**Sparse Estimation with Strongly Correlated Variables**

Robert D. Nowak, University of Wisconsin

This paper studies *ordered weighted* $l_1$ (OWL) norm regularization for sparse estimation problems with strongly correlated variables. We prove sufficient conditions for clustering based on the correlation/colinearity of variables using the OWL norm, of which the so-called OSCAR norm is a particular case. Our results extend previous ones for OSCAR in several ways: for the squared error loss, our conditions hold for the more general OWL norm and under weaker assumptions; we also establish clustering conditions for the absolute error loss, which is, as far as we know, a novel result. Furthermore, we characterize the statistical performance of OWL norm regularization for generative models in which certain clusters of regression variables are strongly (even perfectly) correlated, but variables in different clusters are uncorrelated. We show that if the true p-dimensional signal generating the data involves only s of the clusters, then O(s log p) samples suffice to accurately estimate the signal, regardless of the number of coefficients within the clusters. The estimation of s-sparse signals with completely independent variables requires just as many measurements. In other words, using the OWL we pay no price (in terms of the number of measurements) for the presence of strongly correlated variables.