Category Tree Gaussian Process for Computer Experiments with Many-Category Qualitative Factors and Application to Cooling System Design

Ray-Bing Chen

Department of Statistics & Institute of Data Science, National Cheng Kung University

E-mail: rbchen@ncku.edu.tw

Abstract

In computer experiments, Gaussian process (GP) models are commonly used for emulation. However, when both qualitative and quantitative factors are in the experiments, emulation using GP models becomes challenging. In particular, when the qualitative factors contain many categories in the experiments, existing methods in the literature become cumbersome due to the curse of dimensionality. Motivated by the computer experiments for the design of a cooling system, a new tree-based GP is proposed that emulates computer models with many-category qualitative factors, which we call category tree GP. The proposed method incorporates a tree structure to split the categories of the qualitative factors, and GP or mixed-input GP models are employed for modeling the simulation outputs in the leaf nodes. The splitting rule takes into account the cross-correlations between the categories of the qualitative factors, which have been shown by a recent theoretical study to be a crucial element for improving the prediction accuracy. In addition, a pruning procedure based on the cross-validation error is proposed to ensure the prediction accuracy. The application to the design of a cooling system indicates that the proposed method not only enjoys marked computational advantages and produces accurate predictions, but also provides valuable insights into the cooling system by discovering the tree structure.