EXTREMAL FITS IN REACT CONFIDENCE SETS

RUDOLF BERAN[1]

University of California, Berkeley

REACT estimators use ideas from signal processing, model-selection, and shrinkage to achieve much smaller risk in one-way layouts and other linear models than does the classical least squares estimator. The REACT method can generate automatic scatterplot smoothers that compete well on standard data sets with the best fits obtained by other methods. This paper addresses two further questions: Which features in a REACT estimator are not necessarily present in the true mean vector; and which features of the true mean vector might have been smoothed out by the REACT estimator? We answer both questions by constructing extremal members of a confidence set of asymptotic coverage probability α that is centered at the REACT estimator. The methodology is demonstrated on two data-sets from the smoothing literature.

AMS subject classifications: 62J05, 62G07.
Keywords and phrases: economical basis, risk estimation, adaptation, superefficiency, saturated fit, unsaturated fit.

1 Introduction

Consider the Gaussian linear model in which the $n \times 1$ response vector y has a $N(X\beta, \sigma^2 I_n)$ distribution, the regression parameters β and the variance σ^2 being both unknown. Suppose that the $n \times p$ regression matrix X has full rank $p \leq n$. The least squares estimator of the mean $\eta = \mathrm{E}(y) = X\beta$ is then $\hat{\eta}_{LS} = X(X'X)^{-1}X'y$. Under normalized quadratic loss, the risk of an estimator $\hat{\eta}$ of η is $p^{-1}\mathrm{E}|\hat{\eta} - \eta|^2$. This risk is precisely σ^2 for the least squares fit $\hat{\eta}_{LS}$.

Stein (1956) proved the inadmissibility of the least squares fit $\hat{\eta}_{LS}$ to the Gaussian linear model when dimension of the regression space exceeds 2. This defect in least squares becomes intuitively clear in special cases. Consider the one-way layout with one observation per cell (e.g. a digitized signal observed in white noise) or the two-way layout with one observation per cell (e.g. a digitized image observed in white noise). In such examples, the least squares estimator of the signal η is the raw data y. It is not

[1] This research was supported in part by National Science Foundation Grant DMS95-30492.