Change-point Problems
IMS Lecture Notes – Monograph Series (Volume 23, 1994)

CUBE SPLITTING IN MULTIDIMENSIONAL EDGE ESTIMATION

BY H. G. MÜLLER AND K.-S. SONG
University of California, Davis

Assume noisy measurements are available and that an edge or boundary is given which induces a partition of the domain into two subsets. The regression function on one subset is equal to a constant c_1, on the other subset to a constant c_2. Each measurement is made within a regular pixel. The problem we consider is the estimation of the edge or boundary curve (change curve), for the case that the domain is in $\Re^d$. We propose to seek boundary estimates as maximizers of a weighted squared difference statistic where we maximize over unions of cubes of aggregated pixels. Rates of almost sure convergence of this procedure are established. Its central advantage is its numerical feasibility, as the number of cubes of aggregated pixels to be investigated for inclusion in one of the partitioning sets can be kept small. A numerically efficient "cube splitting" ("CUSP") algorithm is suggested which implements this proposal: Start with an iteratively grown union of big cubes of aggregated pixels to find a first approximate edge/boundary estimate on a coarse level of approximation. Then split those cubes falling near the boundary into smaller cubes and check their allocation to one of the partitioning sets in order to obtain a more refined boundary estimate. This cube splitting (refinement) step may then be iterated until the desired level of resolution is achieved.

1. Introduction. Our main concern in this article is a numerically efficient way of estimating edges, i.e., discontinuities, of a regression function in higher dimensions. We also discuss asymptotic rates of almost sure convergence for one such estimation method. Our basic idea is to aggregate data into larger blocks ("cubes") and to assume that the true "edge" is also anchored on such larger blocks. The edge to be estimated therefore depends on the sample size n, reflecting the intuition that increasing sample size should allow for increasing degrees of resolution of the edge estimate.

The problem of edge estimation in higher dimensions recently found some interest among statisticians (see for instance the various approaches discussed

---

AMS 1991 Subject Classification: Primary 62G05; Secondary 62J02, 62H11

Key words and phrases: Break curve, change-point, discontinuity, image analysis, jump size, multivariate regression, pixels, rate of convergence, symmetric difference.