BUJA, A., HASTIE, T. and TIBSHIRANI, R. (1988). Linear smoothers and additive models (with discussion). *Ann. Statist.* To appear.
CLEVELAND, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *J. Amer. Statist. Assoc.* **74** 829–836.
CLEVELAND, W. S. and DEVLIN, S. (1988). Locally -weighted regression: An approach to regression analysis by local fitting. *J. Amer. Statist. Assoc.* **83** 596–610.
DRAPER, D. (1987). Comment on "Prediction in growth curves" by C. R. Rao. *Statist. Sci.* **2** 454–461.
FRIEDMAN, J. H. and STUETZLE, W. (1981). Projection pursuit regression. *J. Amer. Statist. Assoc.* **76** 817–823.
FRIEDMAN, J. and TIBSHIRANI, R. (1984). The monotone smoothing of scatterplots. *Technometrics* **26** 243–250.
HODGES, J. S. (1987). Uncertainty, policy analysis and statistics (with discussion). *Statist. Sci.* **2** 259–291.
TIBSHIRANI, R. J. (1988). Estimating transformations for regression via additivity and variance stabilization. *J. Amer. Statist. Assoc.* **83** 394–405.
WAHBA, G. (1983). Bayesian "confidence intervals" for the cross-validated smoothing spline. *J. Roy. Statist. Soc. Ser. B* **45** 133–150.

Comment

Grace Wahba

Professor Ramsay is to be congratulated for writing a lively and interesting paper and giving us a handy descriptive tool.

Without at all intending to criticize "eyeball" methods, which play an important role in data analysis, it should be clear that the success of the method depends on the ability of the user to select the number and location of the knots to give a pleasing picture. As the author observes, in the examples given, the results are fairly insensitive to knot location. This is, of course why it is difficult to select knots in the computer by an objective numerical criterion—numerically, that is an ill-posed problem. I would expect that the picture would be different if the number of knots is changed drastically.

Subjective notations of what the answer "ought to" look like appear to play an important role in the proposed method.

Having said this, I would like to raise the issue of "subjective" versus "objective" inference, both of which clearly play a role in statistics. Of course, the dividing line between these types of inference are blurred, every "objective" method has some subjectivity behind it, namely, the statistician had some preconceived framework about the truth when selecting a technique (no matter how "objective" the technique is). Conversely, any good "subjective" method, ideally will display the data in such a way that the "facts" about the truth are helped to come out.

One way of classifying subjective versus objective techniques is the following. A technique may be viewed to be on the objective end of the spectrum, if at least in principle, one could discover its properties on at least some useful class of "truths," by simulating data from various "truths," applying the technique, and studying how well the inference matched the simulated "truth." If I had an "objective" method for constructing confidence intervals or estimating variances, then I could run a big Monte Carlo study and see whether in fact the confidence intervals or estimated variances had an appropriate relation to the simulated "truth."

I am somewhat concerned here with the use of, for example, the "estimated sampling variance" of a. It appears that these estimates are conditioned on certain subjective choices made by the statistician. If I really wanted to claim that these estimates had some objective properties if used in the future, I should do a simulation study, sampling from a population of users who are going to use the eyeball method for choosing the location and number of knots.

On a different tack, I would like to thank Professor Ramsay for his kind reference to my work on smoothing splines and to take this opportunity to compare and contrast smoothing and regression splines. Positivity and monotonicity can also be imposed on smoothing splines (see Villalobos and Wahba (1987) and references cited there), and there is quite a bit of activity in the development of efficient alogrithms for doing this, but, in the absence of user-oriented software, it is work to start from scratch to implement a relatively objective constrained smoothing spline as described in Villalobos and Wahba (1987).

The monotone regression splines, as proposed by Professor Ramsay, appear to be quite accessible to relatively unsophisticated users who know how to call a quadratic programming algorithm.

In examples with larger data sets, smoothing splines do have the ability to resolve finer structure than

Grace Wahba is Bascom Professor of Statistics, Department of Statistics, University of Wisconsin, 1210 West Dayton Street, Madison, Wisconsin 53706. These comments were written while the author was Clare Boothe Luce Visiting Professor of Statistics at Yale University.

regression splines because the optimum number of knots (the smoothing parameter) in regression splines can be quite low—$O(n^{1/5})$ under some assumptions; but, of course, in the applications described here, discovering fine structure is not the goal. Despite this remark, we looked at Table 1 with its data on gas consumption (city gas, y), displacement (x_1) and weight (x_2) and became tempted to try an additive smoothing spline model, to see how the results would compare with the monotone regression splines presented by the author. We have to admit that this temptation was fueled by the recent paper (Gu, Bates, Chen and Wahba, 1988) which provides code allowing the efficient objective choice of several smoothing parameters simultaneously by generalized cross-validation (GCV).

The additive (cubic) smoothing spline is a function $g(z_1, z_2)$ of the form

$$g(z_1, z_2) = \mu + f_1(z_1) + f_2(z_2)$$

with

$$\int f_1(z_1)\, dz_1 = \int f_2(z_2)\, dz_2 = 0$$

and f_1, f_2 in the Sobolev space W_{2m}, which minimizes

$$\frac{1}{n}\sum_{i=1}^{n}\left(y_i - (\mu + f_1(z_1(i)) + f_2(z_2(i)))\right)^2 + \lambda_1 \int (f_1''(z_1))^2\, dz_1 + \lambda_2 \int (f_2''(z_2))^2\, dz_2.$$

(See Hastie and Tibshirani, 1987; Barry, 1986; Wahba, 1986; Gu, Bates, Chen and Wahba, 1988.) The two smoothing parameters λ_1 and λ_2 can be chosen by generalized cross-validation by finding λ_1 and λ_2 to minimize the GCV function.

$$V(\lambda_1, \lambda_2) = \frac{(1/n)\,\|(I - A(\lambda_1, \lambda_2))y\|^2}{(1/n)\ \mathrm{Trace}\ (I - A(\lambda_1, \lambda_2))^2},$$

where $A(\lambda_1, \lambda_2)$ is the influence matrix for the problem. If λ_1 and λ_2 should turn out to be ∞, then a linear model has been fitted.

In preparation for fitting an additive smoothing spline model we first, in Figure 1a plotted the points (x_1, x_2) and observed (as could be ascertained from the text) that x_1 and x_2 are highly correlated. (There are only 41 points visible because of replications). We made the transformation to the canonical variables z_1 and z_2 (using S) which resulted in

$$z_1 = .393x_1 + .919x_2, \qquad z_2 = -.919x_1 + .393x_2.$$

z_2 is plotted against z_1 in Figure 1b.

The "color" coding of the dots codes the response y, with open circles having the smallest values of y, circles with small dots next largest, circles with big dots next and filled circles the largest. Increasing response along the diagonal in Figure 1a is obvious. In Figure 2a we plot y_1 versus z_1, and in Figure 2b we plot y_2 versus z_2. Figures 1b, 2a and 2b are almost "sufficient" for the data. For example, the circled point in Figure 1b is also circled in Figure 2a, so that if one visualizes a cube with the 44 points (y, x_1, x_2) in it, Figure 1b is looking down at the top, Figure 2a is looking into the face perpendicular to the z_2 axis and Figure 2b is looking into the face perpendicular to the z_1 axis.

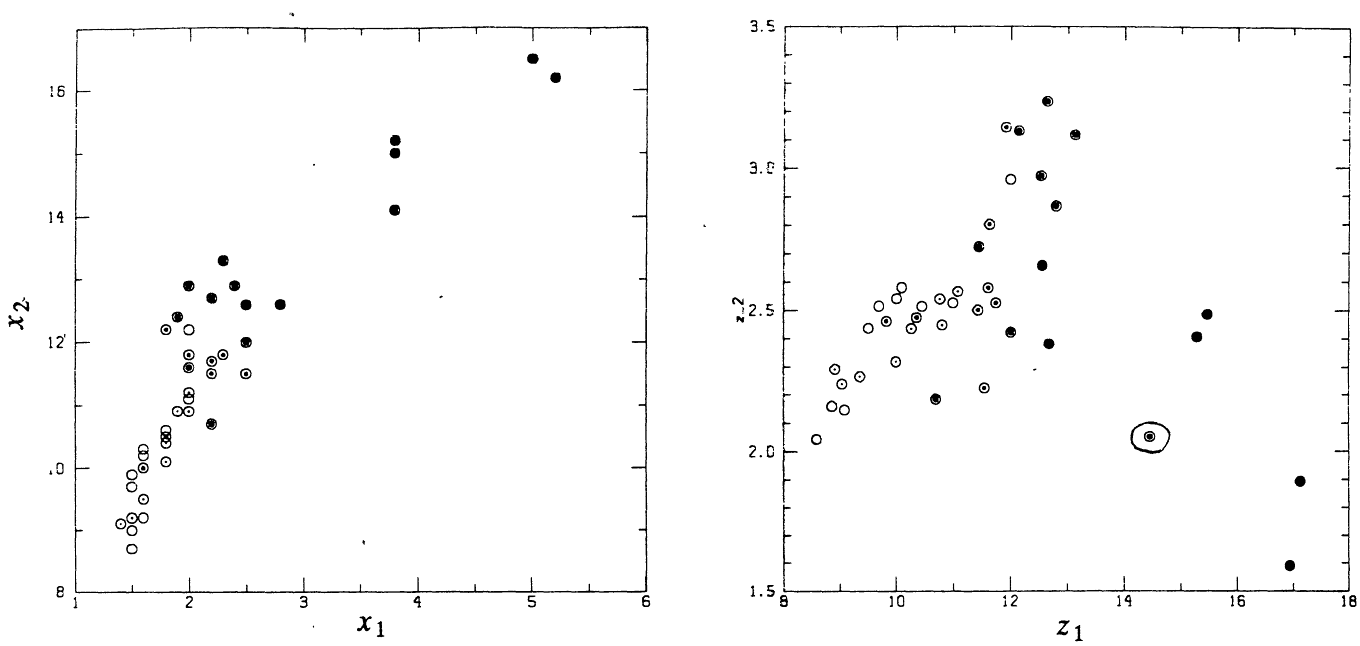

FIG. 1. *The independent variables.* a, *original variables;* b, *canonical variables.*

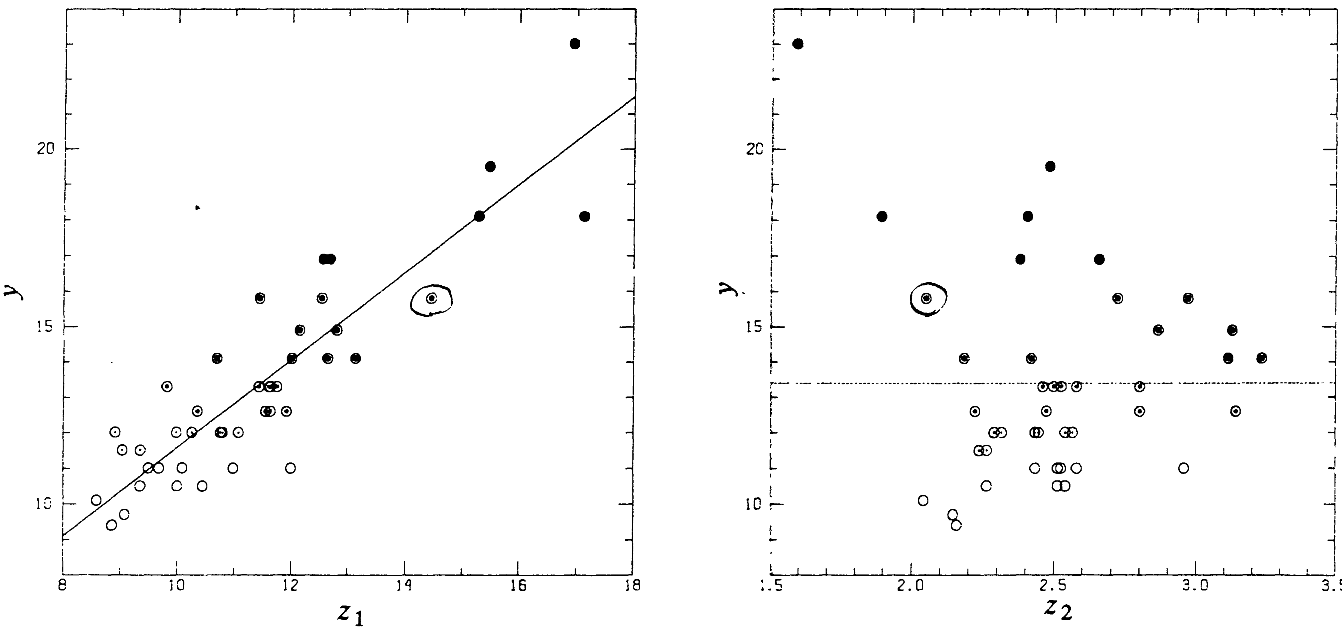

FIG. 2. *The responses.* a, *versus* z_1; b, *versus* z_2.

The line in Figure 2a represents the least squares straight line regression of y onto z_1 and the line in Figure 2b is the mean.

Considering that these data didn't really come from a normal population (etc.) but that the scatter is most likely due to auxiliary design variables not accessible here (rather than "error"), we weren't sure that much more structure would (or even should) be extracted from this data set.

Nevertheless, I persuaded Chong Gu to fit an additive smoothing spline model to this data (covering the rectangle in Figure 1b) and the GCV function $V(\lambda_1, \lambda_2)$ did indeed have a (local) minimum at (∞, ∞), which is equivalent to identifying the model as linear in z_1 and z_2. Unfortunately, V also had a (global) minimum at values of λ_1 and λ_2 which led to a nonsense function which was attempting to follow the data in Figure 2b much more closely than a reasonable person would like. This case and the other "nonsense" pictures generated by other local minima could have been eliminated by eyeball, leaving us with the straight lines in Figure 2. However, this exercise just brought home to us the general difficulty in obtaining "objective" nonparametric function estimates with relatively small data sets (44 is not large for two independent variables) whereas reasonable-looking descriptive estimates like the I-splines or even like the least square lines of Figure 2 can be obtained if human intervention is part of the estimation process. We should, however, be reticent about ascribing "statistical" properties to subjective estimates in examples where they don't really have them.

Again, we thank the author for an interesting contribution.

ADDITIONAL REFERENCES

BARRY, D. (1986). Nonparametric Bayesian regression. *Ann. Statist.* **14** 934–953.

GU, G., BATES, D. M., CHEN, Z. and WAHBA, G. (1988). The computation of GCV functions through Householder tridiagonalization with application to the fitting of interaction spline models. Technical Report 823, Dept. Statistics, Univ. Wisconsin-Madison.

VILLALOBOS, M. and WAHBA, G. (1987). Inequality constrained multivariate smoothing splines with application to the estimation of posterior probabilities. *J. Amer. Statist. Assoc.* **82** 239–248.

WAHBA, G. (1986). Partial and interaction splines for the semiparametric estimation of functions of several variables. In *Computer Science and Statistics: Proceedings of the 18th Symposium on the Interface* (T. E. Boardman, ed.) 75–80. Amer. Statist. Assoc., Washington.