J. Appl. Math. & Informatics Vol. **26***(2008), No. 3 - 4, pp. 679 - 688*
Website: http://www.kcam.biz

MULTIPLE DELETION MEASURES OF TEST STATISTICS IN MULTIVARIATE REGRESSION

KANG-MO JUNG

ABSTRACT. In multivariate regression analysis there exist many influence measures on the regression estimates. However it seems to be few of influence diagnostics on test statistics in hypothesis testing. Case-deletion approach is fundamental for investigating influence of observations on estimates or statistics. Tang and Fung (1997) derived single case-deletion of the Wilks' ratio, Lawley-Hotelling trace, Pillai's trace for testing a general linear hypothesis of the regression coefficients in multivariate regression. In this paper we derived more extended form of those measures to deal with joint influence among observations. A numerical example is given to illustrate the effect of joint influence on the test statistics.

AMS Mathematics Subject Classification : 62J20, 62H99
Key words and phrases : Hypothesis testing, influential cases, multiple case-deletions, multivariate regression.

1. Introduction

The detection of outliers and influential observations in regression analysis has a long history. In multivariate regression Barrett and Ling (1992) proposed the classes of multivariate influence measures using the projection matrix and the analogue of the projection matrix with respect to the residual matrix. However, many diagnostic measures have been proposed for the influence analysis in the context of estimation (Hossain and Naik, 1989). A few works that treat detection of influential observations for test statistics in multivariate analysis are found. Among others, Tang and Fung (1996, 1997) measured influence of observations on test statistics about a general hypothesis in multivariate regression using the local influence approach and the single case-deletion. Jung (2002) proposed the method to detect influential observations on the likelihood ratio test statistic for multivariate normal samples using the influence function.

Received September 7, 2007. Revised November 12, 2007. *Corresponding author.

Testing linear hypothesis is a fundamental step in regression analysis. If this test is of our primary concern, it is not desirable that the test statistic relies heavily on a few anomalous observations. Influence analysis in testing problems is very important because in extreme situations, as will be seen in Section 4 few observations can dominate our conclusion about the hypothesis.

In this work we will investigate influence of observations on testing linear hypothesis using multiple case-deletions. In Section 2 we review four testing statistics, namely the Wilks' ratio, Lawley-Hotelling trace, Pillai's trace and Roy's maximum root for testing a general linear hypothesis about regression coefficients. The change in the test statistics due to deletions of multiple cases is derived in Section 3 and it will be seen to be completely written in terms of statistics without involving the actual deletion of observations.

Furthermore we obtained similar expressions for the multiple case-deletions of the Lawley-Hotelling trace and Pillai's trace in terms of residuals. And the derived deletion measures are strongly related to two terms such as the leverage and residual components which can be a extended form from univariate regression models (Barrett and Ling, 1992). In Section 4 a numerical example is given for illustration and it shows that the conclusion for testing can be reversed by deletion of even a single observation. Furthermore we will see that these influential observations can not be detected from the regression diagnostics for unrestricted and restricted regression estimators. It needs directly the influence analysis for the testing statistics.

2. Test statistics in multivariate regression

Consider the multivariate linear regression model

$$\mathbf{Y}_{n\times p} = \mathbf{X}_{n\times q}\boldsymbol{\beta}_{q\times p} + \boldsymbol{\epsilon}_{n\times p}, \tag{1}$$

where $\mathbf{X}$ is known design matrix of rank q and the regression coefficient $\boldsymbol{\beta}$ is unknown. Furthermore, the rows of the error matrix $\boldsymbol{\epsilon}$ are independent and identically distributed as a p-variate normal distribution with mean zero and covariance matrix $\boldsymbol{\Sigma}$, that is, $\boldsymbol{\epsilon} \sim N_{n\times p}(\mathbf{0}, \mathbf{I}_n \otimes \boldsymbol{\Sigma})$, where the notation $\otimes$ denotes the Kronecker product of matrices. We test a general linear hypothesis for the regression coefficient matrix as

$$H_0 : \mathbf{A}\boldsymbol{\beta} = \mathbf{C},$$

where $\mathbf{A}$ is a known k by q matrix of rank k and $\mathbf{C}$ is a known k by p matrix. It is often to use four multivariate test statistics for testing this hypothesis. They are the Wilks' ratio, Lawley-Hotelling trace, Pillai's trace and Roy's maximum root statistics. The restricted least squares estimate of $\boldsymbol{\beta}$ subject to this hypothesis (Anderson, 1984, Chap. 8) is

$$\mathbf{B}_0 = \mathbf{B} - (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{A}^T\left[\mathbf{A}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{A}^T\right]^{-1}(\mathbf{A}\mathbf{B} - \mathbf{C}),$$

where $\mathbf{B} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$ is the least squares estimate under the unrestricted model.

Let $\mathbf{E} = (\mathbf{Y}-\mathbf{XB})^T(\mathbf{Y}-\mathbf{XB})$ and $\mathbf{E}_0 = (\mathbf{Y}-\mathbf{XB}_0)^T(\mathbf{Y}-\mathbf{XB}_0)$. Here $\mathbf{E}$ and $\mathbf{E}_0$ are the sum of squares and cross-products matrices of residuals under the unrestricted model and the null hypothesis, respectively. And let $\mathbf{H} = \mathbf{E}_0 - \mathbf{E}$, which can be rewritten as

$$\mathbf{H} = (\mathbf{AB} - \mathbf{C})^T\left[\mathbf{A}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{A}^T\right]^{-1}(\mathbf{AB} - \mathbf{C}).$$

If the hypothesis H_0 is true, $\mathbf{E}$ and $\mathbf{H}$ are independently distributed as Wishart distributions $W_p(n-q, \Sigma)$ and $W_p(k, \Sigma)$, respectively.

To test the hypothesis H_0 four test statistics are commonly used as follows.

(i) Wilks' ratio

$$\Lambda = \frac{|\mathbf{E}|}{|\mathbf{E}+\mathbf{H}|} = \frac{|\mathbf{E}|}{|\mathbf{E}_0|}, \tag{2}$$

(ii) Lawley-Hotelling trace

$$T = \mathrm{tr}(\mathbf{HE}^{-1}), \tag{3}$$

(iii) Pillai's trace

$$V = \mathrm{tr}\left[\mathbf{H}(\mathbf{E}+\mathbf{H})^{-1}\right] = \mathrm{tr}(\mathbf{HE}_0^{-1}), \tag{4}$$

(iv) Roy's maximum root

$$R = \text{the largest eigenvalue of } \mathbf{HE}^{-1}. \tag{5}$$

The test statistics are approximately distributed as the F distribution with some degrees of freedom. Under the null hypothesis it is known that

$$\frac{1-\Lambda^{1/t}}{\Lambda^{1/t}} \cdot \frac{rt-2u}{pk}$$

is approximately F distribution with degrees of freedom pk and $rt - 2u$, where

$$\begin{aligned} r &= n - q - \frac{p-k+1}{2} \\ u &= \frac{pk-2}{4} \\ t &= \begin{cases} \sqrt{\dfrac{p^2k^2-4}{p^2+k^2-5}} & \text{if } p^2+k^2-5>0 \\ 1 & \text{otherwise}\ . \end{cases} \end{aligned}$$

And also it is known that

$$\frac{2(s\nu_2+1)}{s^2(2\nu_1+s+1)}T \stackrel{\cdot}{\sim} F(s(2\nu_1+s+1), 2(s\nu_2+1)),$$

$$\frac{2\nu_2+s+1}{2\nu_1+s+1} \cdot \frac{V}{s-V} \stackrel{\cdot}{\sim} F(s(2\nu_1+s+1), s(2\nu_2+s+1)),$$

$$\frac{2\nu_2+2}{2\nu_1+2}R \stackrel{\cdot}{\sim} F(2\nu_1+2, 2\nu_2+2),$$

where $s = min(p,k), \nu_1 = (|p-k|-1)/2, \nu_2 = (n-q-p-1)/2$ (Seber, 1984, pp. 36-40).

Anderson (1984, Chap. 8) discussed the robustness of the four test statistics. All the four test procedures tend to be relatively robust in sense that the errors departures from normality. The limiting distributions of each criterion for nonnormal response random vector $\mathbf{y}_i$ are the same as when $\mathbf{y}_i$ is normal under the conditions of bounded fourth moments. A number of numerical comparisons for test statistics from (2) to (4) have been made. When the population characteristic roots of $|\mathbf{H}-\lambda\mathbf{E}|=0$ were very different, the Lawley-Hotelling trace tended to have the highest power. When the characteristic roots were equal, the Pillai's trace was most powerful. In these situations the Roy's maximum root was least powerful. See Anderson (1984, Chap. 8) for details.

3. Multiple case-deletions

In this section we will derive the change in the four test statistics due to removal of multiple cases. To this end the changes in $\mathbf{E}$ and $\mathbf{E}_0$ will be investigated separately.

Let $I = \{i_1, i_2, \ldots, i_m\}(m < n-p)$ be a set of containing the indices of the m cases to be deleted, where the number m is given. Without loss of generality, let us assume that the m cases to be omitted, are the last m cases, that is, $I = \{n-m+1, n-m+2, \ldots, n\}$. We use the multiple case-deletion method in which the effect of cases corresponding to the index set I on the statistic Z can be studied through $(n-m)(Z-Z_{(I)})$, where $Z_{(I)}$ is the statistic Z based on the sample without cases I. The above quantity is called the sample influence curve. Let $\mathbf{X}_I$ be the submatrix of $\mathbf{X}$ whose m rows are indexed by I and let $\mathbf{X}_{(I)}$ be its complement, the submatrix of $\mathbf{X}$ with $\mathbf{X}_I$ deleted. Thus we can partition $\mathbf{Y}$ and $\mathbf{X}$ as $\mathbf{Y}^T = \left(\mathbf{Y}_{(I)}^T, \mathbf{Y}_I^T\right)$ and $\mathbf{X}^T = \left(\mathbf{X}_{(I)}^T, \mathbf{X}_I^T\right)$, respectively. Thus the multivariate regression model (1) after deleting $\mathbf{X}_I$ and $\mathbf{Y}_I$ becomes

$$\mathbf{Y}_{(I)} = \mathbf{X}_{(I)}\boldsymbol{\beta} + \boldsymbol{\epsilon}_{(I)}, \tag{6}$$

where $\boldsymbol{\epsilon}_{(I)} \sim N_{(n-m)\times p}(\mathbf{0}, \mathbf{I}_{n-m}\otimes\boldsymbol{\Sigma})$.

Based on $\mathbf{X}_{(I)}$ and $\mathbf{Y}_{(I)}$ the least squares estimators of $\boldsymbol{\beta}$ for multivariate regression model is given by

$$\mathbf{B}_{(I)} = \left(\mathbf{X}_{(I)}^T\mathbf{X}_{(I)}\right)^{-1}\mathbf{X}_{(I)}^T\mathbf{Y}_{(I)}.$$

We have known that

$$\left(\mathbf{X}_{(I)}^T\mathbf{X}_{(I)}\right)^{-1} = (\mathbf{X}^T\mathbf{X})^{-1} + (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}_I^T(\mathbf{I}_m-\mathbf{P}_I)^{-1}\mathbf{X}_I(\mathbf{X}^T\mathbf{X})^{-1},$$

where $\mathbf{P}_I = \mathbf{X}_I(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}_I^T$. After some algebra we obtain

$$\mathbf{B}_{(I)} = \mathbf{B} - (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}_I^T(\mathbf{I}_m-\mathbf{P}_I)^{-1}\mathbf{U}_I, \tag{7}$$

where $\mathbf{U}_I = \mathbf{Y}_I - \mathbf{X}_I\mathbf{B}$ can be obtained by partitioning $\mathbf{U} = \mathbf{Y} - \mathbf{X}\mathbf{B}$. Similarly we have

$$\begin{aligned}\mathbf{B}_{0(I)} &= \mathbf{B}_0 - (\mathbf{X}^T\mathbf{X})^{-1}\Big[\mathbf{X}^T\mathbf{X} - \mathbf{A}^T\Big[\mathbf{A}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{A}^T\Big]^{-1}\mathbf{A}\Big] \\ &\qquad \times(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}_I^T(\mathbf{I}_m - \mathbf{P}_{0I})^{-1}\mathbf{U}_{0I},\end{aligned}$$

where $\mathbf{P}_{0I} = \mathbf{X}_I(\mathbf{X}^T\mathbf{X})^{-1}\Big[\mathbf{X}^T\mathbf{X} - \mathbf{A}^T\Big[\mathbf{A}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{A}^T\Big]^{-1}\mathbf{A}\Big](\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}_I^T$ and $\mathbf{U}_{0I} = \mathbf{Y}_I - \mathbf{X}_I\mathbf{B}_0$ is the building block of $\mathbf{U}_0 = \mathbf{Y} - \mathbf{X}\mathbf{B}_0$ corresponding to the index set I.

From the above equation (7) we obtain

$$\begin{aligned}\mathbf{E}_{(I)} &= \Big(\mathbf{Y}_{(I)} - \mathbf{X}_{(I)}\mathbf{B}_{(I)}\Big)^T(\mathbf{Y}_{(I)} - \mathbf{X}_{(I)}\mathbf{B}_{(I)}) \\ &= \Big(\mathbf{Y} - \mathbf{X}\mathbf{B}_{(I)}\Big)^T(\mathbf{Y} - \mathbf{X}\mathbf{B}_{(I)}) - \Big(\mathbf{Y}_I - \mathbf{X}_I\mathbf{B}_{(I)}\Big)^T(\mathbf{Y}_I - \mathbf{X}_I\mathbf{B}_{(I)}) \\ &= \mathbf{E} - \mathbf{U}_I^T(\mathbf{I}_m - \mathbf{P}_I)^{-1}\mathbf{U}_I, \end{aligned} \tag{8}$$

where the last equality can be verified from the facts that

$$\Big(\mathbf{Y} - \mathbf{X}\mathbf{B}_{(I)}\Big)^T\Big(\mathbf{Y} - \mathbf{X}\mathbf{B}_{(I)}\Big) = \mathbf{E} + \mathbf{U}_I^T(\mathbf{I}_m - \mathbf{P}_I)^{-1}\mathbf{P}_I(\mathbf{I}_m - \mathbf{P}_I)^{-1}\mathbf{U}_I$$

and

$$\Big(\mathbf{Y}_I - \mathbf{X}_I^T\mathbf{B}_{(I)}\Big)^T\Big(\mathbf{Y}_I - \mathbf{X}_I^T\mathbf{B}_{(I)}\Big) = \mathbf{U}_I^T(\mathbf{I}_m - \mathbf{P}_I)^{-2}\mathbf{U}_I.$$

The matrix $\mathbf{E}_{(I)}$ in (8) consists of basic building blocks such as residuals and projection matrix available from the full and reduced data with relatively small m cases to the sample size n, and therefore, there is little additional computation. By the decomposition of $\Big[\mathbf{A}(\mathbf{X}_{(I)}^T\mathbf{X}_{(I)})^{-1}\mathbf{A}^T\Big]^{-1}$ in terms of $\Big[\mathbf{A}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{A}^T\Big]^{-1}$ and some tedious algebra, we obtain

$$\begin{aligned}\mathbf{H}_{(I)} &= (\mathbf{A}\mathbf{B}_{(I)} - \mathbf{C})^T\Big[\mathbf{A}\Big(\mathbf{X}_{(I)}^T\mathbf{X}_{(I)}\Big)^{-1}\mathbf{A}^T\Big]^{-1}(\mathbf{A}\mathbf{B}_{(I)} - \mathbf{C}) \\ &= \mathbf{H} + \mathbf{W}_I - \mathbf{W}_{0I},\end{aligned} \tag{9}$$

where $\mathbf{W}_I = \mathbf{U}_I^T(\mathbf{I}_m - \mathbf{P}_I)^{-1}\mathbf{U}_I$ and $\mathbf{W}_{0I} = \mathbf{U}_{0I}^T(\mathbf{I}_m - \mathbf{P}_{0I})^{-1}\mathbf{U}_{0I}$. Note that the matrix $(\mathbf{I}_m - \mathbf{P}_{0I})\mathbf{Y}_I$ is the residual matrix of the reduced data set I under the null hypothesis and the form in (9) is similar to that in (8).

And also it is straightforward to show

$$\begin{aligned}\mathbf{E}_{0(I)} &= \mathbf{H}_{(I)} + \mathbf{E}_{(I)} \\ &= \mathbf{E}_0 - \mathbf{W}_{0I}.\end{aligned} \tag{10}$$

From the above results, we can obtain the multiple case-deletion measures for the first three test statistics described in the previous section. Let $\mathbf{Q}_I = \mathbf{U}_I\mathbf{E}^{-1}\mathbf{U}_I^T$ and $\mathbf{Q}_{0I} = \mathbf{U}_{0I}\mathbf{E}_0^{-1}\mathbf{U}_{0I}^T$.

(i) Wilks' ratio. The determinants of $\mathbf{E}_{(I)}$ in (8) and $\mathbf{E}_{0(I)}$ in (10) become

$$|\mathbf{E}_{(I)}| = |\mathbf{E}||\mathbf{I}_m - \mathbf{P}_I|^{-1}|\mathbf{I}_m - \mathbf{P}_I - \mathbf{Q}_I|$$

and

$$|\mathbf{E}_{0(I)}| = |\mathbf{E}_0||\mathbf{I}_m - \mathbf{P}_{0I}|^{-1}|\mathbf{I}_m - \mathbf{P}_{0I} - \mathbf{Q}_{0I}|.$$

Thus the multiple case-deletion influence measure of Wilks' ratio Λ in (2) is given by

$$\Lambda_{(I)} = \Lambda\left(\frac{|\mathbf{I}_m - \mathbf{P}_I - \mathbf{Q}_I|}{|\mathbf{I}_m - \mathbf{P}_I|} \times \frac{|\mathbf{I}_m - \mathbf{P}_{0I}|}{|\mathbf{I}_m - \mathbf{P}_{0I} - \mathbf{Q}_{0I}|}\right). \tag{11}$$

(ii) Lawley-Hotelling trace. From (8) we have

$$\mathbf{E}_{(I)}^{-1} = \mathbf{E}^{-1} + \mathbf{E}^{-1}\mathbf{R}_I\mathbf{E}^{-1},$$

where $\mathbf{R}_I = \mathbf{U}_I^T(\mathbf{I}_m - \mathbf{P}_I - \mathbf{Q}_I)^{-1}\mathbf{U}_I$. Taking trace of the matrix multiplication of (9) and the above equation yields

$$T_{(I)} = T - \mathrm{tr}\left(\mathbf{W}_{0I}\mathbf{E}^{-1}\right) - \mathrm{tr}\left(\mathbf{W}_{0I}\mathbf{E}^{-1}\mathbf{R}_I\mathbf{E}^{-1}\right) + \mathrm{tr}\left(\mathbf{E}_0\mathbf{E}^{-1}\mathbf{R}_I\mathbf{E}^{-1}\right). \tag{12}$$

(iii) Pillai's trace. In a way similar to (ii), we obtain

$$\mathbf{E}_{0(I)}^{-1} = \mathbf{E}_0^{-1} + \mathbf{E}_0^{-1}\mathbf{R}_{0I}\mathbf{E}_0^{-1},$$

where $\mathbf{R}_{0I} = \mathbf{U}_{0I}^T\left(\mathbf{I}_m - \mathbf{P}_{0I} - \mathbf{Q}_{0I}\right)^{-1}\mathbf{U}_{0I}$. And $\mathbf{H}_{(I)}\mathbf{E}_{0(I)}^{-1}$ in terms of $\mathbf{H}\mathbf{E}_0^{-1}$ can be derived and taking trace gives

$$V_{(I)} = V + \mathrm{tr}\left(\mathbf{W}_I\mathbf{E}_0^{-1}\right) + \mathrm{tr}\left(\mathbf{W}_I\mathbf{E}_0^{-1}\mathbf{R}_{0I}\mathbf{E}_0^{-1}\right) - \mathrm{tr}\left(\mathbf{E}\mathbf{E}_0^{-1}\mathbf{R}_{0I}\mathbf{E}_0^{-1}\right). \tag{13}$$

We can observe that (12) and (13) have similar forms, and they are identical by exchanging the deletion measures of the restricted and unrestricted models.

(iv) Roy's maximum root. Because the eigenvalue of the matrix based on the omission of cases can not be expressed as an explicit form with respect to that of the full data. Therefore we use $R - R_{(I)}$ by direct computation.

Note that Equations (11) to (13) contain the terms of the leverage matrices $\mathbf{P}_I$ and $\mathbf{P}_{0I}$ and the residual matrices $\mathbf{U}_I$ and $\mathbf{U}_{0I}$. And also, in case $m = 1$ the multiple case-deletion influence measures (11) to (13) reduce to the results of Tang and Fung (1997), respectively.

Barrett and Ling (1992) considered two sources for influential cases in estimation of multivariate regression. One is the leverage matrix which can be written in terms of $|(\mathbf{I}_m - \mathbf{P}_I)^a|$, where a is an integer. The other is the residual matrix having $\left|(\mathbf{I}_m - \mathbf{P}_I - \mathbf{Q}_I)^a(\mathbf{I}_m - \mathbf{P}_I)^b\right|$. We can see that Equations (11) to (13) consist of these matrices. However, the deletion measures for test statistics cannot be written directly in simple building blocks of those matrices. It is caused by the fact that the derived deletion measures contain both unrestricted and restricted sources.

TABLE 1. Some case-deletions of the Wilks' ratio test statistic for Rohwer data.

single case		double case		triple case	
cases	$\Lambda - \Lambda_{(i)}$	cases	$\Lambda - \Lambda_{(I)}$	cases	$\Lambda - \Lambda_{(I)}$
14	0.091	14, 31	0.188	8, 14, 31	0.240
31	0.082	8, 31	0.129	14, 20, 31	0.235
5	−0.081	8, 14	0.125	14, 25, 31	0.224
10	−0.044	5, 10	−0.124	14, 21, 31	0.211
21	0.032	20, 31	0.120	14, 28, 31	0.210

4. Numerical example

The multiple case-deletion is applied to the data set collected by Rohwer which is analyzed by Hossain and Naik (1989), Tang and Fung (1996, 1997) and Kim (1995). Thirty-two students from an upper-class, white, residential school, were selected at random for a study of the relation of paired-associated (PA), learning-proficiency tests to children's performances. The independent variables were the sum of the number of items correct out of 20 (on two exposures) to five types of PA tasks. The basic tasks were named (x_1), still (x_2), named still (x_3), named action (x_4), and sentence still (x_5). Children's performances are examined in terms of the Peabody picture vocabulary test (y_1), the student achievement test (y_2), and the Ravin progressive matrices test (y_3).

The model

$$\mathbf{Y}_{32\times 3} = \mathbf{X}_{32\times 6}\boldsymbol{\beta}_{6\times 3} + \boldsymbol{\epsilon}_{32\times 3},$$

with $\boldsymbol{\epsilon}$ is distributed from $N_{32\times 3}(\mathbf{0}, \mathbf{I}_{32} \otimes \boldsymbol{\Sigma})$ and $\boldsymbol{\Sigma}$ is the 3×3 covariance matrix. Consider the null hypothesis $H_0 : \boldsymbol{\beta}_2 = \mathbf{0}$, where $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, ..., \boldsymbol{\beta}_6)$. This hypothesis test whether the basic task of named in the model exists or not. In this case the four test statistics are identical with the value 0.798 and their p-values are 0.138, and all test statistics under the null hypothesis are approximately distributed as the F distribution with 3 and 24 degrees of freedom. Therefore at the 5% level of significance there is insufficient evidence to conclude that $\boldsymbol{\beta}_2 \neq \mathbf{0}$. However, the p-value becomes 0.043 when case 14 is deleted. This implies that case 4 is very influential in testing the hypothesis.

For further analysis we conducted single, double and triple case-deletions for the Wilks' ratio test statistic. The results of some case-deletions are summarized in Table 1 whose rows are arranged in the decreasing order of the absolute difference $|\Lambda - \Lambda_{(I)}|$. The deletion measure $|\Lambda - \Lambda_{(I)}|$ is proportional to the sample influence curve. Also Table 2 presents the p-value of the Wilks' ratio on the reduced data without corresponding cases. Its order is arranged according to the absolute value of p-value differences.

Since the p-value on the full data is 0.138, the case corresponding to the p-value below the reference significance level 0.05 may be influential. The sign of case 5 in the single case-deletion is negative and the associated p-value with the

TABLE 2. The p-value of the Wilks' ratio test statistic on the deleted data set.

single case		double case		triple case	
cases	p-value	cases	p-value	cases	p-value
5	0.389	5, 10	0.608	5, 19, 27	0.782
14	0.042	5, 27	0.572	5, 10, 19	0.746
10	0.258	5, 19	0.490	5, 10, 30	0.743
31	0.049	5, 13	0.478	5, 10, 23	0.710
13	0.214	1, 5	0.469	4, 5, 10	0.697

deletion of case 5 increases to 0.389. Further case-deletions including case 5 can be seen in Table 2. It may need more attention to cases whose p-value changes to below 0.05 by case-deletions. Without case 14 and case 31 the associated p-values changes from 0.138 to 0.043 and 0.049, respectively. Thus cases 14 and 31 are very influential than cases 5 and 10, because the removal of the former cases reverses the conclusion about the test of $H_0 : \boldsymbol{\beta}_2 = \mathbf{0}$ at the 5% significance level.

Tables 1 and 2 show that cases 5, 8, 10, 14 and 31 can be influential for the Wilks' ratio test statistic. In particular cases 5, 14 and 31 are very influential and cases 8 and 10 are little. We can observe that the single case-deletion does not detect case 8, but double and triple case-deletions can detect case 8 as an influential case.

Barrett and Ling (1992) proposed a diagnostic plot for uncovering the sources of multivariate influence. Equation (11) is similar to the residual part of $1/AP_I$ proposed by Andrews and Pregibon (1978). However, the sources of multivariate influence in testing problem cannot be easily obtained. We can observe that the absolute value of $\Lambda - \Lambda_{(I)}$ is proportional to the ratio of $\mathfrak{R}_{\mathfrak{J}}$ and $\mathfrak{R}_{o\mathfrak{J}}$, where $\mathfrak{R}_{\mathfrak{J}} = |\mathbf{I}_m - \mathbf{P}_{\mathfrak{J}} - \mathbf{Q}_{\mathfrak{J}}||\mathbf{I}_m - \mathbf{P}_{\mathfrak{J}}|^{-1}$ and $\mathfrak{R}_{o\mathfrak{J}}$ is similarly defined.

We propose a diagnostic plot $\mathfrak{R}_{\mathfrak{J}}$ versus $\mathfrak{R}_{o\mathfrak{J}}$ for the Wilks' ratio test statistic. When cases are apart from the straight line with slope 1, these cases may be influential for the Wilks' ratio test statistic. Figure 1 is the plot of $\mathfrak{R}_i$ versus $\mathfrak{R}_{oi}$ for the Rohwer data with $m = 1$, the number of deletion. From this figure we can see that the influence of case 5 under the null hypothesis becomes larger than that under the unrestricted case. On the contrary cases 14 and 31 have large influence under the unrestricted case relative to the null hypothesis.

To reveal joint influence the diagnostic plot $\mathfrak{R}_{\mathfrak{J}}$ versus $\mathfrak{R}_{o\mathfrak{J}}$ for $m = 2$ was depicted in Figure 2. The subsets $\{14, 31\}$, $\{5, 25\}$, $\{14, 25\}$ and $\{25, 31\}$ are especially interesting. In Figure 2 the role of case 25 is noticeable for double case-deletions. Case 25 has large joint influence which is not large in individual influence as can be seen in single case-deletion in Table 1 and Figure 1. It shows the source of influence of case 25 which is large under both the unrestricted case and the null hypothesis. We confirmed the joint influence of case 25 in Figure 2.

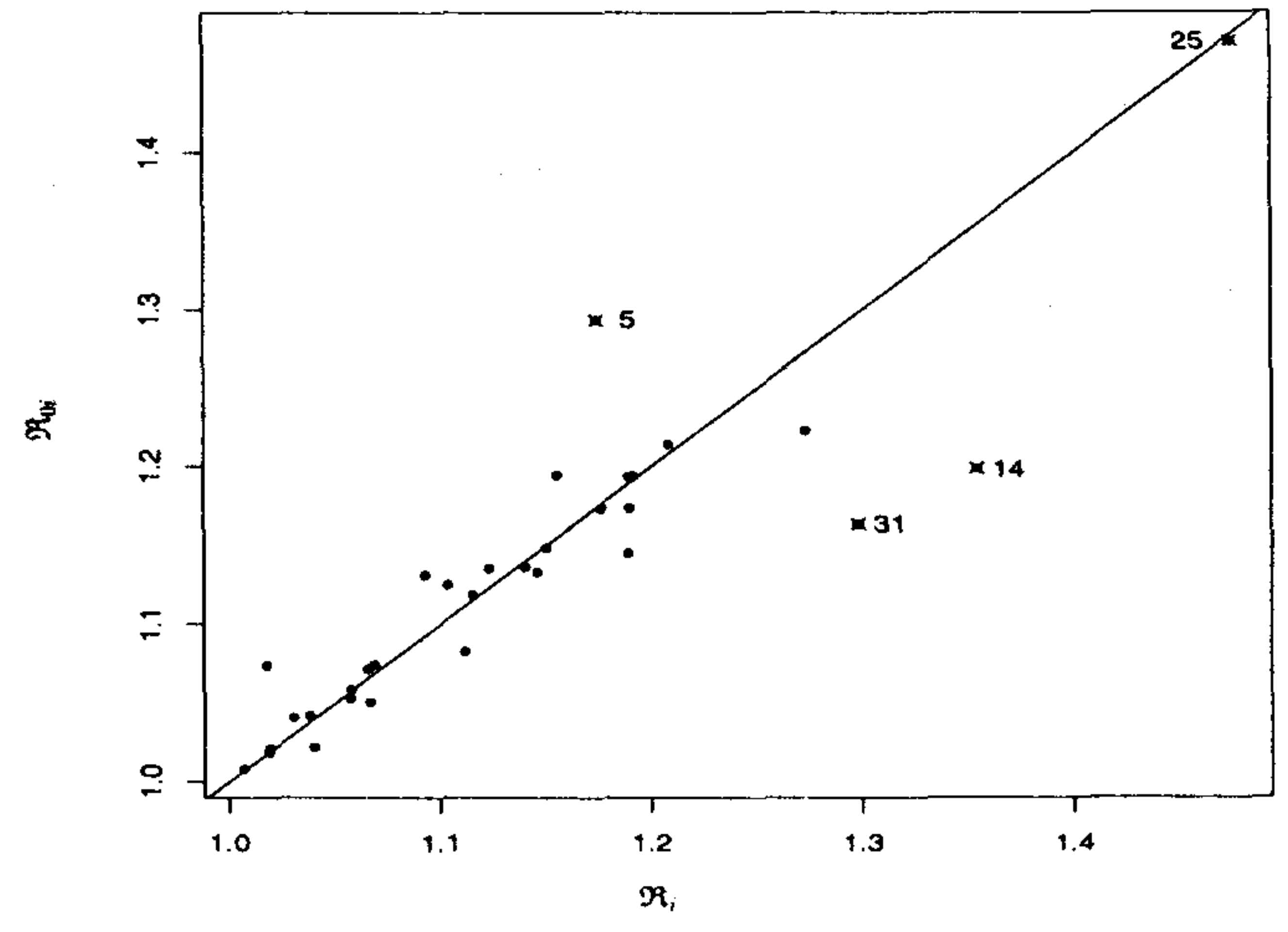

FIGURE 1. Single case influence of $\mathfrak{R}_i$ versus $\mathfrak{R}_{oi}$ for the Rohwer data.

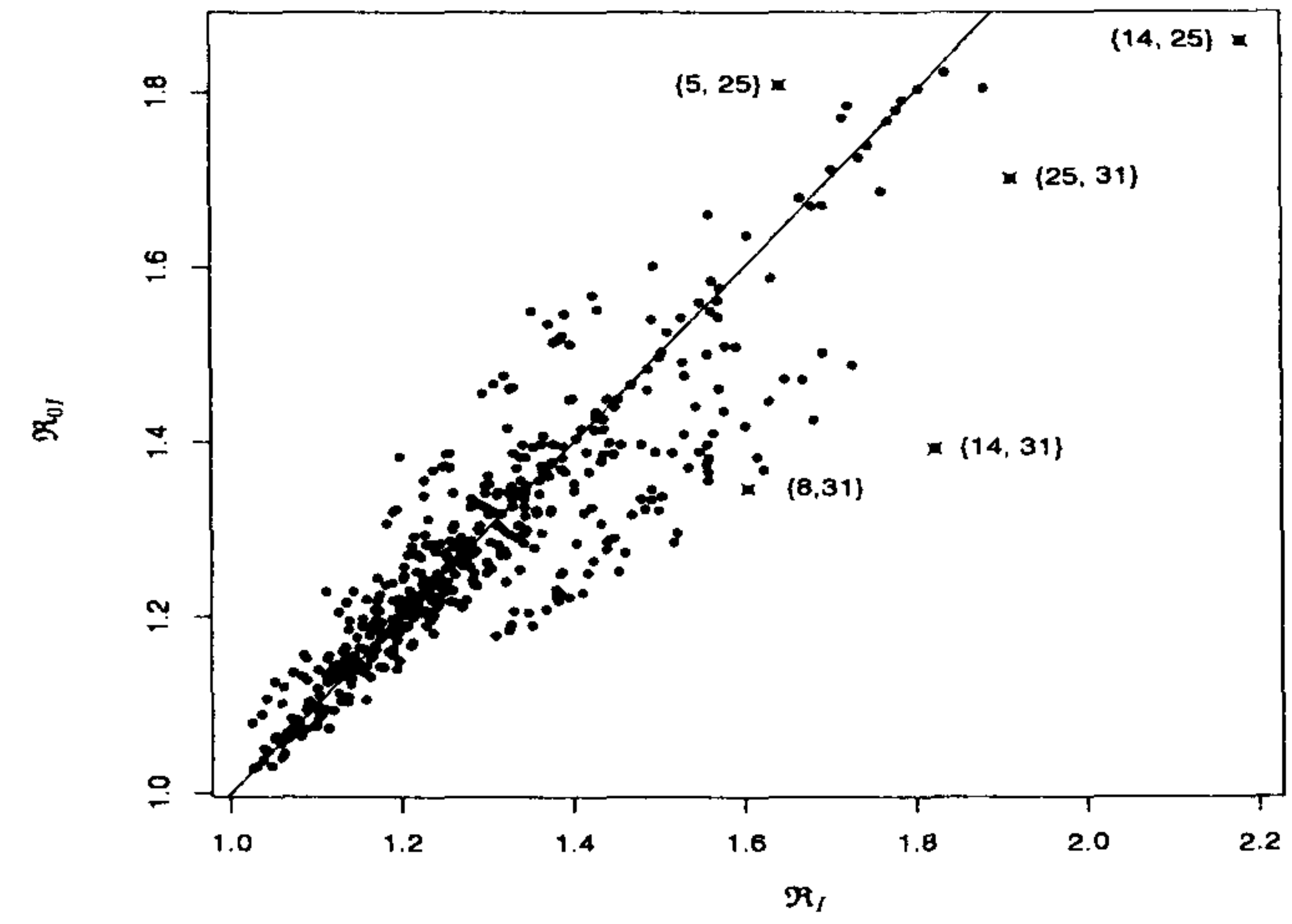

FIGURE 2. Double case influence of $\mathfrak{R}_3$ versus $\mathfrak{R}_{o3}$ for the Rohwer data.

Barrett and Ling (1992) concluded that case 5 has the largest influence as well as the largest leverage in the unrestricted case. But Kim (1995) and Tang and Fung (1997) indicated cases 5 and 14 as influential. We identified all these cases by multiple case-deletions. And we may conclude that cases 5, 14 and 31 are very influential for the general hypothesis $\beta_2 = \mathbf{0}$ and cases 8 and 10 are little. Case 25 has large joint influence with these influential cases. The influence analysis on regression estimates $\mathbf{B}$ and $\mathbf{B}_0$ did not detect influential cases 25 and 31. Hence the example emphasizes the necessity for the direct influence analysis of the Wilks' ratio test statistic, not relying solely on residual analysis and it shows the usefulness of multiple case-deletions for testing problems in multivariate regression. Furthermore the diagnostic plot of $\mathfrak{R}_{\mathfrak{I}}$ versus $\mathfrak{R}_{o\mathfrak{I}}$ helps for uncovering the source of influence.

REFERENCES

1. T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, John Wiley and Sons, New York, 1984.
2. D. F. Andrews and D. Pregibon, *Finding the outliers that matter*, *J. R. Statist. Soc. B*, **40** (1978), 85-93.
3. B. E. Barrett and R. F. Ling, *General classes of influence measures for multivariate regression*, J. Amer. Statist. Assoc., **87** (1992), 184-191.
4. A. Hossain and D. N. Naik, *Detection of influential observations in multivariate regression*, J. of Applied Statistics, **16** (1989), 25-37.
5. K.-M. Jung, *Influence function of the likelihood ratio test statistic for multivariate sample*, Commun. in Statist. A - Theory and Methods, **31** (2002), 1273-1281.
6. M. G. Kim, *Local influence in multivariate regression*, Commun. in Statist. A - Theory and Methods, **24** (1995), 1271-1278.
7. G. A. F. Seber, *Multivariate Observations*, Wiley, New York, 1984.
8. M. K. Tang and W. K. Fung, *First order local influence of test statistics in multivariate regression*, Sankhyā, **58** (1996), 323-337.
9. M. K. Tang and W. K. Fung, *Case-deletion diagnostics for test statistics in multivariate regression*, Austral. J. of Statist., **39** (1997), 345-353.

Kang-Mo Jung received his BS from Seoul National University and Ph.D at KAIST. Since 1997 he has been a professor of Department of Informatics and Statistics at Kunsan National University. His research interests focus on robust statistics.

Department of Informatics and Statistics, Kunsan National University, Kunsan 573-701, Korea

e-mail: kmjung@kunsan.ac.kr