Predictive Subdata Selection for Large-Scale Deterministic Computer Models

Ming-Chung Chang  
*Institute of Statistical Science, Academia Sinica*

**Abstract**

Computer models are implementations of complex mathematical models using computer codes. Tremendous amounts of data generated from computer models are becoming ubiquitous owing to advanced technology. Such data richness, however, may yield an inability to conduct statistical analysis in terms of the time cost. Recently, increased attention has focused on solving this data reduction problem. In this talk, I will introduce a new subdata selection method for large-scale deterministic computer models. In addition to the geometry of the input space, the proposed method takes advantage of the information of the output values and adaptively updates the current subdata with affordable computational cost. Simulated examples and real data analyses are provided.

**Keywords:** Computer experiment, Subsampling, Data reduction.