

LOCAL AND DELETION INFLUENCE

By

A.J. Lawrance

IMA Preprint Series # 731

November 1990

LOCAL AND DELETION INFLUENCE

A.J. LAWRANCE*

Abstract. The local influence approach of Cook (1986) to regression diagnostics is developed and discussed, and compared with Cook's (1977) deletion approach. Similarities between the two approaches are stressed, as well as the more specific nature of the local influence approach, and its ability to handle cases simultaneously. The perturbation ideas of the approach are applied to the linear model assumptions of constant variance and independence, and contrasted to perturbations of the response and explanatory values. The assessment of deletion and local influence is reviewed. Some illustrations are given.

Key words. case deletion, Cook's distance, curvature, diagnostics, linear model, local influence, perturbation, regression

1. INTRODUCTION

A practical and well-established approach to influence analysis in statistical modelling is based on case deletion; Cook (1977) pioneered the idea. The effect or influence of the i th case of the data is measured by a comparison of parameter estimates before and after deletion of the i th case. The approach is also central to the work of Belsley, Kuh and Welsh (1980). The idea of differentiation instead of deletion is prominent in the local influence approach of Cook (1986), as it was in the works by Belsley et. al. (1980) and Pregibon (1981). Local influence is based on the perturbation of a case and not on its total deletion, and employs a differential comparison of parameter estimates before and after perturbation. There are several parallels between the approaches which will be stressed; also there are some important advantages of the local influence approach which will be highlighted. A particular parallel is that case deletion can often be regarded as an infinite perturbation to a single variance, but finite perturbations to other aspects can also be employed to reflect other and more specific concerns.

This paper will begin with an introduction to local influence, inspired by Cook's (1986) work, but mainly in the linear regression setting and with a weaker reliance on geometrical curvatures. Rather, the natural perturbation counterpart of Cook's distance is employed, hopefully making local influence as appealing as deletion influence. Several schemes of perturbation are considered, with some emphasis on the assessment of the effects of perturbations. Two data sets are used to illustrate the theoretical developments; one of these leads to suggesting that a scale modified Cook's distance can be of some advantage.

2. IDEAS OF LOCAL INFLUENCE

We suppose a statistical model with parameters β of interest and a set of n response data y to which the model will be fitted; covariate data and other parameters may also be present. The log likelihood or log profile likelihood, as the case may be, of parameters β will be denoted by $L(\beta|y)$.

*School of Mathematics and Statistics, University of Birmingham, Birmingham, B15 2TT, UK

We next consider perturbations, and three sorts will be logically distinguished: Perturbations to assumptions, perturbations to data values and perturbations to case weights (including deletion with its zero case weight). All models involve simplifying assumptions – the constancy of error variances and the independence of data cases are considered here, the two main ones in regression. When perturbing an assumption, a key requirement is that the parameters of the perturbed model should retain their original meanings and importance. Suppose the perturbations are w ; the w are mathematical quantities introduced into the model, not parameters to be estimated and not data. Parameter estimates, usually maximum likelihood, are then obtained as functions of the perturbation, say as $\hat{\beta}(w)$; the null perturbation is w_0 , with $\hat{\beta}(w_0) = \hat{\beta}$ being the parameter estimate for the unperturbed model. Influence assessment involves the comparison of $\hat{\beta}(w)$ and $\hat{\beta}$; local influence involves this comparison when w departs just slightly from w_0 . This leads to increased mathematical tractability via the use of derivatives with respect to w , instead of differences. One general way of comparing $\hat{\beta}(w)$ and $\hat{\beta}$ is by their likelihood distance apart, defined as

$$(2.1) \quad D(w) = 2\{L(\hat{\beta}|y) - L(\hat{\beta}(w)|y)\}.$$

With local influence, (2.1) is studied differentially or geometrically.

From a diagnostic point of view we are, as the simplest objective, concerned about individual cases; the perturbation w might just involve the i th case. Suppose we find that perturbation to the i th variance in a linear regression model causes large changes to the parameter estimates. This does not tell us much about the i th case of the data. It does obviously tell us that the assumption of the equality of the i th variance to the others is influential as far as parameter estimation is concerned. It does not allow us to infer that the variance of the i th case was different to the rest, for the perturbation was not being estimated, but was assumed. The conclusion is reached that a case is influential if one or more of the model assumptions for this case cannot be varied without important changes to the resulting parameter estimates. The strength of influence of a particular case might be gauged by the number of types of such perturbation which cause large estimate changes. The ideas here might properly be described as “assumption influence” or “model sensitivity”.

It is hard to relate logically the above type of influence to deletion influence, although mathematically there is a connection for standard linear models between an infinite variance and a zero weight for the i th case; a similar view and some misgivings have been expressed by Loynes in discussion of Cook (1986). Perturbation can, however, be applied to aspects of data which will not involve any model modifications. Likelihood distance of changes to parameter estimates does then point to the direct influence of data values; if small changes in particular data values cause large changes to estimates, then data influence is apparently established, in somewhat the same spirit as deletion. It does not seem possible, however, to encompass deletion as a data perturbation operation; deletion influence (or zero case-weighting) is logically distinct from both assumption- and data-influence. Case-weighting gives a third type of influence; it raises interesting questions of interpretation, discussed with Dr. F. Critchley at the IMA workshop, and currently being investigated.

The earlier presentation here has, for initial simplicity, been concerned with individual case influence; much, however, of local influence methodology is concerned with joint influence.

3. JOINT LOCAL INFLUENCE

In Cook's (1986) approach to local influence an important idea is given for the simultaneous handling of perturbations to all cases. Cook suggests regarding the likelihood distance $D(w)$ or some other suitable measure of the distance between $\hat{\beta}(w)$ and $\hat{\beta}$ as a surface in the n -dimensional Euclidean space of w . He then proposes a line across this surface with direction cosines ℓ which passes through the point w_o of the null perturbation; this line is thus of the form $w = w_o + a\ell$, with "a" measuring distance along the line and $|a|$ being the Euclidean magnitude of the perturbation. The direction cosines ℓ are thus associated with the cases of the data, and the n dimensional surface $D(w)$ is replaced by the much more tractable one-dimensional influence curve $\{a, D(w_o + a\ell)\}$. The next key idea is to choose the direction ℓ . This is ideally done in relation to a perturbation of chosen magnitude "a", so that $D(w_o + a\ell)$ is as large as possible; the individual perturbations contributing to "a" which each case should receive are thus selected, the most influential cases being those associated with the larger direction cosines. Further tractability is allowed by requiring "a" to be small and thus needing only to consider the local behaviour of the influence curve $D(w_o + a\ell)$ at w_o . For a scalar parameter β , $D(w)$ may as well be replaced by $\hat{\beta}(w) - \hat{\beta}$, since there is no need to combine the effects of different parameters; it is then clear that the line should be in the direction $\tilde{\ell}$ which maximises or minimises $\partial\hat{\beta}(w_o + a\ell)/\partial a(a=0)$. When D combines parameter estimates, $D(w_o + a\ell)$ may have a minimum at $a=0$, as does the likelihood distance, and then the curvature or second derivative of $D(w_o + a\ell)$ is the appropriate quantity to maximise in order to determine $\tilde{\ell}$. In practice, a small proportion of the cases usually have direction cosines much larger than the rest, and these are the cases of most joint influence. It does not follow, however, that these cases exert practical influence, and ideas for assessing and determining perturbation sizes will be considered in Section 5.

Returning to the likelihood distance form of $D(w_o + a\ell)$, there is some elegant theory for determining $\tilde{\ell}$ due to Cook (1986). The required second derivative of $D(w_o + a\ell)$, which Cook emphasises as the curvature at w_o , and which we will call "Cook's directional curvature" in direction ℓ , evaluates as

$$\begin{aligned}
 (3.1) \quad CU_\ell &= \left. \frac{\partial^2 D(w_o + a\ell)}{\partial a^2} \right|_{a=0} = - \left. \frac{\partial^2 L(\hat{\beta}(w_o + a\ell))}{\partial a^2} \right|_{a=0} \\
 &= - \left. \frac{\partial}{\partial a} \left\{ \frac{\partial L(\hat{\beta}(w_o + a\ell))}{\partial \hat{\beta}(w_o + a\ell)^T} \frac{\partial \hat{\beta}(w_o + a\ell)}{\partial a} \right\} \right|_{a=0}
 \end{aligned}$$

and hence

$$\begin{aligned}
(3.2) \quad CU_\ell &= \ell^T \left\{ \frac{\partial \hat{\beta}(w_o + a\ell)^T}{\partial(w_o + a\ell)} \frac{\partial^2 L(\hat{\beta}(w_o + a\ell))}{\partial \beta(w_o + a\ell) \partial \beta(w_o + a\ell)^T} \frac{\partial \hat{\beta}(w_o + a\ell)}{\partial(w_o + a\ell)^T} \right\} \Bigg|_{a=0} \ell \\
&= \ell^T \frac{\partial \hat{\beta}(w)^T}{\partial w_o} \left\{ -\frac{\partial^2 L(\beta)}{\partial \hat{\beta} \hat{\beta}^T} \right\} \frac{\partial \hat{\beta}(w)}{\partial w_o^T} \ell.
\end{aligned}$$

For the standard linear model with design matrix X and residual mean square error $\hat{\sigma}^2$, there is the particular result

$$(3.3) \quad CU_\ell = \ell^T \frac{\partial \hat{\beta}(w)^T}{\partial w_o} (X^T X) \frac{\partial \hat{\beta}(w)}{\partial w_o^T} \ell / \hat{\sigma}^2.$$

The maximizing of (3.2) over ℓ is a standard eigenvalue problem, for instance, Green & Carroll (1976, p. 213); the largest eigenvalue of the inner $n \times n$ matrix in (3.2) is its maximum, and the corresponding eigenvector is the required $\tilde{\ell}$. Nevertheless, for large n the computations could be prohibitive. Simplification may be possible by writing the observed information matrix in (3.2) as $B^T B$ using the spectral decomposition. Then the $n \times n$ matrix in (3.2) can be written as AA^T , where $A = \{\partial \hat{\beta}(w)^T / \partial w_o\} B^T$ is $n \times p$ and p is the dimension of β ; this has the same positive eigenvalues λ as the $p \times p$ matrix $A^T A$ and its eigenvectors are $Ax / \sqrt{\lambda}$, where x is the eigenvector associated with λ . Since p is usually small, the eigenvalue problem is thus easily solved. It will also be sensible to obtain ℓ for eigenvalues which are nearly as large as the maximum, if there are any. This way of solving large eigenvalue problems is extremely useful in the local influence approach.

4. LOCAL AND DIRECTIONAL VERSIONS OF COOK'S DISTANCE

Since Cook's distance, Cook (1977), has become so well established in deletion influence, it seems worth while to stress parallel local and directional quantities. Cook's distance is concerned with individual cases, and thus in the direction of the i th axis ($i = 1, 2, \dots, n$) in the n dimensional perturbation space. Hence, taking (3.2) with $\ell^T = (0, \dots, 1, \dots, 0)$ where the 1 is in the i th position, and with w_i as the i th element of w , "Cook's curvature" for case i can be defined, as

$$(4.1) \quad \frac{\partial \hat{\beta}(w)^T}{\partial w_i} \left\{ -\frac{\partial^2 L(\beta)}{\partial \hat{\beta} \hat{\beta}^T} \right\} \frac{\partial \hat{\beta}(w)}{\partial w_i} \Bigg|_{w=w_o}.$$

Also, a directional version of Cook's original distance for the perturbation $w_o + a\ell$ in a general direction ℓ can be taken as

$$(4.2) \quad \{\hat{\beta} - \hat{\beta}(w_o + a\ell)\}^T \left\{ -\frac{\partial^2 L(\beta)}{\partial \hat{\beta} \hat{\beta}^T} \right\} \{\hat{\beta} - \hat{\beta}(w_o + a\ell)\} p^{-1},$$

and will be described as the "directional distance". As special cases, take $\ell^T = (0, \dots, 1, \dots, 0)$ for the i th case, as at (4.1), and so obtain

$$(4.3) \quad \{\hat{\beta} - \hat{\beta}(a; i)\}^T \left\{ -\frac{\partial^2 L(\beta)}{\partial \hat{\beta} \hat{\beta}^T} \right\} \{\hat{\beta} - \hat{\beta}(a; i)\} p^{-1}$$

where $\hat{\beta}(a, i) = \hat{\beta}(w_o + al)$; this will be called the “perturbation distance” for case i . Continuing, a “directional local distance”, in analogy to (4.2) can be stated as,

$$(4.4) \quad \left. \frac{\partial \hat{\beta}(w_o + al)^T}{\partial a} \right|_{a=0} \left\{ -\frac{\partial^2 L(\beta)}{\partial \hat{\beta} \hat{\beta}^T} \right\} \left. \frac{\partial \hat{\beta}(w_o + al)}{\partial a} \right|_{a=0} p^{-1},$$

and in this form is seen to be equivalent to (3.2), Cook’s directional curvature, apart from the division by p . Thus Cook’s directional curvature is a natural local and directional analogy to his original “distance”, and as such it should be easily understood. When (4.4) is taken with $\ell^T = (0, \dots, 1, \dots, 0)$, it may sensibly be called the “local distance” for case i , and is a basic quantity of local influence methodology.

5. PROBLEMS OF MEASURING INFLUENCE

For deletion influence, Cook (1977) suggested a way of assessing the importance of the i th case according to whether or not its deletion moved the regression parameter estimate to the edge of the 50% or 10% confidence region based on all the data. Movement to the edge of the 50% confidence level is quoted as “giving cause for concern”, while the generally little movement within a 10% confidence region is said to “give an uncomplicated analysis”. More fully, the attained values of the confidence levels for deletion of each case in turn seems to be suggested, and one-minus these values are described as “descriptive levels of significance”. This led to some objections in a published letter, Obenchain (1977), partly concerning the frequency connotations of the phrase, but Cook responded that nothing of the sort was implied, and confidence level movement simply gave a calibration on a familiar scale. There certainly are difficulties in suggesting a frequency interpretation; in the standard regression cases, Cook’s distance is a function of just the standardized residual and leverage. Its distribution thus depends on the explanatory space via leverage and on the model of the data set. The frequency with which deletion of a case of the data causes parameter estimate movement to the edge of the 50% confidence region will generally be much smaller than 50%; for $n = 20$, $p = 3$ and equal leverage, it is nearly zero; a standardized residual greater than 3.7 is required. Bruce and Martin (1989) reemphasize that the testing framework is not appropriate for influence assessment.

The quoting of 50% levels for F or χ^2 as the cut-off values in packages seems to be misleading people, even assuming the proper understanding of these levels. Belsley et. al. (1980, p. 28) suggested sample size adjusted cut-off values of statistics, implying sample size adjusted confidence interval movement. The difficulties can be ignored by ordering the attained levels and noting for further study any outlying cases, defined by the “natural gap” principle of their separation from the rest. Adherence to fixed cut-off levels does not now have the support of their proposers, at least according to discussions at the workshop. Nevertheless calibration by “descriptive confidence levels” can still be used useful.

The previous discussion applies mainly to Cook’s deletion distance; as far as local influence is concerned, the notion of movement to the edge of confidence

regions cannot be applied. Using the language of curvature, Cook (1986) employed an empirical benchmark figure of 2, but was challenged in subsequent discussion, and it cannot be justified theoretically; similarly for other “local distance” quantities. Actual perturbations need to be involved.

One sensible strategy appears to be to plot $D(w_o + a\tilde{\ell})$, where “ a ” is the size of an actual perturbation and $\tilde{\ell}$ is the maximising local direction, against “ a ” and compare with the upper, say 50% point of the χ^2 distribution on p degrees of freedom. This can give upper and lower bounds on the perturbation size “ a ”, and these values can be judged against subject matter considerations of the data. These plots should not be interpreted frequently, but as calibration of change. However, as Cook’s (1986) Figure 6 shows, a closed region of “ a ” values cannot be guaranteed. When perturbing a single explanatory variable, “ a ” is naturally in units of the variable. When perturbing a combination of variables, Cook’s Figure 6, it is necessary to introduce a scaling of the variables; this is an added complication, and may not anyway be a natural course of action. When perturbing variances, “ a ” is a dimensionless quantity, and comparable across data sets.

6. LOCAL INFLUENCE OF ASSUMPTIONS FOR THE LINEAR MODEL

The key model assumptions of standard linear regression which are examined here with regard to local influence, are

- (i) the constancy of variance of error, and
- (ii) the independence of errors given their constant variance.

The terminology of case weights is not used for (i) for two reasons: first, it seems to divert attention away from the perceived advantage of the local influence approach that it can produce specifically directed diagnostics. Secondly, variance perturbation is in general distinct from weight perturbation, although not for the standard linear model. The results for (i) are mainly due to Cook (1986), while those for (ii) are a contribution to extending the availability of local influence diagnostics. On grounds of specificity there is little appeal in combining (i) and (ii), and anyway, a sensible perturbation scheme is not apparent. Again, for reasons of specificity, the influence is with respect to all the regression coefficients estimates, ignoring the mean square error estimate.

Suppose the response variable y depends on the $p-1$ explanatory variables X_1, X_2, \dots, X_{p-1} ; then the standard linear model is

$$(6.1) \quad y = X\beta + \sigma\epsilon.$$

Here the columns of the $n \times p$ matrix X contain the values of the explanatory variables and there is an initial column of ones, β is the $p \times 1$ vector of regression coefficients, σ^2 is the variance of each response, and ϵ is an $n \times 1$ vector of independent standardized normal variables. The profile loglikelihood for the regression parameters is known to be

$$(6.2) \quad L(\beta) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} + \frac{n}{2} \log n - \frac{n}{2} \log\{(y - X\beta)^T(y - X\beta)\}$$

and has maximized value, when $\hat{\beta} = (X^T X)^{-1} X^T y$, of

$$(6.3) \quad -\frac{n}{2} \log(2\pi) - \frac{n}{2} + \frac{n}{2} \log n - \frac{n}{2} \log(y^T A y),$$

where $A = I - H$, and $H = X(X^T X)^{-1} X^T$ is the leverage matrix.

6.1 Variance Perturbations. The variance of the i th response is perturbed to $(1 + v_i)\sigma^2$, $-1 \leq v_i < \infty$, and the perturbed model is thus

$$(6.4) \quad y = X\beta + \sigma V\epsilon$$

where $V = d(1 + v)$ is the diagonal matrix with entries $1 + v$, $v = (v_1, v_2, \dots, v_n)$; $v = 0$ is the null perturbation. This is the familiar weighted model, and β has the maximum likelihood and weighted least squares estimate given by

$$(6.5) \quad \hat{\beta}(v) = (X^T V^{-1} X)^{-1} X^T V^{-1} y.$$

The basic result as required by (3.2) for the direction of most local influence is given by straightforward calculation as

$$(6.6) \quad \left. \frac{\partial \hat{\beta}(v)}{\partial v^T} \right|_{v=0} = (X^T X)^{-1} X^T d(r),$$

where $d(r)$ is the diagonal matrix of the ordinary residuals $r = y - \hat{\beta}X$. Using (3.3), Cook's directional curvature in direction ℓ (and equivalent directional local distance) becomes

$$(6.7) \quad \ell^T d(r) X (X^T X)^{-1} X^T d(r) \ell / \hat{\sigma}^2.$$

As described after (3.2), the maximum of (6.7) over possible directions ℓ will be the largest eigenvalue of the $n \times n$ matrix between ℓ^T and ℓ of (6.7), and the required $\ell = \tilde{\ell}$ will be the associated eigenvector. Thus with $(X^T X)^{-1}$ written as $B^T B$, $\tilde{\ell}$ is given by $d(r) X B^T x / \sqrt{\lambda}$ where λ and x are the largest eigenvalue and its associated eigenvector of the $p \times p$ matrix $B X^T d(r)^2 X B^T$. Note also that the sum and hence average of the direction cosines in $\tilde{\ell}$ is zero, since $(1, \dots, 1)d(r) = r^T$ and $r^T X = 0$. Index plotting of $\tilde{\ell}$ to highlight jointly influential cases will be illustrated in Section 8.

The diagonal elements of the inner matrix of (6.7) divided by p have been called the local distance for case i in Section 4, and are seen to be $r_i^2 h_i / p \hat{\sigma}^2$, where h_i is the i th leverage: this compares with Cook's distance as $r_i^2 h_i / \{(1 - h_i)^2 p \hat{\sigma}^2\}$. There is an increase by the leverage factor $(1 - h_i)^{-2}$. This is an illustration that high leverage enhances deletion influence more so than it does local influence.

The likelihood distance in the direction $\tilde{\ell}$ in this regression case follows from (2.1) and (6.3) as

$$(6.8) \quad \begin{aligned} D(a, \tilde{\ell}) &= n \log \{ \{y - X\hat{\beta}(a, \tilde{\ell})\}^T \{y - X\hat{\beta}(a, \tilde{\ell})\} / n \hat{\sigma}^2 \} \\ &= n \log \{ 1 + \{ \hat{\beta} - \hat{\beta}(a, \tilde{\ell}) \}^T (X^T X) \{ \hat{\beta} - \hat{\beta}(a, \tilde{\ell}) \} / n \hat{\sigma}^2 \} \\ &= n \log \{ 1 + (p/n) C(a, \tilde{\ell}) \}, \end{aligned}$$

where $C(a, \tilde{\ell})$ is the directional distance (4.2) and $\hat{\beta}(a, \tilde{\ell})$ is calculated using weighted least squares. By asymptotic theory this can be used to obtain the level of a descriptive $100(1-\alpha)\%$ confidence region, by equating it to $\chi_\alpha^2(p)$, the upper $100\alpha\%$ point of the χ^2 distribution with p degrees of freedom; equivalence to Cook's distance, with a directional sense, is achieved from use of the approximation $(n/p)\{\exp \chi_\alpha^2(p)/n-1\}$ to the upper $100\alpha\%$ point of the $F_{p,n-p}$ distribution. Note that variability enters into (6.8) via $\tilde{\ell}$ as well as the parameter estimates. In practice (6.8) can be used to determine a range of perturbation values which will ensure that the parameter estimates will not be moved out of their, say, 10 or 50% descriptive confidence regions; this range should be interpretable as to its subjective importance. The plotting of $D(a, \tilde{\ell})$ against "a" will be illustrated in Section 8.

For more direct comparison of the local and deletion influence approaches, consider perturbation of the just the i th variance to $(1+a)\sigma^2$, and then denote V as V_i ; the following result is employed,

$$(6.9) \quad X(X^T V_i^{-1} X)^{-1} X^T = X(X^T X)^{-1} X^T + \frac{a}{1+a(1-h_i)} \tilde{h}_i^T \tilde{h}_i$$

where \tilde{h}_i is the i th row of $H = X(X^T X)^{-1} X^T$; a reference is Belsley et. al. (1980, p. 66). The i th perturbation distance, as defined at (4.3), becomes

$$(6.10) \quad \left(\frac{a}{1+a(1-h_i)} \right)^2 \frac{h_i r_i^2}{p \hat{\sigma}^2}.$$

For $a \rightarrow \infty$, the infinite perturbation corresponding to deletion, Cook's distance is recovered as

$$(6.11) \quad \frac{1}{(1-h_i)^2} \frac{h_i r_i^2}{p \hat{\sigma}^2}.$$

More usually, but unimportantly, $\hat{\sigma}^2$ would have $n-p$ as its divisor instead of n and be s^2 . It is interesting to note that (6.11) for Cook's distance can also be obtained by a perturbation of $a = -1/\{2(1-h_i)\}$ in (6.10): but this effect is only possible when $h_i < \frac{1}{2}$. Figure 1 in Cook (1986) can be extrapolated to illustrate the effect for likelihood distance. Also, it can be seen from (6.10) that influence at a given descriptive confidence level will be when "a" lies outside an interval determined by a quadratic equation in "a"; this provides a way of calibrating the perturbation distance as discussed in Section 5.

6.2 Perturbation of Independence by Autocorrelation. To assess the importance of the independence assumption in the standard linear model, it is proposed to perturb the variance matrix from the unit diagonal form to a tri-diagonal type representing first order moving average autocorrelation in the index of the data. Because of the local nature of the analysis, and assuming geometric decay in the autocorrelation lag, it is only first order dependence which is relevant.

Thus, in the standard linear model notation of (6.4),

$$(6.13) \quad V = \begin{pmatrix} 1 & \rho_1 & 0 & \cdots & 0 \\ \rho_1 & 1 & \rho_2 & \cdots & 0 \\ 0 & \rho_2 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \rho_{n-1} \\ 0 & 0 & \cdots & \rho_{n-1} & 1 \end{pmatrix}.$$

This means that the covariance between y_i and y_j is $\sigma^2 \rho_i$ for $j = i + 1$ and 0 for $j > i + 1$. As noted at the beginning of Section 6, we prefer to just perturb one assumption at a time, and variances have been assumed constant. Symmetry of the matrix is crucial in the time series sense of the perturbations. The ρ_i are perturbations away from 0, indexing the i th case; note $i = 1, 2, \dots, n - 1$ and not n . We thus assume in the local influence framework that $\rho_i = a\ell_i$ ($1 \leq i \leq n - 1$), with the implied range of “ a ” being $|a| < \min\{1/\max(\text{positive } \ell_i), 1/\max(-\text{negative } \ell_i)\}$. Differentiating the $\hat{\beta}$ of (6.5) in which V of (6.13) is $V(a\ell)$, there is the result

$$(6.14) \quad \left. \frac{\partial \hat{\beta}(a)}{\partial a} \right|_{(a=0)} = (X^T X)^{-1} X^T \begin{pmatrix} 0 & \ell_1 & 0 & 0 & \cdots & 0 \\ \ell_1 & 0 & \ell_2 & 0 & & \vdots \\ 0 & \ell_2 & 0 & \ell_3 & & \\ 0 & 0 & \ell_3 & 0 & \ddots & \\ & & & & \ddots & \ell_{n-1} \\ \vdots & & & \ddots & \ddots & \\ 0 & \cdots & & \ell_{n-1} & 0 & \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix} = (X^T X)^{-1} X^T R \ell$$

where

$$R = \begin{pmatrix} r_2 & 0 & 0 & \cdots & 0 \\ r_1 & r_3 & 0 & & \vdots \\ 0 & r_2 & r_4 & & \\ 0 & 0 & r_3 & \ddots & \\ \vdots & & & \ddots & \\ 0 & \cdots & & r_n & r_{n-1} \end{pmatrix}.$$

of dimension $n \times (n - 1)$. Thus from (4.4), Cook’s directional curvature or equivalent directional local distance becomes

$$(6.15) \quad \ell^T R^T X (X^T X)^{-1} X^T R \ell / \hat{\sigma}^2,$$

which may be maximised to obtain $\tilde{\ell}$, as described in Section 6.1. Unlike the variance case, the sum of the direction cosines in $\tilde{\ell}$ is not zero.

Finally, we consider the effects of correlation perturbations to the $(i, i + 1)$ th cases. Immediately from the i th diagonal element of (6.15), Cook’s curvature for the i th case become

$$(6.16) \quad (h_{ii} r_{i+1}^2 + 2h_{ii+1} r_i r_{i+1} + h_{i+1i+1} r_i^2) / \hat{\sigma}^2,$$

The first and last terms will be large when the leverages and residuals are large; (6.16) becomes larger when the associated off-diagonal element of the leverage matrix is large, and of the same sign as the product of the adjacent residuals. The presence of the off-diagonal term is not surprising when the effect of interest concerns pairs of cases.

7. LOCAL INFLUENCE OF DATA COMPONENTS FOR THE LINEAR MODEL

Concern now is for the sensitivity of parameter estimates to data values used in the estimation, assuming a correct model. The local influence approach is used to perturb data values, a logically distinct operation from perturbing assumptions. The identification of data values upon which the parameter estimates crucially rest, is of course, of much practical importance; it drives the process of data validation and correction. Perturbations to response and individual explanatory variables will be considered; the explanatory case has been fully studied by Cook (1986), although the derivations here are intended to be slightly more explicit. First perturbations to response values are discussed.

7.1 Perturbations to Response Values. Perturbing y to $y + w$ gives $\hat{\beta}$ as $(X^T X)^{-1} X^T (y + w)$ and thus

$$(7.1) \quad \left. \frac{\partial \hat{\beta}(w)}{\partial w} \right|_{w=0} = (X^T X)^{-1} X^T,$$

with Cook's directional curvature from (3.2) as

$$(7.2) \quad \ell^T X (X^T X)^{-1} X^T \ell = \ell^T H \ell.$$

Since H is idempotent, the p non-zero eigenvalues are unity. There is no unique eigenvector in this situation and thus no unique direction maximizing the local curvature. The eigenvectors $\tilde{\ell}$ maximising (7.2) are not functions of the responses. The individual Cook's curvatures are just the leverage coefficients. Also,

$$(7.3) \quad \left. \frac{\partial X \hat{\beta}(w)}{\partial w} \right|_{w=0} = X (X^T X)^{-1} X^T = H,$$

so that the two-way table of fitted values and case derivatives is given by the leverage matrix, perhaps a useful further way to interpret H .

Another way to perturb the response values is multiplicatively; y is perturbed to $\{\text{diag}(1 + a\ell)\}y$; hence

$$(7.4) \quad \left. \frac{\partial \hat{\beta}(a)}{\partial a} \right|_{a=0} = (X^T X)^{-1} X^T D(y) \ell$$

where $D(y)$ is the diagonal matrix with entries y . Hence, Cook's directional curvature is

$$(7.5) \quad \ell^T D(y) H D(y) \ell$$

which may be maximised to produce unique $\tilde{\ell}$; the individual curvatures are then $h_i y_i^2$; the use of such a perturbation scheme remains to be justified.

7.2 Perturbations to Explanatory Variables. The values of the i th explanatory variable x_i , the $(i + 1)$ th column of X , are perturbed to $x_i + w$; as previously, w is taken as $a\ell$ where ℓ is a column vector of direction cosines. Then reasonably straightforward calculation gives

$$(7.6) \quad \left. \frac{\partial \hat{\beta}(a)}{\partial a} \right|_{a=0} = (X^T X)^{-1} \{d_i r^T - \beta_i X^T\} \ell,$$

where β_i is the regression coefficient of x_i , and d_i is a column vector with 1 in the i th position and zeros elsewhere. From (3.2), the directional curvature using (7.6) becomes $\ell^T M \ell$ where

$$(7.7) \quad M = (r d_i^T - \hat{\beta}_i X)(X^T X)^{-1} (d_i r^T - \hat{\beta}_i X^T) / \hat{\sigma}^2.$$

The eigenvalues of the matrix in (7.7) can be determined by writing $(X^T X)^{-1} = B^T B$ as previously, and working with $M = A A^T$ where $A = (r d_i^T - \hat{\beta}_i X) B^T$. The required maximum eigenvalue λ is then that of

$$(7.8) \quad A^T A = B \{ \hat{\beta}_i^2 X^T X + (r^T r) d_i d_i^T \} B^T,$$

and the required eigenvector is $Ax/\sqrt{\lambda}$, where x is the associated eigenvector of $A^T A$. Premultiplying (7.8) by B^T , the required maximum eigenvalue is that of

$$(7.9) \quad \hat{\beta}_i^2 I + r^T r (X^T X)^{-1} d_i d_i^T.$$

There is one non-zero eigenvalue of the matrix $(X^T X)^{-1} d_i d_i^T$ which is conveniently determined as $d_i^T (X^T X)^{-1} d_i$. This latter quantity can be evaluated as $1/\|x_i^*\|^2$ where x_i^* is the residual vector from regressing x_i on the other explanatory variables, and $x_i^*/\|x_i^*\|^2$ is obtainable from $X(X^T X)^{-1} d_i$, the i th row of $X(X^T X)^{-1}$. Thus the one non-zero eigenvalue of (7.9) is $\hat{\beta}_i^2 + (r^T r)/\|x_i^*\|^2$, with associated eigenvector $x_i^*/\|x_i^*\|^2$. Retracing the eigen analysis backwards, (7.7) is found to have a maximum eigenvalue of $\hat{\beta}_i^2 + (r^T r)/\|x_i^*\|^2$, with corresponding normalized eigenvector,

$$(7.10) \quad \tilde{\ell} = (r - \hat{\beta}_i x_i^*) / \sqrt{(r^T r + \hat{\beta}_i^2 \|x_i^*\|^2)}.$$

This is Cook's (1986) important result (4.3), although our notations differ; the elements of $\tilde{\ell}$ are proportional to the vertical distance from the line of slope $\hat{\beta}_i$ in the added variable plot of r against x_i^* .

8. ILLUSTRATION OF DELETION AND LOCAL INFLUENCE

This section is devoted to showing how the methods previously described work on two exemplary data sets. The first such set concerns the record times of 35 Scottish hill races in 1984, and was assembled by Dr. G. Cohen of Edinburgh University; there are two explanatory variables, distance of race and climb. The data has been published and more fully analyzed by A. C. Atkinson in the discussion to Chatterjee and Hadi (1986). The regression of record time on distance and climb additively will be taken as modelling the data. Local influence of variance perturbations will be illustrated with this data set. A second and more classical set of time series data concerned with the consumption of spirits from Durbin and Watson (1951) will be used to illustrate the local influence of independence perturbations. All the calculations can be done with standard software; Minitab macros were used here.

8.1 Scottish Hill Races Data. Figure 1 gives a plot of the explanatory variables, helpful in understanding leverage in the data. Cook's distance and corresponding local distance (Section 6) are compared in Figure 2; they differ only by the leverage factor $(1 - h_i)^{-2}$. Both distances bring cases 7 and 18 for further attention. The numerical values of the distances appear small, but the mean square error

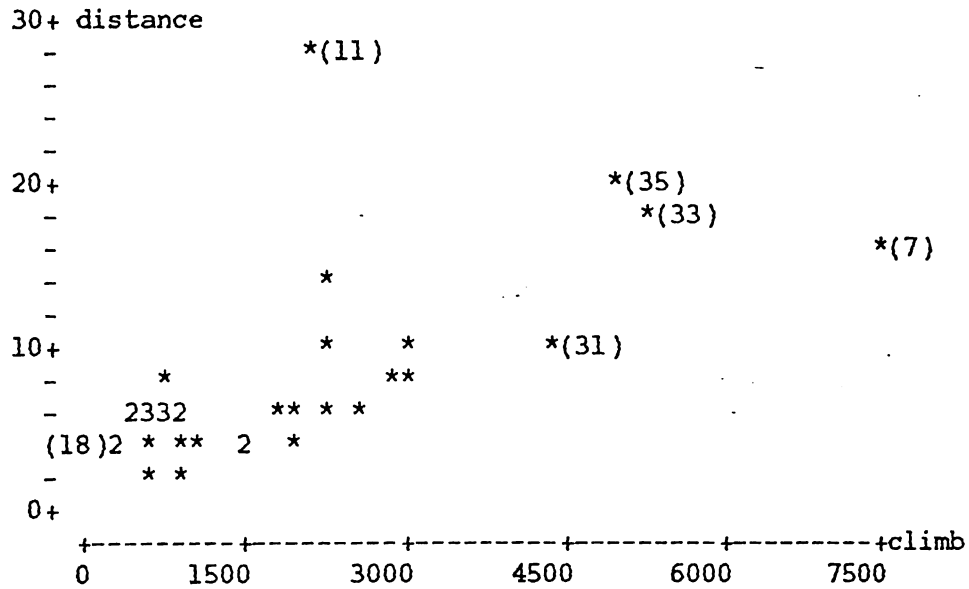


FIGURE 1

Scottish hill races: Distance against climb; case numbers in parentheses.

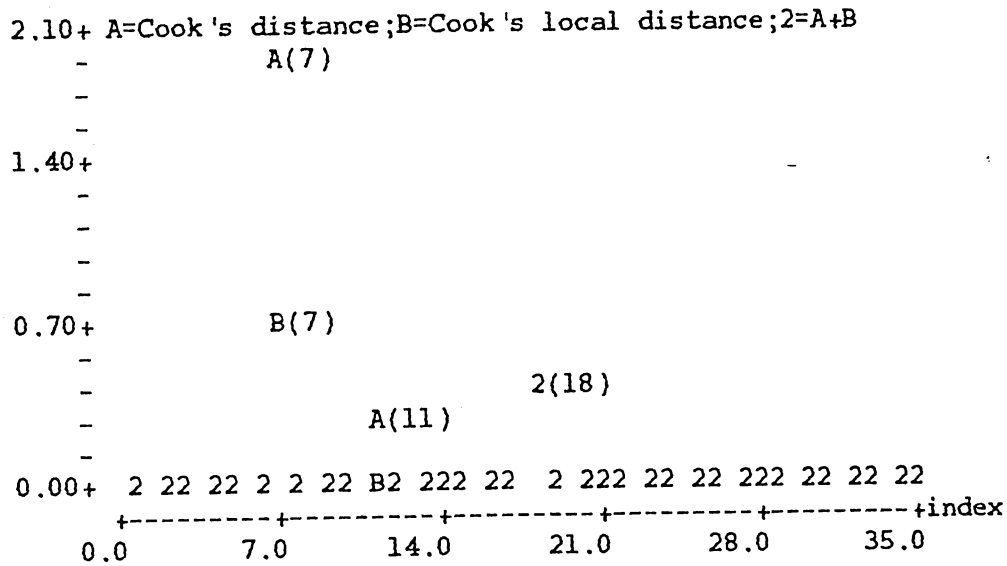


FIGURE 2

Scottish hill races data: Comparison of Cook's distance and corresponding local distance.

s^2 is 215.4 which is much due to the outlying case 18 with a residual of 65.12; this effect will be further discussed in Section 10. The direction cosines for variance perturbations are given in Figure 3; cases 7, 18, 31, 33, 35 show up as being notice-

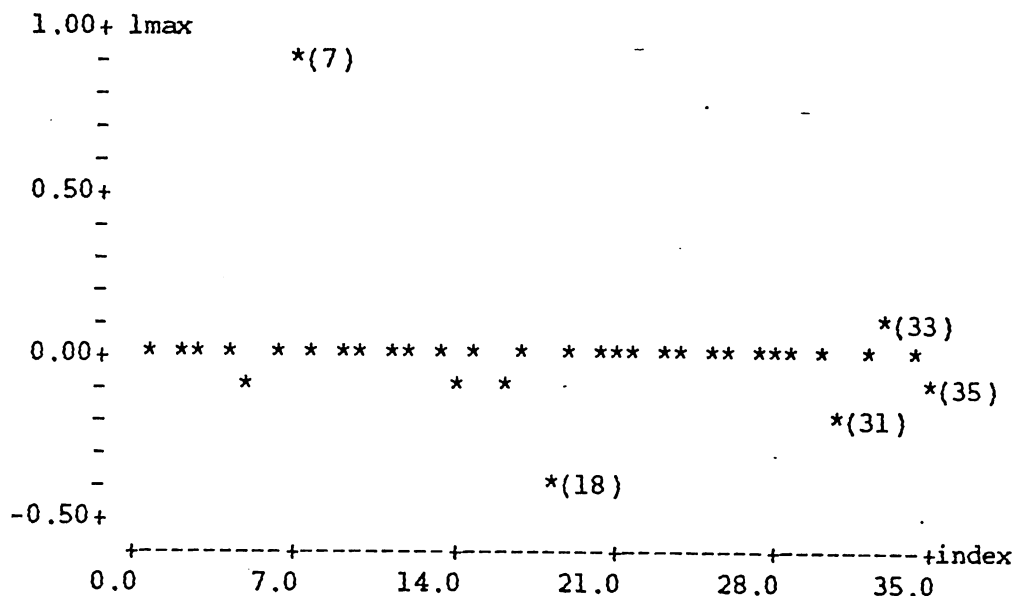


FIGURE 3

Scottish hill races: Direction cosines which locally maximise Cook's curvature.

ably different from zero, but not case 11 which appeared minimally for Cook's deletion distance in Figure 2. With reference to case 18, Atkinson noted very plausibly, in view of its distance of 3 miles, that the record time of 1.18.39 hours should be 18.39 minutes. Its revision removes case 18 from the essentially non-zero direction cosines, and reduces s^2 to 77.50. Case 7 is still outlying with a standardized residual of 4.16, and it is most outlying in the explanatory space, as shown in Figure 1.

Continuing with local influence analysis, a further point from Figure 3 is that cases 7 and 18 have oppositely signed direction cosines, indicating sensitivity to opposite types of perturbation. Cook's local directional distance in this l_{max} direction is 0.81 which may be compared with the largest individual local distance of 0.64 belonging to case 7. More quantitative study of actual perturbation effects is provided by the log likelihood distance plot of Figure 4; here "a" corresponds to variance perturbations of the form $(1 + al_i)\sigma^2$. The rise in the likelihood distance is a little smaller for positive perturbations relative to negative ones; perturbations satisfying $(-0.74 \geq a \geq 1.09)$ will certainly move the parameter estimates out of their 50% confidence region. Values of "a" obtained in this way can be compared across data sets, as absolute measures of sensitivity. Conclusions are, however, still affected by s^2 .

The analysis up to this point, mainly for illustration, has been on the original data; the model used would not be very useful in practice, in spite of its 92% R^2 -

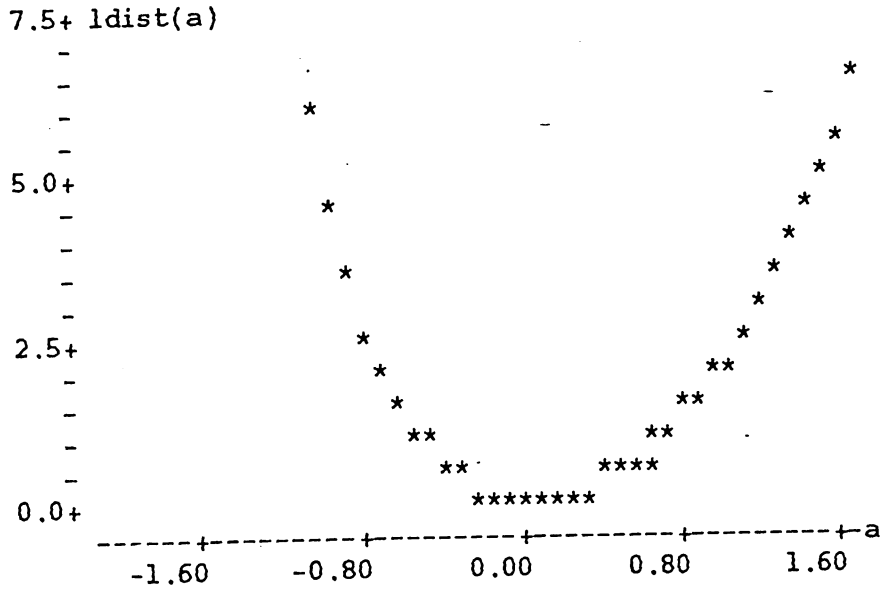


FIGURE 4

Scottish hill races: Likelihood distance for perturbations in the locally optimum direction shown in Figure 3.

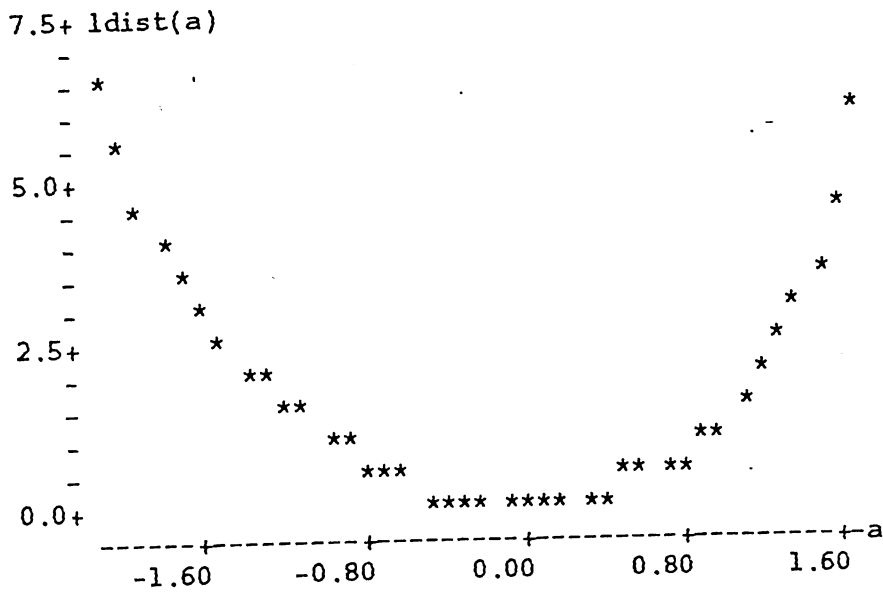


FIGURE 5

Scottish hill races data: Likelihood distance as in Figure 4, but modified data in which case 18's record time has been changed by 1 hour and cases 7 and 33 have been deleted.

-value, because of its high residual mean square of 215.37 mainly due to cases 18, 7 and 33. The next suggestion is to reduce the record time of case 18 by 1 hour and after inspection of the residuals and leverage to omit cases 7 and 33. This produces a model with a still higher R^2 of 99% and a much reduced residual mean square of 22.20, with no evidence of lack of fit. For Cook's distance, cases 11 and 6 with $C_{11} = 0.252$, $C_6 = 0.130$, are the only ones with non-zero values, both small and also based on the small s^2 . For a direct comparison with the previous likelihood distance analysis, Figure 4 may be compared with Figure 5; the latter is much wider, with the 50% values of "a" being given by $(-1.74 \geq a \geq 1.28)$; thus perturbations much larger in magnitude are required for the less sensitive modified data.

8.2 Consumption of Spirits Data. This set of data concerns the consumption of spirits in the U.K. from 1870 to 1938, with the response variable as the annual per capita consumption and per capita income and price of spirits as explanatory variables, all in log form. The data were used by Durbin and Watson (1951) when proposing their Durbin-Watson statistic for testing autocorrelation in regression. The model used here is of the quadratic form employed by Fuller (1976, p. 426),

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 \{(t - 1900)/30\} + \beta_4 \{(t - 1900)/30\}^2 + \epsilon_t$$

where X_{1t} and X_{2t} are the explanatory variables. For this data set, Figure 6 gives direction cosines for independence perturbations, as described in Section 6.2. Roughly the middle half of the data seems to be sensitive. The actual extent of the

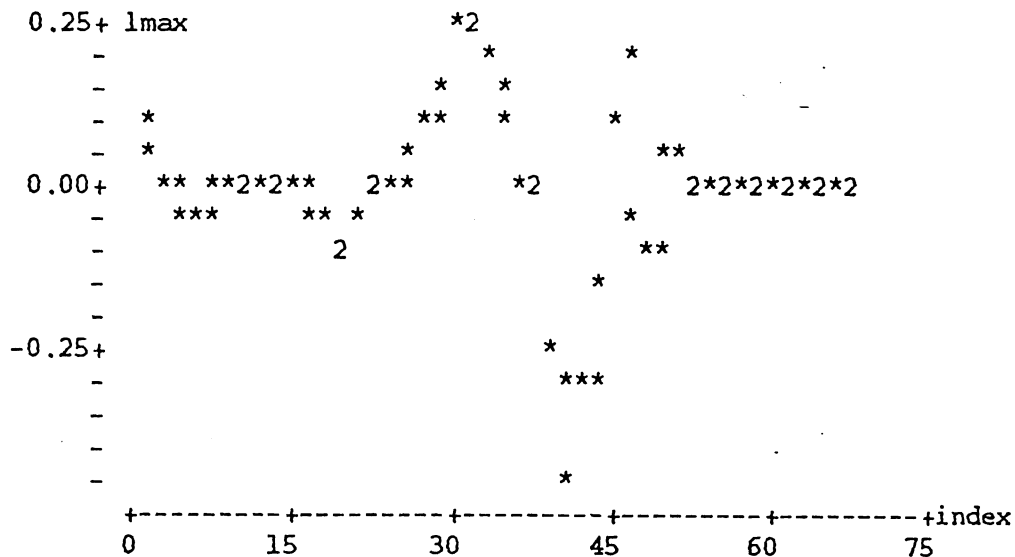


FIGURE 6

Consumption of spirits data: Direction cosines which locally maximise Cook's Curvature or local distance.

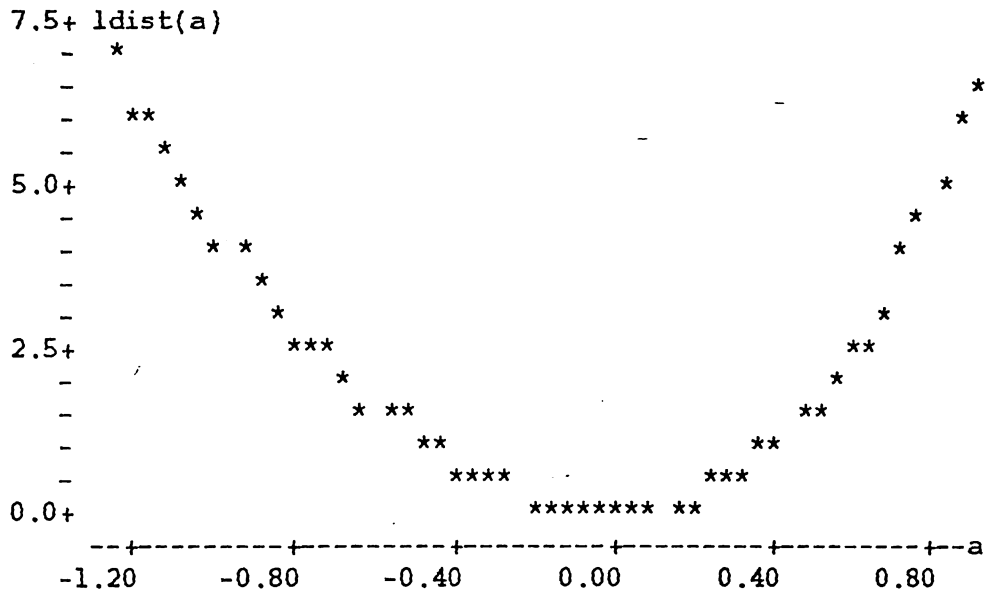


FIGURE 7

Consumption of spirits data: Likelihood distance for perturbations in the locally optimum direction shown in Figure 7.

the effects is assessed by the likelihood distance plot of Figure 7. Perturbations with “ a ” satisfying $(-1.02 \geq a \geq 0.77)$ cause parameter estimates to move outside their original 50% confidence region.

9. SCALE MODIFICATION IN INFLUENCE ASSESSMENT

The Scottish hill races data can be used to illustrate the strong effect of s^2 on the values of influence quantities such as Cook’s distance, local distance and likelihood distance. As used here, these quantities describe the regression parameter β influence, but are scaled by s^2 ; this can itself be greatly increased by a few cases with very large residuals and thus depress the influence assessment, a self-defeating process.

As instances of this effect, suppose the s^2 value of 22.20 is used in respect of all the data; this multiplies Cook’s distance values by nearly 10 and raises cases 11 and 18 in addition to 17 above $F_{3,32}(50\%) = 0.806$. In respect of the likelihood distance, the 50% values of “ a ” change from $(-1.47 \leq a \leq 1.28)$ to $(-0.28 \leq a \leq 0.34)$, a rather dramatic change. Thus, in general it seems sensible to use a robust estimate of s^2 when assessing regression parameter influence; two possible estimates are readily available in Minitab’s undocumented robust regression command `rreg` (see Minitab Newsletter, KeepingTAB 12, 1988). This type of single adjustment seems preferable to case by case deletion adjustments as employed by Belsley et. al. (1980) and Atkinson (1985). In the present instance, Minitab’s estimate based on the Hodges-Lehmann confidence interval reduced the s^2 for all the data by a factor of two.

10. FURTHER TOPICS

The authors' presentation at the IMA workshop also included a preliminary discussion of the masking effect, in relation to both deletion and local influences. Cook's distance was considered and the importance of the leverage quantities $h_{ij}/\sqrt{\{(1-h_i)(1-h_j)\}}$ and the ratio of standardized residuals was evident. The high leverage of a case was found to be some protection against masking. A proposal was made that the opposite effect of masking, when influence is decreased by deletion, be called boosting.

ACKNOWLEDGEMENTS

It is a real pleasure to thank the organizers of the workshop for such a worthwhile event, and to thank the IMA staff for being so helpful and providing such a good environment.

REFERENCES

- ATKINSON, A.C., *Plots, Transformations and Regression*, Oxford: University Press (1985).
- BELSLEY, D.A., KUH, E. AND WELSH, R.E., *Regression diagnostics: Identifying influential data and sources of collinearity*, New York: John Wiley (1980).
- BRUCE, A.G. AND MARTIN, R.D., *Leave-k-out diagnostics for time series* (with discussion), J.R. Statist. Soc. B, 51 (1989), pp. 363-424.
- COOK, R.D., *Detection of influential observations in linear regression*, Technometrics, 19 (1977), pp. 15-18.
- COOK, R.D., *Assessment of local influence* (with discussion), J.R. Statist. Soc. B, 48 (1986), pp. 133-169.
- COOK, R.D. AND WEISBERG, S., *Residuals and Influence in Regression*, New York: John Wiley (1980).
- DURBIN, J. AND WATSON, G.S., *Testing for serial correlation in least squares regression, II*, Biometrika 38 (1951), pp. 159-178.
- FULLER, W.A., *Introduction to Statistical Time Series*, New York, John Wiley (1976).
- GREEN, P.E. AND CARROLL, J.D., *Mathematical Tools for Applied Multivariate Analysis*, New York: Academic Press (1976).
- OBENCHAIN, R.L., Letter to the editor and reply by R.D. Cook, Technometrics, 19 (1977), pp. 348-349.
- PREGIBON, D., *Logistic regression diagnostics*, Ann. Statist., 9 (1981), pp. 705-724.

**Recent IMA Preprints
Title**

#	Author/s	Title
641	Peter J. Olver,	Canonical Forms and Integrability of BiHamiltonian Systems
642	S.A. van Gils, M.P. Krupa and W.F. Langford,	Hopf bifurcation with nonsemisimple 1:1 Resonance
643	R.D. James and D. Kinderlehrer,	Frustration in ferromagnetic materials
644	Carlos Rocha,	Properties of the attractor of a scalar parabolic P.D.E.
645	Debra Lewis,	Lagrangian block diagonalization
646	Richard C. Churchill and David L. Rod,	On the determination of Ziglin monodromy groups
647	Xinfu Chen and Avner Friedman,	A nonlocal diffusion equation arising in terminally attached polymer chains
648	Peter Gritzmann and Victor Klee,	Inner and outer j - Radii of convex bodies in finite- dimensional normed spaces
649	P. Szmolyan,	Analysis of a singularly perturbed traveling wave problem
650	Stanley Reiter and Carl P. Simon,	Decentralized dynamic processes for finding equilibrium
651	Fernando Reitich,	Singular solutions of a transmission problem in plane linear elasticity for wedge-shaped regions
652	Russell A. Johnson,	Cantor spectrum for the quasi-periodic Schrödinger equation
653	Wenxiong Liu,	Singular solutions for a convection diffusion equation with absorption
654	Deborah Brandon and William J. Hrusa,	Global existence of smooth shearing motions of a nonlinear viscoelastic fluid
655	James F. Reineck,	The connection matrix in Morse–Smale flows II
656	Claude Baesens, John Guckenheimer, Seunghwan Kim and Robert Mackay,	Simple resonance regions of torus diffeomorphisms
657	Willard Miller, Jr.,	Lecture notes in radar/sonar: Topics in Harmonic analysis with applica- tions to radar and sonar
658	Calvin H. Wilcox,	Lecture notes in radar/sonar: Sonar and Radar Echo Structure
659	Richard E. Blahut,	Lecture notes in radar/sonar: Theory of remote surveillance algorithms
660	D.V. Anosov,	Hilbert's 21st problem (according to Bolibruch)
661	Stephane Laederich,	Ray–Singer torsion for complex manifolds and the adiabatic limit
662	Geneviève Raugel and George R. Sell,	Navier-Stokes equations in thin 3d domains: Global regularity of solutions I
663	Emanuel Parzen,	Time series, statistics, and information
664	Andrew Majda and Kevin Lamb,	Simplified equations for low Mach number combustion with strong heat release
665	Ju. S. Il'yashenko,	Global analysis of the phase portrait for the Kuramoto–Sivashinsky equation
666	James F. Reineck,	Continuation to gradient flows
667	Mohamed Sami Elbially,	Simultaneous binary collisions in the collinear N–body problem
668	John A. Jacquez and Carl P. Simon,	Aids: The epidemiological significance of two different mean rates of partner-change
669	Carl P. Simon and John A. Jacquez,	Reproduction numbers and the stability of equilibria of SI models for heterogeneous populations
670	Matthew Stafford,	Markov partitions for expanding maps of the circle
671	Ciprian Foias and Edriss S. Titi,	Determining nodes, finite difference schemes and inertial manifolds
672	M.W. Smiley,	Global attractors and approximate inertial manifolds for abstract dissipative equations
673	M.W. Smiley,	On the existence of smooth breathers for nonlinear wave equations
674	Hitay Özbay and Janos Turi,	Robust stabilization of systems governed by singular integro-differential equations
675	Mary Silber and Edgar Knobloch,	Hopf bifurcation on a square lattice
676	Christophe Golé,	Ghost circles for twist maps
677	Christophe Golé,	Ghost tori for monotone maps
678	Christophe Golé,	Monotone maps of $T^n \times R^n$ and their periodic orbits
679	E.G. Kalnins and W. Miller, Jr.,	Hypergeometric expansions of Heun polynomials
680	Victor A. Pliss and George R. Sell,	Perturbations of attractors of differential equations
681	Avner Friedman and Peter Knabner,	A transport model with micro- and macro-structure
682	E.G. Kalnins and W. Miller, Jr.,	A note on group contractions and radar ambiguity functions
683	George R. Sell,	References on dynamical systems
684	Shui-Nee Chow, Kening Lu and George R. Sell,	Smoothness of inertial manifolds
685	Shui-Nee Chow, Xiao-Biao Lin and Kening Lu,	Smooth invariant foliations in infinite dimensional spaces
686	Kening Lu,	A Hartman–Grobman theorem for scalar reaction-diffusion equations
687	Christophe Golé and Glen R. Hall,	Poincaré's proof of Poincaré's last geometric theorem
688	Mario Taboada,	Approximate inertial manifolds for parabolic evolutionary equations via Yosida approxi- mations

- 689 **Peter Rejto and Mario Taboada**, Weighted resolvent estimates for Volterra operators on unbounded intervals
- 690 **Joel D. Avrin**, Some examples of temperature bounds and concentration decay for a model of solid fuel combustion
- 691 **Susan Friedlander and Misha M. Vishik**, Lax pair formulation for the Euler equation
- 692 **H. Scott Dumas**, Ergodization rates for linear flow on the torus
- 693 **A. Eden, A.J. Milani and B. Nicolaenko**, Finite dimensional exponential attractors for semilinear wave equations with damping
- 694 **A. Eden, C. Foias, B. Nicolaenko & R. Temam**, Inertial sets for dissipative evolution equations
- 695 **A. Eden, C. Foias, B. Nicolaenko & R. Temam**, Hölder continuity for the inverse of Mañé's projection
- 696 **Michel Chipot and Charles Collins**, Numerical approximations in variational problems with potential wells
- 697 **Huanan Yang**, Nonlinear wave analysis and convergence of MUSCL schemes
- 698 **László Gerencsér and Zsuzsanna Vágó**, A strong approximation theorem for estimator processes in continuous time
- 699 **László Gerencsér**, Multiple integrals with respect to L -mixing processes
- 700 **David Kinderlehrer and Pablo Pedregal**, Weak convergence of integrands and the Young measure representation
- 701 **Bo Deng**, Symbolic dynamics for chaotic systems
- 702 **P. Galdi, D.D. Joseph, L. Preziosi, S. Rionero**, Mathematical problems for miscible, incompressible fluids with Korteweg stresses
- 703 **Charles Collins and Mitchell Luskin**, Optimal order error estimates for the finite element approximation of the solution of a nonconvex variational problem
- 704 **Peter Gritzmann and Victor Klee**, Computational complexity of inner and outer j -radii of polytopes in finite-dimensional normed spaces
- 705 **A. Ronald Gallant and George Tauchen**, A nonparametric approach to nonlinear time series analysis: estimation and simulation
- 706 **H.S. Dumas, J.A. Ellison and A.W. Sáenz**, Axial channeling in perfect crystals, the continuum model and the method of averaging
- 707 **M.A. Kaashoek and S.M. Verduyn Lunel**, Characteristic matrices and spectral properties of evolutionary systems
- 708 **Xinfu Chen**, Generation and Propagation of interfaces in reaction diffusion systems
- 709 **Avner Friedman and Bei Hu**, Homogenization approach to light scattering from polymer-dispersed liquid crystal films
- 710 **Yoshihisa Morita and Shuichi Jimbo**, ODEs on inertial manifolds for reaction-diffusion systems in a singularly perturbed domain with several thin channels
- 711 **Wenxiong Liu**, Blow-up behavior for semilinear heat equations: multi-dimensional case
- 712 **Hi Jun Choe**, Hölder continuity for solutions of certain degenerate parabolic systems
- 713 **Hi Jun Choe**, Regularity for certain degenerate elliptic double obstacle problems
- 714 **Fernando Reitich**, On the slow motion of the interface of layered solutions to the scalar Ginzburg–Landau equation
- 715 **Xinfu Chen and Fernando Reitich**, Local existence and uniqueness of solutions of the Stefan problem with surface tension and kinetic undercooling
- 716 **C.C. Lim, J.M. Pimbley, C. Schmeiser and D.W. Schwendeman**, Rotating waves for semiconductor inverter rings
- 717 **W. Balsler, B.L.J. Braaksma, J.-P. Ramis and Y. Sibuya**, Multisummability of formal power series solutions of linear ordinary differential equations
- 718 **Peter J. Olver and Chehrzad Shakiban**, Dissipative decomposition of partial differential equations
- 719 **Clark Robinson**, Homoclinic bifurcation to a transitive attractor of Lorenz type, II
- 720 **Michelle Schatzman**, A simple proof of convergence of the QR algorithm for normal matrices without shifts
- 721 **Ian M. Anderson, Niky Kamran and Peter J. Olver**, Internal, external and generalized symmetries
- 722 **C. Foias and J.C. Saut**, Asymptotic integration of Navier–Stokes equations with potential forces. I
- 723 **Ling Ma**, The convergence of semidiscrete methods for a system of reaction-diffusion equations
- 724 **Adelina Georgescu**, Models of asymptotic approximation
- 725 **A. Makagon and H. Salehi**, On bounded and harmonizable solutions on infinite order arma systems
- 726 **San-Yih Lin and Yan-Shin Chin**, An upwind finite-volume scheme with a triangular mesh for conservation laws
- 727 **J.M. Ball, P.J. Holmes, R.D. James, R.L. Pego & P.J. Swart**, On the dynamics of fine structure
- 728 **KangPing Chen and Daniel D. Joseph**, Lubrication theory and long waves
- 729 **J.L. Ericksen**, Local bifurcation theory for thermoelastic Bravais lattices
- 730 **Mario Taboada and Yuncheng You**, Some stability results for perturbed semilinear parabolic equations
- 731 **A.J. Lawrance**, Local and deletion influence
- 732 **Bogdan Vernescu**, Convergence results for the homogenization of flow in fractured porous media