

STATISTICS COLLOQUIUM

Friday, November 13, 2009

Time: 1 PM

Location: McCarthy Hall, Room 380

“Semiparametric Modeling and Adaptive Sampling Using Treed Gaussian Processes”

Herbie Lee

Department of Applied Mathematics and Statistics, UC Santa Cruz

ABSTRACT

This talk presents a flexible methodology which is applicable to a wide range of nonparametric and semiparametric heteroscedastic regression problems. The work was motivated by a collaboration with NASA involving a computational fluid dynamics simulation of a rocket booster as it re-enters the atmosphere. The simulator predicts the flight behavior of the reentry vehicle under a variety of situations, but can take a long time to run. The statistical goal is to map the response surface accurately and efficiently. This entails both a method for nonstationary spatial modeling (to create the map) and a method for adaptive sampling (to choose where to run the simulator to gain the most information to build the map). We provide a general methodology for modeling and adaptive sampling to greatly speed up such surrogate modeling. Binary trees are used to recursively partition the input space, and Gaussian process models are fit within each partition. Trees facilitate non-stationarity and a Bayesian interpretation provides a measure of uncertainty in the sample space which can be used to guide future sampling. The framework is also expanded to address optimization problems (the finding of a global maximum or minimum).