Faculty of Computer Science 2005-2006 Seminar Series

A Branch and Cut Algorithm for the Halfspace Depth Problem

By

Dr. David Bremner & Dan Chen

Wednesday, July 5th, 2006
2:30pm
ITC317

In non-parametric statistics, no assumption is made about the probability distribution of the population, and the test statistics are usually based on the rank of the data. In multivariate data analysis, every data item consists of several elements (i.e. is an n-tuple). The concept of data depth in non-parametric multivariate statistics is the generalization of the univariate rank method. Given a set S of points, the halfspace depth (or rank) k of point p is defined as the minimum number of points of S excluding p contained in any closed halfspace with p on its boundary. The data with the highest rank is considered the centre of the data set, which best describes the set. Computing halfspace depth is NP-hard, and it is a special case of the Maximum Feasible Subsystem problem, which is to find a minimum cardinality set of constraints whose removal makes an infeasible system feasible. In this paper an infeasible linear system is formulated for the halfspace depth problem, and a mixed integer program is formulated with the big-M method. We suggest a branch and cut algorithm. In this algorithm, Chinneck's heuristic algorithm is used to find an upper bound at the beginning and a related technique technique based on sensitivity analysis is used for branch selection. Irreducible Infeasible Subsystem (IIS) hitting set cuts are applied.

This algorithm is implemented with the BCP framework from the COIN (www.coin-or.org) project.

****

STUDENTS ARE ENCOURAGED TO ATTEND

****