approximation, we need n operations to discretize the data into N_B nonempty bins. Thus, the numerical effort for this method is of order $O(n + N_B M)$.

Of course, the WARPing method introduces a discretization bias. The bias may be reduced by joining the obtained discrete step function (see (3.2)), via a polygon. Breuer (1990) has computed for $m(x) = x\sin(2\pi x) + 3x$ and uniform design the MSE as a function of x for both the $\hat{m}_E$ estimator and the WARPed estimator $\hat{m}_{M,K}$.

In Figure 5, the discretization bias is seen to be quite drastic, although we gained in speed of computation. The linear interpolant has a much better bias behavior, as is seen in Figure 6. For this estimator conservative bounds for the numerical discretization error and its effect on $MSE(x)$ can be given and are displayed in Figure 6 as long dashed lines.

ACKNOWLEDGMENT

The work of the first author was in part financially supported by CentER, Tilburg.

Comment

Jeffrey D. Hart

Chu and Marron have provided us with a clear and thorough account of the relative merits of evaluation and convolution type kernel regression estimators. One is left with the impression that neither type of estimator is to be preferred universally over the other. We learn, for example, that the weights of the convolution estimator sometimes have the unsettling behavior exhibited in Figures 6b and 7 of Chu and Marron. The authors make it clear that there are a number of factors, including type of design (fixed or random), design density and nature of underlying regression function, that need to be considered before choosing an estimator type. Having reading their article, I now have a slight preference for $\hat{m}_E$ over $\hat{m}_C$ in the random design case, at least in the absence of any information about the design density or regression curve. When the design points are nonrandom and evenly spaced, I prefer $\hat{m}_C$, since its convolution form appeals to me, and since boundary kernels are easy to construct with $\hat{m}_C$ (see Gasser and Müller, 1979). Below I will mention a modification of $\hat{m}_C$ that I feel is a viable competitor of $\hat{m}_E$ even in the random design case.

The authors' point about the down weighting phenomenon of the convolution estimator is certainly well taken. However, I would like to question an aspect of their comparison of the variances of $\hat{m}_E$ and $\hat{m}_C$. As the authors note in Section 4, the biases of the two estimators are not comparable, the bias of $\hat{m}_E$ being smaller in some cases and that of $\hat{m}_C$ smaller in other cases. It follows that "good" bandwidths for the estimators will generally be different. Why then is it sensible to compare $\mathrm{Var}(\hat{m}_E)$ and $\mathrm{Var}(\hat{m}_C)$ at the same value of h?

A little-used but informative way of comparing the errors of $\hat{m}_E$ and $\hat{m}_C$ is to consider the limiting distribution of

$$\frac{|\hat{m}_E(x) - m(x)|}{|\hat{m}_C(x) - m(x)|}. \tag{1}$$

Unlike an MSE comparison, this approach takes into account the joint behavior of the two estimators. Suppose that Chu and Marron's assumptions (A.1)–(A.5) hold and that the design density is $U(0,1)$. Suppose further that the bandwidths of $\hat{m}_E$ and $\hat{m}_C$ minimize their respective MSEs. Then it can be shown that, for each x, the ratio (1) converges in distribution to

$$\left(\frac{2}{3}\right)^{2/5} \frac{|Z_1 + 1/2|}{|Z_2 + 1/2|} = R \tag{2}$$

as $n \to \infty$, where (Z_1, Z_2) have a bivariate normal distribution with $Z_1 \sim N(0,1)$, $Z_2 \sim N(0,1)$ and

$$\mathrm{Corr}(Z_1, Z_2) = \left(\frac{2}{3}\right)^{3/5} \int K(z) K\left(\left(\frac{2}{3}\right)^{1/5} z\right) dz \Big/ \int K^2 = \rho_K.$$

Jeffrey D. Hart is Associate Professor of Statistics, Texas A&M University, College Station, Texas 77843.