Statistical Papers 41, 211–224 (2000)

Statistical Papers

On local influence for elliptical linear models

Shuangzhe Liu

Received: June 4, 1998; revised version: May 5, 1999

Abstract:

The local influence method plays an important role in regression diagnostics and sensitivity analysis. To implement it, we need the Delta matrix for the underlying scheme of perturbations, in addition to the observed information matrix under the postulated model. Galea, Paula and Bolfarine (1997) has recently given the observed information matrix and the Delta matrix for a scheme of scale perturbations and has assessed of local influence for elliptical linear regression models. In the present paper, we consider the same elliptical linear regression models. We study the schemes of scale, predictor and response perturbations, and obtain their corresponding Delta matrices, respectively. To illustrate the methodology for assessment of local influence for these schemes and the implementation of the obtained results, we give an example.

Keywords: Likelihood displacement, observed information matrix, Delta matrix, regression diagnostics, matrix differential.

AMS Subject classification: 62J05.

1 Introduction

During the last three decades, elliptical distributions-based linear models and multivariate analysis have been developed as stimulating and fruitful fields in statistics and econometrics, see, e.g. Fang and Anderson (1990), Fang and Zhang (1990) and Kollo and Neudecker (1993, 1997). Meanwhile, regression diagnostic techniques useful for many fields have been studied and applied extensively, see, e.g Chatterjee and Hadi (1988) and Peña (1997). The local influence method originated with Cook (1986) has, along with other new methods, been paid considerable attention. We see that the local influence method has an advantage over other methods in several situations, see, e.g. Cook (1997). A comparison of the local influence method with the influence function method and the case deletion method can be found in, e.g. Jung, Kim and Kim (1997). For a useful discussion and some historical notes on the concept of influence, we refer to Farebrother (1992, 1999).

For elliptical linear regression models, Galea, Paula and Bolfarine (1997) has recently assessed a local influence analysis. It establishes the observed information matrix under the postulated model, but deals with only one scheme of scale perturbations. The purpose of the present paper is to consider further studies on local influence for the same elliptical linear models treated in Galea et al. (1997), and to derive Delta matrices for schemes not only of scale perturbations but also of predictor perturbations and response perturbations. The structure of the paper is as follows: Section 2 introduces the elliptical linear models and Section 3 an outline of the local influence method; Sections 4, 5 and 6 each derive Delta matrices for a scheme of scale, predictor and response perturbations; Section 7 gives an example with Ruppert and Carroll's (1980) data to assess local influence by using some of the obtained results.

2 Elliptical linear models

For an introduction to elliptical linear models, we refer to Fang and Zhang (1990). We denote $z \sim El_n(\mu, \Lambda)$, if z is an $n \times 1$ random vector with density function

$$f(z) = |\Lambda|^{-1/2} g[(z-\mu)'\Lambda^{-1}(z-\mu)], \tag{1}$$

where μ is an $n \times 1$ location vector, Λ is an $n \times n$ positive definite scale matrix, $g = g() \geq 0$ is a scalar function (density generator) such that

$$\int_0^\infty u^{n-1} g(u^2) du < \infty.$$

In particular, when μ=0 and $\Lambda = \phi I$ we have the spherical family of densities $z \sim El_n(0, \phi I)$. The class of symmetric distributions includes the normal, Student t- and other distributions.

Consider the following elliptical linear model (see, e.g. Fang and Anderson, 1990 and Galea et al. 1997):

$$y = X\beta + \epsilon, \tag{2}$$

where y is an $n \times 1$ observation vector, $X = (x_1, ..., x_n)'$ is an $n \times p$ model matrix of rank p, β is a $p \times 1$ unknown parameter vector, ϵ is an $n \times 1$ error vector with elliptical distribution $El_n(0, \phi I)$. If g is continuous and decreasing, then the maximum likelihood estimators $\hat{\beta}$ and $\hat{\phi}$

$$\begin{aligned} \hat{\beta} &= (X'X)^{-1}X'y, \\ \hat{\phi} &= \frac{e'e}{u_g}, \end{aligned}$$

where $e = y - X\hat{\beta}$ and u_g maximizes the function $h(u) = u^{n/2}g(u), u \geq 0$. If g is continuous and decreasing, then its maximum u_g exists and is finite and positive; moreover, if g is continuous and differentiable, then u_g is the solution to

$$W(u) + \frac{n}{2u} = 0, \tag{3}$$

where

$$W(u) = \frac{g'(u)}{g(u)}.$$

We know that $W(u) = -\frac{1}{2}$, $W'(u) = 0$ and $u_g = n$ for the normal distribution. For $g(u)$, $W(u)$ and $W'(u)$ of several other elliptical distributions used in assessment of local influence, see Galea et al. (1997).

3 Local influence

Local influence is a method of sensitivity analysis for assessing the influence of small perturbations in a general statistical model. Cook (1986, 1997) introduces the idea with key concepts to implement procedures for local influence analysis. Let $\omega = (\omega_1, ..., \omega_q)'$ denote a $q \times 1$ vector of perturbations confined to some open subset Q of $\Re^q$. Let $L(\theta|\omega)$ and $L(\theta)$ denote the log-likelihood functions of the perturbed and postulated (i.e. unperturbed) models respectively. Assume that the postulated model is nested within the perturbed one and there is such a vector ω_0 that $L(\theta) = L(\theta|\omega_0)$ for all values of θ in the parameter space. Cook (1986) suggests the likelihood displacement

$$LD(\omega) = 2[L(\hat{\theta}) - L(\hat{\theta}_\omega)]$$

to measure the difference between $\hat{\theta}$ and $\hat{\theta}_\omega$ by using the contours of the loglikelihood function $L(\theta)$ for the postulated model, where $\hat{\theta}$ and $\hat{\theta}_\omega$ are the maximum likelihood estimates under the two models respectively.

The geometric normal curvature $C(l)$ can be used to characterize $LD(\omega_0 + tl)$ around $t = 0$, where t is a scalar and l is a direction vector in Q of length 1. The direction of maximum curvature l_{max} shows how to perturb the postulated model to obtain the greatest local change in the likelihood displacement. The curvature in direction l is computed as

$$C_l(\theta) = 2|l'\Delta' H^{-1} \Delta l|, \tag{4}$$

where $||l|| = 1$, $-H = -H_\theta(\hat{\theta})$ is the observed information matrix for the postulated model and $\Delta = \Delta_\theta(\hat{\theta}, \omega_0)$ is the Delta matrix evaluated at $\theta = \hat{\theta}$ and $\omega = \omega_0$:

$$H_\theta = \frac{\partial^2 L(\theta)}{\partial\theta\partial\theta'}, \quad \Delta_\theta = \frac{\partial^2 L(\theta|\omega)}{\partial\theta\partial\omega'}. \tag{5}$$

Thus, l_{max} is the eigenvector corresponding to the largest absolute eigenvalue λ_{max} of $B = \Delta' H^{-1} \Delta$, which should be calculated. The scatter plot of $|l_{max}|$ may be helpful to indicate which observation is most influential.

When $\theta = (\theta_1', \theta_2')'$ and only θ_1 is of interest, we partition H according to the partition of θ and let $B_{22} = diag(0, H_{22}^{-1})$. Then

$$C_l(\theta_1) = 2|l'\Delta'(H^{-1} - B_{22})\Delta l|, \tag{6}$$

and we have to examine the eigenvector l_{max} of $\Delta'(H^{-1} - B_{22})\Delta$ instead.

To obtain H and Δ, we first use the standard matrix differential method, see Magnus and Neudecker (1999), to derive $d_\theta^2 L(\theta) = (d\theta)' H_\theta d\theta$ for the postulated log-likelihood and $d_{\theta\omega}^2 L(\theta|\omega) = (d\theta)'\Delta_\theta d\omega$ for the perturbed log-likelihood with H_θ and Δ_θ defined in (5). We then evaluate $d_\theta^2 L(\theta)$ and $d_{\theta\omega}^2 L(\theta|\omega)$ (rather than H_θ and Δ_θ) at $\theta = \hat{\theta}$ and $\omega = \omega_0$. In Sections 4 through 6, we focus on the elliptical linear models and derive Delta matrices Δ_θ corresponding to the perturbed models of different schemes, respectively.

4 Scale perturbations

For model (2), we have the postulated log-likelihood function

$$L(\theta) = -\frac{n}{2} log\ \phi + log\ g(u), \tag{7}$$

where $\theta = (\beta', \phi)'$, $u = \phi^{-1}\epsilon'\epsilon$, $\epsilon = y - X\beta$ and $\epsilon \sim El_n(0, \phi I)$.

For case-weight perturbations, Galea et al. (1997) present

$$H = \begin{pmatrix} 2W(\hat{u})\hat{\phi}^{-1}X'X & 0 \\ 0 & [\frac{n}{2} + W'(\hat{u})u_g^2 + 2W(\hat{u})u_g]\hat{\phi}^{-2} \end{pmatrix}, \tag{8}$$

$$\Delta = \begin{pmatrix} -2W(\hat{u})\hat{\phi}^{-1}X'D(e) \\ -[W'(\hat{u})u_g + W(\hat{u})]\hat{\phi}^{-2}e'D(e) \end{pmatrix}, \tag{9}$$

and especially, for the normal distribution case

$$H_{nor} = \begin{pmatrix} -\hat{\phi}^{-1}X'X & 0 \\ 0 & -\frac{n}{2}\hat{\phi}^{-2} \end{pmatrix}, \tag{10}$$

$$\Delta_{nor} = \begin{pmatrix} \hat{\phi}^{-1}X'D(e) \\ \frac{1}{2}\hat{\phi}^{-2}e'D(e) \end{pmatrix}, \tag{11}$$

where $u_g = \hat{u} = \hat{\phi}^{-1}e'e$, $e = y - X\hat{\beta} = (e_1, ..., e_n)'$ and $D(e) = diag(e_1, ..., e_n)$.

Now, we study further cases. When ϕ is known, we have the relevant part of the perturbed log-likelihood function

$$L(\theta|\omega) = log\ g(u_w), \tag{12}$$

where $u_w = \phi^{-1}\epsilon'D(\omega)\epsilon$, $\epsilon = y - X\beta$, $\epsilon \sim El_n(0, \phi D^{-1}(\omega))$, $D(\omega) = diag(\omega_1, ..., \omega_n)$ and $\omega = (\omega_1, ..., \omega_n)'$ with $q = n$, where ω_i is the weight of the i-th case ($i = 1, ..., n$). With this scheme, the perturbed model reduces to the postulated model when $\omega = \omega_0$, where $\omega_0 = (1, ..., 1)'$ is of order $n \times 1$.

Taking the differential of $L(\beta|w)$ with respect to β (as ϕ is known), we obtain

$$\begin{aligned} d_\beta L(\beta|w) &= W du_\beta \\ &= -2\phi^{-1}W\epsilon'D(w)Xd\beta. \end{aligned} \tag{13}$$

Then

$$\begin{aligned} d^2_\beta L(\beta|w) &= -2\phi^{-1}W'du_\beta\epsilon'D(w)Xd\beta \\ &\quad +2\phi^{-1}W(d\beta)'X'D(w)Xd\beta, \end{aligned} \tag{14}$$

$$\begin{aligned} d^2_{\beta w} L(\beta|w) &= -2\phi^{-1}W'du_w\epsilon'D(w)Xd\beta \\ &\quad -2\phi^{-1}W(d\beta)'X'D(\epsilon)dw, \end{aligned} \tag{15}$$

where $\epsilon' D(w) X d\beta = (d\beta)' X' D(\epsilon) w$ is used.

Evaluating (14) and (15) at $(\beta, \omega) = (\hat{\beta}, \omega_0)$ and noting that $D(w_0) = I$, $e'X = 0$, $e = \hat{\epsilon}$ and $W = W(\hat{u})$, we obtain

$$\begin{aligned} d^2_\beta L(\beta|w)|_{\beta=\hat{\beta}} &= 2(d\beta)'\phi^{-1}W(\hat{u})X'Xd\beta, & (16) \\ d^2_{\beta w} L(\beta|w)|_{(\beta=\hat{\beta}, w=w_0)} &= -2(d\beta)'\phi^{-1}W(\hat{u})X'D(e)dw, & (17) \end{aligned}$$

and therefore

$$\begin{aligned} H &= 2\phi^{-1}W(\hat{u})X'X, & (18) \\ \Delta &= -2\phi^{-1}W(\hat{u})X'D(e). & (19) \end{aligned}$$

If we consider individual cases where only the weight for the i-th case is perturbed, we define $D(\omega) = diag(1, ..., 1, \omega, 1, ..., 1)$ of order $n \times n$. When only β is of interest, H is given by (18) and

$$\Delta = -2\phi^{-1}W(\hat{u})x_i e_i. \quad (20)$$

Furthermore, the curvature is found to be

$$C_l(\beta) = 4|\phi^{-1}W(\hat{u})e_i^2 x_i'(X'X)^{-1}x_i|. \quad (21)$$

In the normal distribution case (with $W = -\frac{1}{2}$), replacing ϕ by its unbiased estimator, (21) becomes identical to (32) in Cook (1986), which shows the connection between the local influence and the Cook's distance in the simple multiple regression case.

5 Predictor perturbations

First, consider the perturbations in the first column of the predictor matrix: X is replaced by $X + \omega a's$, where $\omega = (\omega_1, ..., \omega_n)'$ is of $n \times 1$, $a = (1, 0, ..., 0)'$ is of $p \times 1$ and s is the (scalar) scale factor. With this scheme, the perturbed model reduces to the postulated model when $\omega = \omega_0 = 0$. The relevant part of the perturbed log-likelihood is

$$L(\theta|\omega) = log\ g(u_w), \tag{22}$$

where $u_w = \phi^{-1}\epsilon'\epsilon$ and $\epsilon = y - X\beta - \omega a' s\beta$.

Taking the differential of $L(\theta|\omega)$ with respect to first $\theta = (\beta', \phi)'$ and then to ω, we have

$$d_\theta L(\theta|\omega) = -W\phi^{-2}d\phi\epsilon'\epsilon - 2W\phi^{-1}(d\beta)'(X + \omega a's)'\epsilon, \tag{23}$$

and

$$\begin{aligned} d^2_{\beta\omega}L(\theta|\omega) &= 4W'\phi^{-2}(d\beta)'(X + \omega a's)'\epsilon\epsilon' d\omega a' s\beta \\ &\quad +2W\phi^{-1}(d\beta)'(X + \omega a's)' d\omega a' s\beta \\ &\quad -2W\phi^{-1}(d\beta)' a s\epsilon' d\omega, \qquad (24) \\ d^2_{\phi\omega}L(\theta|\omega) &= 2W'\phi^{-3}d\phi\epsilon'\epsilon\epsilon' d\omega a' s\beta + 2W\phi^{-2}d\phi\epsilon' d\omega a' s\beta. \qquad (25) \end{aligned}$$

Evaluating (24) and (25) at $(\theta, \omega) = (\hat{\theta}, \omega_0)$ leads to

$$\Delta = \begin{pmatrix} 2W(\hat{u})\hat{\phi}^{-1}s(\hat{\beta}_1 X' - ae') \\ 2[W'(\hat{u})u_g + W(\hat{u})]\hat{\phi}^{-2}s\hat{\beta}_1 e' \end{pmatrix}, \tag{26}$$

where $e = y - X\hat{\beta}$, $X'e = 0$, $u_g = \hat{u} = \hat{\phi}^{-1}e'e$ and $a'\hat{\beta} = \hat{\beta}_1$.

In the normal case, (26) becomes

$$\Delta_{nor} = \begin{pmatrix} \hat{\phi}^{-1}s(ae' - \hat{\beta}_1 X') \\ -\hat{\phi}^{-2}s\hat{\beta}_1 e' \end{pmatrix}. \tag{27}$$

Now, consider the perturbations in all columns of the predictor matrix. The perturbed log-likelihood is constructed with X replaced by $X + \Omega S$, where $\Omega = (\omega_{ij}) = (\omega_1, ..., \omega_j, ..., \omega_p)$ is an $n \times p$ matrix of perturbations, $S = diag(s_1, ..., s_p)$ and s_j $(j = 1, ..., p)$ is the scale factor. The perturbed model reduces to the postulated model when $\omega = \omega_0 = 0$. We obtain

$$\Delta = (\Delta_1, ..., \Delta_p), \tag{28}$$

where

$$
\begin{aligned}
\Delta_j &= \frac{\partial^2 L}{\partial\beta\partial\omega_j'} \\
&= \begin{pmatrix} 2W(\hat{u})\hat{\phi}^{-1}s_j(\hat{\beta}_j X' - a_j e') \\ 2[W'(\hat{u})u_g + W(\hat{u})]\hat{\phi}^{-2}s_j\hat{\beta}_j e' \end{pmatrix},
\end{aligned}
$$

and a_j is a $p \times 1$ vector with one in the j-th position and zeros elsewhere.

Based on H in (8) and Δ in (26), or (28), we can find $B = \Delta' H^{-1}\Delta$, the curvature $C_l(\theta)$ and therefore the maximum direction l_{max}. In particular, using H_{nor} in (10) and Δ_{nor} in (27) we can get B, $C_l(\theta)$ and l_{max} in the normal case.

Again, consider the perturbations in the first column of X. By using (7), (8) and (26), we can write

$$
B = \Delta' H^{-1}\Delta = B_1 + B_2, \tag{29}
$$

where

$$
\begin{aligned}
B_1 &= 2W(\hat{u})\hat{\phi}^{-1}s^2(X\hat{\beta}_1 - ea')(X'X)^{-1}(\hat{\beta}_1 X' - ae'), \\
B_2 &= Cee' \\
C &= \frac{4[W'(\hat{u})u_g + W(\hat{u})]^2\hat{\phi}^{-2}s^2\hat{\beta}_1^2}{\frac{n}{2} + W'(\hat{u})u_g^2 + 2W(\hat{u})u_g}.
\end{aligned}
$$

Then the curvature is $C_l(\theta) = 2|l'(B_1 + B_2)l|$. In particular, if we are interested in only the vector β, the curvature becomes $C_l(\beta) = 2|l'B_1 l|$. Similarly, the curvature for only the scale parameter ϕ is

$$
\begin{aligned}
C_l(\phi) &= 2|l'B_2 l| \\
&= 2|C||l'ee'l|.
\end{aligned}
$$

Then, for the largest curvature, $l_{max} \propto e$, which means that the observations with large absolute values of e_i exercise the most influence on $\hat{\phi}$.

6 Response perturbations

Consider the response perturbations in which y is replaced by the perturbed response $y + \omega s$. The perturbation vector ω is of order $n \times 1$, $\omega_0 = 0$ and s is the (scalar) scale factor. The relevant part of the perturbed log-likelihood is

$$L(\theta|\omega) = log\ g(u_w), \tag{30}$$

where $u_w = \phi^{-1}\epsilon'\epsilon$ and $\epsilon = y + \omega s - X\beta$.

Taking the differential of $L(\theta|\omega)$ with respect to $\theta = (\beta', \phi)'$ we have

$$d_\theta L(\theta|\omega) = -W\phi^{-2}d\phi\epsilon'\epsilon - 2W\phi^{-1}(d\beta)'X'\epsilon. \tag{31}$$

Then

$$\begin{aligned} d^2_{\beta\omega}L(\theta|\omega) &= -4W'\phi^{-2}(d\beta)'X'\epsilon\epsilon' s d\omega - 2W\phi^{-1}(d\beta)'X' s d\omega, & (32)\\ d^2_{\phi\omega}L(\theta|\omega) &= -2W'\phi^{-3}d\phi\epsilon'\epsilon\epsilon' s d\omega - 2W\phi^{-2}d\phi\epsilon' s d\omega. & (33) \end{aligned}$$

From evaluating (32) and (33) at $(\theta, \omega) = (\hat{\theta}, \omega_0)$ it follows that

$$\Delta = \begin{pmatrix} -2W(\hat{u})\hat{\phi}^{-1}X's \\ -2[W'(\hat{u})u_g + W(\hat{u})]\hat{\phi}^{-2}e's \end{pmatrix}, \tag{34}$$

where $e = y - X\hat{\beta}$, $X'e = 0$ and $u_g = \hat{u} = \hat{\phi}^{-1}e'e$.

For the normal case, (34) reduces to

$$\Delta_{nor} = \begin{pmatrix} \hat{\phi}^{-1}X's \\ \hat{\phi}^{-2}e's \end{pmatrix}. \tag{35}$$

Based on H in (8) and Δ in (34) we can find $B = \Delta' H^{-1}\Delta$, the curvature $C_l(\theta)$ and the maximum direction l_{max}. Using H_{nor} in (10) and Δ_{nor} in (35) we can get B, $C_l(\theta)$ and l_{max} in the normal case.

7 Example

The data set of $n = 28$ observations on the salinity of water during the spring in Pamlico Sound, North Carolina is reported and studied by Ruppert and Carroll (1980). It is also examined by Aiktson (1985), Davison and Tsai (1992) and Galea et al. (1997). To illustrate the methodology described and the results obtained in the current paper, we just examine the same data.

The linear regression model for the data is assumed to be

$$y = X\beta + \epsilon, \tag{36}$$

where $X = (1, x_2, x_3, x_4)$, 1 is an 28×1 vector of ones, x_2 is salinity lagged two weeks, x_3 is a dummy variable for the time period, x_4 is river discharge, y is biweekly salinity, and ϵ is assumed to follow a normal distribution or a t-distribution with 3 degrees of freedom.

Both Aiktson (1985) and Davison and Tsai (1992) use the deletion method. Under the normal assumption of errors, Aiktson (1985) finds observations 16 and 5 most influential. Under a t-distribution with 3 degrees of freedom, Davison and Tsai (1992) finds observations 16, 5 and 3 most influential. Using the local influence method under both distributions Galea et al. (1997) specifies only observation 16 as most influential, and therefore comments that the scatter plot of $|l_{max}|$ for $\hat{\theta}$ may be helpful in selecting the less sensitive model with respect to local perturbations in the elliptical linear family.

Based on (29), (8) and (34) we compute $C_l(\theta)$ to obtain the corresponding l_{max} for two cases, both under a t-distribution with 3 degrees of freedom. We present two corresponding scatter plots of $|l_{max}|$. In the first case, we consider the scheme of perturbations of x_1 and find observation 16 most influential, as shown in Figure 1. In the second case, we consider perturbations of y and find observations 16 and 5 most influential, as shown in Figure 2. These two figures suggest accordance with Galea et al.'s (1997) comment.

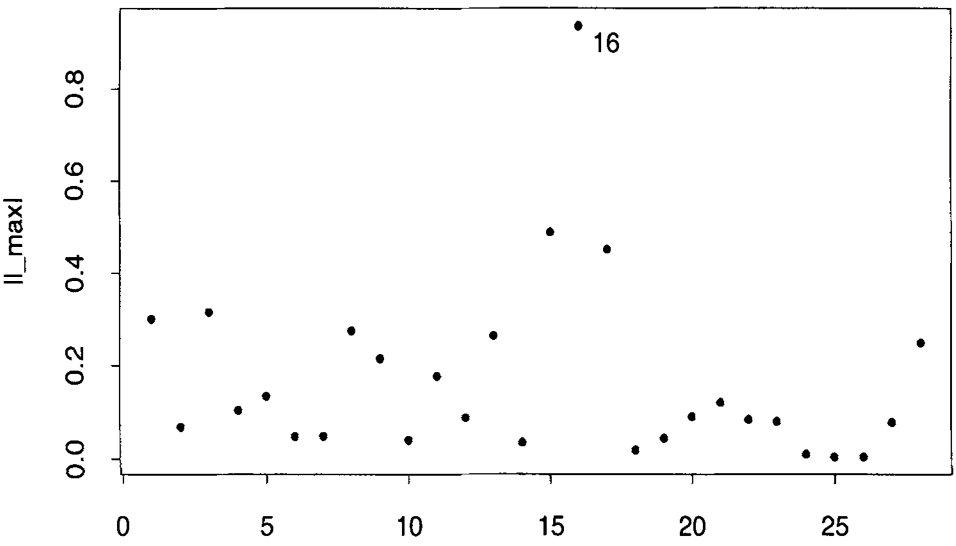

Figure 1

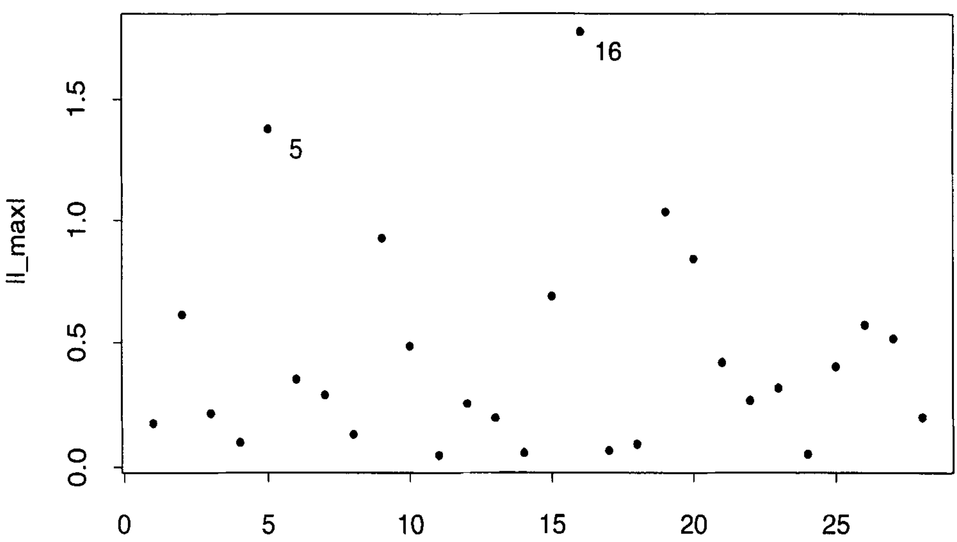

Figure 2

Acknowledgements

The author is very grateful to Professor Götz Trenkler and the referees for their valuable advice and helpful comments on an early version of the paper.

8 References

Aiktson, A.C. (1985) *Plots, Transformations, and Regression*, Clarendon, Oxford.

Chatterjee, S. and Hadi, A. S. (1988) *Sensitivity Analysis in Linear Regression*, Wiley, New York.

Cook, R.D. (1986) Assessment of local influence (with discussion), *J. R. Statist. Soc. B*, 48, 133-169.

Cook, R.D. (1997) Local influence, in *Encyclopedia of Statist. Sciences*, Kotz, S., Read, C.B. and Banks, D.L. eds., Wiley, New York, Update Vol. 1, 380-385.

Davison, A.C. and Tsai, C.-L. (1992) Regression model diagnostics, *Int. Statist. Rev.*, 60, 337-353.

Fang, K.T. and Anderson, T.W. (1990) *Statistical Inferences in Elliptical Contoured and Related Distributions*, Allerton, New York.

Fang, K.T. and Zhang, Y. (1990) *Generalized Multivariate Analysis*, Science Press, Beijing/Springer, Berlin.

Farebrother, R.W. (1992) Relative local influence and the condition number, *Commun. Statist.-Simulation. Comput.*, 21, 707-710.

Farebrother, R.W. (1999) *Fitting Linear Relationships: A History of the Calculus of Observations 1750-1900*, Springer, New York.

Galea, M., Paula, G.A. and Bolfarine, H. (1997) Local influence in elliptical linear regression models, *The Statistician*, 46(1), 71-79.

Jung, K.M., Kim, M.G. and Kim, B.C. (1997) Second order local influence in linear discriminant analysis, *J. Japan. Soc. Comp. Statist.*, 10(1), 1-11.

Kollo, T. and Neudecker, H. (1993) Asymptotics of eigenvalues and unit-length eigenvectors of sample variance and correlation matrices, *J. Multivar. Anal.*, 47, 283-300. Corrigendum, 51, 210.

Kollo, T. and Neudecker, H. (1997) Asymptotics of Pearson-Hotelling principal-component vectors of sample variance and correlation matrices, *Behaviormetrika*, 24(1), 51-69.

Magnus, J.R. and Neudecker, H. (1999) *Matrix Differential Calculus with Applications in Statistics and Econometrics*, second edition, Wiley, Chichester.

Peña, D. (1997) Combining information in statistical modeling, *The Am. Statistician*, 51(4), 326-332.

Ruppert, D. and Carroll, R.J. (1980) Trimmed least squares estimation in the linear model, *J. Am. Statist. Ass.*, 75, 828-838.

Shuangzhe Liu
Institut für Statistik und Ökonometrie
Universität Basel
Holbeinstraße 12
4051 Basel, Switzerland