

Penalised likelihood methods for high-dimensional pattern analysis

Lecturer:

Jing-Hao Xue

Department of Statistical Science, University College London, UK

Abstract:

In recent years penalised likelihood methods, such as the least absolute shrinkage and selection operator (lasso), the elastic net, the smoothly clipped absolute deviation (SCAD), the adaptive lasso and the group lasso, have become popular and been extensively studied in the statistics and machine-learning communities. These methods were developed to achieve both model fitting and feature selection, originally for regression and most recently for classification and clustering. Therefore they are extremely attractive for high-dimensional pattern analysis, in particular when sparsity is present in a so-called 'large-p-small-n' context, the context that the sample size, n , is smaller than the feature dimension, p , and that traditional methods often fail. In this tutorial, we shall review and discuss some established and state-of-the-art penalised likelihood methods for regression, classification and clustering, with respect to their intuitions, methodologies, computational algorithms and Bayesian counterparts.

Course description

This half-day tutorial will review and discuss some established and state-of-the-art penalised likelihood methods for regression, classification and clustering, with respect to their intuitions, methodologies, computational algorithms and Bayesian counterparts.

Syllabus:

1. Introduction to penalised likelihood methods (PLM)
2. PLM for high-dimensional regression (ridge regression, lasso, bridge regression, elastic net, relaxed lasso, adaptive lasso, SCAD, fused lasso, group lasso, etc.)
 - (a) Intuitions and methodologies
 - (b) Computational algorithms

(c) Bayesian counterparts

3. PLM for high-dimensional classification

4. PLM for high-dimensional clustering

About Lecturer:

Jing-Hao Xue received the B.Eng. degree in telecommunication and information systems in 1993 and the Dr.Eng. degree in signal and information processing in 1998, both from Tsinghua University, the M.S. degree in medical imaging and the M.Sc. degree in statistics, both from Katholieke Universiteit Leuven in 2004, and the degree of Ph.D. in statistics from the University of Glasgow in 2008. He has worked in the Department of Statistical Science at University College London, as a Lecturer in Statistics since 2008. His research interests include statistical and machine-learning techniques for pattern recognition, data mining and image processing, in particular supervised, unsupervised and incompletely-supervised learning for complex and high-dimensional data.