Journal of the Korean Statistical Society (2002), **31**: 4, pp 447–458

Local Influence Analysis of the Equicorrelation Model†

Myung Geun Kim[1] and Kang-Mo Jung[2]

ABSTRACT

The influence of observations in the equicorrelation model is investigated using the local influence approach when all parameters or subsets of parameters are of interest. When a parameter of interest is scalar, an analytical form of the local influence measure can be found. We will derive a measure for identifying observations that have a large influence on the test of fitting the equicorrelation model. An example is given for illustration.

Keywords. Equicorrelation model, local influence, test.
AMS 2000 subject classifications. Primary 62J20; Secondary 62H20.

1. Introduction

The local influence method was introduced by Cook (1986) as a general method of assessing the influence of minor perturbations of the model and it can be used for identifying observations that influence the assumptions underlying the model. The local influence method has been adapted to a variety of other models, for example Box-Cox transformation model (Lawrance, 1988), multivariate regression model (Kim, 1995) and maximum likelihood factor analysis model (Jung *et al.*, 1997). It has been an efficient means for obtaining information about the influence of observations. Besides local influence other methods in influence analyses can be found in Barnett and Lewis (1994).

The equicorrelation model, also called the intraclass model has a pattern of equal variances and equal covariances in the covariance matrix. The variables are correlated and every pair of variables has the same correlation coefficient under

Received March 2001; accepted July 2002.
†This work was in part supported by grants from Seowon University.
[1]Department of Applied Statistics, Seowon University, Cheongju, Chungcheongbuk-do 361-742, Korea
[2]Department of Informatics & Statistics, Kunsan National University, Kunsan, Jeollabuk-do 573-701, Korea

the equicorrelation model. The equicorrelation model is used for checking the validity of standard ANOVA approach to repeated measures design (Rencher, 1995). Also, the common correlation coefficient is used for measuring the agreement between quantitative measures in epidemiological studies and Giraudeau *et al.* (1996) studied the single case deletion effect on the common correlation coefficient. There are few or no works on diagnostics in the equicorrelation model.

In this work we will study the influence in the equicorrelation model using the local influence method. In Section 2 the likelihood equations for estimating the model parameters are reviewed. In Section 3 we will review the local influence method and then derive local influence measures for the equicorrelation model when all parameters or subsets of parameters are of interest. In particular, when a parameter of interest is a scalar, an analytical form of the local influence measure can be found. In Section 4 we will derive a measure for identifying observations that have a large influence on the likelihood ratio test of fitting the equicorrelation model, following the results of Lawrance (1988). In Section 5 an example is given for illustration.

2. Preliminaries

Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ be a random sample from a p-variate normal distribution

$$f(\mathbf{x}) = | 2\pi\mathbf{\Sigma} |^{-1/2} \exp\{-(1/2)(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\},$$

where the covariance matrix $\mathbf{\Sigma}$ has the following equicorrelation form

$$\mathbf{\Sigma} = \sigma^2\{(1-\rho)\mathbf{I}_p + \rho\mathbf{1}_p\mathbf{1}_p^T\},$$

where $\sigma > 0$ and $0 < \rho < 1$. We write $\mathbf{1}_p$ as the $p \times 1$ vector with all elements equal to 1. Let $\bar{\mathbf{x}} = (1/n)\sum_{i=1}^n \mathbf{x}_i$ and $\mathbf{S} = (1/n)\sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T$. In what follows, we use the hat notation to denote its maximum likelihood estimator of a parameter. Then the maximum likelihood estimators of $\boldsymbol{\mu}$, σ^2 and ρ are given by

$$\begin{aligned} \hat{\boldsymbol{\mu}} &= \bar{\mathbf{x}}, \\ \hat{\sigma}^2 &= \frac{1}{p}\mathrm{tr}(\mathbf{S}), \\ \hat{\rho} &= \frac{2}{p(p-1)\hat{\sigma}^2}\sum_{i<j} s_{ij}, \end{aligned}$$

where s_{ij} is the $(i,j)^{th}$ element of $\mathbf{S}$. This result will be used in Section 4. More details can be found in Rencher (1995).

3. Local Influence

In this section we will review the local influence introduced by Cook (1986) and then derive local influence measures for the equicorrelation model when all parameters or subsets of parameters are of interest.

Let $\mathbf{w} = (w_1, \ldots, w_n)^T$ be an $n \times 1$ vector of perturbations. We consider the perturbed model in which the u^{th} observation $\mathbf{x}_u$ is perturbed according to

$$\mathbf{x}_u \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}/w_u) \tag{3.1}$$

for $u = 1, \ldots, n$. When the w_u are set equal to one, the perturbed model reduces to the unperturbed model.

Let $\boldsymbol{\theta}$ be the $(p+2) \times 1$ vector of parameters formed by stacking σ, ρ and $\boldsymbol{\mu}$. We denote the log-likelihoods for the unperturbed and perturbed models by $L(\boldsymbol{\theta})$ and $L(\boldsymbol{\theta}|\mathbf{w})$, respectively. The likelihood displacement $LD(\mathbf{w})$ is defined by $2[L(\hat{\boldsymbol{\theta}}) - L(\hat{\boldsymbol{\theta}}_w)]$, where $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\theta}}_w$ are the maximum likelihood estimators of $\boldsymbol{\theta}$ under the unperturbed and perturbed models, respectively. The surface of interest is formed by the $(n+1) \times 1$ vector of the values $\mathbf{w}$ and $LD(\mathbf{w})$ as $\mathbf{w}$ varies over a certain space. Define the $(p+2) \times n$ matrix

$$\boldsymbol{\Delta} = \frac{\partial^2 L(\boldsymbol{\theta} \mid \mathbf{w})}{\partial \boldsymbol{\theta} \partial \mathbf{w}^T}$$

evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ and $\mathbf{w} = \mathbf{1}_n$, and the $(p+2) \times (p+2)$ matrix

$$\ddot{\mathbf{L}} = \frac{\partial^2 L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}$$

evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$. Let $\ddot{\mathbf{F}}$ be the $n \times n$ matrix defined by

$$\ddot{\mathbf{F}} = \boldsymbol{\Delta}^T (\ddot{\mathbf{L}})^{-1} \boldsymbol{\Delta}.$$

Let $\boldsymbol{l}_{max}$ be the eigenvector corresponding to the largest absolute eigenvalue of $-2\ddot{\mathbf{F}}$ and let $\mathbf{1}_{(i)}$ be the $i \times 1$ vector with its i^{th} element equal to 1 and the others being zero. Then the largest absolute eigenvalue is the maximum curvature of the curve which is the portion of the surface cut out by the plane spanned by the vectors $\mathbf{1}_{(n+1)}$ and $(\boldsymbol{l}_{max}^T, 0)^T$ (Cook, 1986, pp. 138–139). Observations corresponding to the large elements of the first direction vector $\boldsymbol{l}_{max}$ are influential.

3.1. All parameters are of interest

The determinant and inverse of $\mathbf{\Sigma}$ are given by

$$\begin{aligned} |\,\mathbf{\Sigma}\,| &= \sigma^{2p}(1-\rho)^{p-1}\{1+(p-1)\rho\}, \\ \mathbf{\Sigma}^{-1} &= \frac{1}{(1-\rho)\sigma^2}\left\{\mathbf{I}_p - \frac{\rho}{1+(p-1)\rho}\mathbf{1}_p\mathbf{1}_p^T\right\}, \end{aligned}$$

respectively. Ignoring unimportant terms, the log-likelihood function for the un perturbed model is

$$\begin{aligned} L(\boldsymbol{\mu},\sigma,\rho) &= -\frac{n}{2}\{2p\log(\sigma)+(p-1)\log(1-\rho)+\log(1+(p-1)\rho)\} \\ &\quad -\frac{1}{2(1-\rho)\sigma^2}\sum_{u=1}^{n}(\mathbf{x}_u-\boldsymbol{\mu})^T(\mathbf{x}_u-\boldsymbol{\mu}) \\ &\quad +\frac{\rho}{2(1-\rho)\{1+(p-1)\rho\}\sigma^2}\sum_{u=1}^{n}\{\mathbf{1}_p^T(\mathbf{x}_u-\boldsymbol{\mu})\}^2. \end{aligned}$$

Using the identity that

$$\begin{aligned} \sum_{u=1}^{n}\{\mathbf{1}_p^T(\mathbf{x}_u-\bar{\mathbf{x}})\}^2 &= \mathbf{1}_p^T\left\{\sum_{u=1}^{n}(\mathbf{x}_u-\bar{\mathbf{x}})(\mathbf{x}_u-\bar{\mathbf{x}})^T\right\}\mathbf{1}_p = n\mathbf{1}_p^T\mathbf{S}\mathbf{1}_p \\ &= np\hat{\sigma}^2\{1+(p-1)\hat{\rho}\}, \end{aligned}$$

we have the following partial derivatives evaluated at $\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}$ to get $\ddot{\mathbf{L}}$

$$\begin{aligned} \frac{\partial^2 L(\boldsymbol{\mu},\sigma,\rho)}{\partial\boldsymbol{\mu}\partial\boldsymbol{\mu}^T} &= -n\hat{\mathbf{\Sigma}}^{-1}, \\ \frac{\partial^2 L(\boldsymbol{\mu},\sigma,\rho)}{\partial\boldsymbol{\mu}\partial\sigma} &= \mathbf{0}, \\ \frac{\partial^2 L(\boldsymbol{\mu};\sigma,\rho)}{\partial\boldsymbol{\mu}\partial\rho} &= \mathbf{0}, \\ \frac{\partial^2 L(\boldsymbol{\mu},\sigma,\rho)}{\partial\sigma^2} &= \frac{np}{\hat{\sigma}^2}-\frac{3n}{(1-\hat{\rho})\hat{\sigma}^4}\mathrm{tr}(\mathbf{S}) \\ &\quad +\frac{3\hat{\rho}}{(1-\hat{\rho})\{1+(p-1)\hat{\rho}\}\hat{\sigma}^4}\sum_{u=1}^{n}\left\{\mathbf{1}_p^T(\mathbf{x}_u-\bar{\mathbf{x}})\right\}^2 \\ &= -\frac{2np}{\hat{\sigma}^2}, \\ \frac{\partial^2 L(\boldsymbol{\mu},\sigma,\rho)}{\partial\sigma\partial\rho} &= \frac{n}{(1-\hat{\rho})^2\hat{\sigma}^3}\mathrm{tr}(\mathbf{S}) \\ &\quad -\frac{\{1+(p-1)\hat{\rho}^2\}}{(1-\hat{\rho})^2\{1+(p-1)\hat{\rho}\}^2\hat{\sigma}^3}\sum_{u=1}^{n}\left\{\mathbf{1}_p^T(\mathbf{x}_u-\bar{\mathbf{x}})\right\}^2 \end{aligned}$$

$$
\begin{aligned}
&= \frac{np(p-1)\hat\rho}{(1-\hat\rho)\{1+(p-1)\hat\rho\}\hat\sigma}, \\
\frac{\partial^2 L(\boldsymbol{\mu},\sigma,\rho)}{\partial\rho^2} &= \frac{n(p-1)}{2(1-\hat\rho)^2} + \frac{n(p-1)^2}{2\{1+(p-1)\hat\rho\}^2} - \frac{n}{(1-\hat\rho)^3\hat\sigma^2}\mathrm{tr}(\mathbf{S}) \\
&\quad + \frac{(p-1)^2\hat\rho^3 + 3(p-1)\hat\rho - p + 2}{(1-\hat\rho)^3\{1+(p-1)\hat\rho\}^3\hat\sigma^2}\sum_{u=1}^{n}\left\{\mathbf{1}_p^T(\mathbf{x}_u-\bar{\mathbf{x}})\right\}^2 \\
&= -\frac{np(p-1)\{1+(p-1)\hat\rho^2\}}{2(1-\hat\rho)^2\{1+(p-1)\hat\rho\}^2}.
\end{aligned}
$$

Ignoring unimportant terms, the log-likelihood function for the perturbed model is

$$
L(\boldsymbol{\mu},\sigma,\rho\mid\mathbf{w}) = \frac{p}{2}\sum_{u=1}^{n}\log(w_u) - \frac{n}{2}\log\mid\boldsymbol{\Sigma}\mid - \frac{1}{2}\sum_{u=1}^{n} w_u(\mathbf{x}_u-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}_u-\boldsymbol{\mu}).
$$

The matrix $\boldsymbol{\Delta}$ in $\ddot{\mathbf{F}}$ is obtained by using the following partial derivatives evaluated at $\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}$ and $\mathbf{w}=\mathbf{1}_n$,

$$
\begin{aligned}
\frac{\partial^2 L(\boldsymbol{\mu},\sigma,\rho\mid\mathbf{w})}{\partial\boldsymbol{\mu}\partial w_u} &= \hat{\boldsymbol{\Sigma}}^{-1}(\mathbf{x}_u-\bar{\mathbf{x}}), \\
\frac{\partial^2 L(\boldsymbol{\mu},\sigma,\rho\mid\mathbf{w})}{\partial\sigma\partial w_u} &= \frac{1}{(1-\hat\rho)\hat\sigma^3}(\mathbf{x}_u-\bar{\mathbf{x}})^T(\mathbf{x}_u-\bar{\mathbf{x}}) \\
&\quad - \frac{\hat\rho}{(1-\hat\rho)\{1+(p-1)\hat\rho\}\hat\sigma^3}\left\{\mathbf{1}_p^T(\mathbf{x}_u-\bar{\mathbf{x}})\right\}^2 \\
&= \frac{1}{\hat\sigma}(\mathbf{x}_u-\bar{\mathbf{x}})^T\hat{\boldsymbol{\Sigma}}^{-1}(\mathbf{x}_u-\bar{\mathbf{x}}), \\
\frac{\partial^2 L(\boldsymbol{\mu},\sigma,\rho\mid\mathbf{w})}{\partial\rho\partial w_u} &= -\frac{1}{2(1-\hat\rho)^2\hat\sigma^2}(\mathbf{x}_u-\bar{\mathbf{x}})^T(\mathbf{x}_u-\bar{\mathbf{x}}) \\
&\quad + \frac{1+(p-1)\hat\rho^2}{2(1-\hat\rho)^2\{1+(p-1)\hat\rho\}^2\hat\sigma^2}\left\{\mathbf{1}_p^T(\mathbf{x}_u-\bar{\mathbf{x}})\right\}^2 \\
&= \frac{1}{2\hat\rho(1-\hat\rho)\{1+(p-1)\hat\rho\}\hat\sigma^2}(\mathbf{x}_u-\bar{\mathbf{x}})^T(\mathbf{x}_u-\bar{\mathbf{x}}) \\
&\quad - \frac{1+(p-1)\hat\rho^2}{2\hat\rho(1-\hat\rho)\{1+(p-1)\hat\rho\}}(\mathbf{x}_u-\bar{\mathbf{x}})^T\hat{\boldsymbol{\Sigma}}^{-1}(\mathbf{x}_u-\bar{\mathbf{x}}).
\end{aligned}
$$

Theorem 1. *Under the perturbation scheme (3.1), the $n\times n$ matrix $\ddot{\mathbf{F}}$ can be written as*

$$
-2\ddot{\mathbf{F}} = \frac{1}{n}\left\{2\mathbf{H} + \frac{1}{p}\mathbf{h}\mathbf{h}^T + \frac{1}{p(p-1)\hat\rho^2}\left(\frac{1}{\hat\sigma^2}\mathbf{e}-\mathbf{h}\right)\left(\frac{1}{\hat\sigma^2}\mathbf{e}-\mathbf{h}\right)^T\right\}, \tag{3.2}
$$

where $\mathbf{H}=\widetilde{\mathbf{X}}\hat{\boldsymbol{\Sigma}}^{-1}\widetilde{\mathbf{X}}^T$, $\mathbf{h}$ and $\mathbf{e}$ are $n\times 1$ vectors formed by the diagonal elements of $\mathbf{H}$ and $\mathbf{E}=\widetilde{\mathbf{X}}\widetilde{\mathbf{X}}^T$, respectively, and $\widetilde{\mathbf{X}}$ is the centered matrix of the $n\times p$ data matrix $\mathbf{X}$, that is, $\widetilde{\mathbf{X}}=\mathbf{X}-\mathbf{1}_n\bar{\mathbf{x}}^T$.

Proof. First, the matrix $\ddot{\mathbf{L}}$ can be partitioned as

$$\ddot{\mathbf{L}} = \begin{bmatrix} \mathbf{L}_{11} & \mathbf{O} \\ \mathbf{O} & -n\hat{\mathbf{\Sigma}}^{-1} \end{bmatrix},$$

where

$$\mathbf{L}_{11} = \begin{bmatrix} \frac{\partial^2 L(\boldsymbol{\mu},\sigma,\rho)}{\partial \sigma^2} & \frac{\partial^2 L(\boldsymbol{\mu},\sigma,\rho)}{\partial \rho \partial \sigma} \\ \frac{\partial^2 L(\boldsymbol{\mu},\sigma,\rho)}{\partial \sigma \partial \rho} & \frac{\partial^2 L(\boldsymbol{\mu},\sigma,\rho)}{\partial \rho^2} \end{bmatrix}.$$

A little algebra yields

$$\ddot{\mathbf{L}}^{-1} = \begin{bmatrix} \mathbf{L}_{11}^{-1} & \mathbf{O} \\ \mathbf{O} & -\frac{1}{n}\hat{\mathbf{\Sigma}} \end{bmatrix}, \tag{3.3}$$

where

$$\mathbf{L}_{11}^{-1} = -\frac{1}{np} \begin{bmatrix} \frac{1}{2}\{1+(p-1)\hat{\rho}^2\}\hat{\sigma}^2 & \hat{\rho}(1-\hat{\rho})\{1+(p-1)\hat{\rho}\}\hat{\sigma} \\ \hat{\rho}(1-\hat{\rho})\{1+(p-1)\hat{\rho}\}\hat{\sigma} & \frac{2}{p-1}(1-\hat{\rho})^2\{1+(p-1)\hat{\rho}\}^2 \end{bmatrix}.$$

Next, the matrix $\mathbf{\Delta}$ can be partitioned as $\mathbf{\Delta}^T = [\mathbf{\Delta}_1^T, \mathbf{\Delta}_2^T]$, where $\mathbf{\Delta}_2 = \hat{\mathbf{\Sigma}}^{-1}\widetilde{\mathbf{X}}^T$ and $\mathbf{\Delta}_1$ is the $2 \times n$ matrix whose u^{th} column vector is

$$\left[\frac{\partial^2 L(\boldsymbol{\mu},\sigma,\rho|\mathbf{w})}{\partial \sigma \partial w_u}, \frac{\partial^2 L(\boldsymbol{\mu},\sigma,\rho|\mathbf{w})}{\partial \rho \partial w_u}\right]^T.$$

Then we have

$$\ddot{\mathbf{F}} = \mathbf{\Delta}_1^T \mathbf{L}_{11}^{-1} \mathbf{\Delta}_1 - \frac{1}{n}\mathbf{\Delta}_2^T \hat{\mathbf{\Sigma}} \mathbf{\Delta}_2.$$

Therefore, the $(u,v)^{th}$ element of $\mathbf{\Delta}_1^T \mathbf{L}_{11}^{-1} \mathbf{\Delta}_1$ is

$$-\frac{1}{2np}h_{uu}h_{vv} - \frac{1}{2np(p-1)\hat{\rho}^2}\left(\frac{1}{\hat{\sigma}^2}e_{uu} - h_{uu}\right)\left(\frac{1}{\hat{\sigma}^2}e_{vv} - h_{vv}\right),$$

where h_{uv} and e_{uv} are the $(u,v)^{th}$ element of $\mathbf{H}$ and $\mathbf{E}$, respectively. This completes the proof. □

In Theorem 1 we can see that the $(u,v)^{th}$ element of $\mathbf{H}$ is just the squared Mahalanobis distance between observations $\mathbf{x}_u$ and $\mathbf{x}_v$ under the equicorrelation model. It seems that we cannot find a closed form of the eigenvector $\boldsymbol{l}_{max}$ of $-2\ddot{\mathbf{F}}$.

3.2. Subsets of parameters are of interest

Suppose that $\boldsymbol{\theta}$ is partitioned into $\boldsymbol{\theta}^T = (\boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T)$, where only the subset $\boldsymbol{\theta}_1$ is of interest. Cook (1986) obtained the normal curvature for $\boldsymbol{\theta}_1$ as the eigenvalue of $-2\ddot{\mathbf{G}}$ with

$$\ddot{\mathbf{G}} = \boldsymbol{\Delta}^T(\ddot{\mathbf{L}}^{-1} - \mathbf{B}_{22})\boldsymbol{\Delta}, \tag{3.4}$$

where $\mathbf{B}_{22} = \begin{bmatrix} \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{L}_{22}^{-1} \end{bmatrix}$ and $-\mathbf{L}_{22}$ is the information matrix of $\boldsymbol{\theta}_2$.

When only σ is of interest, we have in (3.4)

$$\mathbf{L}_{22}^{-1} = \begin{bmatrix} \frac{\partial^2 L}{\partial \rho^2} & \frac{\partial^2 L}{\partial \boldsymbol{\mu}^T \partial \rho} \\ \frac{\partial^2 L}{\partial \boldsymbol{\mu} \partial \rho} & \frac{\partial^2 L}{\partial \boldsymbol{\mu} \partial \boldsymbol{\mu}^T} \end{bmatrix}^{-1} = \begin{bmatrix} \left(\frac{\partial^2 L}{\partial \rho^2}\right)^{-1} & \mathbf{O} \\ \mathbf{O} & -\frac{1}{n}\hat{\boldsymbol{\Sigma}} \end{bmatrix},$$

where L denotes $L(\boldsymbol{\mu}, \sigma, \rho)$. Thus, we obtain

$$\ddot{\mathbf{L}}^{-1} - \mathbf{B}_{22} = \begin{bmatrix} \mathbf{L}_* & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix},$$

where

$$\mathbf{L}_* = -\frac{1}{np} \begin{bmatrix} \frac{1}{2}(1 + (p-1)\hat{\rho}^2)\hat{\sigma}^2 & \hat{\rho}(1-\hat{\rho})(1 + (p-1)\hat{\rho})\hat{\sigma} \\ \hat{\rho}(1-\hat{\rho})(1 + (p-1)\hat{\rho})\hat{\sigma} & \frac{2\hat{\rho}^2(1-\hat{\rho})^2(1+(p-1)\hat{\rho})^2}{1+(p-1)\hat{\rho}^2} \end{bmatrix}.$$

Therefore, $-2\ddot{\mathbf{G}}$ becomes

$$-2\ddot{\mathbf{G}}_\sigma = \frac{1}{np\{1 + (p-1)\hat{\rho}^2\}\hat{\sigma}^4}\mathbf{e}\mathbf{e}^T, \tag{3.5}$$

where $\mathbf{E}$ is defined in Theorem 1. Let $\mathbf{A} = c\,\mathbf{a}\mathbf{a}^T$ for a $p \times 1$ vector $\mathbf{a}$ and a constant c. The eigenvalues of A are $c\,\mathbf{a}^T\mathbf{a}$ and zeros. Furthermore, the normalized eigenvector of $\mathbf{A}$ corresponding to $c\,\mathbf{a}^T\mathbf{a}$ is $\mathbf{a}/\sqrt{\mathbf{a}^T\mathbf{a}}$. Therefore, $\boldsymbol{l}_{max}$ for $-2\ddot{\mathbf{G}}_\sigma$ is $\mathbf{e}/\sqrt{\mathbf{e}^T\mathbf{e}}$.

Next consider the case in which only ρ is of interest. We rearrange $\boldsymbol{\Delta}$ and $\ddot{\mathbf{L}}$ so that the terms related to ρ appear in the left upper corner. Then similarly to the previous case we obtain

$$-2\ddot{\mathbf{G}}_\rho = \frac{1}{np(p-1)\hat{\rho}^2}\left(\frac{1}{\hat{\sigma}^2}\mathbf{e} - \mathbf{h}\right)\left(\frac{1}{\hat{\sigma}^2}\mathbf{e} - \mathbf{h}\right)^T. \tag{3.6}$$

As in $-2\ddot{\mathbf{G}}_\sigma$, $\boldsymbol{l}_{max}$ for $-2\ddot{\mathbf{G}}_\rho$ becomes the normalized vector of $\mathbf{e}/\hat{\sigma}^2 - \mathbf{h}$.

When both parameters σ and ρ are of interest, we get $\ddot{\mathbf{G}}_{\sigma,\rho} = \mathbf{\Delta}_1^T \ddot{\mathbf{L}}_{11}^{-1} \mathbf{\Delta}_1$, where $\mathbf{\Delta}_1$ and $\ddot{\mathbf{L}}_{11}$ are defined in proof of Theorem 1. It can be rewritten as

$$-2\ddot{\mathbf{G}}_{\sigma,\rho} = \frac{1}{np}\mathbf{h}\mathbf{h}^T + \frac{1}{np(p-1)\hat{\rho}^2}\left(\frac{1}{\hat{\sigma}^2}\mathbf{e} - \mathbf{h}\right)\left(\frac{1}{\hat{\sigma}^2}\mathbf{e} - \mathbf{h}\right)^T.$$

In this case a closed form of $\boldsymbol{l}_{max}$ is not available.

4. Influence on a Test of Fitting the Equicorrelation Model

The test of the hypothesis

$$H_0 : \mathbf{\Sigma} = \sigma^2\{(1-\rho)\mathbf{I}_p + \rho \mathbf{1}_p \mathbf{1}_p^T\}$$

is usually performed by the likelihood ratio statistic given by

$$T = \frac{|\,\mathbf{S}\,|}{\hat{\sigma}^{2p}(1-\hat{\rho})^{p-1}\{1+(p-1)\hat{\rho}\}}.$$

Then

$$T_* = -\left\{n - 1 - \frac{p(p+1)^2(2p-3)}{6(p-1)(p^2+p-4)}\right\}\ln T \tag{4.1}$$

is approximately distributed as a chi-squared distribution with $\{p(p+1)/2 - 2\}$ degrees of freedom. The null hypothesis H_0 would be rejected for a significantly large value of T_* (Rencher, 1995, p. 277).

To investigate the influence of observations on the likelihood ratio statistic T, we consider the perturbed statistic $T(\mathbf{w})$ under the perturbation scheme (3.1). Differentiating $L(\boldsymbol{\mu}, \sigma, \rho|\mathbf{w})$ with respect to each parameter yields

$$\begin{aligned}\hat{\sigma}^2(\mathbf{w}) &= \mathrm{tr}(\mathbf{S}(\mathbf{w}))/p,\\ \hat{\rho}(\mathbf{w}) &= \frac{2}{p(p-1)\hat{\sigma}^2(\mathbf{w})}\sum_{i<j} s_{ij}(\mathbf{w}),\end{aligned}$$

where $s_{ij}(\mathbf{w})$ is the $(i,j)^{th}$ element of $\mathbf{S}(\mathbf{w})$ and

$$\begin{aligned}\mathbf{S}(\mathbf{w}) &= \sum_{u=1}^{n} w_u(\mathbf{x}_u - \hat{\boldsymbol{\mu}}(\mathbf{w}))(\mathbf{x}_u - \hat{\boldsymbol{\mu}}(\mathbf{w}))^T,\\ \hat{\boldsymbol{\mu}}(\mathbf{w}) &= \sum_{u=1}^{n} w_u \mathbf{x}_u \Big/ \sum_{u=1}^{n} w_u.\end{aligned}$$

Since the denominator of T is the determinant of $\hat{\mathbf{\Sigma}}$, the perturbed statistic $T(\mathbf{w})$ becomes $|\mathbf{S}(\mathbf{w})|/|\hat{\mathbf{\Sigma}}(\mathbf{w})|$, where $\hat{\mathbf{\Sigma}}(\mathbf{w}) = \hat{\sigma}^2(\mathbf{w})[\{1-\hat{\rho}(\mathbf{w})\}\mathbf{I}_p + \hat{\rho}(\mathbf{w})\mathbf{1}_p\mathbf{1}_p^T]$. Thus we have a surface $(\mathbf{w}^T, T(\mathbf{w}))$ from which an influence measure can be obtained (Lawrance, 1988). In this case the first derivative at $\mathbf{w} = \mathbf{1}_n$ does not vanish and thus provides valuable information about the local behaviour of $T(\mathbf{w})$. Note that for the likelihood displacement described in Section 3, the first derivative becomes zero.

First, we obtain the first order derivative of $\mathbf{S}(\mathbf{w})$ with respect to w_u evaluated at $\mathbf{w} = \mathbf{1}_n$ as $\mathbf{S}_u = (\mathbf{x}_u - \bar{\mathbf{x}})(\mathbf{x}_u - \bar{\mathbf{x}})^T/n$. Since

$$\frac{\partial}{\partial w_u}|\mathbf{S}(\mathbf{w})| = |\mathbf{S}(\mathbf{w})| \left(\mathrm{tr}[\mathbf{S}(\mathbf{w})^{-1} \frac{\partial}{\partial w_u} \mathbf{S}(\mathbf{w})] \right),$$

we have

$$\left. \frac{\partial}{\partial w_u}|\mathbf{S}(\mathbf{w})| \right|_{\mathbf{w}=\mathbf{1}_n} = \frac{1}{n}|\mathbf{S}|(\mathbf{x}_u - \bar{\mathbf{x}})^T \mathbf{S}^{-1}(\mathbf{x}_u - \bar{\mathbf{x}}).$$

Second, we will derive the first order derivatives of the perturbed estimators $\hat{\sigma}^2(\mathbf{w})$ and $\hat{\rho}(\mathbf{w})$. By the chain rule of differentiation (Eq. (14) of Cook, 1986), from the proof of Theorem 1 we obtain

$$\begin{aligned} (\dot{\sigma}, \dot{\rho}) &\equiv \left(\left. \frac{\partial \hat{\sigma}(\mathbf{w})}{\partial \mathbf{w}} \right|_{\mathbf{w}=\mathbf{1}_n}, \left. \frac{\partial \hat{\rho}(\mathbf{w})}{\partial \mathbf{w}} \right|_{\mathbf{w}=\mathbf{1}_n} \right) = -(\mathbf{L}_{11}^{-1} \mathbf{\Delta}_1)^T \\ &= \frac{1}{np} \left[\frac{1}{2\hat{\sigma}} \mathbf{e}, \frac{(1-\hat{\rho})\{1+(p-1)\hat{\rho}\}}{(p-1)\hat{\rho}} \left(\frac{1}{\hat{\sigma}^2} \mathbf{e} - \mathbf{h} \right) \right]. \end{aligned}$$

We can see that the normalized vectors of $\dot{\sigma}$ and $\dot{\rho}$ are the same as the eigenvector corresponding to the largest eigenvalue of $-2\ddot{\mathbf{G}}_\sigma$ and $-2\ddot{\mathbf{G}}_\rho$, respectively. When the parameter of interest is a scalar, the likelihood displacement approach and the first order derivative of the perturbed estimator provide the same influence information.

Similarly to the first order derivative of $|\mathbf{S}(\mathbf{w})|$, a little algebra gives

$$\left. \frac{\partial}{\partial w_u}|\hat{\mathbf{\Sigma}}(\mathbf{w})| \right|_{\mathbf{w}=\mathbf{1}_n} = \frac{1}{n}|\hat{\mathbf{\Sigma}}|(\mathbf{x}_u - \bar{\mathbf{x}})^T \hat{\mathbf{\Sigma}}^{-1}(\mathbf{x}_u - \bar{\mathbf{x}}).$$

Therefore we obtain the first order derivative of $T(\mathbf{w})$ with respect to w_u evaluated at $\mathbf{w} = \mathbf{1}_n$ as

$$T_u = \frac{T}{n}(\mathbf{x}_u - \bar{\mathbf{x}})^T \left(\mathbf{S}^{-1} - \hat{\mathbf{\Sigma}}^{-1} \right) (\mathbf{x}_u - \bar{\mathbf{x}}). \tag{4.2}$$

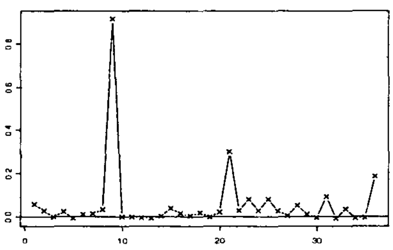
A large absolute value of T_u indicates that $\mathbf{x}_u$ is influential in testing H_0.

5. A Numerical Example

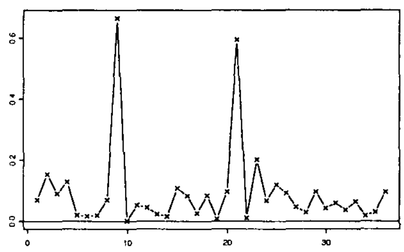
In this section, we consider the cost data which consists of 36 measurements on the per-mile cost of three variables: fuel, repair and capital. This data set is taken from Johnson and Wichern (1992, p. 276). When the covariance matrix does not have any structure, Bacon-Shone and Fung (1987) analyzed the data set and concluded that observations 9 and 21 are possible outliers.

First we will check whether this data set follows the equicorrelation model using the test statistic T_* in (4.1). The hypothesis that the data set follows the equicorrelation model would be rejected for a significantly large value of T_*. For the cost data, the value of T_* is 9.11 and the p-value is 0.058. Thus we would not reject the assumption of equicorrelation model at any significance level less than 0.058.

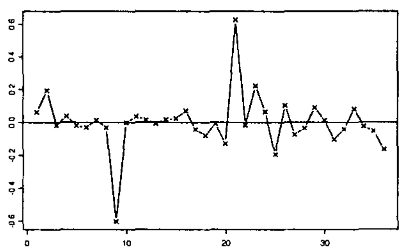
Next the results by the local influence method discussed in Section 3 are summarized in Figure 5.1. The y-axis in Figure 5.1 indicates the element of



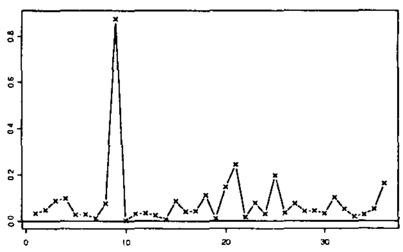
(a) *all parameters are of interest*



(b) *only σ is of interest*



(c) *only ρ is of interest*



(d) *both σ and ρ are of interest*

FIGURE 5.1 *Local influence method for parameters*

l_{max} corresponding to each case. When all parameters are of interest, Figure 5.1 (a) shows that observation 9 is most influential and that observations 21 is a little influential. When σ or ρ is of interest, Figure 5.1 (b) and (c) show that observations 9 and 21 are most influential in both figures. When both σ and ρ are of interest, Figure 5.1 (d) shows that the result is similar to that in Figure 5.1 (a).

Finally the index plot of the first order derivatives in (4.2) for the perturbed test statistic $T(\mathbf{w})$ (or $T_*(\mathbf{w})$), which is depicted in Figure 5.2, shows that only observation 9 is most influential on the test statistic. This result is supported by the case deletions of the test statistic T_* summarized in Table 5.1. Numbers in Table 5.1 are arranged in decreasing order of $|T_u|$ or $|T_{*(u)} - T_*|$, where $T_{*(u)}$ denotes the test statistic T_* after deleting observation u. Table 5.1 shows that the change in the value of the test statistic T_* due to single case deletions has its maximum value for the deletion of observation 9 and the next for the deletion of observation 20. However, its magnitude due to deletion of observation 20 is relatively small to that of observation 9. Furthermore after the deletion of observation 9, the p-value for the test statistic based on the remaining sample becomes 0.80 and thus the hypothesis that the data set follows the equicorrelation model would not be rejected at any reasonable significance level, whereas for the deletion of observation 20 the p-value is 0.023. It indicates that observation 9 has a large influence on the test but others do not.

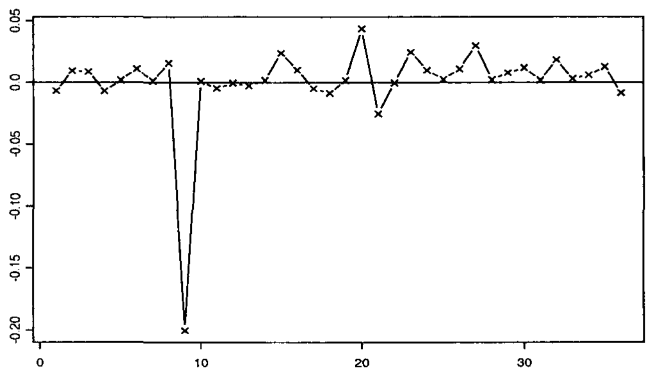


FIGURE 5.2 *A local influence measure for test statistic of equicorrelation model*

TABLE 5.1 *Comparison of T_u and case deletion result for T_**

u	T_u	u	$T_{*(u)} - T_*$
9	−0.200	9	−7.453
20	0.044	20	2.160
27	0.030	27	1.218
21	−0.025	21	−1.058

This example shows that the local influence method provides useful information about the influence of observations and that it can be a useful diagnostic tool.

REFERENCES

Bacon-Shone, J. and Fung, W. K. (1987). "A new graphical method for detecting single and multiple outliers in multivariate data", *Journal of the Royal Statistical Society*, **C36**, 153–162.

Barnett, V. and Lewis, T. (1994). *Outliers in Statistical Data*, 3rd ed., John Wiley & Sons, New York.

Cook, R. D. (1986). "Assessment of local influence (with discussions)", *Journal of the Royal Statistical Society*, **B48**, 133–169.

Giraudeau, B., Mallet, A. and Chastang, C. (1996). "Case influence on the intraclass correlation coefficient estimate", *Biometrics*, **52**, 1492–1497.

Johnson, A. J. and Wichern, D. W. (1992). *Applied Multivariate Statistical Analysis*, 3rd ed., Prentice-Hall, Englewood Cliffs.

Jung, K. -M., Kim, M. G. and Kim, B. C. (1997). "Local influence in maximum likelihood factor analysis", *Computational Statistics & Data Analysis*, **24**, 483–491.

Kim, M. G. (1995). "Local influence in multivariate regression", *Communications in Statistics-Theory and Methods*, **24**, 1271–1278.

Lawrance, A. J. (1988). "Regression transformation diagnostics using local influence", *Journal of the American Statistical Association*, **83**, 1067–1072.

Rencher, A. C. (1995). *Methods of Multivariate Analysis*, Wiley, New York.