数学与系统科学研究院

计算数学所学术报告

报告人： Prof. Marc Teboulle

(Tel–Aviv University, Israel)

报告题目：

A Unified Continuous Optimization Framework for Center–based Clustering Methods

邀请人： 袁亚湘研究员

报告时间： 2008 年 4 月 11 日(周五)

上午 10:00—11:00

报告地点： 科技综合楼三层 311

计算数学所报告厅

Abstract:

Clustering consists of grouping or classifying data into similar objects. It is a fundamental technique in unsupervised machine learning and arise in a wide spectrum of research fields such as, astrophysics, statistics, biology, computer vision, data compression in image processing, and information retrieval, to name just a few. The interdisciplinary nature of the clustering problem has generated a very large body of literature in cluster analysis, with many formulations of the problem, and numerous clustering algorithms which are based on various disparate motivations and approaches.

In this talk, we present a unifying framework for center based clustering algorithms from an optimization theory perspective. Starting with the standard nonconvex and nonsmooth formulation of the clustering problem, which minimizes the sum of a finite collection of ``min'' functions, we present a systematic way for designing center based clustering methods. Our approach builds on two fundamental mathematical concepts: convex asymptotic functions, and the so-called nonlinear means of Hardy,

Littlewood and Polya. This allows us to derive a generic smoothing iterative scheme for clustering, which computationally, is as simple as the popular k-means algorithm. We analyze its properties, and establish its convergence. We then demonstrate that most well known center based clustering algorithms, which were derived either heuristically, or/and which have emerged from intuitive analogies in physics, statistical techniques, and from information theoretic perspectives, can be simply recovered, analyzed and extended through the proposed framework.

欢迎大家参加！