



THE UNIVERSITY OF CHICAGO

Department of Statistics

STATISTICS COLLOQUIUM

NOUREDDINE EL KAROUI

Department of Statistics
University of California, Berkeley

On High-Dimensional Robust Regression and Inefficiency of Maximum Likelihood Methods

THURSDAY, May 28, 2015, at 3:00 PM
Room 112, 5727 S. University Avenue

ABSTRACT

I will discuss the behavior of widely used statistical methods in the high-dimensional setting where the number of observations, n , and the number of predictors, p , are both large. I will present limit theorems about the behavior of the corresponding estimators, their asymptotic risks etc. The results apply not only to robust regression estimators, but also Lasso-type estimators and many much more complicated problems.

Many surprising statistical phenomena occur: for instance, maximum likelihood methods are shown to be (grossly) inefficient, and loss functions that should be used in regression are shown to depend on the ratio p/n . This means that dimensionality should be explicitly taken into account when performing simple tasks such as regression.

More generally, we'll see that intuition based on results obtained in the small p , large n setting leads to misconceptions and the use of suboptimal procedures. It also turns out that inference is possible in this setting. We'll also see that the geometry of the design matrix plays a key role in these problems and use this fact to disprove claims of universality of some of the results.

Mathematically, the tools needed mainly come from random matrix theory, measure concentration and convex analysis. Based on several papers, including some which are joint work with Derek Bean, Peter Bickel, Chinghay Lim and Bin Yu.

For further information and about building access for persons with disabilities, please contact Kirsten Wellman at 773.702.8333 or send email (kwellman@galton.uchicago.edu). If you wish to subscribe to our email list, please visit the following website: <https://lists.uchicago.edu/web/arc/statseminars>.