Statistica Sinica **35** (2025), 1-18
doi:https://doi.org/10.5705/ss.202023.0151

MULTIVARIATE CALIBRATIONS WITH AUXILIARY INFORMATION

Pei Feng Tong, Song Xi Chen* and Cheng Yong Tang

Peking University, Tsinghua University and Temple University

Abstract: We investigate multivariate calibrations from a modern perspective with a focus on incorporating auxiliary variables and handling complex data dependencies with random effects. By introducing auxiliary variables, the roles of the variables in multivariate calibration problems are no longer restricted to being either response or explanatory, which offers much flexibility and adaptability to a broader range of practical problems. Our analysis reveals that a new shrinkage approach, that connects the conventional generalized least squares and the inverse regression approaches, offers much improved performance. To accommodate complex dependence in contemporary studies, we develop a computationally efficient expectation-maximization algorithm for solving multivariate calibration problems with random effects. The shrinkage approach shows promising performance in numerical simulations and an empirical study.

Key words and phrases: Inverse regression, linear mixed-effect models, multivariate calibration, multivariate response variables, shrinkage estimation.

1. Introduction

Regression analysis is a class of foundational statistical tools for modeling and predictions. In regression analysis, the response variables are modeled by a collection of explanatory variables. Practically, interests may arise from the opposite direction, concerning some or all of the variables whose roles are explanatory rather than response in some established regression models. Such problems, referred to as calibrations or inverse predictions in the literature, are often seen in practical investigations in areas including economics, sociology, earth sciences and analytical chemistry (Marden et al., 2018; Yun et al., 2019; Wei et al., 2021). In this study, the term "calibration" particularly refers to inversely predicting some or all of the explanatory variables in the framework of regression models. It is worth noting that the term "calibration" is employed in other contexts such as to describe the adjustment that correcting the coverage probabilities in interval estimations and the alignment of parameters in generic models like regression quantiles (Kuleshov, Fenner and Ermon, 2018; Fasiolo et al., 2021).

Multivariate calibration problems have been extensively investigated in the

*Corresponding author. E-mail: sxchen@tsinghua.edu.cn