Department of Decision Sciences

Statistics Seminars

Penalized likelihood estimation and inference in high-dimensional logistic regression

Ioannis Kosmidis

University of Warwick

Thursday, 22nd February 2024
12:00 pm **Room 3-E4-SR03** Via Roentgen 1 Milano

Abstract

In recent years, there has been a surge of interest in estimators and inferential procedures that exhibit optimal asymptotic properties in high-dimensional logistic regression when the number of covariates grows proportionally as a fraction ($\kappa \in (0,1)$) of the number of observations. In this seminar, we focus on the behaviour of a class of maximum penalized likelihood estimators, employing the Diaconis-Ylvisaker prior as the penalty.

Building on advancements in approximate message passing, we analyze the aggregate asymptotic behaviour of these estimators when covariates are normal random variables with arbitrary covariance. This analysis enables us to eliminate the persistent asymptotic bias of the estimators through straightforward rescaling for any value of the prior hypertuning parameter. Moreover, we derive asymptotic pivots for constructing inferences, including adjusted Z-statistics and penalized likelihood ratio statistics.

Unlike the maximum likelihood estimate, which only asymptotically exists in a limited region on the plane of κ versus signal strength, the maximum penalized likelihood estimate always exists and is directly computable via maximum likelihood routines. As a result, our asymptotic results remain valid even in regions where existing maximum likelihood results are not obtainable, with no overhead in implementation or computation.

The dependency of the estimators on the prior hyper-parameter facilitates the derivation of estimators with zero asymptotic bias and minimal mean squared error. We will explore these estimators' shrinkage properties, substantiate our theoretical findings with simulations and applications, and present evidence for conjectures with different penalties, such as Jeffreys' prior and non-normal covariate distributions.

Department of Decision Sciences
Via Röntgen 1 – 20136 Milano

Tel. 02 5836.5632
Fax 02 5836.5630