

STATISTICS COLLOQUIUM

MONDAY, FEBRUARY 01, 2010

TALK: 4:00 PM — SCIENCE CENTER ROOM 705

RECEPTION: 5:15 PM — SCIENCE CENTER, 7TH FLOOR

“Optimization in Some High-Dimensional Statistical Problems”

Hua Zhou

Postdoctoral Fellow

Department of Human Genetics

UCLA

ABSTRACT

Modern high-throughput datasets from data mining, genomics, and imaging demand high-dimensional models with ten to hundreds of thousands of parameters. This poses challenges to the classical techniques of optimization such as Newton's method and Fisher's scoring algorithm, which involves storing, computing, and inverting large Hessian or information matrix. If parameter constraints and parameter bounds intrude, then the algorithms require further modification. Although numerical analysts have devised numerous remedies and safeguards, these all come at a cost of greater implementation complexity.

In this talk, I first describe the minorization-maximization (MM) principle which generalizes the celebrated expectation-maximization(EM) algorithm and offers a versatile weapon for attacking optimization problems of this sort. Then I present two devices for accelerating MM algorithms for high-dimensional problems: a quasi-Newton scheme and a parallel computing method based on inexpensive graphics processing units (GPUs). Both offer one to two orders of magnitude improvement over the naive algorithm. Applications include positron emission tomography, nonnegative matrix factorization, a movie rating algorithm, and multidimensional scaling. Time permitting, I will also present several variations of deterministic annealing that tend to avoid inferior modes and find the dominant mode in some non-convex statistical problems.