Hall, Peter

On projection pursuit regression. (English) Zbl 0698.62041

Ann. Stat. 17, No. 2, 573-588 (1989).

A mathematical model of projection pursuit regression based on kernel estimation (of "marginal density of errors in projection direction") is proposed and the necessary mathematical calculus for projection approximation of the target function $G(X) = EY$ is built up. The main results are explicit formulae for bias and error about the mean in orientation estimates and curve estimates. They show that the estimate of orientation (of projection) has most its error in the form of bias.

They also prove that the kernel-based projection pursuit regression does estimate corresponding projections with convergence rates identical to those known from one-dimensional estimation, namely $O(h^r) = O((nh)^{-1/2})$ (h being the bandwidth of kernel estimator). The estimator of G(x) based on projections to the direction θ (say $\hat{G}_\theta(x)$) is required to minimize

$$\hat{S}(\theta) = n^{-1} \sum_{k=1}^{n} (Y_k - \hat{G}_\theta(x_k))^2$$

(where $(Y_k, X_k)_{k=1}^n$, $Y_k \in R$, $X_k \in R^p$ are data). The main idea how to construct $\hat{G}_\theta(x)$ is to do it through estimating θ by $\hat{\theta}$ which minimizes

$$\tilde{S}(\theta) = n^{-1} \sum_{k=1}^{n} (Y_k - \hat{G}_\theta^k(x_k))^2$$

where $\hat{G}_\theta^k(x_k)$ is the nonparametric (kernel) estimate of G(x) based on all points $(x_i)_{i=1}^n$ except of x_k. At the end, some alternative approaches (with randow window, two-stage algorithm etc.) are discussed.

Reviewer: J.Á.Víšek

MSC:

62G05 Nonparametric estimation
62H99 Multivariate analysis
62H05 Characterization and structure theory for multivariate probability distributions; copulas

Keywords:

nonparametric regression; projection pursuit regression; kernel estimation; marginal density of errors in projection direction; projection approximation; explicit formulae; bias; orientation estimates; curve estimates; convergence rates; randow window; two-stage algorithm

Full Text: DOI