PRINCETON UNIVERSITY
OPERATIONS RESEARCH & FINANCIAL ENGINEERING SEMINAR

Convex Regularization Algorithms for Learning Large Incomplete Matrices

Rahul Mazumder
Stanford University

Abstract

In many applications measured data can be represented in a matrix $X_{m \times n}$, for which only a relatively small number of entries are observed. The task is to “complete" the matrix based on the observed entries and has been dubbed the matrix completion problem. This is a fundamental problem arising for example in recommender systems (for eg the “Netflix" prize) and collaborative filtering; internet and electronic-commerce industries; image processing and micro-array imputation, among others. In practice, the matrix dimensions frequently range from hundreds of thousands to even a million. Low-rank matrix modeling is often a model of choice for these problems but is computationally intractable. Using nuclear norm as the convex surrogate of the combinatorially hard rank constraint we develop a simple and efficient algorithm SOFT-IMPUTE which minimizes the reconstruction error subject to a regularization on the nuclear norm. With warm-starts this allows us to efficiently compute an entire regularization path of solutions on a grid of values of the regularization parameter. The computationally intensive part of our algorithm is in computing a low-rank SVD of a large dense matrix. Using problem structure, we show how to circumvent this computational bottleneck and demonstrate algorithmic scalability to problems of the Netflix size, and even larger— all of which can be achieved within a very reasonable time. We address convergence properties and computational complexities of SOFT-IMPUTE and establish connections with other popular variants of matrix factorization. We show how our proposed framework can be adapted to more general regularized low-rank modeling frameworks, discussing specific applications in image processing, computer vision and collaborative filtering.

Tuesday, February 14, 2012
Sherrerd Hall, Room 101
4:30 PM