Composite Difference-Max Programs for Modern Statistical Estimations

**ABSTRACT** - Many modern statistical estimation problems are defined by three major components: a statistical model that postulates the dependence of an output variable on the input features; a loss function measuring the error between the observed output and the model predicted output; and a regularizer that controls the overfitting and/or variable selection in the model. We study the sampled version of this generic statistical estimation problem where the model parameters are estimated by empirical risk minimization, which involves the minimization of the empirical average of the loss function at the data points weighted by the model regularizer. In our setup we allow all three component functions to be of the difference-of-convex type and illustrate them with a host of commonly used examples, including those in continuous piecewise affine regression and in deep learning with piecewise affine activation functions. We describe a non-monotone majorization-minimization (MM) algorithm for solving the unified nonconvex, nondifferentiable optimization problem which is formulated as a specially structured composite dc program of the pointwise max type, and present convergence results to a directional stationary solution. An efficient semismooth Newton method is proposed to solve the dual of the MM subproblems. Numerical results are presented to demonstrate the effectiveness of the proposed algorithm and the superiority of continuous piecewise affine regression over the standard linear model. The talk is based on joint work with Jong-Shi Pang (USC) and Bodhisattva Sen (Columbia University).

**SPEAKER BIO** – Ying Cui is currently a Postdoctoral Scholar - Research Associate in the Daniel J. Epstein Department of Industrial and Systems Engineering at University of Southern California, working with Professor Jong-Shi Pang. She completed her Ph.D in Department of Mathematics from National University of Singapore in 2016. Her research focuses on theoretical foundations and computational methods for large scale semidefinite programming, and modern nonconvex nondifferentiable optimization with applications in statistical estimations, operations research and machine learning.