



THE UNIVERSITY OF CHICAGO

Department of Statistics STATISTICS COLLOQUIUM

ERNST WIT

Statistics and Probability
University of Groningen

Sparse GLMs (With a Bit of Differential Geometry)

WEDNESDAY, May 15, 2013, at 12:00 PM

110 Eckhart Hall, 5734 S. University Avenue

ABSTRACT

Sparsity is an essential feature of many contemporary data problems. Remote sensing, various forms of automated screening and other high-throughput measurement devices collect a lot of information, typically about few independent statistical subjects or units. In certain cases it is reasonable to assume that the underlying process generating the data is itself sparse, in the sense that only a few of the measured variables are involved in the process. We propose an explicit method of monotonically decreasing sparsity for outcomes that can be modelled by an exponential family. In our approach we generalize a so-called equiangular condition in a generalized linear model. Although the geometry involves the Fisher information in a way that is not obvious in the simple regression setting, the equiangular condition turns out to be equivalent with an intuitive condition imposed on the Rao score test statistics. In certain special cases the method can be tweaked to obtain L_1 -penalized GLM solution paths, but that's not the point. The method itself defines sparsity more directly. Although the computation of the solution paths is not trivial, the method compares favorably to other path following algorithms.

This is joint work with Luigi Augugliaro.