

HARVARD UNIVERSITY
DEPARTMENT OF STATISTICS



Tel. 617-495-5496
Fax. 617-496-8057
statdept@stat.harvard.edu
<http://www.stat.harvard.edu/>

Science Center 7th Fl.
One Oxford Street
Cambridge, MA 02138

STATISTICS COLLOQUIUM

MONDAY, NOVEMBER 24, 2014

TALK: 4:15 PM — SCIENCE CENTER RM. 705

RECEPTION: 3:50 PM — SCIENCE CENTER, 7TH FLOOR

“Matrix Completion and Matrix Concentration”

Lester Mackey

**Department of Statistics / Department of Computer Science
Stanford University**

ABSTRACT

The goal in matrix completion is to recover a matrix from a small subset of noisy entries. Web-scale instances of this problem include collaborative filtering for recommendation and link prediction in social networks. Many modern matrix completion algorithms provide strong recovery guarantees but scale poorly due to the repeated and costly computation of truncated SVDs. To address this deficiency, in the first part of this talk, I will introduce Divide-Factor-Combine (DFC), a parallel divide-and-conquer framework for boosting the scalability of a matrix completion algorithm while retaining its theoretical guarantees. Our experiments demonstrate the near-linear to super-linear speed-ups attainable with this approach, and our analysis shows that DFC enjoys high-probability recovery guarantees comparable to those of its base algorithm.

Fundamental to our analysis – and to the analyses of many matrix completion procedures – are matrix concentration inequalities that characterize the fluctuations of a random matrix about its mean. In the second part of this talk, I will show how Stein’s method of exchangeable pairs can be used to derive concentration inequalities for matrix-valued random elements. When applied to a sum of independent random matrices, this approach yields matrix generalizations of the classical inequalities due to Hoeffding, Bernstein, Khintchine, and Rosenthal. The same technique delivers bounds for sums of dependent random matrices and more general matrix functionals of dependent random elements.