ELSEVIER Linear Algebra and its Applications 354 (2002) 159–174

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

# Local influence in multivariate elliptical linear regression models

Shuangzhe Liu

*School of Mathematical Sciences, Australian National University, Canberra, ACT 0200, Australia*

Received 14 April 2000; accepted 16 November 2001

Submitted by H.J. Werner

**Abstract**

Local influence is a method of sensitivity analysis for assessing the influence of small perturbations in a general statistical model. In the present paper, this popular method is applied to multivariate elliptical linear regression models. Several schemes of perturbation, including perturbations in case-weights, explanatory variables and response variables are considered. The observed information matrix under the postulated model and Delta matrices under the corresponding perturbed models are derived. Assessment of local influence is made.


*AMS classification:* 62J05

*Keywords:* Diagnostics; Likelihood displacement; Local influence; Matrix differential; Multivariate elliptical regression

## 1. Introduction

The multivariate normal distribution, normal distribution-based linear models and multivariate analysis have played an essential role in statistics, see e.g. [1]. However, there are findings the normality assumption does not cope well in certain situations. Alternative distributions are then needed to consider and one choice is the elliptical distribution family. As known, studies and applications of non-normal elliptical distributions and statistical inference have been progressing rapidly during the last

*E-mail address:* lius@maths.anu.edu.au (S. Liu).


PII: S0024-3795(01)00585-7

two decades. For an extensive review on the non-normal cases, we refer to [2]. For a relevant paper, see e.g. [16,17], and for a comprehensive volume, see e.g. [8,13].

Meanwhile, sensitivity analysis and diagnostic techniques have been paid considerable attention, see e.g. [4]. Among other techniques, the local influence method with key concepts and applications to a simple normal distribution-based linear model has been popularized by Cook's landmark contribution [5]. The concept of the influence of an observation and some historical notes have been given in a very recently published monograph by Farebrother [10]. A useful discussion on local influence and its connection with other concepts can be found in [9,24,26]. A comparison of the local influence method with the influence function method and the case deletion method is presented by Jung et al. [14]. For local influence in ridge regression, see e.g. [25]. For Bayesian local influence, see e.g. [22]. Recently, the local influence method has been applied to multivariate normal linear regression models by Kim [15] and Fung and Tang [11]. Local influence analysis for elliptical linear regression models in the univariate case has been made by Galea et al. [12] and Liu [19]. However, local influence analysis for non-normal elliptical linear regression models in the multivariate case has not yet been developed.

In this paper, a general framework is introduced so that the local influence method can be applied to elliptical linear regression models in the multivariate case. As seen, the results obtained can reduce to those for the normal distribution assumption studied by Kim [15] and Fung and Tang [11]. In Section 2, the multivariate elliptical linear models and maximum likelihood estimation are introduced. In Section 3, the local influence method with the concepts of the observed information matrix and the so-called Delta matrix are outlined. In Section 4, the observed information matrix under the postulated model is given. In Sections 5, 6 and 7, Delta matrices under the perturbed models of perturbations in case-weights, explanatory variables and response variables are derived, respectively. Based on these results, an assessment of local influence is made. Remarks are given in Section 8.

## 2. Elliptical linear models

In this section, we briefly introduce the notations to be used in the paper. Early studies on maximum likelihood estimators for elliptical distributions are in [3] and some of the other papers collected in [7]. For further details on elliptical matrix distributions, linear models and maximum likelihood estimators, see e.g. [8,13].

Let $U = (u_1, \ldots, u_n)'$ be an $n \times p$ $(n \geqslant p)$ data matrix, where $u_1, \ldots, u_n$ can be viewed as a sample from a $p$-dimensional population. Consider $U$, following an elliptical matrix distribution such that

$$U \sim \mathrm{EM}_{np}(0, \Sigma, I_n, \psi), \tag{1}$$

where $\Sigma$ is a $p \times p$ positive definite scale matrix, $I_n$ is an $n \times n$ identity matrix and $\psi$ is the characteristic generator. That is, $u_1, \ldots, u_n$ are uncorrelated and their joint distribution is elliptically contoured and absolutely continuous. If $U$ has finite first and second moments, then $E(\operatorname{vec} U') = 0$ and $\operatorname{Var}(\operatorname{vec} U') = -2\phi(0) I_n \otimes \Sigma$, where vec indicates the vectorization operator, which stacks the columns of a matrix one underneath the other, $\phi$ is the first derivative of $\psi$, and $\otimes$ indicates the Kronecker product. Suppose that (1) has a density of the form

$$f(U) = |\Sigma|^{-n/2} g\big(\operatorname{tr} U \Sigma^{-1} U'\big), \tag{2}$$

where $g$ is the known density generator.

This is an extension of the univariate distribution $\mathrm{El}_n(0, \alpha I_n)$ used by Galea et al. [12] and Liu [19] where $p = 1$, $\Sigma$ becomes a positive scalar and $\alpha > 0$ is the scale parameter.

Consider the following model:

$$Y = XB + U, \tag{3}$$

where $Y$ is an $n \times p$ observation matrix, $X$ is an $n \times m$ model matrix of full column rank, $B$ is an $m \times p$ unknown parameter matrix, and $U$ is an $n \times p$ error matrix as defined in (1) with $\Sigma$ unknown, in general.

Clearly, (3) covers the model under the normality investigated by Kim [15] and Fung and Tang [11] as a special case. Note that a multivariate model such as (3) cannot be simply reduced to an existing univariate model by vectorization. This is because in the multivariate model the second moment (when it exists) is a particular positive definite matrix $-2\phi(0) I_n \otimes \Sigma$, and obviously in the univariate model studied by Galea et al. [12] and Liu [19] a counterpart, say of $\alpha I_n$, is inadequate.

Assume that $h(z) = z^{np/2} g(z)$, $z \geqslant 0$, has a finite maximum at $z = z_g > 0$. Then we can see that the maximum likelihood estimators of $B$ and $\Sigma$ in model (3) are as follows:

$$\hat{B} = (X'X)^{-1} X'Y, \tag{4}$$

$$\begin{aligned}\hat{\Sigma} &= \frac{p}{z_g}(Y - X\hat{B})'(Y - X\hat{B}) \\ &= \frac{p}{z_g} Y'\big(I_n - X(X'X)^{-1}X'\big)Y \\ &= \frac{p}{z_g} \hat{U}'\hat{U},\end{aligned} \tag{5}$$

where

$$\hat{U} = \big(I_n - X(X'X)^{-1}X'\big)Y. \tag{6}$$

For (2)–(6), see e.g. [13, Chapter 9]. We see that $\hat{\Sigma}$ in (5) is dependent on $\hat{B}$ in (4); for the relation of $\hat{\Sigma}$ to $\hat{B}$ under the normality assumption, we refer to [1, Section 8.2]. For additional background when $p > 1$, see [8]. In particular, if $p = 1$, then model (3) changes to the univariate case studied by Galea et al. [12] and Liu [19]. Accordingly, (4) and (5) above reduce to (3) in [12].

Moreover, for continuous and differentiable $g$, we define

$$G = G(z) = \frac{\partial \ln g(z)}{\partial z} = \frac{g'(z)}{g(z)}, \tag{7}$$

$$F = F(z) = \frac{\partial G(z)}{\partial z}. \tag{8}$$

We find for the multivariate normal distribution

$$g(z) = c_1 \exp\left(-\frac{z}{2}\right),$$
$$G(z) = -\frac{1}{2},$$
$$F(z) = 0,$$
$$z_g = np,$$

where $c_1$ is a normalizing constant and $z \geqslant 0$.

For the multivariate $t$ distribution with $r$ degrees of freedom

$$g(z) = c_2 \left(1 + \frac{z}{r}\right)^{-(np+r)/2},$$
$$G(z) = -\frac{np+r}{2r}\left(1 + \frac{z}{r}\right)^{-1},$$
$$F(z) = \frac{np+r}{2r^2}\left(1 + \frac{z}{r}\right)^{-2},$$
$$z_g = np,$$

where $c_2$ is a normalizing constant.

We refer to [13, Section 2.7] for $g(z)$, and [12] for $g(z)$, $G(z)$ and $F(z)$ when $p = 1$, used in the assessment of local influence of several elliptical distributions.

## 3. Local influence

To implement the procedures for local influence analysis we introduce several concepts and the method. For a recent introduction, see [6]. Let $\omega = (\omega_1, \dots, \omega_q)'$ denote a $q \times 1$ vector of perturbations confined to some open subset of $\mathscr{R}^q$ and $\omega_0$ denote a no-perturbation vector. Let $\theta$ indicate an $r \times 1$ vector of parameters of interest. Let $L(\theta)$ and $L(\theta, \omega) = L(\theta|\omega)$ denote the log-likelihood functions of the postulated (i.e. unperturbed) and the perturbed models, respectively. Note that $L(\theta) = L(\theta, \omega_0)$. The idea of the local influence method is to investigate how much the estimates are affected by those corresponding perturbations. The likelihood displacement $\mathrm{LD}(\omega)$ is useful to measure the distance between $\hat{\theta}$ and $\hat{\theta}_\omega$, which are the maximum likelihood estimates under the two models, respectively. It is given by

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

Define

$$H_\theta = \frac{\partial^2 L(\theta)}{\partial\theta\,\partial\theta'}, \quad \varDelta_\theta = \frac{\partial^2 L(\theta,\omega)}{\partial\theta\,\partial\omega'}, \tag{9}$$

where $H_\theta$ is the $r \times r$ Hessian matrix and $\varDelta_\theta$ is an $r \times q$ matrix.

Then, based on (9) evaluated at $\theta = \hat{\theta}$ and $\omega = \omega_0$ we can find $-H = -H_\theta(\hat{\theta})$, the observed information matrix for the postulated model and $\varDelta = \varDelta_\theta(\hat{\theta}, \omega_0)$, the so-called Delta matrix for the perturbed model. Actually, in Sections 4–7 we obtain $H$ and $\varDelta$ in a more efficient way as in [19]: We first use the matrix *differential* method (instead of *derivatives*), see e.g. [8, Section 1.5.4] and [21], to derive $\mathrm{d}^2_\theta L(\theta) = (\mathrm{d}\theta)' H_\theta \mathrm{d}\theta$ for the postulated log-likelihood and $\mathrm{d}^2_{\theta\omega} L(\theta|\omega) = (\mathrm{d}\theta)'\varDelta_\theta \mathrm{d}\omega$ for the perturbed log-likelihood with $H_\theta$ and $\varDelta_\theta$ defined in (9). Then we evaluate $\mathrm{d}^2_\theta L(\theta)$ and $\mathrm{d}^2_{\theta\omega} L(\theta|\omega)$ at $\theta = \hat{\theta}$ and $\omega = \omega_0$ to obtain $H$ and $\varDelta$; we do not need *explicit* expressions of $H_\theta$ and $\varDelta_\theta$.

Based on $\mathrm{LD}(\omega)$, Cook [5] shows that the curvature in direction $l$ is

$$C_l(\theta) = 2\left|l'\varDelta' H^{-1}\varDelta l\right|, \tag{10}$$

where $l$ is a $q \times 1$ vector of unit length. That is, $C_l(\theta)$ is the local influence on the estimation of $\theta$ of perturbing the postulated model; Large values of $C_l(\theta)$ indicate sensitivity to the induced perturbations in direction $l$.

We can then carry out our local influence analysis by finding $M = \varDelta' H^{-1}\varDelta$, its largest absolute eigenvalue $\lambda_{\max}$ and the associated eigenvector $l_{\max}$. If the absolute value of the $i$th element of $l_{\max}$ is the largest, then the $i$th observation of the data may be most influential. A nice way to examine this is to make an indexed scatter plot of $l_{\max}$. The plot may indicate which observations are more influential than the others.

When $\theta = (\theta_1', \theta_2')'$ and only $\theta_1$ is of interest, we partition $H$ according to the partition of $\theta$. Let

$$H = \begin{pmatrix} H_{11} & H_{21}' \\ H_{21} & H_{22} \end{pmatrix}, \quad A_{22} = \begin{pmatrix} 0 & 0 \\ 0 & H_{22}^{-1} \end{pmatrix}. \tag{11}$$

Then

$$C_l(\theta_1) = 2\left|l'\varDelta'(H^{-1} - A_{22})\varDelta l\right|, \tag{12}$$

and therefore we have to examine the eigenvector $l_{\max}$ associated with the largest eigenvalue $\lambda_{\max}$ of $M_1 = \varDelta'(H^{-1} - A_{22})\varDelta$.

## 4. Information matrix

Consider $\theta = (b', s')'$ with $b = \operatorname{vec} B$ and $s = \operatorname{vech} \Sigma$, where $b$ is an $mp \times 1$ vector, $s$ is an $(p+1)p/2 \times 1$ vector, vech denotes the vectorization operator which eliminates all supradiagonal elements of the matrix, and $\theta$ is an $r \times 1$ vector ($r = mp + (p+1)p/2$). The postulated log-likelihood of model (3) is

$$L = L(\theta) = -\frac{n}{2}\ln|\Sigma| + \ln g(z), \tag{13}$$

where $z = \operatorname{tr} U'U\Sigma^{-1}$ and $U = Y - XB$.

Taking the differentials of $L$ with respect to $B$, we obtain

$$\mathrm{d}_b L = G(\mathrm{d}_b z) = -2G \operatorname{tr} \Sigma^{-1}U'X(\mathrm{d}B), \tag{14}$$

$$\mathrm{d}_b^2 L = 4F \operatorname{tr}(\mathrm{d}B)'X'U\Sigma^{-1} \operatorname{tr} \Sigma^{-1}U'X(\mathrm{d}B) + 2G \operatorname{tr}(\mathrm{d}B)'X'X(\mathrm{d}B)\Sigma^{-1}. \tag{15}$$

Since $\hat{U} = (I - X(X'X)^{-1}X')Y$ as given in (6), and then $\hat{U}'X = 0$, we have

$$\begin{aligned}\mathrm{d}_b^2 L\big|_{\theta=\hat{\theta}} &= 2\hat{G} \operatorname{tr}(\mathrm{d}B)'X'X(\mathrm{d}B)\hat{\Sigma}^{-1} \\ &= 2\hat{G}(\mathrm{d}\operatorname{vec}B)'(\hat{\Sigma}^{-1} \otimes X'X)(\mathrm{d}\operatorname{vec}B),\end{aligned} \tag{16}$$

where $\hat{G} = G(\hat{z})$ with $\hat{z} = z_g$, and $\hat{\Sigma}$ is the same as in (5).

From (14) it follows that

$$\mathrm{d}_{sb}^2 L = -2(\mathrm{d}_s G) \operatorname{tr} \Sigma^{-1}U'X(\mathrm{d}B) - 2G \operatorname{tr}(\mathrm{d}_s\Sigma^{-1})U'X(\mathrm{d}B), \tag{17}$$

and then using (17) and $\hat{U}'X = 0$ leads to

$$\mathrm{d}_{sb}^2 L\big|_{\theta=\hat{\theta}} = 0. \tag{18}$$

Taking the differentials of $L$ in (13) with respect to $\Sigma$, we get

$$\begin{aligned}\mathrm{d}_s L &= -\frac{n}{2}\operatorname{tr} \Sigma^{-1}(\mathrm{d}\Sigma) + G(\mathrm{d}_s z) \\ &= -\frac{n}{2}\operatorname{tr} \Sigma^{-1}(\mathrm{d}\Sigma) - G \operatorname{tr} \Sigma^{-1}U'U\Sigma^{-1}(\mathrm{d}\Sigma),\end{aligned} \tag{19}$$

$$\begin{aligned}\mathrm{d}_s^2 L = {}& \frac{n}{2}\operatorname{tr}(\mathrm{d}\Sigma)\Sigma^{-1}(\mathrm{d}\Sigma)\Sigma^{-1} + F \operatorname{tr}(\mathrm{d}\Sigma)\Sigma^{-1}U'U\Sigma^{-1} \operatorname{tr} \Sigma^{-1}U'U\Sigma^{-1}(\mathrm{d}\Sigma) \\ & + 2G \operatorname{tr}(\mathrm{d}\Sigma)\Sigma^{-1}U'U\Sigma^{-1}(\mathrm{d}\Sigma)\Sigma^{-1},\end{aligned} \tag{20}$$

then

$$\begin{aligned}\mathrm{d}_s^2 L\big|_{\theta=\hat{\theta}} &= \frac{n}{2}(\mathrm{d}\operatorname{vech}\Sigma)'D'\big(\hat{\Sigma}^{-1} \otimes \hat{\Sigma}^{-1}\big)D(\mathrm{d}\operatorname{vech}\Sigma) \\ &\quad + \hat{F}(\mathrm{d}\operatorname{vech}\Sigma)'D'D\operatorname{vech}\big(\hat{\Sigma}^{-1}\hat{U}'\hat{U}\hat{\Sigma}^{-1}\big) \\ &\quad \times \operatorname{vech}'\big(\hat{\Sigma}^{-1}\hat{U}'\hat{U}\hat{\Sigma}^{-1}\big)D'D(\mathrm{d}\operatorname{vech}\Sigma) \\ &\quad + 2\,\hat{G}(\mathrm{d}\operatorname{vech}\Sigma)'D'\big(\hat{\Sigma}^{-1} \otimes \hat{\Sigma}^{-1}\hat{U}'\hat{U}\hat{\Sigma}^{-1}\big)D(\mathrm{d}\operatorname{vech}\Sigma) \\ &= \left(\frac{n}{2} + \frac{2z_g\hat{G}}{p}\right)(\mathrm{d}\operatorname{vech}\Sigma)'D'\big(\hat{\Sigma}^{-1} \otimes \hat{\Sigma}^{-1}\big)D(\mathrm{d}\operatorname{vech}\Sigma) \\ &\quad + \frac{z_g^2\hat{F}}{p^2}(\mathrm{d}\operatorname{vech}\Sigma)'D'D\operatorname{vech}(\hat{\Sigma}^{-1})\operatorname{vech}'(\hat{\Sigma}^{-1}) \\ &\quad \times D'D(\mathrm{d}\operatorname{vech}\Sigma),\end{aligned} \tag{21}$$

where $\hat{G} = G(\hat{z})$, $\hat{F} = F(\hat{z})$ and $\hat{U}'\hat{U} = \hat{z}\hat{\Sigma}/p$ with $\hat{z} = z_g$, and $D$ is the $p^2 \times (p + 1)p/2$ duplication matrix with $\operatorname{vec}\Sigma = D \operatorname{vech}\Sigma$. For properties of $D$, see e.g. [21].

Hence, it follows from (16), (18) and (21) that

$$H = \begin{pmatrix} 2\hat{G}(\hat{\Sigma}^{-1} \otimes X'X) & 0 \\ 0 & H_s \end{pmatrix}, \tag{22}$$

where

$$\begin{aligned} H_s = {}& \left(\frac{n}{2} + \frac{2z_g\hat{G}}{p}\right) D'\big(\hat{\Sigma}^{-1} \otimes \hat{\Sigma}^{-1}\big)D \\ & + \frac{z_g^2\hat{F}}{p^2} D'D \operatorname{vech}\big(\hat{\Sigma}^{-1}\big)\operatorname{vech}'\big(\hat{\Sigma}^{-1}\big)D'D. \end{aligned} \tag{23}$$

For the normal distribution case, where $\hat{G} = -\frac{1}{2}$, $\hat{F} = 0$ and $z_g = np$, we obtain from (22) and (23)

$$H_{\text{nor}} = -\begin{pmatrix} \hat{\Sigma}^{-1} \otimes X'X & 0 \\ 0 & \frac{n}{2}D'(\hat{\Sigma}^{-1} \otimes \hat{\Sigma}^{-1})D \end{pmatrix}. \tag{24}$$

For $H$ and $H_{\text{nor}}$ when $p = 1$ (and therefore $\Sigma$ and $D = 1$ are both scalars), we refer to [12] and [5], respectively.

## 5. Perturbation in case-weights

We consider three situations in order.

### 5.1. Full parameters

In this case, we define $W = \operatorname{diag}(w_1, \ldots, w_n)$ of case-weights of perturbation for model (3), where diag indicates a diagonal matrix and $w_i$ is the weight of the $i$th case ($i = 1, \ldots, n$). Under this perturbation scheme, we have $U \sim \mathrm{EM}_{np}(0, \Sigma, W^{-1}, \psi)$, with $W$ (replacing $W_0 = I_n$) in the perturbed model; the perturbed model reduces to the postulated model when $W = W_0$.

First, we consider $\Sigma$ as unknown, and both $B$ and $\Sigma$ are of interest. Due to the definition of $\Delta_\theta$ which involves $\theta$ and $w$ jointly, we need only the relevant part $L_w$ of the perturbed log-likelihood $L(\theta, w)$

$$L_w = L_w(\theta, w) = \ln g(z_w), \tag{25}$$

where $\theta = (b', s')'$, $b = \operatorname{vec} B$, $s = \operatorname{vech}\Sigma$, $z_w = \operatorname{tr} U'WU\Sigma^{-1}$, $w = (w_1, \ldots, w_n)'$ for $q = n$, $W = \operatorname{diag}(w_1, \ldots, w_n)$, $W_0 = I_n$, $U = Y - XB$ and $U \sim \mathrm{EM}_{np}(0, \Sigma, W^{-1}, \psi)$.

Taking the differentials of $L_w$, first with respect to $W$ and then $B$, we have

$$\mathrm{d}_w L_w = G(\mathrm{d}_w z_w) = G \operatorname{tr} U\Sigma^{-1}U'(\mathrm{d}W) \tag{26}$$

and

$$\begin{aligned}\mathrm{d}^2_{bw} L_w = & -2F \operatorname{tr}(\mathrm{d}B)'X'WU\Sigma^{-1} \operatorname{tr} U\Sigma^{-1}U'(\mathrm{d}W) \\ & - 2G \operatorname{tr}(\mathrm{d}B)'X'(\mathrm{d}W)U\Sigma^{-1}.\end{aligned} \tag{27}$$

As $W_0 = I_n$ and $X'\hat{U} = 0$, we have

$$\mathrm{d}^2_{bw} L_w|_{\theta=\hat{\theta}, w=w_0} = -2\hat{G}(\mathrm{d}\operatorname{vec}B)'\big(\hat{\Sigma}^{-1}\hat{U}' \otimes X'\big)J(\mathrm{d}w), \tag{28}$$

where $J$ is the $n^2 \times n$ selection matrix with $\mathrm{d}\operatorname{vec} W = J\,\mathrm{d}w$. For properties and applications of $J$, see e.g. [18,20].

Also, taking the differential of (26) with respect to $\Sigma$ we have

$$\begin{aligned}\mathrm{d}^2_{sw} L_w = & -F \operatorname{tr}(\mathrm{d}\Sigma)\Sigma^{-1}U'WU\Sigma^{-1} \operatorname{tr} U\Sigma^{-1}U'(\mathrm{d}W) \\ & - G \operatorname{tr}(\mathrm{d}\Sigma)\Sigma^{-1}U'(\mathrm{d}W)U\Sigma^{-1},\end{aligned} \tag{29}$$

and therefore

$$\begin{aligned}& \mathrm{d}^2_{sw} L_w\big|_{\theta=\hat{\theta}, w=w_0} \\ & = -\hat{F}\ (\mathrm{d}\operatorname{vech}\Sigma)'D'D\operatorname{vech}\big(\hat{\Sigma}^{-1}\hat{U}'\hat{U}\hat{\Sigma}^{-1}\big)\operatorname{vech}'\big(\hat{U}\hat{\Sigma}^{-1}\hat{U}'\big)D'J(\mathrm{d}w) \\ & \quad - \hat{G}(\mathrm{d}\operatorname{vech}\Sigma)'D'\big(\hat{\Sigma}^{-1}\hat{U}' \otimes \hat{\Sigma}^{-1}\hat{U}'\big)J(\mathrm{d}w) \\ & = -\frac{z_g\hat{F}}{p}\ (\mathrm{d}\operatorname{vech}\Sigma)'D'D\operatorname{vech}(\hat{\Sigma}^{-1})\operatorname{vech}'(\hat{U}\hat{\Sigma}^{-1}\hat{U}')D'J(\mathrm{d}w) \\ & \quad - \hat{G}\ (\mathrm{d}\operatorname{vech}\Sigma)'D'\big(\hat{\Sigma}^{-1}\hat{U}' \otimes \hat{\Sigma}^{-1}\hat{U}'\big)J(\mathrm{d}w),\end{aligned} \tag{30}$$

where $\hat{G} = G(\hat{z})$, $\hat{F} = F(\hat{z})$, $\hat{U}'\hat{U} = \hat{z}\hat{\Sigma}/p$ with $\hat{z} = z_g$, $D$ is the $p^2 \times (p+1)p/2$ duplication matrix as in (21) and $J$ is the $n^2 \times n$ selection matrix as in (28).

Then, (28) and (30) lead to

$$\Delta = \begin{pmatrix} -2\hat{G}\ \big(\hat{\Sigma}^{-1}\hat{U}' \otimes X'\big)J \\ \Delta_s \end{pmatrix}, \tag{31}$$

where

$$\begin{aligned}\Delta_s = & -\frac{z_g\hat{F}}{p}\ D'D\operatorname{vech}\big(\hat{\Sigma}^{-1}\big)\operatorname{vech}'\big(\hat{U}\hat{\Sigma}^{-1}\hat{U}'\big)D'J \\ & - \hat{G}\ D'\big(\hat{\Sigma}^{-1}\hat{U}' \otimes \hat{\Sigma}^{-1}\hat{U}'\big)J.\end{aligned} \tag{32}$$

For the normal distribution case ($\hat{G} = -\frac{1}{2}$ and $\hat{F} = 0$), we have

$$\Delta_{\text{nor}} = \begin{pmatrix} (\hat{\Sigma}^{-1}\hat{U}' \otimes X')J \\ \frac{1}{2}\, D'(\hat{\Sigma}^{-1}\hat{U}' \otimes \hat{\Sigma}^{-1}\hat{U}')J \end{pmatrix}. \tag{33}$$

For $\Delta$ and $\Delta_{\text{nor}}$ when $p = 1$, see [12] and [5], respectively.

Hence, we can calculate $M$ and $C_l(\theta)$ as defined in (10). Using (22) and (31), we obtain

$$\begin{aligned} M &= \Delta' H^{-1} \Delta \\ &= 2\hat{G} J'\big(\hat{U}\hat{\Sigma}^{-1} \otimes X\big)\big(\hat{\Sigma}^{-1} \otimes X'X\big)^{-1}\big(\hat{\Sigma}^{-1}\hat{U}' \otimes X'\big)J + \Delta_s' H_s^{-1} \Delta_s \\ &= 2\hat{G}\hat{U}\hat{\Sigma}^{-1}\hat{U}' \odot X(X'X)^{-1}X' + \Delta_s' H_s^{-1}\Delta_s, \end{aligned} \tag{34}$$

where $H_s$ is the same as in (23), $\Delta_s$ is the same as in (32), $\odot$ indicates the Hadamard product, which links with the Kronecker product via $J$, see e.g. [18].

Second, we consider $\Sigma$ as known and only $B$ is of interest. Clearly, (34) leads to

$$C_l(\theta) = 4\big|l'(\hat{G}\hat{U}\Sigma^{-1}\hat{U}' \odot X(X'X)^{-1}X')l\big|. \tag{35}$$

For the normal distribution case with $\Sigma$ unknown and both $\Sigma$ and $B$ of interest, using (24) and (33) we establish that

$$\begin{aligned} M_{\text{nor}} &= \Delta_{\text{nor}}' H_{\text{nor}}^{-1} \Delta_{\text{nor}} \\ &= -J'\big(\hat{U}\hat{\Sigma}^{-1} \otimes X\big)\big(\hat{\Sigma}^{-1} \otimes X'X\big)^{-1}\big(\hat{\Sigma}^{-1}\hat{U}' \otimes X'\big)J \\ &\quad - \frac{1}{2n} J'\big(\hat{U}\hat{\Sigma}^{-1} \otimes \hat{U}\hat{\Sigma}^{-1}\big)D\big[D'(\hat{\Sigma}^{-1} \otimes \hat{\Sigma}^{-1})D\big]^{-1} \\ &\quad \times D'\big(\hat{\Sigma}^{-1}\hat{U}' \otimes \hat{\Sigma}^{-1}\hat{U}'\big)J \\ &= -\hat{U}\hat{\Sigma}^{-1}\hat{U}' \odot X(X'X)^{-1}X' - \frac{1}{2n}\hat{U}\hat{\Sigma}^{-1}\hat{U}' \odot \hat{U}\hat{\Sigma}^{-1}\hat{U}'. \end{aligned} \tag{36}$$

We now show that our $M_{\text{nor}}$ is the same as (the two different but identical expressions of) $M_{\text{nor}}$ obtained by Kim [15] and Fung and Tang [11], respectively, and therefore our $C_l(\theta)$ is the same as theirs. Clearly, the first part of the right-hand side of (36) is the same as the first part of the expression of $M_{\text{nor}}$ of [15], and therefore is identical to that of [11]. The second part of the right-hand side of (36) is derived as

$$\begin{aligned} &-\frac{1}{2n} J'(\hat{U}\hat{\Sigma}^{-1} \otimes \hat{U}\hat{\Sigma}^{-1})D[D'(\hat{\Sigma}^{-1} \otimes \hat{\Sigma}^{-1})D]^{-1}D'(\hat{\Sigma}^{-1}\hat{U}' \otimes \hat{\Sigma}^{-1}\hat{U}')J \\ &\quad = -\frac{1}{2n} J'(\hat{U}\hat{\Sigma}^{-1} \otimes \hat{U}\hat{\Sigma}^{-1})DD^{+}(\hat{\Sigma} \otimes \hat{\Sigma})D'^{+}D'(\hat{\Sigma}^{-1}\hat{U}' \otimes \hat{\Sigma}^{-1}\hat{U}')J \\ &\quad = -\frac{1}{2n} J'(\hat{U}\hat{\Sigma}^{-1} \otimes \hat{U}\hat{\Sigma}^{-1})(\hat{\Sigma} \otimes \hat{\Sigma})(\hat{\Sigma}^{-1}\hat{U}' \otimes \hat{\Sigma}^{-1}\hat{U}')J \\ &\quad = -\frac{1}{2n} J'(\hat{U} \otimes \hat{U})(\hat{\Sigma}^{-1} \otimes \hat{\Sigma}^{-1})(\hat{U}' \otimes \hat{U}')J & (37) \\ &\quad = -\frac{1}{2n}\hat{U}\hat{\Sigma}^{-1}\hat{U}' \odot \hat{U}\hat{\Sigma}^{-1}\hat{U}', & (38) \end{aligned}$$

where we use the following properties of $D$ and $J$, see e.g. [21]:

$$\left[D'(\hat{\Sigma}^{-1} \otimes \hat{\Sigma}^{-1})D\right]^{-1} = D^{+}(\hat{\Sigma} \otimes \hat{\Sigma})D'^{+},$$
$$D'^{+}D' = DD^{+},$$
$$DD^{+}\left(\hat{\Sigma}^{-1}\hat{U}' \otimes \hat{\Sigma}^{-1}\hat{U}'\right)J = \left(\hat{\Sigma}^{-1}\hat{U}' \otimes \hat{\Sigma}^{-1}\hat{U}'\right)J.$$

By noting that

$$\begin{aligned} &U = (u_1, \ldots, u_n)', \\ &(\hat{U}' \otimes \hat{U}')J = (\hat{u}_1 \otimes \hat{u}_1, \ldots, \hat{u}_n \otimes \hat{u}_n), \end{aligned} \tag{39}$$

we see that (37) is the same as the second part of $M_{\text{nor}}$ of [11]; expression (38) is simpler.

When $\Sigma$ is known and only $B$ is of interest, we have $\hat{G} = -\frac{1}{2}$ and

$$C_{\text{nor}}(\theta) = 2\left|l'(\hat{U}\Sigma^{-1}\hat{U}' \odot X(X'X)^{-1}X')l\right|.$$

### 5.2. Subset of regression parameters

In some situations, only a subset, say $B_1$, of the rows of the parameter matrix $B = (B_1', B_2')'$ is of interest, and the columns of $X = (X_1, X_2)$ are rearranged correspondingly to the partition of $B$, where $B_1$ is of order $m_1 \times p$, $B_2$ is of $m_2 \times p$, $X_1$ is of $n \times m_1$ and $X_2$ is of $n \times m_2$ $(m_1 + m_2 = m)$. We denote $\theta = (\theta_1', \theta_2')' = (\operatorname{vec}' B_1', \operatorname{vec}' B_2', s')' = (\operatorname{vec}' B', s')'$, where $\theta_1 = \operatorname{vec} B_1'$ and $\theta_2 = (\operatorname{vec}' B_2', s')'$ with $s = \operatorname{vech} \Sigma$.

To use (12) we rewrite model (3) as

$$Y = XB + U = X_1B_1 + X_2B_2 + U, \tag{40}$$

and then

$$\begin{aligned} \operatorname{vec} Y' &= (X \otimes I_p)\operatorname{vec} B' + \operatorname{vec} U' \\ &= (X_1 \otimes I_p)\operatorname{vec} B_1' + (X_2 \otimes I_p)\operatorname{vec} B_2' + \operatorname{vec} U'. \end{aligned} \tag{41}$$

Note that

$$\operatorname{vec} B = K'\operatorname{vec} B', \tag{42}$$

where $K$ is the $mp \times mp$ commutation matrix, see e.g. [21].

Based on (9) and (42), we obtain

$$\mathrm{d}_b^2 L\big|_{\theta=\hat{\theta}} = (\mathrm{d}\operatorname{vec} B)' H_b (\mathrm{d}\operatorname{vec} B) = (\mathrm{d}\operatorname{vec} B')' K H_b K' (\mathrm{d}\operatorname{vec} B'), \tag{43}$$
$$\mathrm{d}_{bw}^2 L_w\big|_{\theta=\hat{\theta}, w=w_0} = (\mathrm{d}\operatorname{vec} B)' \Delta_b dw = (\mathrm{d}\operatorname{vec} B')' K \Delta_b dw, \tag{44}$$

where $H_b$ and $\Delta_b$ are the corresponding (to $B$) parts of $H$ and $\Delta$, respectively.

We see from (43) and (44) that $H_{b\text{new}} = KH_bK'$ and $\Delta_{b\text{new}} = K\Delta_b$. Hence, by virtue of $H$ in (22), we obtain

$$H_{\text{new}} = \begin{pmatrix} 2\hat{G}\left(X'X \otimes \hat{\Sigma}^{-1}\right) & 0 \\ 0 & H_s \end{pmatrix} \tag{45}$$

and

$$H_{22\text{new}} = \begin{pmatrix} 2\hat{G}\left(X_2'X_2 \otimes \hat{\Sigma}^{-1}\right) & 0 \\ 0 & H_s \end{pmatrix}, \tag{46}$$

where $H_s$ is the same as (23).

By $\varDelta$ in (31), we obtain

$$\varDelta_{\text{new}} = \begin{pmatrix} -2\hat{G}\, K\left(\hat{\Sigma}^{-1}\hat{U}' \otimes X'\right)J \\ \varDelta_s \end{pmatrix}, \tag{47}$$

where $\varDelta_s$ is the same as in (32).

Especially, for the normal distribution case we have

$$H_{\text{nornew}} = -\begin{pmatrix} X'X \otimes \hat{\Sigma}^{-1} & 0 \\ 0 & \frac{n}{2} D'\left(\hat{\Sigma}^{-1} \otimes \hat{\Sigma}^{-1}\right)D \end{pmatrix}, \tag{48}$$

$$H_{\text{nor22new}} = -\begin{pmatrix} X_2'X_2 \otimes \hat{\Sigma}^{-1} & 0 \\ 0 & \frac{n}{2} D'\left(\hat{\Sigma}^{-1} \otimes \hat{\Sigma}^{-1}\right)D \end{pmatrix}, \tag{49}$$

and

$$\varDelta_{\text{nornew}} = \begin{pmatrix} K\left(\hat{\Sigma}^{-1}\hat{U}' \otimes X'\right)J \\ \frac{1}{2}\, D'\left(\hat{\Sigma}^{-1}\hat{U}' \otimes \hat{\Sigma}^{-1}\hat{U}'\right)J \end{pmatrix}. \tag{50}$$

Based on (12), (45), (46) and (47), we obtain

$$\begin{aligned} M_1 &= \varDelta_{\text{new}}'\left(H_{\text{new}}^{-1} - A_{22}\right)\varDelta_{\text{new}} \\ &= 2\hat{G}J'\left(\hat{U}\hat{\Sigma}^{-1} \otimes X\right)K'(E \otimes \hat{\Sigma})K\left(\hat{\Sigma}^{-1}\hat{U}' \otimes X'\right)J \\ &= 2\hat{G}\hat{U}\hat{\Sigma}^{-1}\hat{U}' \odot XEX', \end{aligned} \tag{51}$$

where $E$ is an $m \times m$ positive semidefinite matrix defined as

$$E = (X'X)^{-1} - \begin{pmatrix} 0 & 0 \\ 0 & (X_2'X_2)^{-1} \end{pmatrix} \geqslant 0. \tag{52}$$

It thus follows that

$$C_l(\theta_1) = 4\left|l'(\hat{G}\hat{U}\hat{\Sigma}^{-1}\hat{U}' \odot XEX')l\right|. \tag{53}$$

Especially, for the normal distribution ($\hat{G} = -\frac{1}{2}$)

$$C_{\text{nor}}(\theta_1) = 2\left|l'(\hat{U}\hat{\Sigma}^{-1}\hat{U}' \odot XEX')l\right|.$$

### 5.3. Individual cases

Comparing with the case deletion, we consider only the situation where the weight for the $i$th case is perturbed, $i = 1, \ldots, n$. Therefore we write $W = \mathrm{diag}(1, \ldots, w_i, \ldots 1)'$, where $w_i$ is located as the $i$th diagonal element. Suppose $\Sigma$ is known and only $B$ is of interest, we have

$$H = 2\hat{G}\,\Sigma^{-1} \otimes X'X, \tag{54}$$

$$\Delta = -2\hat{G}\,(\Sigma^{-1}\hat{U}' \otimes X')R_i, \tag{55}$$

where $R_i$ is an $n^2 \times 1$ vector with one in the $(n(i-1)+i)$th position and zeros elsewhere.

The curvature is simplified as

$$\begin{aligned} C_l(\mathrm{vec}\,B) &= 4\left|l'(\hat{G}R_i'(\hat{U}\Sigma^{-1}\hat{U}' \otimes X(X'X)^{-1}X')R_i)l\right| \\ &= 4|\hat{G}|\hat{u}_i'\Sigma^{-1}\hat{u}_i P_{ii}, \end{aligned} \tag{56}$$

where $u_i$ is the $i$th column of $U'$ and $P_{ii}$ is the $i$th diagonal element of $X(X'X)^{-1}X'$.

For the normal distribution case, (56) becomes

$$C_{\mathrm{nor}} = 2\hat{u}_i'\Sigma^{-1}\hat{u}_i P_{ii}, \tag{57}$$

which is the same as (13) of [11], and has a connection with Cook's distance in the simpler mulptiple regression case; see [5].

## 6. Perturbation in explanatory variables

If perturbation in the explanatory variables is of interest, the perturbed log-likelihood is constructed with $X$ replaced by $X_w = X + WS$, where $W = (w_{ij}) = (w_1, \ldots, w_j, \ldots, w_m)$ is an $n \times m$ matrix of perturbations, $S = \mathrm{diag}(s_1, \ldots, s_m)$, $s_j$ $(j = 1, \ldots, m)$ is the scale factor and $W_0 = 0$. We obtain the relevant part of the perturbed log-likelihood

$$L_w = \ln\, g(z_w), \tag{58}$$

where $\theta = (b', s')'$, $b = \mathrm{vec}\,B$, $s = \mathrm{vech}\,\Sigma$, $z_w = \mathrm{tr}\,U_w'U_w\Sigma^{-1}$ and $U_w = Y - (X + WS)B$.

By taking the differential of $L_w$ with respect to $W$, we get

$$\mathrm{d}_w L_w = -2G\,\mathrm{tr}\,\Sigma^{-1}U_w'(\mathrm{d}W)SB. \tag{59}$$

Taking the differentials of $\mathrm{d}_w L_w$ with respect to $B$ and $\Sigma$, we get $\mathrm{d}^2_{bw}L_w$ and $\mathrm{d}^2_{sw}L_w$, respectively. We evaluate them as follows:

$$\begin{aligned} \mathrm{d}^2_{bw}L_w\big|_{\theta=\hat{\theta},w=w_0} &= 2\hat{G}\,\mathrm{tr}(\mathrm{d}B)'X'(\mathrm{d}W)S\hat{B}\hat{\Sigma}^{-1} \\ &\quad - 2\hat{G}\,\mathrm{tr}(\mathrm{d}B)\hat{\Sigma}^{-1}\hat{U}'(\mathrm{d}W)S, \end{aligned} \tag{60}$$

$$\mathrm{d}^2_{sw} L_w \big|_{\theta=\hat{\theta}, w=w_0} = 2\hat{F} \operatorname{tr} \hat{\Sigma}^{-1} \hat{U}' \hat{U} \hat{\Sigma}^{-1} (\mathrm{d}\Sigma) \operatorname{tr} \hat{\Sigma}^{-1} \hat{U}' (\mathrm{d}W) S B + 2\hat{G} \operatorname{tr}(\mathrm{d}\Sigma) \hat{\Sigma}^{-1} \hat{U}' (\mathrm{d}W) S B \hat{\Sigma}^{-1}. \tag{61}$$

Hence, we obtain

$$\Delta = \begin{pmatrix} \Delta_b \\ \Delta_s \end{pmatrix}, \tag{62}$$

where

$$\Delta_b = 2\hat{G}\big(\hat{\Sigma}^{-1} \hat{B}' S \otimes X'\big) - 2\hat{G} K' (S \otimes \hat{\Sigma}^{-1} \hat{U}'), \tag{63}$$

$$\Delta_s = \frac{2 z_g \hat{F}}{p} D' D \operatorname{vech}(\hat{\Sigma}^{-1}) \operatorname{vech}'\big(\hat{U} \hat{\Sigma}^{-1} \hat{B}' S\big) + 2\hat{G} D' \big(\hat{\Sigma}^{-1} \hat{B}' S \otimes \hat{\Sigma}^{-1} \hat{U}'\big), \tag{64}$$

$z_g$ is the same as in (5) and $S$ is the same as defined above.

Using $H$ in (22) and $\Delta$ in (62), we obtain $M$ in (10)

$$\begin{aligned} M = {} & 2\hat{G} S \hat{B} \hat{\Sigma}^{-1} \hat{B}' S \otimes X (X'X)^{-1} X' \\ & - 2\hat{G}\big(S \hat{B} \hat{\Sigma}^{-1} \hat{U}' \otimes X (X'X)^{-1} S\big) K' \\ & - 2\hat{G} K \big(\hat{U} \hat{\Sigma}^{-1} \hat{B}' S \otimes S (X'X)^{-1} X'\big) \\ & + 2\hat{G} S (X'X)^{-1} S \otimes \hat{U} \hat{\Sigma}^{-1} \hat{U}' + \Delta_s' H_s^{-1} \Delta_s, \end{aligned} \tag{65}$$

where $H_s$ is the same as in (23) and $\Delta_s$ is the same as in (64).

If $\Sigma$ is known and only $B$ is of interest, (65) leads to

$$\begin{aligned} M_B = {} & 2\hat{G} S \hat{B} \Sigma^{-1} \hat{B}' S \otimes X (X'X)^{-1} X' \\ & - 2\hat{G}\big(S \hat{B} \Sigma^{-1} \hat{U}' \otimes X (X'X)^{-1} S\big) K' \\ & - 2\hat{G} K \big(\hat{U} \Sigma^{-1} \hat{B}' S \otimes S (X'X)^{-1} X'\big) \\ & + 2\hat{G} S (X'X)^{-1} S \otimes \hat{U} \Sigma^{-1} \hat{U}' \end{aligned} \tag{66}$$

and

$$C_l(\operatorname{vec} B) = 2\big|l' M_B l\big|. \tag{67}$$

For the normal distribution case with both $B$ and $\Sigma$ of interest, from (24), (63) and (64) we get

$$\begin{aligned} M_{\text{nor}} = {} & -S \hat{B} \hat{\Sigma}^{-1} \hat{B}' S \otimes X (X'X)^{-1} X' + \big(S \hat{B} \hat{\Sigma}^{-1} \hat{U}' \otimes X (X'X)^{-1} S\big) K' \\ & + K \big(\hat{U} \hat{\Sigma}^{-1} \hat{B}' S \otimes S (X'X)^{-1} X'\big) - S (X'X)^{-1} S \otimes \hat{U} \hat{\Sigma}^{-1} \hat{U}' \\ & - \frac{2}{n} \big(S \hat{B} \hat{\Sigma}^{-1} \otimes \hat{U} \hat{\Sigma}^{-1}\big) D D^{+} (\hat{B}' S \otimes \hat{U}'), \end{aligned} \tag{68}$$

and therefore

$$C_l(\theta) = 2\left|l' M_{\text{nor}} l\right|, \tag{69}$$

which corresponds to (18) of [11].

## 7. Perturbation in response variables

For perturbation in the response variables, the perturbed log-likelihood is constructed with $Y$ replaced by $Y_w = Y + WS$, where $W = (w_{ij}) = (w_1, \ldots, w_j, \ldots, w_p)$ is an $n \times p$ matrix of perturbations, $S = \operatorname{diag}(s_1, \ldots, s_p),\ s_j\ (j = 1, \ldots, p)$ is the scale factor and $W_0 = 0$. We have the relevant part of the perturbed log-likelihood

$$L_w = \ln g(z_w), \tag{70}$$

where $\theta = (b', s')'$, $b = \operatorname{vec} B$, $s = \operatorname{vech} \Sigma$, $z_w = \operatorname{tr} U_w' U_w \Sigma^{-1}$ and $U_w = Y + WS - XB$.

We obtain

$$\mathrm{d}^2_{bw} L_w\big|_{\theta=\hat{\theta}, w=w_0} = -2\hat{G} \operatorname{tr}(\mathrm{d}B)' X' (\mathrm{d}W) S \hat{\Sigma}^{-1} \tag{71}$$

and

$$\begin{aligned} \mathrm{d}^2_{sw} L_w\big|_{\theta=\hat{\theta}, w=w_0} = & -2\hat{F} \operatorname{tr}(\mathrm{d}\Sigma)\hat{\Sigma}^{-1}\hat{U}'\hat{U}\hat{\Sigma}^{-1} \operatorname{tr}\hat{\Sigma}^{-1}\hat{U}'(\mathrm{d}W)S \\ & - 2\hat{G} \operatorname{tr}(\mathrm{d}\Sigma)\hat{\Sigma}^{-1}\hat{U}'(\mathrm{d}W)S\hat{\Sigma}^{-1}. \end{aligned} \tag{72}$$

Then

$$\Delta = \begin{pmatrix} \Delta_b \\ \Delta_s \end{pmatrix}, \tag{73}$$

where

$$\Delta_b = -2\hat{G}\big(\hat{\Sigma}^{-1} S \otimes X'\big), \tag{74}$$

$$\begin{aligned} \Delta_s = & -\frac{2 z_g \hat{F}}{p} D' D \operatorname{vech}(\hat{\Sigma}^{-1}) \operatorname{vec}'\big(\hat{U}\hat{\Sigma}^{-1} S\big) \\ & - 2\hat{G} D'\big(\hat{\Sigma}^{-1} S \otimes \hat{\Sigma}^{-1}\hat{U}'\big), \end{aligned} \tag{75}$$

$z_g$ is the same as in (5) and $S$ is the same as defined above.

Hence, we use $H$ in (22) and $\Delta$ in (73) to derive $M$ in (10) as

$$M = 2\hat{G} S \hat{\Sigma}^{-1} S \otimes X(X'X)^{-1} X' + \Delta_s' H_s^{-1} \Delta_s, \tag{76}$$

where $H_s$ and $\Delta_s$ are the same as in (23) and (75), respectively.

For the normal distribution case, from (24), (74) and (75) we get

$$\begin{aligned} M_{\text{nor}} = & -S\hat{\Sigma}^{-1} S \otimes X(X'X)^{-1} X' \\ & - \frac{2}{n}\big(S\hat{\Sigma}^{-1} \otimes \hat{U}\hat{\Sigma}^{-1}\big) D D^{+} (S \otimes \hat{U}'), \end{aligned} \tag{77}$$

and therefore

$$C_l(\theta) = 2\left|l' M_{\text{nor}} l\right|, \tag{78}$$

which is (20) of [11].

## 8. Remarks

Instead of presenting an index plot to illustrate the methodology described in this paper, we refer to two sets of plots. The first is that given by Galea et al. [12] and Liu [19] for the data set reported by Ruppert and Carrol [23] on the salinity of water during the spring in Pamlico Sound, North Carolina in the univariate elliptical linear regression models. The second is given by Kim [15] and Fung and Tang [11] for the data set collected by W.D. Rohwer on children's performances in the multivariate linear regression models under the normality assumption. We note that the local influence method suggests accord with those provided by the case deletion and other methods (in the univariate case), see e.g. [12]. However, different observations need special attention under different perturbation schemes of the local influence analysis (in the multivariate case), as explained by Fung and Tang [11].

## Acknowledgement

The author would like to thank the associate editor Hans Joachim Werner and the referees for their helpful comments on an early version of the paper. He is also very grateful to his colleagues Joe Gani, Peter Hall and George Styan for their encouragement. The paper was presented at 7th International Workshop on Matrices and Statistics, in celebration of T.W. Anderson's 80th birthday, Fort Lauderdale, Florida (by title), and at 45th Annual Meeting of the Australian Mathematical Society, Canberra.

## References

[1] T.W. Anderson, An Introduction to Multivariate Statistical Analysis, second ed., Wiley, New York, 1984.
[2] T.W. Anderson, Nonnormal multivariate distribution: inference based on elliptically contoured distributions, in: C.R. Rao (Ed.), Multivariate Analysis: Future Directions, Elsevier, Amsterdam, 1993.
[3] T.W. Anderson, K.T. Fang, On the theory of multivariate elliptically contoured distributions and their applications, in: K.T. Fang, T.W. Anderson (Eds.), Statistical Inference in Elliptically Contoured and Related Distributions, Allerton Press, New York, 1990, pp. 1–23.
[4] S. Chatterjee, A.S. Hadi, Sensitivity Analysis in Linear Regression, Wiley, New York, 1988.
[5] R.D. Cook, Assessment of local influence (with discussion), J. Roy. Statist. Soc. Ser. B 48 (1986) 133–169.
[6] R.D. Cook, Local influence, in: S. Kotz, C.B. Read, D.L. Banks (Eds.), Encyclopedia of Statistical Sciences Update, vol. 1, Wiley, New York, 1997, pp. 380–385.

[7] K.T. Fang, T.W. Anderson, Statistical Inferences in Elliptically Contoured and Related Distributions, Allerton Press, New York, 1990.

[8] K.T. Fang, Y. Zhang, Generalized Multivariate Analysis, Science Press, Beijing and Springer, Berlin, 1990.

[9] R.W. Farebrother, Relative local influence and the condition number, Comm. Statist. Simulation Comput. 21 (1992) 707–710.

[10] R.W. Farebrother, Fitting Linear Relationships: A History of the Calculus of Observations 1750–1900, Springer, New York, 1999.

[11] W.K. Fung, M.K. Tang, Assessment of local influence in multivariate regression analysis, Comm. Statist. Theory Methods 26 (4) (1997) 821–837.

[12] M. Galea, G.A. Paula, H. Bolfarine, Local influence in elliptical linear regression models, The Statistician 46 (1) (1997) 71–79.

[13] A.K. Gupta, T. Varga, Elliptically Contoured Models in Statistics, Kluwer Academic Publishers, Dordrecht, 1993.

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

[15] M.G. Kim, Local influence in multivariate regression, Comm. Statist. Theory Methods 24 (1995) 1271–1278.

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

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

[18] S. Liu, Contributions to Matrix Calculus and Applications in Econometrics, Tinbergen Institute Research Series no. 106, Thesis Publishers, Amsterdam, 1995.

[19] S. Liu, On local influence in elliptical linear regression models, Statistical Papers 41 (2000) 211–224.

[20] J.R. Magnus, Linear Structures, Griffin's Statistical Monographs, No. 42, Edward Arnold, London and Oxford University Press, New York, 1988.

[21] J.R. Magnus, H. Neudecker, Matrix Differential Calculus with Applications in Statistics and Econometrics, revised ed., Wiley, Chichester, 1999.

[22] J.X. Pan, K.T. Fang, E.P. Liski, Bayesian local influence for the growth curve model with Rao's simple covariance structure, J. Multivar. Anal. 58 (1) (1996) 55–81.

[23] D. Ruppert, R.J. Carroll, Trimmed least squares estimation in the linear model, J. Amer. Statist. Assoc. 75 (1980) 828–838.

[24] B. Schwarzmann, A connection between local influence analysis and residual diagnostics, Technometrics 33 (1991) 103–104.

[25] L. Shi, X.R. Wang, Local influence in ridge regression, Comput. Statist. Data Anal. 31 (3) (1999) 341–353.

[26] B.C. Wei, Y.Q. Hu, W.K. Fung, Generalized leverage and its applications, Scand. J. Statist. 25 (1) (1998) 25–37.