Statistica Sinica **34** (2024), 1-28
doi:https://doi.org/10.5705/ss.202022.0103

MULTIVARIATE VARYING-COEFFICIENT MODELS VIA TENSOR DECOMPOSITION

Fengyu Zhang[#], Ya Zhou[#], Kejun He* and Raymond K. W. Wong

Renmin University of China, Chinese Academy of Medical Sciences and Peking Union Medical College, Renmin University of China and Texas A&M University

Abstract: Multivariate varying-coefficient models are popular statistical tools for analyzing the relationship between multiple responses and covariates. Nevertheless, estimating large numbers of coefficient functions is challenging, especially with limited samples. In this work, we propose a reduced-dimension model based on the Tucker decomposition that unifies several existing models. In addition, we use sparse predictor effects, in the sense that only a few predictors are related to the responses, to achieve an interpretable model and sufficiently reduce the number of unknown functions to be estimated. These dimension-reduction and sparsity considerations are integrated into a penalized least squares problem on the constraint domain of third-order tensors. To compute the proposed estimator, we propose a block updating algorithm based on the alternating direction method of multipliers and manifold optimization. We also establish the oracle inequality for the prediction risk of the proposed estimator. A real data set from the Framingham Heart Study is used to demonstrate the good predictive performance of the proposed method.

Key words and phrases: Dimensionality reduction, group Lasso, polynomial splines, sparsity, Tucker low rank.

1. Introduction

Varying-coefficient models (VCMs, Hastie and Tibshirani (1993)) are popular structured regression models that have reasonably flexible nonparametric components and can be estimated well with a moderate amount of data (Ruppert, Wand and Carroll (2003)). In VCMs, the regression coefficients of the predictors vary with an observable exposure variable. VCMs have been studied extensively in literature and are widely used in practice; see, for example, Hoover et al. (1998), Huang, Wu and Zhou (2002), Park et al. (2015), and the references therein. For settings with a large number of predictors (possibly larger than the sample size), Wang, Li and Huang (2008) use basis function expansions and the smoothly clipped absolute deviation (SCAD) penalty to address the problem of variable selection. Wei, Huang and Li (2011) and Lian (2012) apply an adaptive group least absolute shrinkage and selection operator (lasso) and spline function

[#]The first two authors contributed equally to this work.
*Corresponding author.