CBMF : A Clustering Approach to Binary Matrix Factorization

Dr. Peng Jiming
University of Illinois at Urbana-Champaign

Date : 27 May 2013 (Monday)
Time : 2.00pm – 3.00pm
Venue : MAS Executive Classroom 2, MAS-03-07
School of Physical and Mathematical Sciences

In general, binary matrix factorization (BMF) refers to the problem of finding a matrix product of two binary low rank matrices such that the difference between the matrix product and a given binary matrix is minimal. As an important tool in dimension reduction for high-dimensional data sets with binary attributes, BMF has been widely and successfully used in various applications. In this talk, we first introduce two new constrained BMF models (called CBMF) and discuss their relation to other dimensional reduction models such as unconstrained BMF (UBMF). Then we propose alternating update procedures for CBMF. In every iteration of the proposed procedure, we solve a specific binary linear programming (BLP) problem to update the involved matrix argument. By exploring the interrelation between the BLP subproblem and clustering, we develop both efficient deterministic and randomized approximation algorithms for CBMF. An effective algorithm for UBMF is also developed. We conclude the talk by presenting numerical results obtained from applying the new models and algorithms on data mining applications in bioinformatics and document clustering.

Speaker Biography

Jiming Peng is an assistant professor in the department of industrial and enterprise system engineering, University of Illinois at Urbana-Champaign. He received his PhD degree in 2001 from Delft University of Technology, the Netherlands. Then he joined McMaster University in Canada and moved to Illinois in 2006. His research covers several branches in the field of mathematical programming such as interior-point methods for linear conic optimization, approximation/relaxation to large scale mixed integer nonlinear programming, sparse/low rank solutions in optimization, as well as optimization modeling and algorithm design with applications from data mining and financial engineering. So far he has published a research monograph and more than 50 papers in major optimization journals or top CS/IEEE conferences. For his research works in optimization, he received the Stieltjes prize for the best thesis in mathematics from Stieltjes institute, the Netherlands (2001) and the premier research excellence award from Ontario, Canada (2003). He was also selected as one of the three finalists for the Tucker prize awarded by the mathematical optimization society in 2003. He has been serving as an associate editor for the journal “optimization letters” since 2006.

Host: Division of Mathematical Sciences, School of Physical and Mathematical Sciences