Classification of longitudinal data through a semiparametric mixed-effects model based on lasso-type estimators

Cristian Meza[⋆]

CIMFAV-Universidad de Valparaíso

Email: cristian.meza@uv.cl

Abstract: We propose a classification method for longitudinal data. The Bayes classifier is classically used to determine a classification rule where the underlying density in each class needs to be well modeled and estimated. This work is motivated by a real data set of hormone levels measured at the early stages of pregnancy that can be used to predict *normal* versus *abnormal* pregnancy outcomes. The proposed model, which is a semiparametric linear mixed-effects model (SLMM), is a particular case of the semiparametric nonlinear mixed-effects class of models (SNMM) in which finite dimensional (fixed effects and variance components) and infinite dimensional (an unknown function) parameters have to be estimated. In SNMM's maximum likelihood estimation is performed iteratively alternating parametric and nonparametric procedures. However, if one can make the assumption that the random effects and the unknown function interact in a linear way, more efficient estimation methods can be used. Our contribution is the proposal of a unified estimation procedure based on a penalized EM-type algorithm. The Expectation and Maximization steps are explicit. In this latter step, the unknown function is estimated in a nonparametric fashion using a lasso-type procedure. A simulation study and an application on real data are performed.

[⋆]Joint work with A. Arribas-Gil (U. Carlos III de Madrid), R. De la Cruz (PUC) and E. Lebarbier (AgroParisTech).