Multivariate Analysis
MATH 4424
Fall HKUST
http://www.math.ust.hk/ makchen/Math347/Math347.htm

Instructor, TA and access for help.

Instructor: Kani Chen.

All students are welcome to get help from the instructor. In particular, they are *strongly encouraged* to ask questions in class and talk to the instructor right after class. Other accesses for help are:

1). email to the instructor at "makchen@ust.hk", or drop by the instructor's office at Room 3477 anytime, or call the instructor at 2358-7425;

Course Description.

A course for students to equip themselves with multivariate statistical analysis, an important statistical skill for data analysis. Topics to be explored include the principal component analysis, factor analysis, canonical analysis and classification.

Intended Learning Outcomes

By completing the course, students should be able to

deal with basic statistical problems about mean comparison for multivariate data;

find out underlying factors that affect variations of multivariate data;

find out connections between two sets of variables;

correctly classify new subjects into given classes;

develop basic computational skills using R.

Textbook.

"Applied multivariate statistical analysis", by Richard Johnson and Dean Wichern, (6th edition , Pearson Prentice Hall 2007), ISBN: 0135143500.

Reference Books:

1). A book of elementary level (easier than the textbook): Multivariate data analysis, 5th edition, by Hair, Anderson, Tatham and Black.

2). A book of advanced level (more sophisticated than the textbook): Multivariate analysis, by Muirhead.

Assessment:

Homework 20%, Midterm 30%, Final 50%. Homework shall be assigned once every other week. Both exams are closed-book exams, except for a formula sheet prepared and provided by the instructor. **Midterm Exam on Oct 24, Week 8**.

Prerequisites:

Linear algebra (Math 111). Statistical Inference (Math 243).

Course Schedule:

Week 1. (Chapters 1-3) Review: Aspects of multivariate analysis: data organization, data description— quantitative methods and graphic methods; basics of matrix algebra and fundamental statistics and their properties.

Week 2-3. (Chapter 4). Multivariate normal distribution: Properties of multivariate normal distribution; sample mean and variances; Normality assumption related data analysis.

Week 4-5. (Chapter 5). Test of hypothesis on normal means–Hotelling's T^2 and likelihood ratio tests; Confidence region and simultaneous inferences;

Week 6-8. (Chapter 6). Paired comparisons; repeated measures design, Comparison for two populations; one-way MANOVA.

Week 9. (Chapter 7). Multivariate linear regression: Linear model, LSE and inferences, multivariate linear model.

Week 10-11. (Chapter 8). Principal components: population principal components, summarizing and graphing of principal components.

Week 12. (Chapter 9). Factor analysis: orthogonal factor model, estimation, factor rotation and factor scores, a general strategy.

Week 13. (Chapter 10). Canonical correlation analysis: canonical variates and correlations, interpreting canonical variables, sample canonical variates and correlations.

Week 14. (Chapter 11). Discrimination and classification.

Week 15. Final.

Remark: We may or may not cover cluster analysis, depending on the progress of teaching.