

Linear Regression with Limited Observation

We consider the most common variants of linear regression, including Ridge, Lasso and Support-vector regression, in a setting where the learner is allowed to observe only a fixed number of attributes of each example at training time. We present simple and efficient algorithms for these problems: for Lasso and Ridge regression they need the same total number of attributes (up to constants) as do full-information algorithms, for reaching a certain accuracy. For Support-vector regression, we require exponentially less attributes compared to the state-of-the-art. By that, we resolve an open problem recently posed by Cesa-Bianchi, Shalev-Shwartz and Shamir (2010).

We complement our algorithms for Ridge and Lasso regression with lower bounds that imply that our algorithmic results cannot be improved in general.

Experiments show the theoretical bounds to be justified by superior performance compared to the state-of-the-art.

Joint work with Elad Hazan.