University of California, Irvine
Statistics Seminar

Orthogonality Meets Marginalization: Fast and Flexible Gaussian Processes for Large Correlated Data

Mengyang Gu
UC Santa Barbara

4 p.m., Thursday, May 7, 2020
Join via Zoom: https://uci.zoom.us/j/700695554
Meeting ID: 700 695 554

Join via Skype for Business:
https://uci.zoom.us/skype/700695554

We will first introduce the generalized probabilistic principal component analysis (GPPCA) to study the latent factor model for multiple correlated outcomes, where each factor is modeled by a Gaussian process. Our method generalizes the previous probabilistic formulation of PCA (PPCA) by providing the closed-form maximum marginal likelihood estimator of the factor loadings and predictive distribution, when factors are Gaussian processes. In the second half of the talk, we will discuss the extension of the GPPCA on large incomplete matrices of correlated data. We decompose the Gaussian random field with multi-dimensional input domain into the product of orthogonal components with lower dimensional inputs. As various Gaussian processes with one dimensional input can be written as stochastic differential equations, we are able to apply the continuous-time Kalman filter to compute the likelihood and predictive distribution with linear computational complexity without any approximation. Analysis of simulated data and real spatial data, spatio-temporal data and functional data confirms excellent performance of our approach.