

USC Department of Mathematics
PROBABILITY & STATISTICS SEMINAR

3:30 PM, Friday 27.Mar.09
146 Kaprielian Hall

Julien Emile-Geay
Department of Earth Sciences
USC

Imputation of missing values in geophysical datasets: An improved, data-adaptive regularization scheme

Geophysical data are often incomplete, by virtue of the scale and non-repeatability of the underlying physical measurements. Several attempts have been made to impute the missing values, the most successful to date being a regularized version of the Expectation-Maximization algorithm (RegEM) [Schneider, 2001]. So far, two regularization schemes have been offered: Ridge regression (AKA Tikhonov regularization) and Truncated Total Least Squares (TTLS). In this talk I will describe iTTLS, a new regularization scheme for the RegEM algorithm. The method combines the most desirable features of previously implemented schemes: Like the ridge-regression method, it is data-adaptive; like the truncated total least squares method, it preserves exactly the variance of the leading eigenvalues of the covariance matrix. These two properties jointly give rise to an objective imputation method whose accuracy greatly exceeds that of ridge-regression over a broad range of incompleteness levels. An application to the infilling of global, monthly sea-surface temperature data over 1850-2008 is presented, using synthetic and real-world data.