Mathematics > Statistics Theory

Estimation in high-dimensional linear models with deterministic design matrices

Jun Shao, Xinwei Deng

(Submitted on 5 Jun 2012)

Because of the advance in technologies, modern statistical studies often encounter linear models with the number of explanatory variables much larger than the sample size. Estimation and variable selection in these high-dimensional problems with deterministic design points is very different from those in the case of random covariates, due to the identifiability of the high-dimensional regression parameter vector. We show that a reasonable approach is to focus on the projection of the regression parameter vector onto the linear space generated by the design matrix. In this work, we consider the ridge regression estimator of the projection vector and propose to threshold the ridge regression estimator when the projection vector is sparse in the sense that many of its components are small. The proposed estimator has an explicit form and is easy to use in application. Asymptotic properties such as the consistency of variable selection and estimation and the convergence rate of the prediction mean squared error are established under some sparsity conditions on the projection vector. A simulation study is also conducted to examine the performance of the proposed estimator.

Comments:	Published in at this http URL the Annals of Statistics (this http URL) by the Institute of Mathematical Statistics (this http URL)
Subjects:	**Statistics Theory (math.ST)**
Journal reference:	Annals of Statistics 2012, Vol. 40, No. 2, 812-831
DOI:	10.1214/12-AOS982
Report number:	IMS-AOS-AOS982
Cite as:	**arXiv:1206.0847 [math.ST]**
	(or **arXiv:1206.0847v1 [math.ST]** for this version)

Submission history

From: Jun Shao [view email]
[v1] Tue, 5 Jun 2012 09:05:01 GMT (897kb)

Which authors of this paper are endorsers?