MATH48011 - 2011/2012

General Information

- Title: Linear Models with Nonparametric Regression
- Unit code: MATH48011
- Credits: 15
- **This course unit may not be taken as well as MATH38011.**
- Prerequisites: MATH20701, MATH20802, MATH20812 is helpful background but not essential.
- Co-requisite units: None
- School responsible: Mathematics
- Members of staff responsible: Prof. J. Pan and Dr A Donev

Specification

Aims

- To familiarise students with the methodology and applications of standard techniques of regression analysis, and analysis of variance and covariance.
- To introduce students to different methods in nonparametric regression.
- To explore some of the wide range of real-life situations occurring in the fields of agriculture, biology, engineering, industrial experimentation, medical and social sciences that can be investigated using Linear Statistical Models.

Brief Description of the unit

In many areas of science, technology, social science and medicine one often wishes to explore the relationship between one observable random response and a number of 'factors' which may influence simultaneously the response. The techniques developed to study such relationships fall in three broad categories:

1. Regression Analysis where the influence of the factors is quantitative;
2. Analysis of Variance where each factor's influence is qualitative; and
3. Analysis of Covariance where both qualitative and quantitative factors are present.

However, these three valuable techniques can be studied together as special cases of a unified theory of Linear Models. The course starts with a study of estimation and hypothesis testing in the general linear problem. Once the principles and techniques are established practical applications in the three types of analysis are examined in greater detail.

Nonparametric regression provides a very flexible approach to exploring the relationship between a response and an associated covariate but without having to specify a parametric model. The different techniques available are essentially based on forms of local averaging controlled by the value of a smoothing parameter. In this part of the module we will study a few different techniques, along with their statistical properties. We will also look briefly at how such estimators can be used in more inferential procedures.

Learning Outcomes

On successful completion of this course unit students will be able to

- Fit linear models and comment on the adequacy of the fit;
- identify and apply appropriate transformations of either the response or of the covariates;
- check for and identify colinearities, and understand the implications of their presence;
- produce confidence intervals for linear combinations and ratios of linear combinations of the model parameters;
- formulate hypotheses in terms of the model parameters and construct test procedures for testing such hypotheses.
- Be familiar with different approaches to estimating a nonparametric regression model;
- understand the role of the smoothing parameter in constructing a nonparametric regression curve;
- know how to derive the asymptotic properties of the chosen estimator;
- be able to use such estimators to check certain parametric assumptions.
- Use the statistical software R or S-PLUS to analyze real data using both parametric and nonparametric techniques.

Future topics requiring this course unit

This course is naturally related to some other 3rd year courses, including MATH38082 *Analysis of Designed Experiments*, MATH38052 *Generalised Linear Models*.

Syllabus

Linear Models – taught by Prof Jianxin Pan

1. General Linear Models: Least squares estimators (l.s.e) and their properties. Residuals and residual sum of squares. Leverage. Distribution of l.s.e and of the residual sum of squares. [5]
2. The general linear hypothesis. Extra sum of squares, sequential sum of squares, partial sum of squares. The test statistic of the general linear hypothesis and its distribution. Confidence intervals and prediction intervals. [5]
3. Linear regression: Simple regression, multiple regression, dummy variables and analysis of covariance. [6]
4. Analysis of Variance. One and two way analysis of variance. Use of comparisons. Interactions. [6]

Nonparametric Regression - taught by Dr Peter Foster

1. Least squares regression, local averaging, the Priestley-Chao estimator. [2]
2. Local polynomial kernel regression. [3]
3. Choosing the value of the smoothing parameter. [1]
4. Variability bands, checking the validity of a parametric regression model. [3]
5. Introduction to spline regression. [2]

Textbooks

- Draper, D. N. R. and Smith, H., *Applied Regression*, J. Wiley 1998.
- Weisberg, S., *Applied Linear Regression* J. Wiley 2005
- Montgomery, D. C. and Peck, E. A., *Introduction to Linear Regression Analysis*, J. Wiley 2001.
- Rawlings, J. O., *Applied Regression Analysis: A Research Tool*, Wadsworth and Brooks/Cole 1998.
- Bowman, A. W. and Azzalini, A. *Applied Smoothing Techniques for Data Analysis*. Oxford University Press (1998)
- Wand, M.P. and Jones, M.C. *Kernel Smoothing*. Chapman and Hall (1995)
- Eubank, R.L. *Spline Smoothing and Nonparametric Regression*. Marcel Dekkar (1999)
- Hardle, W. *Applied Nonparametric Regression*. Cambridge University Press (1991)

Teaching and learning methods

Two lectures and one examples class each week. In addition students should expect to spend at least seven hours each week on private study for this course unit.

Assessment

Coursework: weighting 20%
End of semester examination: two and a half hours weighting 80%

Arrangements