

## DISCUSSION: THE DANTZIG SELECTOR: STATISTICAL ESTIMATION WHEN $p$ IS MUCH LARGER THAN $n$

BY T. TONY CAI<sup>1</sup> AND JINCHI LV

*University of Pennsylvania and Princeton University*

Professors Candès and Tao are to be congratulated for their innovative and valuable contribution to high-dimensional sparse recovery and model selection. The analysis of vast data sets now commonly arising in scientific investigations poses many statistical challenges not present in smaller scale studies. Many of these data sets exhibit sparsity where most of the data corresponds to noise and only a small fraction is of interest. The needs of this research have excited much interest in the statistical community. In particular, high-dimensional model selection has attracted much recent attention and has become a central topic in statistics. The main difficulty of such a problem comes from collinearity between the predictor variables. It is clear from the geometric point of view that the collinearity increases as the dimensionality grows.

A common approach taken in the statistics literature is the penalized likelihood, for example, Lasso (Tibshirani [11]) and adaptive Lasso (Zou [12]), SCAD (Fan and Li [7] and Fan and Peng [9]) and nonnegative garrote (Breiman [1]). Commonly used algorithms include LARS (Efron, Hastie, Johnstone and Tibshirani [6]), LQA (Fan and Li [7]) and MM (Hunter and Li [10]). In the present paper, Candès and Tao take a new approach, called the Dantzig selector, which uses  $\ell_1$ -minimization with regularization on the residuals. One promising fact is that the Dantzig selector solves a linear program, usually faster than the existing methods. In addition, the authors establish that, under the Uniform Uncertainty Principle (UUP), with large probability the Dantzig selector mimics the risk of the oracle estimator up to a logarithmic factor  $\log p$ , where  $p$  denotes the number of variables.

We appreciate the opportunity to comment on several aspects of this article. Our discussion here will focus on four issues: (1) connection to sparse signal recovery in the noiseless case; (2) the UUP condition and identifiability of the model; (3) computation and model selection; (4) minimax rate.

**1. Sparse signal recovery.** The “large  $p$ , small  $n$ ” regression problem considered in this paper can be viewed as a generalization of the classical linear algebra problem in which one wishes to solve the linear equation

$$(1) \quad y = X\beta,$$

---

Received January 2007.

<sup>1</sup>Supported in part by NSF Grant DMS-06-04954.