Estimation and inference for high dimensional, doubly-structured regression models

Classical multiple regression models lack a unique solution in the high dimensional (*p* > *n*) case. This issue has been addressed by popular regularized regression models, such as the LASSO and Ridge regression. However, these models are unable to incorporate additional structural information into the fitting procedure, which may be a disadvantage when variables in the model are highly correlated. Kernel Penalized Regression (KPR) is an extension of Ridge regression that fits models with respect to variable and subject structure. We have developed an R software package for fitting KPR models and performing inference on the estimated coefficients. To fit model tuning parameters, we have implemented both cross validation and Restricted Maximum Likelihood estimation via a mixed model formulation. An analysis of microbiome bacterial abundance data with the KPR model showed that the inclusion of structural data can improve model interpretability and statistical power. Our package provides a simple interface for fitting, interpreting, and performing inference on KPR models, as well as functions for data processing and preparation.