The Korean Communications
in Statistics Vol. 4, No. 2, 1997
pp. 565 - 579

On a Robust Test for Parallelism of Regression Lines against Ordered Alternatives [1)]

Moon Sup Song[2)] and Jin Ho Kim[3)]

Abstract

A robust test is proposed for the problem of testing the parallelism of several regression lines against ordered alternatives. The proposed test statistic is based on a linear combination of *one-step pairwise GM-estimators*. We compare the performance of the proposed test with that of the other tests through a Monte Carlo simulation. The results of the simulation study show that the proposed test has stable levels, good empirical powers in various circumstances, and particularly higher empirical powers under the presence of extreme outliers or leverage points.

1. Introduction

We are interested in the problem of testing the equality of slopes of several regression lines against the alternatives that the slopes are in increasing (or decreasing) order of magnitude. In practical situations such as biology and ecology, the problem of ordered alternatives in slope parameters could arise. Adichie (1976) and Rao and Gore (1984) stated these situations. For example, a biologist may be interested in knowing whether the rate of dependence of infection on exposure is the same for groups of rats of increasing ages.

Consider the k simple regression lines

$$y_{ij}=\alpha_i+\beta_i x_{ij}+\varepsilon_{ij}, \quad i=1,\cdots,k\ ;\ j=1,\cdots,n_i, \tag{1.1}$$

where the α_i's and β_i's are unknown regression parameters, the x_{ij}'s are known constants, and the ε_{ij}'s are independent and identically distributed (iid) random variables with a continuous symmetric distribution function F with finite variance σ^2. Here, the α_i's are nuisance parameters and the β_i's are the slope parameters of interest.

We are concerned with the problem of testing the parallelism of regression lines

---

1) The present studies were supported in part by the Seoul National University Daewoo Research Fund.
2) Professor, Department of Statistics, Seoul National University, Seoul, 151-742, Korea.
3) Researcher, Korea Research Institute of Standards and Science, Taejon, 305-606, Korea.

$$H_0 : \beta_1 = \beta_2 = \cdots = \beta_k = \beta \text{ (unknown)} \tag{1.2}$$

against the ordered alternatives

$$H_1 : \beta_1 \le \beta_2 \le \cdots \le \beta_k \text{ (with at least one strict inequality).} \tag{1.3}$$

In the case of $k=2$, the classical t-test is a most powerful test at normal distribution. But, the t-test is very sensitive to the assumption of normality, especially to outliers. As an alternative of t-test, some nonparametric tests have been considered by Hollander (1970) and Potthoff (1974), among others.

For the parallelism of general k regression lines against ordered alternatives, parametric or nonparametric tests have been considered by Adichie (1976), Rao and Gore (1984), and Sin (1993) among others. Adichie (1976) proposed some parametric and nonparametric tests. Parametric tests are the likelihood ratio (LR) test and the scores test based on a linear combination of maximum likelihood estimators (MLE) of slopes. Nonparametric tests are rank analogues of the parametric tests. Rao and Gore (1984) also proposed distribution-free tests for the same problem. But they assumed that the design points are equispaced and each line has the same number of observations.

However, it is questionable whether these tests perform well in the presence of outliers, although some tests are distribution-free. Especially, none of these have been considered in the case of outliers in x-direction or leverage points. We consider a robust test based on the *one-step pairwise GM-estimator* (Song and Kim, 1997) for the parallelism of k regression lines. We expect that the test is reasonably powerful and robust, particularly in the presence of bad leverage points.

The parametric and nonparametric tests proposed by Adichie (1976) are briefly reviewed in Section 2. It is known that the procedure proposed by Rao and Gore (1984) does not appear to be better than the Adichie's test (Jee, 1989; Song, Kim, Jeon, and Park, 1989) and the empirical powers of Adichie's LR test and its rank version are similar to those of the scores test and its rank analogue, respectively (Jee, 1989; Lee, 1990). We hence include only the Adichie's parametric and nonparametric tests based on scores to compare with our proposed test.

In Section 3 we propose the test statistic based on a linear combination of one-step pairwise GM-estimators. We compare the performance of the proposed test with that of others through a Monte Carlo study in Section 4. The results of the simulation study show that the proposed test has stable levels, good empirical powers in various circumstances, and particularly higher empirical powers under the presence of extreme outliers or leverage points.

2. Adichie's Tests Based on Scores

Adichie (1976) proposed some parametric and nonparametric tests based on scores.

Parametric test is a linear combination of maximum likelihood estimators of slopes. Nonparametric test is a rank analogue of the parametric test. To introduce Adichie's tests, we denote some notations. Let, for $i=1,2,\cdots,k,$

$$\bar{x}_i=\sum_j x_{ij}/n_i;\quad \bar{y}_i=\sum_j y_{ij}/n_i,$$

$$w_i^2=\sum_j (x_{ij}-\bar{x}_i)^2;\quad W^2=\sum_i w_i^2;\quad r_i=w_i^2/W^2,$$

$$\tilde{\beta}_i=\sum_j (x_{ij}-\bar{x}_i)y_{ij}/w_i^2;\quad \tilde{\beta}=\sum_i r_i\tilde{\beta}_i;\quad N=\sum_i n_i \tag{2.1}$$

The *parametric scores test statistic*, proposed by Adichie (1976), is defined as

$$S=\sum_i C_i\tilde{\beta}_i,\quad (\sum_i C_i=0). \tag{2.2}$$

It is a linear combination of the MLE of β_i's under H_1. Adichie (1976) showed that S has a normal distribution with mean $\sum_i C_i\beta_i$ and variance $\sigma^2\sum_i(C_i^2/w_i^2)$ at normal distribution. When the scores C_i are nondecreasing, the test rejects H_0 for large values of S.

The power of S-test which was derived by Adichie (1976), depends on the scores C_i. When the alternative β_i's are preassigned, optimum scores which maximize the power are given by

$$C_i=w_i^2(\beta_i-\sum_j r_j\beta_j). \tag{2.3}$$

However, we can not use these scores, since the β_i's are not usually specified under the alternatives. Adichie (1976) suggested, instead of β_i in (2.3), to use

$$Q_i=w_1^2+\cdots+w_{i-1}^2+(w_i^2/2). \tag{2.4}$$

That is, the scores become

$$C_i=w_i^2(Q_i-\sum_j r_jQ_j). \tag{2.5}$$

When σ^2 is unknown, S can be studentized to yield S_t which has a Student t distribution with $(N-2k)$ degrees of freedom. The statistic S_t is defined by

$$S_t=[\sum_i(C_i^2/w_i^2)]^{-1/2}S/\tilde{\sigma}, \tag{2.6}$$

where

$$\tilde{\sigma}^2=\sum_i\sum_j[y_{ij}-\bar{y}_i-\tilde{\beta}_i(x_{ij}-\bar{x}_i)]^2/(N-2k).$$

Adichie (1976) also proposed a nonparametric test as a rank version of the test statistic S in (2.2). To introduce the rank version, consider the statistic

$$T_{n_i}=\sum_i\sum_j(x_{ij}-\bar{x}_j)R_{ij}^*/(n_i+1)w_i^2,$$

where R_{ij}^* is the rank of the jth residual, $y_{ij}-\tilde{\beta}(x_{ij}-\bar{x}_i)$, among the ith group of residuals with some estimator $\bar{\beta}$ of the common slope β. To derive an estimator of the common slope β, Adichie (1976) defined the statistic

$$T(\beta) = \sum_i \sum_j (x_{ij} - \bar{x}_i) R_{ij} / W^2,$$

where R_{ij} is the rank of $y_{ij} - \beta x_{ij}$ based on n_i observations within the ith group of samples. Then, the estimator $\bar{\beta}$ is defined by

$$\bar{\beta} = (\bar{\beta}_1 + \bar{\beta}_2)/2,$$

where

$$\bar{\beta}_1 = \sup\{\beta : T(\beta) > 0\} \text{ and } \bar{\beta}_2 = \inf\{\beta : T(\beta) < 0\}.$$

The *rank analogue of the scores test*, proposed by Adichie (1976), is of the form

$$S_R = \sum_i C_i T_{n_i}, \quad (\sum_i C_i = 0). \tag{2.7}$$

Adichie (1976) showed also that, under H_0, $W \cdot S_R$ is asymptotically $N(0, \upsilon^2)$, where $\upsilon^2 = (1/12) \sum_i (C_i^2 / r_i)$. We reject the null hypothesis H_0 for large values of S_R.

3. The Pairwise GM-Estimator and the Proposed Test Statistic

3.1 The One-Step Pairwise GM-Estimator

To construct a test statistic based on a linear combination of *one-step pairwise GM-estimators* proposed by Song and Kim (1997), we briefly introduce the pairwise GM-estimator.

In the subsection, we consider the multiple linear regression model

$$y_i = \alpha + \boldsymbol{x}_i^T \boldsymbol{\beta} + \varepsilon_i, \quad i = 1, 2, \cdots, n, \tag{3.1}$$

where $(\boldsymbol{x}_i, y_i), i = 1, 2, \cdots, n$, is a sequence of iid random variables with distribution function $F(\boldsymbol{x}, y)$ and density $f(\boldsymbol{x}, y)$, $\boldsymbol{x}_i$ is a $p \times 1$ vector of explanatory variables, and $(\alpha, \boldsymbol{\beta}^T)^T$ is a $(p+1) \times 1$ vector of parameters. Assume that the errors (ε_i's) are iid, independent of $\boldsymbol{x}$, symmetric about 0, and have a finite variance σ^2. Given an estimator $(\hat{\alpha}, \hat{\boldsymbol{\beta}})$, the residuals are denoted by $r_i(\hat{\alpha}, \hat{\boldsymbol{\beta}}) = y_i - \hat{\alpha} - \boldsymbol{x}_i^T \hat{\boldsymbol{\beta}}$ and $r_i(\hat{\boldsymbol{\beta}}) = y_i - \boldsymbol{x}_i^T \hat{\boldsymbol{\beta}}$.

The *pairwise GM-estimator* is defined by the solution of the (vector) equation

$$g = \binom{n}{2}^{-1} \sum\sum_{i<j} \eta\left(\boldsymbol{x}_i, \boldsymbol{x}_j, \frac{r_j(\boldsymbol{\beta}) - r_i(\boldsymbol{\beta})}{\sqrt{2}\sigma} \right) (\boldsymbol{x}_j - \boldsymbol{x}_i) = \mathbf{0}, \tag{3.2}$$

where the η-function is defined by

$$\eta\left(\mathbf{x}_i, \mathbf{x}_j, \frac{r_j(\beta)-r_i(\beta)}{\sqrt{2}\sigma}\right) = w_{ij}\,\psi\left(\frac{r_j(\beta)-r_i(\beta)}{\sqrt{2}\sigma}\right)$$

with an odd and bounded function ψ and $2\sigma^2 = Var\{r_j(\beta)-r_i(\beta)\}$. Song and Kim (1997) proposed the weight w_{ij} to downweight the leverage points and also outliers simultaneously as follows:

$$\begin{aligned} w_{ij} &= w\left(\mathbf{x}_i, \mathbf{x}_j, \frac{r_j(\beta)-r_i(\beta)}{\sqrt{2}\sigma}\right) \\ &= \begin{cases} v(\mathbf{x}_i)v(\mathbf{x}_j), & |\{r_j(\beta)-r_i(\beta)\}/\sqrt{2}\sigma| < a \\ 0, & otherwise, \end{cases} \end{aligned} \tag{3.3}$$

where a is a constant and $v(\mathbf{x})$ is a measure of leverageness defined as

$$v(\mathbf{x}) = \min\left[1, \left\{\frac{b}{(\mathbf{x}-\mathbf{m}_x)^T C_x^{-1}(\mathbf{x}-\mathbf{m}_x)}\right\}^{r/2}\right], \tag{3.4}$$

which was considered by Simpson, Ruppert, and Carroll (1992). Note that $\mathbf{m}_x$ and C_x are the minimum volume ellipsoid (MVE) estimators of location and covariance of $\mathbf{x}$, respectively, and b is a quantile of the chi-squared distribution with p degrees of freedom.

Let $\widehat{\beta}_0$ be an initial estimator of β such as the least trimmed squares (LTS) estimator, which was proposed by Rousseeuw (1984). Let $\hat{\sigma}_0$ be a robust estimator of σ such as $\hat{\sigma}_0 = 1.4826 \times MAD\{r_i(\widehat{\beta}_0)\}$, where MAD is the median absolute deviation defined as

$$MAD\{r_i(\widehat{\beta}_0)\} = med_i\{\,|r_i(\widehat{\beta}_0) - med_j(r_j(\widehat{\beta}_0))|\}.$$

The *one-step pairwise GM-estimator* (Song and Kim, 1997) based on $\widehat{\beta}_0$ is defined, by taking a first-order Taylor-series expansion of the left side of (3.2) about β, as follows:

$$\widehat{\beta} = \widehat{\beta}_0 + \sqrt{2}\,\hat{\sigma}_0\,\widehat{H}^{-1}\,\hat{\mathbf{g}}, \tag{3.5}$$

where

$$\hat{\mathbf{g}} = \binom{n}{2}^{-1} \sum\sum_{i<j} \eta\left(\mathbf{x}_i, \mathbf{x}_j, \frac{r_j(\widehat{\beta}_0)-r_i(\widehat{\beta}_0)}{\sqrt{2}\,\hat{\sigma}_0}\right)(\mathbf{x}_j-\mathbf{x}_i)$$

and

$$\widehat{H} = \binom{n}{2}^{-1} \sum\sum_{i<j} \hat{w}_{ij}\,\psi'\left(\frac{r_j(\widehat{\beta}_0)-r_i(\widehat{\beta}_0)}{\sqrt{2}\,\hat{\sigma}_0}\right)(\mathbf{x}_j-\mathbf{x}_i)(\mathbf{x}_j-\mathbf{x}_i)^T \tag{3.6}$$

with

$$\begin{aligned} \hat{w}_{ij} &= w\left(\mathbf{x}_i, \mathbf{x}_j, \frac{r_j(\widehat{\beta}_0)-r_i(\widehat{\beta}_0)}{\sqrt{2}\hat{\sigma}_0}\right) \\ &= \begin{cases} v(\mathbf{x}_i)v(\mathbf{x}_j), & |\{r_j(\widehat{\beta}_0)-r_i(\widehat{\beta}_0)\}/\sqrt{2}\,\hat{\sigma}_0| < a \\ 0, & otherwise. \end{cases} \end{aligned} \tag{3.7}$$

Here ψ' denotes $\psi'(t)=(\partial/\partial t)\psi(t)$.

Remark 1. The one-step pairwise estimator inherits the breakdown point of an initial estimator and has a bounded influence function. So, it can obtain the asymptotic breakdown point of .5.

We now state a theorem concerning the asymptotic normality of the one-step pairwise GM-estimator $\hat{\boldsymbol{\beta}}$. The following assumptions are used in the theorem.

(A1) ψ is odd and bounded with properties such that ψ' & ψ'' exist, and $|\psi'(t)|$ & $|\psi''(t)|$ are bounded.

(A2) $\binom{n}{2}^{-1}\sum\sum_{i<j} w_{ij}\psi'\left(\frac{r_j(\boldsymbol{\beta})-r_i(\boldsymbol{\beta})}{\sqrt{2}\sigma}\right)(\boldsymbol{x}_j-\boldsymbol{x}_i)(\boldsymbol{x}_j-\boldsymbol{x}_i)^T \xrightarrow{p} H$, where

$$H=\int\int w_{ij}\psi'\left(\frac{r_2(\boldsymbol{\beta}(F))-r_1(\boldsymbol{\beta}(F))}{\sqrt{2}\sigma}\right)(\boldsymbol{x}_2-\boldsymbol{x}_1)(\boldsymbol{x}_2-\boldsymbol{x}_1)^T dF(\boldsymbol{x}_2,y_2)dF(\boldsymbol{x}_1,y_1) \quad (3.8)$$

and H is a positive definite $p\times p$ matrix.

(A3) $\|w(\boldsymbol{x}_i,\boldsymbol{x}_j,t)(\boldsymbol{x}_j-\boldsymbol{x}_i)\|$ is bounded for any t and $\boldsymbol{x}$.

(A4) $w_{ij}=w\left(\boldsymbol{x}_i,\boldsymbol{x}_j,\frac{r_j(\boldsymbol{\beta})-r_i(\boldsymbol{\beta})}{\sqrt{2}\sigma}\right)$ is an even function of $r_j(\boldsymbol{\beta})-r_i(\boldsymbol{\beta})$ such that as $n\to\infty$, (a) $\binom{n}{2}^{-1}\sum\sum_{i<j}\left(\frac{r_j(\boldsymbol{\beta})-r_i(\boldsymbol{\beta})}{\sqrt{2}\sigma}\right)^2 w_{ij}\|\boldsymbol{x}_i-\boldsymbol{x}_j\|=O_p(1)$,

(b) $\binom{n}{2}^{-1}\sum\sum_{i<j} w_{ij}\|\boldsymbol{x}_i-\boldsymbol{x}_j\|^3=O_p(1)$,

(c) $\frac{1}{n}\sum_j\left(\frac{r_j(\boldsymbol{\beta})-r_i(\boldsymbol{\beta})}{\sqrt{2}\sigma}\right) w_{ij}\|\boldsymbol{x}_i-\boldsymbol{x}_j\|=O_p(1)$, for any i, and

(d) $\frac{1}{n}\sum_j w_{ij}\|\boldsymbol{x}_i-\boldsymbol{x}_j\|^2=O_p(1)$, for any i.

(A5) For any $1\le i\le n$,

$$\frac{1}{n}\sum_j\eta\left(\boldsymbol{x}_i,\boldsymbol{x}_j,\frac{r_j(\boldsymbol{\beta})-r_i(\boldsymbol{\beta})}{\sqrt{2}\sigma}\right)(\boldsymbol{x}_j-\boldsymbol{x}_i) \xrightarrow{p} \int\eta\left(\boldsymbol{x}_i,\boldsymbol{x}_j,\frac{r_j(\boldsymbol{\beta})-r_i(\boldsymbol{\beta})}{\sqrt{2}\sigma}\right)(\boldsymbol{x}_j-\boldsymbol{x}_i)dF(\boldsymbol{x}_j,y_j).$$

Theorem 3.1. Assume that (A1) ~ (A5) and $\binom{n}{2}^{-1}\sum\sum_{i<j}\|\boldsymbol{x}_j-\boldsymbol{x}_i\|=O_p(1)$ hold. Suppose $n^{1/2}(\hat{\boldsymbol{\beta}}_0-\boldsymbol{\beta})=O_p(1)$ and $n^{1/2}(\hat{\sigma}_0-\sigma)=O_p(1)$ with $\sigma>0$, then

$$n(2\hat{\sigma}_0^{\,2})^{-1}[\hat{H}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})]^T\hat{E}^{-1}[\hat{H}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})] \xrightarrow{d} \chi_p^2$$

where $\hat{H}$ is defined by (3.6) and

$$\hat{E}=4\frac{1}{n}\sum_i\left\{\frac{1}{n}\sum_j\eta\left(\boldsymbol{x}_i,\boldsymbol{x}_j,\frac{r_j(\widehat{\beta}_0)-r_i(\widehat{\beta}_0)}{\sqrt{2}\,\hat{\sigma}_0}\right)(\boldsymbol{x}_j-\boldsymbol{x}_i)\right\}\times\left\{\frac{1}{n}\sum_j\eta\left(\boldsymbol{x}_i,\boldsymbol{x}_j,\frac{r_j(\widehat{\beta}_0)-r_i(\widehat{\beta}_0)}{\sqrt{2}\,\hat{\sigma}_0}\right)(\boldsymbol{x}_j-\boldsymbol{x}_i)\right\}^T$$

Proof. We denote $r_i(\boldsymbol{\beta})=y_i-\boldsymbol{x}_i^T\boldsymbol{\beta}$ and $r_i(\widehat{\boldsymbol{\beta}}_0)=y_i-\boldsymbol{x}_i^T\widehat{\boldsymbol{\beta}}_0$ by ε_i and r_i, respectively, for notational simplicity.

We have, by Theorem 3.4 and Theorem 3.5 of Song and Kim (1997),

$$n^{1/2}(\sqrt{2}\,\hat{\sigma}_0)^{-1}\hat{H}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})\xrightarrow{d}N(\mathbf{0},E),$$

where

$$E=4\int\left\{\int\eta\left(\boldsymbol{x}_1,\boldsymbol{x}_2,\frac{\varepsilon_2-\varepsilon_1}{\sqrt{2}\sigma}\right)(\boldsymbol{x}_2-\boldsymbol{x}_1)dF(\boldsymbol{x}_2,y_2)\right\}\times\left\{\int\eta\left(\boldsymbol{x}_1,\boldsymbol{x}_2,\frac{\varepsilon_2-\varepsilon_1}{\sqrt{2}\sigma}\right)(\boldsymbol{x}_2-\boldsymbol{x}_1)dF(\boldsymbol{x}_2,y_2)\right\}^T dF(\boldsymbol{x}_1,y_1). \tag{3.9}$$

We now consider the asymptotic behaviour of $\hat{E}$. Let $d_i(\boldsymbol{\beta},\hat{\sigma}_0)=\sum_j\eta(\boldsymbol{x}_i,\boldsymbol{x}_j,(\varepsilon_j-\varepsilon_i)/\sqrt{2}\,\hat{\sigma}_0)(\boldsymbol{x}_j-\boldsymbol{x}_i)$. We obtain the following equations by taking a first-order Taylor series expansion about $\widehat{\boldsymbol{\beta}}_0$ and $\sqrt{2}\sigma$, respectively:

$$d_i(\boldsymbol{\beta},\hat{\sigma}_0)=\sum_j\eta\left(\boldsymbol{x}_i,\boldsymbol{x}_j,\frac{r_j-r_i}{\sqrt{2}\,\hat{\sigma}_0}\right)(\boldsymbol{x}_j-\boldsymbol{x}_i)+(\sqrt{2}\,\hat{\sigma}_0)^{-1}\sum_j\eta'\left(\boldsymbol{x}_i,\boldsymbol{x}_j,\frac{r_j-r_i}{\sqrt{2}\,\hat{\sigma}_0}\right)\times(\boldsymbol{x}_j-\boldsymbol{x}_i)(\boldsymbol{x}_j-\boldsymbol{x}_i)^T(\widehat{\boldsymbol{\beta}}_0-\boldsymbol{\beta})+O_p(n^{-1/2}), \tag{3.10}$$

$$d_i(\boldsymbol{\beta},\hat{\sigma}_0)=\sum_j\eta\left(\boldsymbol{x}_i,\boldsymbol{x}_j,\frac{\varepsilon_j-\varepsilon_i}{\sqrt{2}\sigma}\right)(\boldsymbol{x}_j-\boldsymbol{x}_i)-(\sqrt{2}\,\hat{\sigma}_0-\sqrt{2}\sigma)(\sqrt{2}\,\hat{\sigma}_0)^{-1}\times\sum_j\eta'\left(\boldsymbol{x}_i,\boldsymbol{x}_j,\frac{\varepsilon_j-\varepsilon_i}{\sqrt{2}\sigma}\right)\left(\frac{\varepsilon_j-\varepsilon_i}{\sqrt{2}\sigma}\right)(\boldsymbol{x}_j-\boldsymbol{x}_i)+O_p(n^{-1/2}). \tag{3.11}$$

Equating (3.10) and (3.11), under the assumptions on $\widehat{\boldsymbol{\beta}}_0$, $\hat{\sigma}_0$, (A1), and (A4), after some simplification, $\hat{E}$ can be represented as follows.

$$\hat{E}=4\frac{1}{n}\sum_i\left\{\frac{1}{n}\sum_j\eta\left(\boldsymbol{x}_i,\boldsymbol{x}_j,\frac{\varepsilon_j-\varepsilon_i}{\sqrt{2}\sigma}\right)(\boldsymbol{x}_j-\boldsymbol{x}_i)\right\}\times\left\{\frac{1}{n}\sum_j\eta\left(\boldsymbol{x}_i,\boldsymbol{x}_j,\frac{\varepsilon_j-\varepsilon_i}{\sqrt{2}\sigma}\right)(\boldsymbol{x}_j-\boldsymbol{x}_i)\right\}^T+O_p(n^{-1/2})$$

Hence by (A5), $\hat{E}\xrightarrow{p}E$ and the theorem follows. ■

In the case of simple linear regression, the pairwise GM-estimator $\hat{\beta}$ of the slope parameter has asymptotically a normal distribution. As a corollary of Theorem 3.1, we have the following results.

Corollary 3.1. Assume that the conditions in Theorem 3.1 are satistified. Then, in the case of simple regression, we have the following asymptotic normality:

$$\sqrt{n}(\hat{\beta}-\beta)\,/\,(2\hat{\sigma}_0^{\,2}\hat{\Sigma}) \xrightarrow{d} N(0,1)$$

where $\hat{\Sigma}=\hat{E}/\hat{H}^2$ with

$$\hat{E}=4\frac{1}{n}\sum_{i=1}^{n}\left\{\frac{1}{n}\sum_{j=1}^{n}\eta\left(x_i, x_j, \frac{r_j(\hat{\beta}_0)-r_i(\hat{\beta}_0)}{\sqrt{2}\,\hat{\sigma}_0}\right)(x_j-x_i)\right\}^2,$$

$$\hat{H}=\frac{1}{n}\sum_{i=1}^{n}\left\{\frac{1}{n}\sum_{j=1}^{n}\hat{w}_{ij}\psi'\left(\frac{r_j(\hat{\beta}_0)-r_i(\hat{\beta}_0)}{\sqrt{2}\,\hat{\sigma}_0}\right)(x_j-x_i)^2\right\}.$$

3.2 The Proposed Test Statistic

From Corollary 3.1, we can construct a test statistic for testing the the parallelism of k regression lines (1.1). The proposed test statistic is defined as follows.

$$Z_0=\frac{\sum_{i=1}^{k}C_i\hat{\beta}_i}{\left(2\sum_{i=1}^{k}\hat{\sigma}_{0i}^{\,2}\,C_i^2\hat{\Sigma}_i/n_i\right)^{1/2}}, \tag{3.12}$$

where $C_i's$ are scores to be determined satisfying $\sum_{i=1}^{n}C_i=0$, $\hat{\sigma}_{0i}$ is the scale estimator for the ith line based on the initial estimator $\hat{\beta}_0$, and $\hat{\Sigma}_i=\hat{E}_i\,/\,\hat{H}_i^{\,2}$ with

$$\hat{E}_i=4\frac{1}{n_i}\sum_{j=1}^{n_i}\left\{\frac{1}{n_i}\sum_{k=1}^{n_i}\eta\left(x_{ij}, x_{ik}, \frac{r_k(\hat{\beta}_{0i})-r_j(\hat{\beta}_{0i})}{\sqrt{2}\,\hat{\sigma}_{0i}}\right)(x_{ik}-x_{ij})\right\}^2$$

and

$$\hat{H}_i=\frac{1}{n_i}\sum_{j=1}^{n_i}\left\{\frac{1}{n_i}\sum_{k=1}^{n_i}\hat{w}_{jk}\psi'\left(\frac{r_k(\hat{\beta}_{0i})-r_j(\hat{\beta}_{0i})}{\sqrt{2}\,\hat{\sigma}_{0i}}\right)(x_{ik}-x_{ij})^2\right\}.$$

Here $\hat{\beta}_{0i}$, $\hat{\sigma}_{0i}$, and $\hat{\beta}_i$ are an initial estimator of β_i, a robust estimator of σ, and the pairwise GM-estimator of β_i in the ith line, respectively. Since Z_0 has asymptotically a normal distribution, we reject H_0 in favour of H_1 if

$$Z_0\ge z_\alpha,$$

where z_α is an upper α-quantile of the standard normal distribution.

The scores must be chosen to maximize the asymptotic power of the test. The power of Z_0 against any given $\beta_1\le\cdots\le\beta_k$ is given by

$$\Phi\left(\sum_{i=1}^{k} C_i\beta_i / (2\sigma^2 \sum_{i=1}^{k} C_i^2\Sigma_i / n_i)^{1/2} - z_\alpha\right),$$

where Φ is the standard normal distribution function and $\Sigma_i = E_i / H_i^2$. Here H_i and E_i are the scalar version of H in (3.8) and E in (3.9), respectively, in the ith line. Thus, the optimum scores, which can be derived as Adichie (1976), are

$$C_i = n_i \Sigma_i^{-1}(\beta_i - \sum_{j=1}^{k} t_j\beta_j), \tag{3.13}$$

where $t_i = n_i\Sigma_i^{-1}(\sum_{j=1}^{k} n_j\Sigma_j^{-1})^{-1}$. These scores are similar to those in (2.3). Therefore, we propose to use

$$C_i = m_i^2(D_i - \sum_{j=1}^{k} \rho_j D_j), \tag{3.14}$$

where

$$m_i^2 = n_i \times med_j\{|x_{ij} - med_j(x_{ij})|\}^2,$$
$$D_i = m_1^2 + \cdots + m_{i-1}^2 + (m_i^2/2),$$

and

$$\rho_j = m_j^2 / \sum_{i=1}^{k} m_i^2.$$

Note that m_i^2 is a robust version of w_i^2 and has a high breakdown point for design points.

4. A Small-Sample Monte Carlo Study

In this section we compare the small-sample behaviors of the proposed test Z_0 and the Adichie's tests. The measures of performances evaluated are empirical powers and significance levels.

4.1 A Robust Scale Estimator

To use Z_0 in (3.12) as a test statistic, we have to choose a robust scale estimator $\hat{\sigma}$. Computing robust GM-estimators of regression parameters, the scale estimator based on the median absolute deviation, $\hat{\sigma}_0 = 1.4826 \times MAD\{r_i(\hat{\beta}_0)\}$, is most widely used. But, it is well known that for any initial estimator $\hat{\beta}_0$, MAD usually underestimates the scale in normal case. Thus, for example, Rousseeuw and Leroy (1987) considered

$$\hat{\hat{\sigma}} = 1.4826\left(1 + \frac{5}{n-2}\right)\sqrt{med\{r_i^2(\hat{\hat{\beta}}_0)\}},$$

where $\widehat{\widehat{\beta}}_0$ is the least median of squares (LMS) estimator, and used the standardized residual $r_i/\widehat{\widehat{\sigma}}$ to identify outliers.

Using the LTS estimator as an initial estimator, we compare the following three scale estimators through a Monte Carlo study.

i) $\hat{\sigma}=1.4826\ MAD\{r_i\}$, (4.1)

where $r_i=y_i - \hat{\alpha}_0 - \hat{\beta}_0 x_i$ and $(\hat{\alpha}_0, \hat{\beta}_0)$ is the LTS estimator of (α, β).

ii) $\hat{\sigma}^*=1.4826(1+4/(n-1))\times MAD\{r_i\}$. (4.2)

iii) $\tilde{\sigma}=\left(\frac{\sum a_i r_i^2}{\sum a_i - 1}\right)^{1/2}$, with $a_i=\begin{cases}1, & \text{if } |r_i|/\hat{\sigma}^* < 2.5 \\ 0, & \textit{otherwise.}\end{cases}$ (4.3)

Remark 2. i) The correction factor $(1+4/(n-1))$ in the definition of $\hat{\sigma}^*$ is obtained through a Monte Carlo simulation for various distributions as in Rousseeuw and Leroy (1987). ii) All estimators have a 50% breakdown point.

To evaluate small-sample properties of the three scale estimators we performed a Monte Carlo simulation study in simple regression model $y_i=\alpha+\beta x_i+\varepsilon_i$, $i=1,\cdots,n$. In all cases, we set $x_i=i$. For simplicity, we let $\alpha=0$, $\beta=1$. The number of observations are $n=10$, 20, 30, and 100. In each case 1,000 trials are performed. The error term ε_i's are randomly generated from normal. The performance measures evaluated are empirical means (MEAN) and mean squared errors (MSE).

Table 4.1 shows the results of the empirical study to compare the performances of the three scale estimators in (4.1) ~ (4.3). In the case of $n=10$, the estimators tend to underestimate

Table 4.1. Empirical MEAN and MSE of estimators of σ

	$\hat{\sigma}$	$\hat{\sigma}^*$	$\tilde{\sigma}$
	Empirical MEAN		
n= 10	0.6672	0.9637	0.8737
n= 20	0.8505	1.0296	0.9789
n= 30	0.8911	1.0140	0.9839
n=100	0.9706	1.0098	0.9807
	Empirical MSE		
n= 10	0.4540	0.4476	0.4155
n= 20	0.2848	0.2950	0.2517
n= 30	0.2238	0.2230	0.1967
n=100	0.1214	0.1230	0.1049

the scale. As the sample size increases, the bias is reduced. However, for $\hat{\sigma}$ the bias is still significant for moderately large samples. $\tilde{\sigma}$ and $\hat{\sigma}^*$ have better performances than $\hat{\sigma}$ in MEAN. In terms of MSE $\tilde{\sigma}$ is better than $\hat{\sigma}$ and $\hat{\sigma}^*$. $\hat{\sigma}$ and $\hat{\sigma}^*$ are very similar in MSE. In terms of bias, $\hat{\sigma}^*$ is the best among the three estimators. We thus choose $\hat{\sigma}^*$ as an estimator of the scale parameter in defining the test statistic Z_0.

4.2 Results of a Monte Carlo Study

We now consider the problem of testing the parallelism of three regression lines against ordered alternatives. We thus consider the case of $k=3$. The sample size of each line is 20. We set $\alpha_i=0$ for each i. To construct ordered alternatives, we set the equally-spaced slopes given by

$$\beta_i=\beta_0+(i-1)m\delta, \quad i=1,2,3,$$

where δ is the standard deviation of the MLE of β from the combined sample, and $m=0,1,2,3$. The increment of the values of m indicates the change of slopes from the null parameters space to divergent alternatives. β_0 is set to be 1.

To compare the small-sample behavior of the proposed tests, we consider the following four situations:

Case 1) No leverage points and no outliers.
Case 2) No leverage points but with some outliers only in y-direction.
Case 3) Some bad leverage points, i.e., outliers both in x- and y-directions.
Case 4) Some bad leverage points and some good leverage points.

For Case 1, the design point x's are fixed with $(1, 2, \cdots, 20)$ and the error term ε_i's are randomly generated from the standard normal $N(0,1)$.

For Case 2, the x's are fixed as in Case 1. But the error term ε_i's are randomly generated from the double exponential and two contaminated normals : $CN(0.1,3)$ and $CN(0.2,5)$, where the distribution function of $CN(\varepsilon,\sigma)$ is given by $F(x)=(1-\varepsilon)\Phi(x)+\varepsilon\Phi(x/\sigma)$.

For Case 3, we first generate the same x's and y's as in Case 1. To make some bad leverage points, we choose two points randomly in each line, and replace these (x_{ij}, y_{ij}) by $(30+x_{ij}/20, y_{ij})$.

For Case 4, first we generate the same x's and ε_i's as in Case 1. To make some good leverage points, after replacing two randomly selected x_{ij} by $(30+x_{ij}/20)$, we obtain y_{ij}

by (1.1). To make some bad leverage points, we choose two points randomly from the remaining samples not selected as good leverage points, and replace these (x_{ij}, y_{ij}) by $(30 + x_{ij}/20, y_{ij})$.

In computing the one-step GM-estimator, Huber's ψ with the tuning constant $c=1.5$ is used. To compute the initial estimator $\widehat{\beta}_0$, the LTS estimator is used. We have to calculate the weights in (3.7). In simple linear regression model, the leverageness $v(x_{lk})$ of the jth observation in the ith line can be written as

$$v(x_{ij}) = \min\left\{1, \left(\frac{b}{(x_{ij} - m_{x_i})\, C_{x_i}^{-1} (x_{ij} - m_{x_i})} \right)^{r/2} \right\}, \tag{4.4}$$

where $m_{x_i} = med_j\{x_{ij}\}$ and $C_{x_i} = (1.4826 MAD_j\{x_{ij}\})^2$ (Naranjo and Hettmansperger, 1994). $r=2$ and $b=\chi^2(1, 0.975)$, and $\widehat{w}_{ij}$ with $a=2.7$ in (3.7) are also used.

The simulation was performed on a personal computer with a Pentium 150 MHz processor by using S-PLUS (Ver. 3.2 Release 1 for MS Windows 3.1 : 1994). The normal and uniform variates were generated by the S-PLUS function *rnorm* and *runif*, respectively. And LTS estimates were calculated by S-PLUS functions *lsfit*.

1,000 trials for each experiment are performed. For each setting, the values of the test statistics are calculated and compared with their respective critical values at significance levels of $\alpha=0.05$ and $\alpha=0.10$. The results are summarized in Table 4.2. The empirical powers at $m=0$ indicate the empirical significance levels.

By observing Table 4.2, we find that the proposed test has stable empirical levels in various situations. As expected, the parametric test S_t performs well with respect to empirical level and power in the normal case. However, the parametric test S_t has unstable empirical levels in the presence of leverage points. For medium tailed distributions such as normal, double exponential, and $CN(0.1, 3)$ distributions, the empirical powers of the proposed test are competitive to those of others. Moreover, the proposed test is better in power than other competitors for heavy tailed distribution such as $CN(0.2, 5)$. On the other hand, in the presence of leverage points, other competitors have unstable empirical levels, that is, very sensitive to high leverage points, but the proposed test has still stable empirical level and good performances. Therefore, the proposed test is considerably robust to leverage points as well as substantial outliers for testing the parallelism of regression lines.

Table 4.2. Empirical Powers (Replication=1,000)

m	S_t	S_R	Z_0
Case 1 : No leverage points, $\varepsilon \sim N(0,1)$ ($\delta = 0.02$)			
	$\alpha = 0.05$		
0	0.049	0.045	0.056
1	0.174	0.155	0.150
2	0.408	0.389	0.343
3	0.675	0.645	0.582
	$\alpha = 0.10$		
0	0.092	0.096	0.092
1	0.283	0.271	0.267
2	0.555	0.536	0.482
3	0.786	0.779	0.709
Case 2 : No leverage points, $\varepsilon \sim DE(0,1)$ ($\delta = 0.02$)			
	$\alpha = 0.05$		
0	0.041	0.046	0.051
1	0.151	0.159	0.179
2	0.294	0.337	0.356
3	0.473	0.533	0.560
	$\alpha = 0.10$		
0	0.086	0.096	0.106
1	0.245	0.257	0.265
2	0.420	0.469	0.486
3	0.617	0.690	0.692
Case 2 : No leverage points, $\varepsilon \sim CN(0.1,(0,3))$ ($\delta = 0.025$)			
	$\alpha = 0.05$		
0	0.050	0.042	0.045
1	0.150	0.177	0.167
2	0.398	0.444	0.429
3	0.665	0.724	0.687
	$\alpha = 0.10$		
0	0.101	0.091	0.095
1	0.259	0.276	0.274
2	0.547	0.593	0.564
3	0.781	0.843	0.804

Table 4.2. (Continued)

m	S_t	S_R	Z_0
Case 2 : No leverage points, $\varepsilon \sim CN(0.2, (0,5))$ ($\delta = 0.03$)			
	$\alpha = 0.05$		
0	0.051	0.042	0.054
1	0.137	0.172	0.196
2	0.247	0.424	0.461
3	0.419	0.691	0.734
	$\alpha = 0.10$		
0	0.098	0.100	0.103
1	0.216	0.285	0.300
2	0.370	0.581	0.586
3	0.562	0.818	0.824
Case 3 : 10% Bad leverage points, ($\delta = 0.03$)			
	$\alpha = 0.05$		
0	0.153	0.068	0.050
1	0.194	0.141	0.218
2	0.267	0.308	0.563
3	0.262	0.449	0.828
	$\alpha = 0.10$		
0	0.212	0.118	0.092
1	0.243	0.223	0.337
2	0.330	0.434	0.694
3	0.319	0.583	0.916
Case 4 : 10% Bad and 10% Good leverage points, ($\delta = 0.02$)			
	$\alpha = 0.05$		
0	0.150	0.073	0.048
1	0.182	0.189	0.183
2	0.227	0.348	0.469
3	0.222	0.517	0.706
	$\alpha = 0.10$		
0	0.213	0.133	0.100
1	0.258	0.287	0.301
2	0.291	0.472	0.601
3	0.297	0.614	0.827

References

[1] Adichie, J.N. (1976). Testing Parallelism of Regression Lines against Ordered Alternatives, *Communications in Statistics - Theory and Methods*, A5 (11), 985-997.

[2] Hollander, M. (1970). A Distribution-Free Test for Parallelism, *Journal of the American Statistical Association*, 65, 387-394.

[3] Jee, E.S. (1989). *A Nonparametric Test for the Parallelism of Regression Lines against Ordered Alternatives*, Ph.D. Dissertation, Department of Computer Science and Statistics, Seoul National University, Seoul, Korea.

[4] Lee, K.H. (1990). *On a Nonparametric Test for the Parallelism of Several Regression Lines against Ordered Alternatives*, Ph.D. Dissertation, Department of Computer Science and Statistics, Seoul National University, Seoul, Korea.

[5] Naranjo, J. D. and Hettmansperger, T.P. (1994). Bounded Influence Rank Regression, *Journal of the Royal Statistical Society, Series B*, 56(1), 209-220.

[6] Potthoff, R. F. (1974). A Nonparametric Test of Whether Two Simple Regression Lines are Parallel, *The Annals of Statistics*, 2, 295-310.

[7] Rao, K.S.M. and Gore, A.P. (1984). Testing Concurrence and Parallelism of Several Sample Regressions against Ordered Alternatives, *Mathematische Operationsforschung und Statistik, Series Statistics*, 15, 43-50.

[8] Rousseeuw, P.J. (1984). Least Median of Squares Regression. *Journal of the American Statistical Association*, 79, 871-880.

[9] Rousseeuw, P.J. and Leroy, A.M. (1987). *Robust Regression and Outlier Detection*, John Wiley & Sons, New York.

[10] Simpson, D.G., Ruppert, D., and Carroll, R.J. (1992). On One-Step GM Estimates and Stability of Inferences in Linear Regression, *Journal of the American Statistical Association*, 87, 439-450.

[11] Sin, B.S. (1993). *A Study on Tests for the Parallelism of Several Regression Lines against Ordered Alternatives*, Ph.D. Dissertation, Department of Computer Science and Statistics, Seoul National University, Seoul, Korea.

[12] Song, M.S. and Kim, J.H. (1997). A Study on a One-Step Pairwise GM-Estimator in Linear Models, *Journal of the Korean Statistical Society*, 26, 1-22.

[13] Song, M.S., Kim, J.H., Jeon, J.W., and Park, C.S (1989). On a Nonparametric Test for Parallelism against Ordered Alternatives, *Journal of the Korean Society for Quality Control*, 17, 70-80.

[14] Statistical Sciences (1994). *S-PLUS for Windows User's Manual*, Statistical Sciences, Inc., Seattle.