

Syllabus for Multivariate Analysis
(26:630:670)

Rutgers Business School Newark • PHD in Management Program

Professor J. Douglas Carroll

Fall 2008

- PHONE: (973) 353-5814 (Office)
- E-MAIL: dcarroll@rci.rutgers.edu
- OFFICE HOURS: 4-5PM Tuesdays or by appointment. E-mail communication is strongly encouraged.
- OFFICE: 125 Management Education Center
- Class meets: Tuesdays, 5:30-8:20PM in MEC117
- Text: *Analyzing Multivariate Data* by James M. Lattin, J. Douglas Carroll, and Paul E. Green (2003). Belmont, CA: Duxbury Press.
- Optional Supplementary Text: Carroll, J. D., & Green, P. E. (1997). *Mathematical Tools for Applied Multivariate Analysis*. San Diego, CA: Academic Press (with contributions by A. D. Chaturvedi).

Teaching Assistant: Shuojia (Nancy) Guo

Office Hours: By appointment

E-mail: nancyguo@pegasus.rutgers.edu

Class 1.

Chapter 1: Introduction. Overview of book and course.

Chapter 2: Review of basics of vector and matrix algebra and applicability to multivariate analysis (MVA).

Classes 2+3.

Chapter 2 and supplementary material.

Review of Singular Value Decomposition (SVD) of general matrices, Eigenstructure (eigenvector and eigenvalue) decomposition of square matrices, and application of eigenstructure decomposition of square symmetric “product moment” matrices to calculation of SVD of a general matrix. Statistical applications of eigenstructure and SVD analysis of various matrices; use of SVD to determine rank, compute determinant, and define inverse (or various generalized inverses, including the right or left pseudo-inverse) of various special matrices.

Classes 4+5.

Chapter 3: Regression Analysis.

Overview of multiple linear regression analysis and its applications. Both ordinary regression analysis and partial regression analysis, in which the regression of one set of variables on another after the effect of a third set of variables has been “partialled out,” or subtracted, will be dealt with, as well as methods of testing significance of various regression models, or of more complex models as compared to less complex ones.

Class 6.

Chapter 4: Principal Components Analysis.

Overview of principal components analysis (PCA). Relationship to multivariate normal distribution and to eigenstructure of covariance or correlation matrix and to SVD of original (mean centered or standardized) matrix of observations by variables. The latter discussion will introduce the notion of component scores (closely related to that of factor scores in factor analysis).

Class 7.

Chapter 5: Exploratory Factor Analysis.

Will primarily focus on principal axis form of exploratory factor analysis (FA) and its relationship to PCA. Will include a discussion of computation of factor scores as well as factor loadings, as well as various approaches to estimation of “communalities”.

Classes 8+9.

Chapter 7: Multidimensional Scaling.

Metric and nonmetric models and methods of “two-way” multidimensional scaling (MDS). Three-way MDS; specifically the INDSCAL model and the SINDSCAL method of three-way (or individual differences) MDS, and its implementation. Methods of multidimensional analysis of preferential choice (or other “dominance”) data. MDS as an alternative to PCA or exploratory FA, as well as a method of analysis of directly judged or otherwise measured proximities (similarities or dissimilarities, measures of association, concurrence, overlap, friendship or other indicators of affinity, or other measures of “closeness”) of a set of stimuli or other objects (e.g., products or brands in marketing studies).

Please Note:

Take home midterm exam will be distributed in Class 7. Will be due by time of meeting of **Class 9**. Alternatively, some students may substitute a project in which a substantive data set relevant to that student’s current research or planned dissertation research is analyzed via methods discussed in this course. Substitution of such a project will require permission of Prof. Carroll, and, if permitted, the project will be due not later than the meeting of **Class 11**.

Class 10.

Chapter 8: Cluster Analysis.

Emphasis will be placed on hierarchical clustering, specifically single, complete and average linkage, Ward's method, and K^{th} nearest neighbor clustering, based on direct or derived measures of similarity or dissimilarity, and on K -means clustering for partitioning based on a standard objects by variables multivariate data matrix.

Classes 11+12.

Chapter 11: Analysis of Variance, and generalizations.

A review of standard analysis of variance (ANOVA), plus an introduction of ANCOVA (analysis of covariance), MANOVA (Multivariate ANOVA) and MANCOVA (Multivariate ANCOVA) will all be covered. It will be shown that ANOVA can be viewed as a special case of multiple linear regression analysis, with

“dummy” independent variables, while ANCOVA is a special case of partial regression analysis. Both MANOVA and MANCOVA will be shown to be formulatable as special cases of canonical correlation analysis.

Class 13.

Chapter 6: Confirmatory Factor Analysis.

This class will deal with confirmatory factor analysis (CFA), focusing on the maximum likelihood approach to fitting CFA models to data. CFA will be compared and contrasted to PCA and EFA.

Final exam: Will be given in last class.

Grading: Grades will be based on midterm exam or project (40%), final exam (40%), as well as class participation and homework (20%).