Statistics 925: Multivariate Statistics

Syllabus, Fall 2011

Classes: Mon/Wed 1:30–3:00 p.m., in G86 JMHH

Instructor: Zongming Ma
Email: zongming@wharton.upenn.edu
Office: 468 JMHH
Office hours: Wed 3–4 p.m., or by appointment

Course Overview

This is a course that prepares PhD students in statistics for research in multivariate statistics and high dimensional statistical inference.

In the first part of the course, we focus on classical multivariate statistics. Topics include the multivariate normal distribution and the Wishart distribution; estimation and hypothesis testing of mean vectors and covariance matrices; principal component analysis, canonical correlation analysis and discriminant analysis; etc.

In the second part of the course, we shift gear to high dimensional statistics. Topics include the Marchenko-Pastur law, the Tracy-Widom law, regularized estimation of high-dimensional covariance and precision matrices, nonparametric hypothesis testing in high dimensions, high-dimensional principal component analysis, high-dimensional discriminant analysis, etc.

Course prerequisites are Stat 550 and linear algebra. Familiarity with basic asymptotic theory will be helpful, but is not required.

Textbook and References

There is no required textbook for the course. The following two books are recommended:

- *Multivariate Analysis*, by K.V. Mardia, J.T. Kent, and J.M. Bibby. Academic Press, 1979.
- *An Introduction to Multivariate Statistical Analysis, 3rd Ed.*, by T.W. Anderson, Wiley, 2003.

Either one makes an excellent reference for future work. The Lippincott Library has both books on reserve on the library use only reserve shelves. Students can ask for them under the author's last name.

Course Requirements

There will be occasional homework problems and no exam. Students will be expected to make a presentation in the later part of the course. A list of possible topics for presentation will be made available. Evaluation will be based on homework completion, presentation, and class participation.

Course Agenda

Part 1: Classical Multivariate Statistics

Lecture 1: Multivariate Normal Distribution

Lecture 2: Kronecker Product and Matrix Normal Distribution

Lecture 3: Jacobian and Exterior Differential Forms

Lecture 4: Wishart Distribution I: Density

Lecture 5: Wishart Distribution II: Properties

Lecture 6: Wishart Distribution III: Eigenvalues

Lecture 7: Hotelling's T^2 Test

Lecture 8: Basic Principles of Testing: Likelihood Ratio Tests and Union Intersection Tests

Lecture 9: Hypothesis Testing for Multivariate Distributions I: One Sample

Lecture 10: Hypothesis Testing for Multivariate Distributions II: Multiple Samples

Lecture 11: Principal Component Analysis

Lecture 12: Discriminant Analysis

Lecture 13: Canonical Correlation Analysis

Part 2: High-dimensional Statistics

Lecture 14: Introduction to Random Matrix Theory: Wigner's Semi-circle Law

Lecture 15: Stieltjes Transform and the Marchenko-Pastur Law

Lecture 16: Extreme Eigenvalues of Wishart Matrices: The Tracy-Widom Distributions

Lecture 17: Regularized Estimation of Covariance Matrices I: The Bandable Case

Lecture 18: Regularized Estimation of Covariance Matrices II: The Sparse Case

Lecture 19: Regularized Estimation of Precision Matrices

Lecture 20: Principal Component Analysis in High Dimensions

Lecture 21: Discriminant Analysis in High Dimensions

Lecture 22: Hypothesis Testing in High Dimensions I: Mean

Lecture 23: Hypothesis Testing in High Dimensions II: Covariance

Lecture 24: Hypothesis Testing in High Dimensions III: Linear Hypotheses

The remaining lectures are devoted to student presentations.