Stat 05: Introduction to Statistical Shrinkage Techniques

Dr Amar Ahmad

Course Description:

This course provides an in-depth exploration of statistical shrinkage techniques, a powerful and flexible statistical methodology used to improve parameter estimation and prediction in various fields, including machine learning, finance, and biostatistics. Students will learn the principles, applications, and implementation of shrinkage techniques, including ridge regression, lasso regression, Bayesian methods, and empirical Bayes estimation. The course will also cover practical examples and real-world applications.

Prerequisites:
- Basic knowledge of statistics and regression analysis.
- Familiarity with the statistical software package R.
- Understanding of linear algebra concepts is beneficial but not required.

Course Objectives:
By the end of this course, students should be able to:
1. Understand the principles of statistical shrinkage techniques.
2. Implement ridge regression and lasso regression for parameter estimation.
3. Apply Bayesian methods and empirical Bayes estimation.
4. Select appropriate shrinkage techniques for different data scenarios.
5. Evaluate the benefits and limitations of shrinkage methods through practical examples.

Course Outline:

Module 1: Introduction to Shrinkage Techniques
- Motivation for shrinkage methods.
- Bias-variance trade-off.
- Overview of ridge and lasso regression.

Module 2: Ridge Regression
- Formulation and penalty term in ridge regression.
- Ridge parameter tuning and cross-validation.
- Interpretation of ridge regression results.

Module 3: Lasso Regression
- Formulation and penalty term in lasso regression.
- Lasso parameter tuning and feature selection.
- Interpretation of lasso regression results.

Module 4: Mathematical Foundations
- L1 (lasso) and L2 (ridge) penalties.
- Elastic Net penalty as a combination of L1 and L2.
- Objective function and optimization.

Module 5: Shrinkage Techniques in High-Dimensional Data
- Shrinkage in high-dimensional settings.
- Elastic Net regression.
- Group lasso and structured shrinkage.

Module 6: Time-Series and Panel Data Applications
- Shrinkage techniques for time series data.
- Panel data analysis with shrinkage methods.

Module 7: Real-World Applications
- Case studies and practical examples.
- Shrinkage techniques in finance, biology, and machine learning.

Module 8: Shrinkage Software Tools
- Implementing shrinkage techniques in R.
- Libraries and packages for shrinkage.

Assessment Methods:
1. Quizzes and Homework Assignments: Assessing understanding of theoretical concepts.
2. Final Exam: Covering material from the entire course.
3. Group Projects: Applying hypothesis testing to real-world data analysis.
4. Class Participation: Active engagement in discussions and activities.

Grading:
- Quizzes and Homework Assignments: 30%
- Final Exam: 30%
- Group Projects: 30%
- Class Participation: 10%

Textbook:
- Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). The elements of statistical learning: data mining, inference, and prediction. 2nd ed. New York, Springer.