976).

Some authors include these more difficult terms rather than avoiding them (consider Ingram 1998). When technical terms are used to explain a common term such as "factor," students may be overwhelmed by what appears to be alien language. Ingram (1998) defined a factor as "a vector which is weighted proportionally to the amount of the total variance which it represents. The factor loadings are the elements in the factor vector. The sum of the squares of these loadings should equal the eigenvalue." Understandably, students reading this may have difficulty with these definitions as they attempt to relate "factor" to "vector," "total variance," "loadings," and "eigenvalue."

2.2. Mathematical Approach

In the mathematical approach, factor analysis is taught within the context of the linear model (Harman 1976; Kim and Mueller 1978; Joreskog and Sorbom 1979; Gorsuch 1983; Basilevsky 1994). One difficulty with this approach is that while both regression and factor analysis result in weighted linear combinations of variables, the differences in mathematical terms used for the two procedures fail to help the learner integrate these procedures under the umbrella concept of the linear model. In regression, the weight of the linear combination is called a "coefficient" or a "beta weight" while in factor analysis this weight is called a "loading." With the exception of Kim and Mueller (1978), the texts reviewed did not emphasize the relationships among the preceding terms, and students are unlikely to make the necessary connections themselves.

Eigenvectors and eigenvalues are concepts central to the topic of factor analysis. In some introductory statistics texts, an overly mathematical discussion of these terms may be confusing for students or even researchers who do not have a strong mathematics background. For instance, Hagle (1995, p. 89) contains the following explanation: "X is called an eigenvector (characteristic vector; eigen is German for characteristic) if there exists a nonzero vector X n*1 such that A n*n X n*1 = LX n*1. This scalar L is called an eigenvalue of A n*n." One would be hard pressed to find an intermediate student that could make much sense of this equation.

2.3. Geometric Approach

Several topics such as orthogonality are spatially oriented. A text-based explanation would define "orthogonal" as "uncorrelated," but the learner may have difficulties understanding this statement. On the other hand, a spatial representation of two perpendicular vectors is clear (Gorsuch 1983).

A number of reviewed texts (see, for example, Harman 1976; Comrey and Lee 1992; Basilevsky 1994; Wulder 1998) mentioned that factor analysis is sensitive to an ill-conditioned correlation matrix, which is a manifestation of multicollinearity. None of these texts utilize graphical representations to explain

conditioning and multicollinearity. In a simplistic sense, multicollinearity is the opposite of orthogonality. Perhaps the omission of graphical representation of the former is based upon the assumption that the student has learned the concept of orthogonality, but this is not necessarily the case. Multicollinearity is more comprehensible if orthogonality is understood.

The geometric approach relies upon the concept of subject space as a means of visualization of spatial relationships. Many textbooks using the geometric approach (Pedhazur and Schelkin 1991; Comrey and Lee 1992) begin with matrix algebra and then plot vectors in a coordinate system. In this context, it is difficult to convert the matrix algebra information to a representation in person space.

In addition, the only text reviewed explaining factor analysis in terms of variable space and vector space is *Applied Factor Analysis in the Natural Sciences* by Reyment and Joreskog (1993). No other textbook reviewed uses the terms "subject space" or "person space." Instead vectors are presented in "Euclidean space" (Joreskog and Sorbom 1979), "Cartesian coordinate space" (Gorsuch 1983), "factor space" (Comrey and Lee 1992; Reese and Lochmüller 1998), and "n-dimensional space" (Krus 1998). The first two phrases do not adequately distinguish vector space from variable space. A scatterplot representing variable space is also a Euclidean space or a Cartesian coordinate space. The third is tautological. Stating that factors are in factor space may be compared to stating that Americans are in America. The phrase does not provide additional information. "N-dimensional space" is closer to the meaning of subject space since in subject space the number of dimensions is equal to the number of subjects. On the other hand, the notation "n" could mean either the number of subjects or just any number.

3. Software Development

Three sources of information were used to plan and develop the animated tutorial. First, the strengths and weaknesses of the three approaches discussed above were taken into consideration. Second, in order to discover which aspects of factor analysis were most in need of elaboration, a survey was administered to a group of graduate students from various disciplines, who were already familiar with the concept of regression. The current multimedia program uses regression as a metaphor for factor analysis since the linear model subsumes both. Finally, a beta version of the tutorial developed with regard to the three approaches and to the information gained from the survey was given to a second group consisting of twelve students from various disciplines. These students were not exposed to the survey in order to avoid pre-conception of the subject matter. The "think aloud" protocol was used to capture information on the students' understanding of the instruction and thus functioned as formative evaluation. The instruction was modified where indicated.

3.1. Integration of the Three Approaches to Teaching Factor Analysis

In the conceptual approach, where practical uses of factor analysis are emphasized, misconceptions may arise when explanations of terms are omitted. In order to be useful to courses taking this approach, the current multimedia program begins with an explanation of basic terms such as "space" and "variance." This ensures that readers do not impose their own non-technical understanding onto statistical terminology.

In the mathematical approach, factor analysis is taught within the context of the linear model. An overly mathematical discussion of such concepts as the difference between linear regression and factor analysis, or of eigenvectors and eigenvalues, may confuse intermediate students. To remedy the first weakness, the multimedia program under discussion uses regression as a metaphor for factor analysis, since the linear model subsumes both. In an attempt to remove the second conceptual block, the program

uses animated graphics to illustrate eigenvectors and eigenvalues in the context of the subject space.

The geometric approach lends itself easily to graphical representations. Surprisingly, many textbooks lack such representations. In particular, multicollinearity was identified as a topic that could usefully be illustrated in contrast with orthogonality, since the former is more comprehensible if the latter is understood. To fill this conceptual gap, the multimedia program series designed for this study contains a module addressing multicollinearity and employing graphical illustration. Moreover, without clearly distinguishing subject space from variable space, an explanation of vectors may be difficult to follow. The current multimedia project is based upon the belief that starting from variable space and then relating subject space to variable space is an easier path. Finally, the program shows both spaces at once using the biplot for illustration.

In summary, the three conventional approaches were adopted and enhanced in the development of the program. The conceptual approach was used with further explanations of some common terms such as "factor, " "space," "model," and "rotation." The mathematical approach was used to compare and contrast regression and factor analysis in the context of weighted linear combinations. Lastly, the geometric approach was applied to help learners transit easily from variable space into subject space. This combination of approaches ensures that instructors using any approach may use the tutorial with benefit to students in their classes.

3.2. Development of the Survey

A survey was developed by a panel consisting of one statistician and two instructional designers, with content validity established by two experts in the field. In order to widen the scope of the generalization, data were collected using a Web-enabled database server. Responses to the survey came from graduate students in various disciplines. An invitation to participate in the study was sent to three student listserv groups owned by two different universities. Both the email message and the Web page explicitly spelled out that only graduate students who had taken at least two statistics courses were qualified to participate in this study. Twenty-five graduate students responded to the survey; no one was disqualified. Among all respondents, nine are males (36%) and fourteen are females (64%). On the average, respondents have previously taken 4.94 undergraduate and graduate statistics courses. Respondents came from a wide variety of academic backgrounds that include a Bachelor's or Master's degree in education, mass communication, mathematics, engineering, psychology, sociology, economics, and others. Table 1 gives the areas of study represented.

Table 1. Disciplines of Respondents to the Survey

Graduate programs	Frequency	Percentage
Education	9	36%
Social sciences	6	24%
Unknown	5	20%
Physical sciences & engineering	3	12%
Business	1	4%

Nursing	1	4%
	25	100%

In the survey no time constraints were set. The survey contained five short-answer questions, one multiple-choice question, and three identify questions on concepts regarding factor analysis and linear models (see the Appendix for a copy of the survey form).

The survey confirmed the researchers' suspicion that most students confuse the definition of "factor" in factor analysis with that of "factor" in factorial analysis. In factor analysis there are no dependent or independent variables, yet in answering Question 7 twenty-five percent of respondents referred to factors as predictors, independent variables, or causes. Only eight percent of the participants could answer the question correctly while all others gave irrelevant answers.

The survey also verified the researchers' assertion that many students failed to conceptualize factor analysis under the premise of weighted linear combinations. In Question 10 eighty-eight percent were not able to conceptualize the connections between weight, coefficient, and loading. Sixteen percent could not distinguish weights from variables.

Responses to Questions 9 and 9b reveal that the difference between eigenvectors and regression lines is another area of major confusion. Twelve percent of the participants misidentified eigenvectors as regression lines, thirty-two percent as "regression vectors," and twelve percent as "eigenlines," which do not exist.

3.3. "Think Aloud" Protocol Analysis

Twelve subjects with differing levels of computer and statistical literacy were asked to perform a "think aloud" protocol as they navigated through the beta program. Subjects were videotaped individually as they completed the process. Participation was voluntary and thus subjects could leave the study at any time without penalty. Nonetheless, all subjects completed their sessions. Since most people were unaccustomed to operating a computer while thinking aloud, a researcher demonstrated the "think aloud" protocol with another software package before each subject began. The recordings were analyzed and coded for common difficulties regarding user interface as well as statistical understanding. The program was then revised in accordance with the findings. Subjects found that the program helped them to clearly distinguish regression lines and eigenvectors. Their comments also indicated that most subjects could easily follow the instruction, in particular the step-by-step manner in which it was presented.

4. Description of the Completed Program

As has been indicated above, initial analysis led the researchers to develop a tutorial focusing on the goals of distinguishing subject space and variable space, understanding eigenvectors and eigenvalues, and understanding both regression and factor analysis in the context of the linear model. Regression is a topic that most intermediate statistics students have studied. As the survey results indicate, this prior knowledge is a source of misconceptions since eigenvectors in subject space are often misperceived as regression lines in variable space. This misperception provides an opportunity to use regression as a basis of comparison in explaining the differences between variable space and subject space. Regression

becomes a metaphor with which to illustrate factor analysis.

The multimedia program incorporating this approach was developed using Macromedia Director® as a remedy for the problems discussed above. The target audience for this program is graduate students in all disciplines who have learned the basic concepts of regression. It is important to note that the program was developed as a supplemental tool to conventional textbooks and lectures, rather than a standalone self-teaching module. We do not expect that students will become experts by finishing the tutorial. Instead, the function of this program is to clarify common misconceptions and to provide a general overview of factor analysis.

The multimedia program begins with a presentation of regression in variable space, then shows the user how information can be converted from variable space to subject space. Properties of regression are used to explain the properties of factor analysis as shown in Figures 1a and 1b and Table 2.
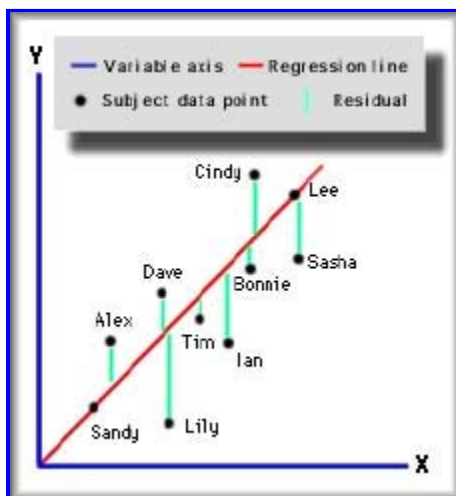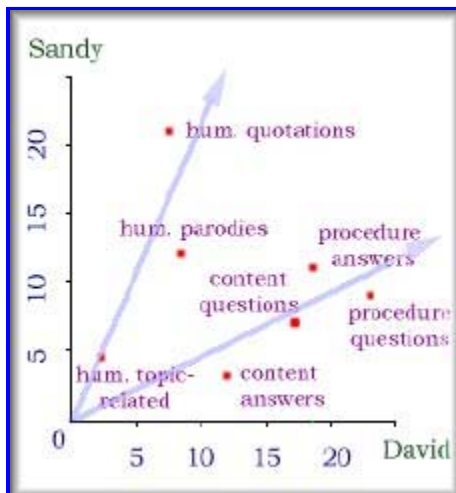

Figure 1a. Regression Represented in Subject Space

Figure 1b. Factor Analysis Represented in Subject Space

Table 2. Mapping Between Variable Space and Subject Space

	Variable space	**Subject space**
Graphical representation	The axes are variables whereas the data points are people.	The axes are people whereas the data points are variables.
Reduction	The purpose of regression analysis is to reduce a large number of people's responses into a small manageable number of trends called **regression lines**.	The purpose of factor analysis is to reduce a large number of variables into a small manageable number of factors which are represented by **eigenvectors**.
Fit	This reduction of people's responses is essentially to make the scattered data form a meaningful pattern. To find the pattern in variable space we "fit" the regression line to the people's responses. In statistical terminology we call it **the best fit**.	In subject space we look for the fit between the variables and the factors. We want each variable to "load" into the factor most related to it. In statistical terminology we call this **factor loading**.
Criterion	In regression we sum the squares of residuals and make the best fit based on **the theory of least squares**. These are the criteria used to make the reduction and the fit.	In factor analysis we sum the squares of factor loadings to get the eigenvalue. The sizes of the eigenvalues determine how many factors are "extracted" from the variables.
Structure	In regression we want the regression line to represent the trend for as many points as possible.	In factor analysis the eigenvalue is geometrically expressed with the eigenvector. We want the eigenvector to represent variability for as many points as possible. In statistical terminology we call this **simple structure**, which will be explained later.
Equation	In regression the relationship between the outcome variable and the predictor variables can be expressed in a **weighted linear combination** such as $Y = a + b_1 X_1 + b_2 X_2 + e$.	In factor analysis the relationship between the latent variable (factor) and the observed variables can also be expressed in a **weighted linear combination** such as $Y = b_1 X_1 + b_2 X_2$. Note that there is no intercept in the equation.

The meaning of a vector in subject space can be more easily understood if the learner can relate the vector to a person in variable space. Regression can thus be used as a metaphor to enhance understanding of the relationship between regression and factor analysis.

The multimedia program uses graphics and animation to illustrate both subject and variable space. Two graphing techniques that combine these types of space are the Coneplot (Dawkins 1992), which is available in S-Plus®, and the Gabriel biplot (Gabriel 1981; Gower and Hand 1996), which is available in JMP® and SAS/Insight®. Neither type of plot is directly related to factor analysis. Coneplots are primarily used for spotting multiple-dimensional outliers and discriminating between clusters. Gabriel biplots are intended for principal component analysis (PCA). Of the two, Gabriel biplots were chosen for three reasons. First, the basic objectives and principles of PCA and factor analysis are very similar except for the fact that the latter addresses communality. Second, Gabriel biplots can illustrate both regression lines and eigenvectors, which is in line with our instructional strategy. Finally, for beginning and intermediate statistics students, Coneplots may appear complicated and counterintuitive (see Figure 2) whereas Gabriel biplots are clear and self-explanatory (see Figure 3).
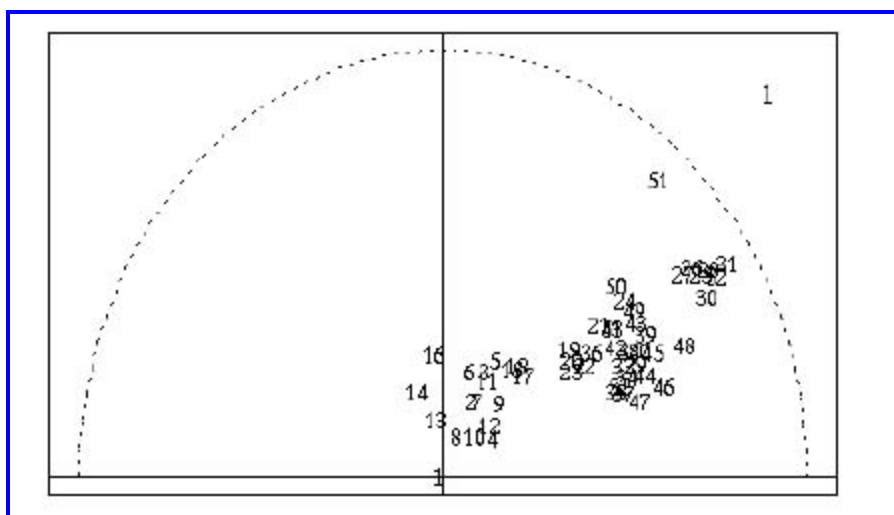

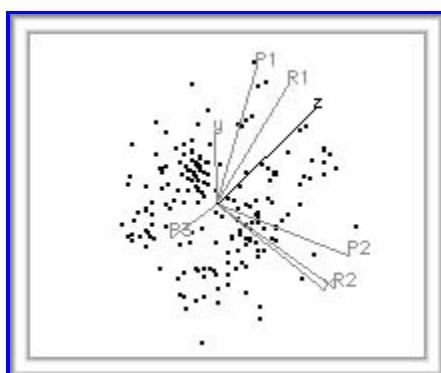
Figure 2. Example of a Coneplot.

Figure 3. Example of a Biplot.

The biplot has two limitations associated with a graphical approach to data analysis. First, a biplot uses only partial information from the singular value composition, which is a variance-maximizing

transformation of the data matrix. In other words, it gives an approximation of the data rather than showing all data. Second, a bipolot is based on an assumption that the structure underlying the data is linear. If the data structure does not conform to linearity, a biplot will show a distorted view of the data (Jacoby 1998). Realizing these shortcomings, the authors do not endorse the biplot as a data analysis tool in our tutorial, but use it only as a teaching tool.

5. Conclusions

Conventional pedagogical approaches were developed on the assumption that certain terminology will be understood by learners, but this is not necessarily true. A parallel may be drawn to the evolution of the personal computer. The computer industry began to realize the confusion caused by the command-line syntax and the proliferation of error messages during computer operation. Computer user interfaces have been redesigned to be more user-friendly.

By the same token, statisticians should consider renaming certain terms or expanding on the explanations of those terms. In particular, they should explain relationships among the terms and possible integration of these terms under the linear model. Further, conventional teaching methods are confined to limited computing resources. With the advance of high-power computers, visualizing eigenvectors in subject space is an easier path for students to conceptualize factor analysis.

The computer-based multimedia program current at the time of publication of this article can be downloaded from an overview Web page that is located on the *JSE* Web site at ww2.amstat.org/publications/jse/yu/factor_analysis.html. The version maintained by the authors can be found at the Web site: seamonkey.ed.asu.edu/~alex/multimedia/factor_analysis.html.

The authors have also prepared a Web document that presents much of the program content. A version of this document current at the time of the publication of this article can be viewed at ww2.amstat.org/publications/jse/yu/biplot.html. The version maintained by the authors can be found at seamonkey.ed.asu.edu/~alex/computer/sas/biplot.html.

The program has been distributed to statistics instructors through the Internet. Feedback from instructors is collected as ongoing and informal evaulation. The multimedia program reflects our pursuit of enhancing statistical education. Use of the application and dialogue on this topic are encouraged.

Acknowledgements

Appendix: Statistical Concepts Survey

Q1. Current Major:

Q2. Undergraduate Major:

Q3. Gender: Male Female

Q4. Age:

Q5. Number of previous statistics courses:

Q6. Course title:

If you took statistics courses at ASU, please give the prefix/number, or title (e.g. COE 502 Introduction to Quantitative Methods)

Attempt to answer each question to the best of your knowledge. For each question, if you don't know what a word or a concept means, tell what you think it means. Thank you!

Q7: In the context of research, what does the term "factor" refer to?

Q8: Define and describe "factor analysis."

Q9: In the following graph, the lines P1, P2, R1, and R2 are all examples of the same statistical concept. Which one?

   a. Regression lines
   b. Eigenvectors
   c. Regression vectors
   d. Eigenlines
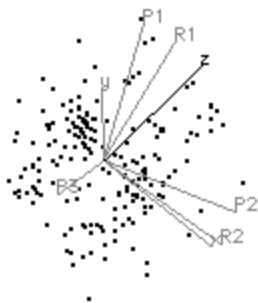
Q9B: Please explain your answer in Question 9:

Q10: In statistical analysis, what does the term "weight" refer to?

Q11: Which components in the following linear equation are "weights"? You can choose more than one choice.

$Y = a1X1 + a2X2 + a3X3 + a4X4 + e$

Y ___ a1 ___ X1 ___ a2 ___ X2 ___ a3 ___ X3 ___ a4 ___ X4 ___ E ___

Q12: What is a coefficient?

Q13: Which components in the following linear equation are "coefficients"? You can choose more than one choice.

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + e$$

Y ___ a ___ b1 ___ X1 ___ b2 ___ X2 ___ b3 ___ X3 ___ b4 ___ X4 ___ e ___

Q14: In factor analysis, what does the term "loading" refer to?

Q15: Which components in the following linear equation are "loadings"? You can choose more than one choice.

$$Y = b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + e$$

Y ___ b_1 ___ X_1 ___ b_2 ___ X_2 ___ b_3 ___ X_3 ___ b_4 ___ X_4 ___ e ___

References

Basilevsky, A. (1994), *Statistical Factor Analysis and Related Methods: Theory and Applications*, New York: John Wiley and Sons.

Comrey, A. L., and Lee, H. B. (1992), *A First Course in Factor Analysis* (2nd ed.), Hillsdale, NJ: Lawrence Erlbaum Associates.

Dawkins, B. P. (1992), *Investigating the Geometry of a P-dimensional Data Set*, Wellington, NZ: The Institute of Statistics and Operations Research.

Gabriel, K. R. (1981), "Biplot Display of Multivariate Matrices for Inspection of Data and Diagnose," in *Interpreting Multivariate Data*, ed. V. Barnett, London: John Wiley and Sons.

Gorsuch, R. L. (1983), *Factor Analysis* (2nd ed.), Hillsdale, NJ: Lawrence Erlbaum Associates.

Gower, J. C., and Hand, D. J. (1996), *Biplots*, London: Chapman and Hall.

Hagle, T. (1995), *Basic Math for Social Scientists*, Thousand Oaks, CA: Sage Publications.

Harman, H. H. (1976), *Modern Factor Analysis* (3rd ed.), Chicago, IL: The University of Chicago Press.

Huberty, C. (1994), "Why Multivariate Analyses?," *Educational and Psychological Measurement*, 54, 620-627.

Ingram, P. (1998), "Multi-Variate Statistics" [On-line]. (137.111.98.10/users/pingram/mmvar.html)

Jacoby, W. G. (1998), *Statistical Graphics for Visualizing Multivariate Data*, Thousand Oaks, CA: Sage Publications.

Joreskog, K. G., and Sorbom, D. (1979), *Advances in Factor Analysis and Structural Equation Models*, Cambridge, MA: ABT Books.

Kim, J. O., and Mueller, C. W. (1978), *Introduction to Factor Analysis: What It is and How to Do It?*,

Newbury Park, CA: Sage Publications.

Krus, D. (1998), *Visual Statistics with Multimedia*, Tempe, AZ: Cruise Scientific.

Pedhazur, E. J., and Schmelkin, L. P. (1991), *Measurement, Design, and Analysis: An Integrated Approach*, Hillsdale, NJ: Lawrence Erlbaum Associates.

Reese, C. E., and Lochmüller, C. H. (1998), "Introduction to Factor Analysis" [Online]. (www.chem.duke.edu/~reese/tutor1/factucmp.html)

Reyment, R., and Joreskog, K. G. (1993), *Applied Factor Analysis in the Natural Sciences*, Cambridge, UK: Cambridge University Press.

Someren, M. W., Barnard, Y. F., and Sandberg, J. A. C. (1994), *The Think Aloud Method : A Practical Guide to Modelling Cognitive Processes*, San Diego, CA: Academic Press.

Tabachnick B., and Fidell, L. S. (1996), *Using Multivariate Statistics* (3rd ed.), New York: Harper Collins College Publishers.

Thapalia, F. (1998), "Multivariate Statistics: An Introduction" [Online]. (trochim.human.cornell.edu/tutorial/flynn/multivar.htm)

Wulder, M. (1998), "Principal Components and Factor Analysis" [Online]. (www.pfc.forestry.ca/landscape/inventory/wulder/mvstats/pca_fa.html)

Chong Ho Yu
Arizona State University
331 West Musket Place
Chandler, AZ 85248
USA
alex@asu.edu

Sandra Andrews
Arizona State University
PO Box 870101
Tempe, AZ 85287
USA
sandya@asu.edu

David Winograd
State University of Potsdam
44 Pierrepont Ave.
Potsdam, NY 13676
USA
winogrdm@potsdam.edu

Angel Jannasch-Pennell
Arizona State University

P.O. Box 870101
Tempe AZ 85287
USA
angel@asu.edu

Samuel A. DiGangi
Arizona State University
P.O. Box 870101
Tempe, AZ 85287
USA
sam@asu.edu