Summer Course on High Dimensional Statistical Inference

Tony Cai and Ming Yuan

University of Pennsylvania and University of Wisconsin at Madison

This course will cover high-dimensional statistical inference with the focus on the recovery of high dimensional sparse signals and the estimation of large matrices. These and other related problems have attracted much recent interest in a range of fields including statistics, applied mathematics and electrical engineering. We will discuss in detail the penalized and constrained l_1 minimization methods and give a unified and elementary analysis on sparse signal recovery in three settings: noiseless, bounded noise and Gaussian noise. This course will also present the latest results on optimal estimation of large covariance/precision matrices. More specifically, the course will cover the following topics:

1. Compressed sensing: recovery of sparse signals in the noiseless case;
2. High-dimensional linear regression: LASSO, Dantzig Selector;
3. Construction of compressed sensing matrices;
4. Estimation of general matrices;
5. Estimation of covariance matrices;
6. Estimation of precision matrices.

Time permitting, high dimensional linear discriminant analysis will also be discussed at the end.

References:

Bickel, P. J., Ritov, Y. and Tsybakov, A. B. (2009). Simultaneous analysis of Lasso and Dantzig selector, *The Annals of Statistics*, 37, 1705-1732.

Cai, T. (1999). Adaptive wavelet estimation: a block thresholding and oracle inequality approach. *The Annals of Statistics* 27, 898-924.

Cai, T. and Jiang, T. (2011). Limiting laws of coherence of random matrices with applications to testing covariance structure and construction of compressed sensing matrices. *The Annals of Statistics* 39, 1496-1525.

Cai, T. and Jiang, T. (2012). Phase transition in limiting distributions of coherence of high-dimensional random matrices. *J. Multivariate Analysis* 107, 24-39.

Cai, T. and Liu, W. (2011). Adaptive thresholding for sparse covariance matrix estimation. *Journal of the American Statistical Association* 106, 672-684.

Cai, T. and Liu, W. (2011). A direct estimation approach to sparse linear discriminant analysis. *Journal of the American Statistical Association* 106, 1566-1577.

Cai, T., Liu, W. and Luo, X. (2011). A constrained l1 minimization approach to sparse precision matrix estimation. *Journal of the American Statistical Association* 106, 594-607.

Cai, T., Wang, L. and Xu, G. (2010). Shifting inequality and recovery of sparse signals. *IEEE Transactions on Signal Processing*, 58, 1300-1308.

Cai, T., Wang, L. and Xu, G. (2010). Stable recovery of sparse signals and an oracle inequality. *IEEE Transactions on Information Theory* 56, 3516-3522.

Cai, T. and Yuan, M. (2012). Adaptive covariance matrix estimation through block thresholding. *The Annals of Statistics* 40, 2014-2042.

Cai, T., Zhang, C.-H. and Zhou, H. (2010). Optimal rates of convergence for covariance matrix estimation. *The Annals of Statistics* 38, 2118-2144.

Cai, T. and Zhou, H. (2012). Optimal rates of convergence for sparse covariance

matrix estimation. *The Annals of Statistics*, to appear.

Candes, E. T. and Tao, T. (2007). The Dantzig selector: statistical estimation when p is much larger than n (with discussion), *The Annals of Statistics* 35, 2313-2351.

Ravikumar, P., Wainwright, M., Raskutti, G. and Yu, B. (2011). High-dimensional covariance estimation by minimizing l1-penalized log-determinant divergence. *Electronic Journal of Statistics* 5, 935-980.

Yuan, M. (2009). Sparse inverse covariance matrix estimation via linear programming. *Journal of Machine Learning Research*, 11, 2261?286.

Yuan, M. and Lin, Y. (2007). Model selection and estimation in the Gaussian graphical model. *Biometrika*, 94, 19?5.