COMPUTER-AID GRAPHICS TO TEACH EIGENVALUES AND EIGENVECTORS

Jaime B. Curts
Department of Curriculum and Instruction
The University of Texas Pan American, Edinburg, TX, USA
jbcurts@utpa.edu

This paper presents a model for teaching vector orthogonal rotation with the use of a computer-aid statistical graphics software. The rotation transformations are defined as: x' = x cosθ + y sinθ and y' = -x sinθ + y cosθ. The purpose is to facilitate students' understanding of a rotation matrix by allowing them to explore and plot different values of theta (θ) given a point P. Students will learn that the maximum variance is reached when θ = 45 degrees. They will also examine the graphical and statistical properties of these transformations, in particular that the variability along y' is the largest and that x' and y' are uncorrelated (orthogonal). Finally students explore how to determine the angle theta so that the variability of a set of observations along the y' axis is maximized and x' is orthogonal to y'.

Most of the rationale for using rotation in statistical analysis can be traced in the early and mid- century papers related to the need for methods for data reduction (i.e., Principal Components and Factor Analysis) and to simplify the analysis of complex multivariate data structures (i.e, multivariate). The goal of techniques like Principal Component Analysis (PCA) are to extract the most important information from the data; compress and simplify the size and description of the data; and analyze the structure of the observations and the variables. To accomplish the aforementioned goal, students are typically exposed first to the eigenvector-eigenvalue decomposition of a data covariance matrix and vector orthogonal rotation, and it is not surprising that many students have real difficulties with definitions of such concepts.

The teaching and learning of vector orthogonal rotation is an important concept rooted in geometry and statistics that requires understanding the representation of a high-dimensional dataset by a linear low- dimensional subspace. This process requires students' geometric, matrix and algebraic understanding rather than procedural abilities. Experience has shown that students generally achieve well in procedural computations, while their performance in issues requiring geometric understanding is less than satisfactory. The root of the problem is associated to a deficient learning and intuitively meaning of quantities (i.e., eigenvalues) and students' misconceptions and schema of basic concepts like linearly independent, orthogonal, and uncorrelated variables. This presentation aims to provide a framework for non-mathematical graduate students needing to understand rotation using computer-aided visualization tools and help abstract statistical concepts in a concrete manner. A model for teaching vector orthogonal rotation is illustrated with the use of a computer-aid statistical graphics software. It assumes the learner has some basic knowledge of statistical terms (mean, variance-covariance z-scores, and correlation) and basic geometry about a 2-D vector space; in particular the notion of x,y rotation transformation defined as: x' = x cos (θ) + y sin (θ) and y' = -x sin (θ) + y cos (θ).

The purpose is to facilitate students' understanding of a rotation matrix by allowing them to explore and plot different values of Theta (θ) given a point P (x,y). It starts with a simulation data set with two random variables (x,y) with known mean and variance and transforming the data into Z-scores with a mean of 0 and a variance of 1. This allows the total variance for our two-variable standardized dataset to add to 2.0. Next step is to plot the standardized scores, examine the correlation, and discuss if overlying another axis through the origin may help describe or simplify the information shown on the graph. Using basic principles of geometry, this axis has an angle of θ, and all perpendicular projection of a point on the new axis intersects at a point whose distance from the origin can be expressed as x' = x cos (θ) + y sin (θ). A new axis orthogonal to x' will contain the projection of observations defined by y' = -x sin (θ) + y cos (θ).

The idea now is to use an electronic spreadsheet or the calculation functions of a statistical software and create several columns with several values of theta, calculate the values of x' and y' for each of the data points of the simulated two variable set, and for each column calculate the mean , variance, and the cumulative variance compared to the total of the original data set (which

In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014), Flagstaff, Arizona, USA.* Voorburg, The Netherlands: International Statistical Institute. iase-web.org

should be 2.0). It is convenient to guide the students to produce a range of results using values of theta going from small ($\theta = 5^\circ$) up to $\theta = 90^\circ$

Once results are generated, students will learn that the maximum variance is reached when $\theta = 45^\circ$ degrees and how it decrease when the angle is less or greater than 45˚.They will also examine the graphical and statistical properties of these transformations, in particular that the variability along x' is the largest and that x' and y' are uncorrelated (orthogonal). Finally students explore how to determine the angle theta so that the variability of a set of observations along the y' axis is maximized and x' is orthogonal to y'. This can be easily be done by plotting parallel boxplots of the variance for the raw, standardized and the transformed data for each theta angle selected.

The methodology suggested in this presentation will allow students to compare the variances of the raw and transformed variables. They will conclude that the variances of the transformed variables sum to 2.0 which will correspond to the total variation of the raw data. However, the most important aspect is for them to see is the parsimony of the process: obtaining linear combinations –through the transformation process- that will yield the first one to account for the largest share of the total variance, the second –orthogonal to the first- with less variance, etc.

With the later in mind, students can be introduced to the idea of PCA is simply a linear recombination of original variables into a new set of variables, each of which is orthogonal to one another (and thus correlation between them is cero). All of the variance in the original dataset is simply re-allocated among the new measures to facilitate interpretation of the data set. The aim of PCA is to explain as much of the variance of the observed variables as possible using few composite variables (usually referred to as components).

Now it is possible to present the concept of eigenvalue (λ) to represent the amount of variance associated with each component. If the eigenvalues are added, the resulting total should be the total variance in the correlation matrix (i.e., the addition $\lambda_1 + \lambda_2 + ... + \lambda_n$ should be equal to n, the total number of raw variables). The percentage of explained variance of each component can be easily computed as the corresponding eigenvalue divided by the total variance.

With more advance quantitative students the discussion can continue asking how to determine x' and y' axis so that the variability of the observations along x' is maximized and orthogonal to y'. It will entail having the students writing and plotting variance-covariance matrices, rewriting PCA relationship in matrix notation, calculate the eigenvectors of the covariance matrix, etc. Research on student thinking about eigenvalues and eigenvectors suggest first to develop a based on the fundamental methods and unifying concepts of linear algebra that fosters a basic theoretical and practical understanding of the field from first principles. This presentation presents the idea of developing first a "feeling" for the intuitive sense of eigenvalues with a flexible vector-based environment and tools which students can use for meaningful learning of such concepts.