



# Group least squares regression for linear models with strongly correlated predictor variables

Min Tsao<sup>1</sup>

Received: 3 September 2021 / Revised: 15 February 2022 / Accepted: 1 June 2022 /  
Published online: 26 July 2022

© The Institute of Statistical Mathematics, Tokyo 2023, corrected publication 2022

## Abstract

Traditionally, the main focus of the least squares regression is to study the effects of individual predictor variables, but strongly correlated variables generate multicollinearity which makes it difficult to study their effects. To resolve the multicollinearity issue without abandoning the least squares regression, for situations where predictor variables are in groups with strong within-group correlations but weak between-group correlations, we propose to study the effects of the groups with a group approach to the least squares regression. Using an all positive correlations arrangement of the strongly correlated variables, we first characterize group effects that are meaningful and can be accurately estimated. We then discuss the group approach to the least squares regression through a simulation study and demonstrate that it is an effective method for handling multicollinearity. We also address a common misconception about prediction accuracy of the least squares estimated model.

**Keywords** Strongly correlated predictor variables · Multicollinearity · Group effects · Linear models · Least squares regression

---

✉ Min Tsao  
mtsao@uvic.ca

<sup>1</sup> Department of Mathematics and Statistics, University of Victoria, Victoria, BC V8W 2Y2, Canada