DENSITY ESTIMATION VIA OPTIMAL SEGMENTATION

J. Scargle[1], B. Jackson[2]
(1) NASA Ames Research Center
(2) San Jose State University
Jeffrey.D.Scargle@NASA.GOV

May 11, 2005

Abstract

Many density estimation and signal processing problems can be solved by maximizing the fitness of a segmented model over all possible partitions of the data interval. For one dimensional problem, we have discovered [1,2] a simple but powerful algorithm that searches the exponentially large space of partitions of N data points in time $O(N^2)$. This algorithm is guaranteed to find the exact global optimum, automatically determines the model order (the number of segments), has a convenient real-time mode, and solves a surprising variety of problems in signal detection and characterization, density estimation (including the construction of data-adaptive histograms), cluster analysis and classification. This talk will emphasize recent work extending the formalism to higher dimensional data spaces, and error analysis using bootstrap and model-averaging procedures.

References:

[1] B. Jackson, J. Scargle, et al., IEEE Signal Processing Letters, **12**, 105 (2005).

[2] J. Scargle, et al., "Studies in Astronomical Time Series Analysis: VI. Optimum Partition of the Interval: Bayesian Blocks, Histograms and Triggers," in preparation (http://astrophysics.arc.nasa.gov/ jeffrey/).

Key Words: OPTIMAL SEGMENTATION, DENSITY ESTIMATION, BAYESIAN BLOCKS