



An Analysis of Random Design Linear Regression

Daniel Hsu, Sham M. Kakade, Tong Zhang

(Submitted on 13 Jun 2011)

The random design setting for linear regression concerns estimators based on a random sample of covariate/response pairs. This work gives explicit bounds on the prediction error for the ordinary least squares estimator and the ridge regression estimator under mild assumptions on the covariate/response distributions. In particular, this work provides sharp results on the "out-of-sample" prediction error, as opposed to the "in-sample" (fixed design) error. Our analysis also explicitly reveals the effect of noise vs. modeling errors. The approach reveals a close connection to the more traditional fixed design setting, and our methods make use of recent advances in concentration inequalities (for vectors and matrices). We also describe an application of our results to fast least squares computations.

Subjects: **Statistics Theory (math.ST)**; Artificial Intelligence (cs.AI); Learning (cs.LG); Machine Learning (stat.ML)

Cite as: **arXiv:1106.2363 [math.ST]**
(or **arXiv:1106.2363v1 [math.ST]** for this version)

Submission history

From: Daniel Hsu [[view email](#)]
[v1] Mon, 13 Jun 2011 01:08:48 GMT (25kb)

Which authors of this paper are endorsers?

Link back to: [arXiv](#), [form interface](#), [contact](#).

Download:

- [PDF](#)
- [PostScript](#)
- [Other formats](#)

Current browse context:

math.ST

< [prev](#) | [next](#) >

[new](#) | [recent](#) | [1106](#)

Change to browse by:

- cs
 - [cs.AI](#)
 - [cs.LG](#)
- math
- stat
 - [stat.ML](#)

References & Citations

- [NASA ADS](#)

Bookmark([what is this?](#))

