

Stat 586: Nonparametric Regression

Instructor: Guoyi Zhang, gzhang123@gmail.com Office hours: MWF 15:00pm-16:00pm in SMLC 318

Teaching Assistant:

Class Time/Place: 14:00am-14:50am MWF - DSH 233

Prerequisites: Stat 440/540 Regression Analysis

Text: Nonparametric Regression and Spline Smoothing (Eubank 1999) ISBN:0-8247-9337-4

Objective: This course is intended to provide theoretical and practical knowledge of modern smoothing and nonparametric regression methodologies. By the end of course, the students are expected to be familiar with the commonly used smoothing techniques including related software and algorithms.

Topics:

Kernel estimators, smoothing splines, density estimation, additive models, higher dimensional splines, and model selection. Special treatment is given to the important problem of choosing the tuning parameters, such as smoothing parameters in smoothing splines and bandwidth parameter in kernel smoothers.

Computing: R will be used in this class.

Reference Books:Smoothing and Regression: Approaches, Computation, and Application (Schimek 2000),Generalized Additive Models (Hastie and Tibshirani, 1990),Spline Models for Observational Data (Wahba 1990);Nonparametric Regression and Generalized Linear Models (Green and Silverman, 1994), Smoothing Spline ANOVA Models (Gu 2002)

Grading: Homework, 50% (bi-weekly homework); Final Project, 50%.

Teaching Plan:

1 Introduction

- General introduction on smoothing and nonparametric models
- Single predictor (scatterplot) smoothing:
- bias-variance tradeoff, selection of smoothing parameters, degrees of freedom
- R package

2 Additive Models

- additive models, generalized additive models, semiparametric models, modeling interactions
- model fitting (backfitting algorithm), model tuning
- projection pursuit regression, multivariate adaptive regression splines (MARS)

3 Smoothing Splines

- theory of reproducing kernel Hilbert space (RKHS), roughness penalty approach
- smoothing splines revisited, thin-plate splines, tensor product splines and smoothing spline ANOVA
- smoothing parameter selection, generalized CV, Bayesian interpretations and inference
- model/variable selection, LASSO, component selection and smoothing operator (COSSO)

4 Kernel Density Smoothers/Estimators

- univariate kernel density estimation, multivariate kernel density estimation
- bandwidth selection, MSE and MISE criteria, cross validation
- asymptotic theory, selecting kernels, computational algorithms and programs