Threshold Boundary Logistic Regression for Binary Data

Chih-Hao Chang[1], Shih-Feng Huang[2]

[1] National Chengchi University [2]National Central University

Abstract

This talk introduces Threshold Boundary Logistic Regression (TBLR) for binary outcomes, which links covariates to both the logistic component and a linear or nonlinear threshold that partitions the feature space for distinct logistic models. We develop an iterative two-stage sample-splitting estimator that recasts the non-differentiable likelihood into an optimization: threshold parameters are obtained by minimizing a weighted classification loss, while logistic parameters are updated by likelihood maximization. Under suitable conditions, we establish consistency, oracle-optimal convergence rates, and asymptotic normality. Computation uses both Mixed-Integer Programming (MIP) and Weighted SVM (WSVM) as solvers: for linear boundaries we warm-start MIP with a WSVM solution—improving estimation at extra runtime—whereas nonlinear boundaries are solved by WSVM only. Simulations and real-data analyses illustrate finite-sample behavior and feasibility in nonlinear regimes. We further outline an extension to count responses—Threshold Boundary Poisson Regression (TBPR)—adapting the two-stage scheme to the Poisson likelihood and log link, and we will present preliminary empirical analyses that demonstrate applicability and the modeling workflow for count data.

Keyword: Maximum likelihood (ML), mixed integer programming (MIP), two-stage sample-splitting, weighted support vector machine (WSVM).