MODIFIED PRINCIPAL COMPONENTS REGRESSION

by

Huan-Ter Wu

Dissertation submitted to the Graduate Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Statistics

APPROVED:

Dr. John W. White, Chairman

Dr. Jesse C. Arnold

Dr. Raymond H. Myers

Dr. Klaus Hinkelmann

Dr. Marion R. Reynolds, Jr.

August, 1979

Blacksburg, Virginia

ACKNOWLEDGMENTS

The author expresses his sincere appreciation to his major advisor, Dr. John W. White, for suggesting the area of research in this dissertation and for his patience and guidance during the course of this study. Also the author expresses his sincere gratitude to Dr. Raymond H. Myers for co-reading this manuscript and for his many helpful comments for improving the dissertation.

Sincere appreciation is extended to Dr. Jesse C. Arnold, Dr. Klaus H. Hinkelmann, and Dr. Marion R. Reynolds, Jr., for their many beneficial comments and suggestions and for serving on the dissertation committee.

The author is also grateful to for her assistance with the computer and for efficient typing of this manuscript.

The author wishes to thank his parents for their continuing support and encouragement throughout his education. Finally, the author would like to express his appreciation to his wife, , for her patience and understanding during this endeavor.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Multiple linear regression is a procedure for estimating the relationship between a set of independent variables and a dependent variable. Estimation of the regression coefficients of the independent variables is often performed by using least squares estimators. When certain assumptions are satisfied, the least squares estimator has minimum variance in the class of linear unbiased estimators of the regression coefficients. However, this property does not guarantee that its variance is small, particularly when near linear relationships exist among the independent variables, a situation referred to as multicollinearity.

In this situation of multicollinearity, the variance of the least squares estimators can be very large and hence the estimators tend to be unreliable. It is difficult to isolate and interpret confidently the separate effects of variables involved in multicollinearities. Alternative estimation procedures are often required to provide more reliable coefficient estimates under the condition of multicollinearity. Biased estimation procedures provide alternative methods.

One biased estimation procedure which has been proposed in the statistical literature is principal components regression (Massy, 1965)

(Mansfield, 1975). Principal components regression attempts to remove the damaging effects of multicollinearities on the least squares estimator by an orthogonal transformation of the data, and by deleting the components associated with small latent roots of X'X. This is done to obtain more stable estimates of the regression coefficients. The major disadvantage of the principal components regression is that it may result in loss of information when the deleted component has predictive value because components associated with small latent roots may be useful as predictors.

The objective of this paper is to present a compromise method, modified principal component regression, in which the components associated with small latent roots are dampened but are not completely deleted. It provides a reasonable compromise choice when the orientation of the vector of the regression coefficients and the latent vectors associated with small latent roots is unknown.

In Chapter II, we explore the effect of multicollinearities on the least squares estimates of the coefficients in multiple linear regression and present methods of detecting multicollinearities. We also discuss principal components regression, and other biased estimation procedures and point out some of their advantages and disadvantages. Modified principal components regression is discussed in Chapter III. Certain properties of the estimator are derived and discussed. In Chapter IV, we derive a t-statistic for the test of significance of the individual coefficients and the confidence interval for the individual coefficients using a modified principal components estimator. We also construct an F-statistic for testing any general linear hypothesis and the confidence region for the linear functions of the unknown regression coefficients based on a modified principal components estimator. In Chapter V, we present the results of a Monte Carlo simulation. The performance of the modified principal components estimator is evaluated by the mean squared error criterion and is compared to the principal components estimator and the least squares estimator. Summary and conclusions are presented in the final chapter.

CHAPTER II

MULTICOLLINEARITY PROBLEMS AND BIASED ESTIMATION PROCEDURES

2.1 Classical Linear Regression Model

The multiple linear regression model can be written as

$$y_i = \beta_0^* + \beta_1^* x_{1i}^* + \ldots + \beta_p^* x_{pi}^* + \varepsilon_i \qquad i = 1, 2, \ldots, n \qquad (2.1.1)$$

where y_i: the response for the ith observation (dependent variable).

x_{ji}^*: the value of jth regressor variable for the ith observation (independent variable).

β_j^*: unknown regression coefficients to be estimated, $j = 0, 1, 2, \ldots, p$.

ε_i: an unobservable random error.

In order to minimize roundoff error and to increase computational accuracy, we transform the regression model (2.1.1) by letting

$$x_{ji} = \frac{(x_{ji}^* - \bar{x}_j^*)}{\sqrt{\sum_{i=1}^{n} (x_{ji}^* - \bar{x}_j^*)^2}} \qquad \text{where } \bar{x}_j^* = \frac{1}{n} \sum_{i=1}^{n} x_{ji}^*$$

Thus, the equation (2.1.1) becomes

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_p x_{pi} + \varepsilon_i \qquad i = 1, 2, \ldots, n. \qquad (2.1.2)$$

where

$$\beta_0 = \beta_0^* + \sum_{j=1}^{p} \beta_j^* \bar{x}_j^*$$

$$\beta_j = \beta_j^* \sqrt{\sum_{i=1}^{n} (x_{ji}^* - \bar{x}_j^*)^2}$$

The n equations in (2.1.2) can be written in matrix notation as

$$\underline{y} = \beta_0 \underline{1} + X\underline{\beta} + \underline{\varepsilon} \qquad (2.1.3)$$

$\underline{y}$: a n×1 vector of observed responses.

X: a n×p matrix of known values of the regressor variables.

$X = [\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_p]$

$\underline{1}$: a n×1 vector of ones.

β_0: the intercept parameter whose value is unknown.

$\underline{\beta}$: a p×1 vector of unknown parameters to be estimated.

$\underline{\varepsilon}$: a n×1 unobservable random error vector.

In the equation (2.1.3), the regressor variables have been standardized so that $\underline{x}_j'\underline{1} = 0$, and $\underline{x}_j'\underline{x}_j = 1$ for j = 1, 2, ..., p. Thus, X'X will have the form of a correlation matrix.

To complete the specification of the regression model, we add the following basic assumptions:

(i) The elements of $\underline{\varepsilon}$ are unobservable errors with $E(\underline{\varepsilon}) = 0$, and $E(\underline{\varepsilon}\underline{\varepsilon}') = \sigma^2 I_n$.

(ii) Each of the regressor variables is nonstochastic with values fixed in repeated samples.

(iii) The matrix X has rank $p < n$.

The first assumption states that ε_i are variables with zero expectation and that they have constant variance σ^2. This property is referred to as homoscedasticity. Assumption (i) also implies that the ε_i are pairwise uncorrelated. Assumption (ii), that each of the regressor variables is nonstochastic, means that the only source of variation in the $\underline{y}$ vector is variation in the $\underline{\varepsilon}$ vector and the properties of estimators and tests are conditional upon X. Assumption (iii) states that the number of observations exceeds the number of parameters to be estimated and that no exact linear relationships exist among the x variables.

The following notation will be used throughout this paper:

$\ell_1 \leq \ell_2 \leq \dots \leq \ell_p$ are latent roots of X'X.

$\underline{V}_1, \underline{V}_2, \dots, \underline{V}_p$ are the corresponding orthonormal latent vectors of X'X.

$V = (\underline{V}_1, \underline{V}_2, \dots, \underline{V}_p)$ is a p×p matrix of unit length latent vectors of X'X, and $\Lambda = \text{diag}\ (\ell_1, \ell_2, \dots, \ell_p)$.

2.2 Least Squares Estimation

The relationship (2.1.3) with assumptions (i) through (iii) describes the classical linear regression model. Our problem is to obtain estimates of the unknown parameters in this model. The least squares estimators of the regression coefficients are given by

$$\hat{\beta}_0 = \bar{y}\ ,$$

and

$$\hat{\underline{\beta}} = (X'X)^{-1}\, X'\, \underline{y}\ . \tag{2.2.1}$$

The equation (2.2.1) can be written in another form as

$$\underline{\hat{\beta}} = \sum_{j=1}^{p} \ell_j^{-1} c_j \underline{V}_j \quad \text{where} \quad c_j = \underline{V}_j' X' \underline{y} \ . \tag{2.2.2}$$

Since $$V'(X'X)^{-1}V = \Lambda^{-1}$$

thus, $$(X'X)^{-1} = V\Lambda^{-1}V' = \sum_{j=1}^{p} \ell_j^{-1} \underline{V}_j \underline{V}_j' \ .$$

The variance-covariance matrix of the least squares estimator is given by

$$\text{Var } \underline{\hat{\beta}} = \sigma^2 (X'X)^{-1} = \sigma^2 C \tag{2.2.3}$$

where $C = (X'X)^{-1} = (c_{ij})$.

Thus, $\text{Var}(\hat{\beta}_j) = \sigma^2 c_{jj}$ and $\text{Cov}(\hat{\beta}_i, \hat{\beta}_j) = \sigma^2 c_{ij}$.

Equation (2.2.3) can also be written as

$$\text{Var } \underline{\hat{\beta}} = \sigma^2 \sum_{j=1}^{p} \ell_j^{-1} \underline{V}_j \underline{V}_j' \tag{2.2.4}$$

since $(X'X)^{-1} = \sum_{j=1}^{p} \ell_j^{-1} \underline{V}_j \underline{V}_j'$.

The mean squared error of the least squares estimator is

$$\begin{aligned} \text{MSE}(\underline{\hat{\beta}}) &= E(\underline{b} - E\underline{b})' (\underline{b} - E\underline{b}) \quad \text{where } \underline{b} = \underline{\hat{\beta}} \\ &= E \sum_{j=1}^{p} (b_j - Eb_j)^2 = \sum_{j=1}^{p} \text{Var } b_j \ , \quad \text{(where } b_j\text{'s are elements of } \underline{b}) \\ &= \text{tr Var } \underline{\hat{\beta}} = \sigma^2 \text{ tr } (X'X)^{-1} \\ &= \sigma^2 \sum_{j=1}^{p} \ell_j^{-1} \text{tr } \underline{V}_j \underline{V}_j' = \sigma^2 \sum_{j=1}^{p} \ell_j^{-1} \ . \end{aligned} \tag{2.2.5}$$

The least squares estimator of the regression coefficients is unbiased and has minimum variance among the class of unbiased linear estimators (Graybill, 1976). However, the restriction to unbiased estimators can result in a very large variance when the matrix of independent variables is ill-conditioned.

2.3 Effects of Multicollinearity

Multicollinearity can be defined as the presence of a near linear dependence among the set of column vectors $\underline{x}_j$, of the matrix X; that is, multicollinearity exists if

$$\sum_{j=1}^{p} a_j \underline{x}_j \simeq 0 \ . \qquad (2.3.1)$$

for some set of constant a_1, a_2, a_3, ..., a_p, not all zero (Mason, Gunst, and Webster, 1975).

One of the assumptions underlying the classical linear regression model is that matrix X has rank $p < n$; that is, no exact linear dependence exists among the regressor variables and the number of observations exceed the number of parameters to be estimated. When some of the regressor variables are perfectly correlated with any other regressor variables, the classical least squares estimates are not unique. This describes an exact multicollinearity.

Exact multicollinearity is not very frequent in practical application. The multicollinearity problems of greatest concern arise from situations where near linear relationship exists between two or more regressor variables but not a perfect linear relationship; that is, the cases where (2.3.1) is approximately zero are our concern in this paper.

Johnston (1972) clearly presents the main consequences of multicollinearity as:

1. The precision of estimation falls so that it becomes very difficult, if not impossible, to disentangle the relative influences of the various x variables. This loss of precision has three aspects: specific estimates may have very large errors; these errors may be highly correlated, one with another; and the sampling variances of the coefficients will be very large.

2. Investigators are sometimes led to drop variables incorrectly from an analysis because their coefficients are not significantly different from zero, but the true situation may be not that a variable has no effect but simply that the set of sample data has not enabled us to pick it up.

3. Estimates of coefficients become very sensitive to particular sets of sample data, and the addition of a few more observations can sometimes produce dramatic shifts in some of the coefficients.

These problems which arise from the presence of multicollinearities can be demonstrated as follows:

Strong multicollinearities among the regressor variables cause $X'X$ to be nearly singular, this implies that the determinant of $(X'X)$ is very small, so that elements of $(X'X)^{-1}$, and, therefore, some of the variances and covariances of the estimated regression coefficients in the equation (2.2.3) are very large.

To illustrate this further, we can write the diagonal elements of $C = (X'X)^{-1}$ as

$$c_{jj} = (1 - R_j^2)^{-1} \qquad j = 1, 2, \ldots, p \qquad (2.3.2)$$

(Mason, Gunst, and Webster, 1975)

where R_j^2 is the coefficient of determination of the least squares regression of x_j on the remaining p-1 regressor variables. Suppose a single linear relationship of the form (2.3.1) holds approximately among the first $s \leq p$ regressor variables. If x_j is involved in the multicollinearity, then R_j^2 is very close to one and c_{jj} will be very large. Since $\mathrm{Var}(\hat{\beta}_j) = \sigma^2 c_{jj}$, the variance of the estimator of regression coefficients of x_j is also very large. Mason, Gunst, and Webster (1975) also demonstrated that if both x_i and x_j are involved in the multicollinearity, c_{ij} $(i \neq j)$ will tend to be large in absolute value.

$$\text{Since} \qquad \underline{\hat{\beta}} = (X'X)^{-1}X'\underline{y} = C\,X'\underline{y}$$

$$\text{and} \qquad \hat{\beta}_j = \sum_{i=1}^{p} c_{ji}(\underline{x}_i'\underline{y}) \qquad j = 1, 2, \ldots, p$$

where $\hat{\beta}_j$ is the jth element of $\underline{\hat{\beta}}$, we see that if x_j is one of the variables involved in the multicollinearity, $\hat{\beta}_j$ will tend to be very large in absolute value. This is due primarily to the relationship between the regressor variables, and not to the relationship between the response variable and the regressor variables. Variables with large estimates of coefficients are not necessarily strong in predictive ability. The signs of the estimates of the regression coefficients are not necessarily indicative of a regressor variable having a positive or negative effect on the dependent variable. This phenomenon can be explained by the fact that the elements of $\underline{\hat{\beta}}$ corresponding to regressors involved in multicollinearities tend to be dominated by the first few terms of equation (2.2.2). For example, if X'X contains one strong multicollinearity, the smallest latent root of X'X becomes very small so that ℓ_1^{-1} becomes several times larger than the remaining ℓ_j^{-1}. Therefore, the values of $\hat{\beta}_j$, the elements of $\underline{\hat{\beta}}$ in equation (2.2.2), corresponding to regressors involved in multicollinearity are approximately equal to the corresponding elements of the first term of $\underline{\hat{\beta}}$ in equation (2.2.2) (Gunst and Mason, 1977).

2.4 Detecting the Presence of the Multicollinearity

Several techniques for detecting the presence of multicollinearities have been proposed in the statistical literature (Mason, Webster, and Gunst, 1975). Two of the most useful methods will be discussed here.

1. Examine the diagonal elements of $(X'X)^{-1}$.

As mentioned earlier, if a near linear relationship exists among x_j and a subset of the remaining columns of X, c_{jj} in equation (2.3.2) will become very large. Large values of some of the c_{jj} indicate that the variables corresponding to them are involved in a multicollinearity. However, this method gives no information about the nature of the dependence relationship.

The variance inflation factors (V.I.F.), defined by Marquardt (1970), are the diagonal elements of $(X'X)^{-1}$, when the X matrix is scaled so that X'X is in the form of a correlation matrix. The variance inflation factor associated with each coefficient represents the amount by which the variance of that coefficient is inflated by the correlation between variables.

From equation (2.3.2), the variance inflation factor of x_j can be written as

$$V.I.F.(x_j) = c_{jj} = (1 - R_j^2)^{-1} \qquad j = 1, 2, \ldots, p. \quad (2.4.1)$$

If x_j is involved in a multicollinearity, R_j^2 can be close to one and the corresponding variance inflation factor will be very large. If $R_j^2 = 0.9$, V.I.F.$(x_j) = 10$ and hence variance inflation factors of 10 or more are considered an indication of the presence of one or more strong multicollinearities (Marquardt (1970)).

2. Examine the smallest latent roots and the corresponding latent vector of X'X.

A small latent root of X'X indicates the near singularity of the X'X matrix. The number of small latent roots indicates the number of linear relationships present. The closer the smallest latent root is to zero, the stronger is the linear dependency among the columns of the X matrix. The elements of the corresponding latent vector of X'X are the coefficients of these linear relationship, the large elements indicating which variables are involved in the multicollinearity (Mason, Gunst, and Webster, 1975).

No precise rules have been set up to determine what is a small latent root since this decision depends on such factors as uniformity of all latent roots and the number of regressor variables. However, Webster, Gunst, and Mason (1974) indicated that latent roots of size 0.05 or smaller are reliable indicators of a near singularity.

Marquardt and Snee (1975) suggested that the variance inflation factor can be used as a measure of how close the smallest latent root is to zero.

2.5 Biased Estimation: Alternative Procedures to Least Squares Estimation

Since least squares estimators of regression coefficients may be unreliable when multicollinearities exist, we need alternative methods of estimation that enable us to remove the effects of the multicollinearity. Often the large variance of the unbiased least squares estimates can be greatly reduced by the addition of a small amount of bias, resulting in estimators with smaller mean squared error. Several alternative procedures to least squares estimation of regression coefficients have been proposed in the statistical literature.

Among these proposed alternatives are ridge regression, principal components regression, and latent root regression. All of these are biased estimation procedures which are useful when multicollinearity is present in the data. Some of the properties of each alternative procedure will also be discussed.

2.5.1 Ridge Regression

One form of biased estimation of the regression parameters is ridge regression proposed by Hoerl and Kennard (1970a, b). A motivating factor for ridge regression is the fact that the least squares estimates of regression coefficients tend to be overestimates of $\underline{\beta}$ when multicollinearities exist among the regressor variables.

Therefore, the Euclidean distance from $\hat{\underline{\beta}}$ to $\underline{\beta}$ will tend to be large. This can easily be seen as follows:

Let L_1 = distance from $\hat{\underline{\beta}}$ to $\underline{\beta}$.

$$L_1^2 = (\hat{\underline{\beta}} - \underline{\beta})'(\hat{\underline{\beta}} - \underline{\beta})$$

$$E(L_1^2) = \sigma^2 \sum_{j=1}^{p} \ell_j^{-1} = MSE(\hat{\underline{\beta}})$$

where $\ell_1 \leq \ell_2 \leq \cdots \leq \ell_p$ are latent roots of X'X (Hoerl and Kennard, 1970a).

If strong multicollinearities are present, one or more of the latent roots of X'X will be small, and the expected value of squared distance between $\hat{\underline{\beta}}$ and $\underline{\beta}$ becomes large. Ridge regression seeks a reduction in the length of the vector of estimated coefficients by adding small quantities to the diagonal elements of X'X. The ridge regression estimator can be written as

$$\hat{\underline{\beta}}_R = (X'X + kI)^{-1}X'\underline{y} \qquad (2.5.1)$$

where k is a positive constant (Hoerl and Kennard, 1970a). With the appropriate choice of k, the average value of the squared distance between the vector of estimated coefficients and the vector of true coefficient, $\underline{\beta}$, can be reduced.

Many properties of the ridge regression estimator have been derived by Hoerl and Kennard (1970a); we will discuss some of the important properties here.

The ridge regression estimator, $\hat{\underline{\beta}}_R$, is a linear transformation of the least squares estimator; the transformation depends only on X and k.

$$\hat{\underline{\beta}}_R = G_1\hat{\underline{\beta}},$$

where $\quad G_1 = (X'X + kI)^{-1} X'X = (I + k(X'X)^{-1})^{-1}$.

It follows that the ridge estimator is a biased estimator of $\underline{\beta}$, for $k \neq 0$.

$$E(\hat{\underline{\beta}}_R) = G_1\underline{\beta}, \text{ where } G_1 = (I + k(X'X)^{-1})^{-1} \ . \qquad (2.5.2)$$

The variance of the ridge regression estimator is given by

$$\mathrm{Var}(\hat{\underline{\beta}}_R) = \sigma^2(X'X + kI)^{-1}X'X(X'X + kI)^{-1} \ . \qquad (2.5.3)$$

The mean squared error of the ridge regression estimator is given by the following, where b_j^R are the elements of $\hat{\underline{\beta}}_R$,

$$\begin{aligned} \mathrm{MSE}(\hat{\underline{\beta}}_R) &= E(\hat{\underline{\beta}}_R-\underline{\beta})'(\hat{\underline{\beta}}_R-\underline{\beta}) \\ &= \sum_{j=1}^{p} \mathrm{Var}\ b_j^R + \sum_{j=1}^{p} (\text{bias of } b_j^R)^2 \\ &= \sigma^2 \sum_{j=1}^{p} \ell_j(\ell_j+k)^{-2} + k^2\underline{\beta}'(X'X+kI)^{-2}\underline{\beta} \ . \qquad (2.5.4) \end{aligned}$$

The variance term is a monotonic decreasing function of k. The bias term is a monotonic increasing function of k. The limiting value of the squared bias is $\underline{\beta}'\underline{\beta}$ as $k \to \infty$ (Hoerl and Kennard, 1970a).

Hoerl and Kennard (1970a) also demonstrated that there exists a range of values of k for which the ridge regression estimator has smaller mean squared error than the least squares estimator. Unfortunately, the range depends upon the unknown parameters, $\underline{\beta}$ and σ^2.

Consequently, Hoerl and Kennard recommended that k be determined from a ridge trace.

Hoerl and Kennard (1970b),defined the ridge trace as a two-dimensional plot of the b_j^R, j = 1, 2, ..., p, and the residual sum of squares versus k. To illustrate the procedure, an example with two variables is given as follows:

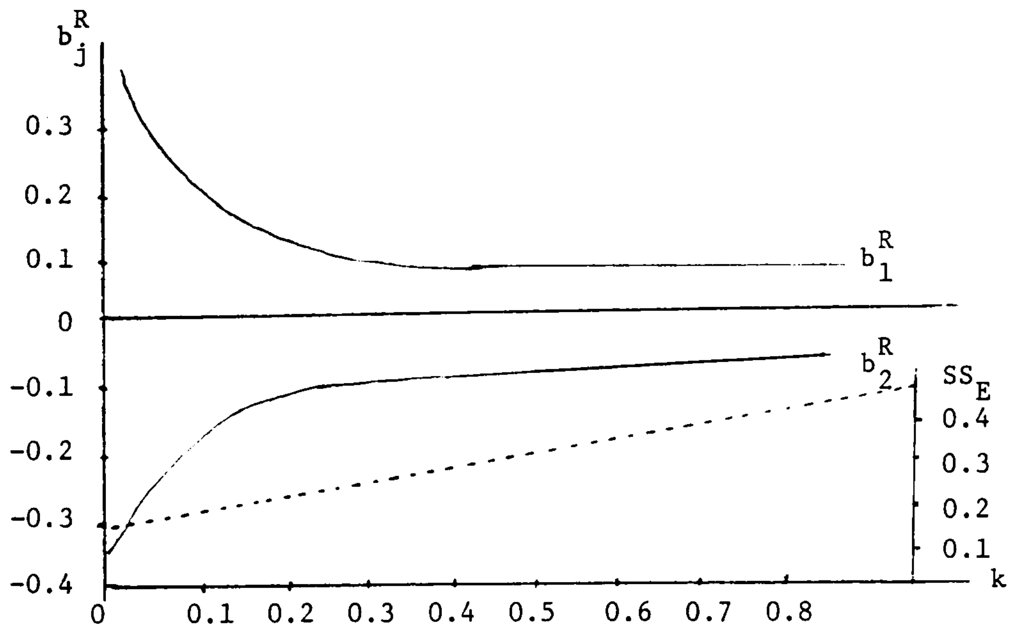

Fig. 2.5.1 Ridge Trace, Two Factor Example

Upon examination of the ridge trace in Figure 2.5.1 above, we can see that the system "stabilizes" at a value of k in the interval (0.2, 0.4). Therefore, we could choose k = 0.2 as a stable point solution since at this point the ridge trace indicates that the system is reasonably stable and there is not a large increase in the bias or residual sum of squares.

Since k is estimated from the data, the mean, variance, and mean squared error properties of the ridge regression estimator as derived by Hoerl and Kennard are no longer valid. The properties of the ridge regression estimator when k is estimated from the data are unknown.

Hoerl and Kennard (1970 a, b) also give a general form of the ridge regression estimator

$$\hat{\underline{\beta}}_R = (X'X + K^*)^{-1}X'\underline{y}$$

where $K^* = \text{diag}(k_1, k_2, \ldots, k_p)$. The ridge regression estimator in equation (2.5.1) is a special case of this general form when k_j are equal for all j.

A modification of ridge regression, partitioned ridge regression, in which some of β_j's are estimated by least squares estimators and other β_j's by ridge regression estimators, discussed by Farebrother (1978), is based on

$$\hat{\underline{\beta}}_R = (X'X + kA)^{-1}X'\underline{y} \qquad k > 0 \qquad (2.5.5)$$

where

$$A = \begin{pmatrix} 0 & 0 \\ 0 & A_{22} \end{pmatrix}$$

and A_{22} is a $p_2 \times p_2$ positive semidefinite matrix. We partition the matrix X into $X = (X_1, X_2)$, where X_2 has p_2 columns and X_1 has $p_1 = p - p_2$ columns, and partition $\hat{\underline{\beta}}_R$ similarly,

$$\hat{\underline{\beta}}_R = \begin{pmatrix} \underline{b}_1^R \\ \underline{b}_2^R \end{pmatrix} .$$

Thus, equation (2.5.5) becomes

$$\begin{pmatrix} \underline{b}_1^R \\ \underline{b}_2^R \end{pmatrix} = \begin{pmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2+kA_{22} \end{pmatrix}^{-1} \begin{pmatrix} X_1'\underline{y} \\ X_2'\underline{y} \end{pmatrix}$$

then,

$$\underline{b}_2^R = [X_2'M_1X_2 + kA_{22}]^{-1}X_2'M_1\underline{y}$$

$$\underline{b}_1^R = (X_1'X_1)^{-1}X_1' \ (\underline{y} - X_2\underline{b}_2^R)$$

where $M_1 = I - X_1(X_1'X_1)^{-1}X_1'$.

Most of the discussion of ridge regression in the literature centers around the choice of the optimal values of k in the ridge regression estimator. Hoerl, Kennard, and Baldwin (1975) suggested that a reasonable choice for k is an estimate of $\frac{p\sigma^2}{\underline{\beta}'\underline{\beta}}$ since MSE($\hat{\underline{\beta}}_R$) is minimized if $X'X = I$ and $k = \frac{p\sigma^2}{\underline{\beta}'\underline{\beta}}$. They present the results of a simulation study and conclude that the ridge estimator with this choice of k has a probability greater than 0.5 of producing estimates with a smaller mean squared error than the least squares estimator. Hoerl and Kennard (1976) suggest an iterative method for selecting the biasing parameter k in the ridge regression estimator; that is, they consider a sequence of estimates of $\underline{\beta}$ and k as:

$$\underline{\hat{\beta}},\ k_0 = \frac{p\hat{\sigma}^2}{\underline{\hat{\beta}}'\underline{\hat{\beta}}},\ \underline{\hat{\beta}}_R(k_0),\ k_1 = \frac{p\hat{\sigma}^2}{[\underline{\hat{\beta}}_R(k_0)]'[\underline{\hat{\beta}}_R(k_0)]},\ \underline{\hat{\beta}}_R(k_1),\ \ldots\ .$$

The sequence is stopped if $\frac{k_{i+1}-k_i}{k_i} < \delta$, where δ is a very small positive constant. They conclude that the method has a smaller estimated mean squared error than the least squares estimator or ridge regression estimator with single iteration estimate of $k = \frac{p\hat{\sigma}}{\underline{\hat{\beta}}'\underline{\hat{\beta}}}$. Hemmerle (1974) obtained a non-iterative, closed form solution for the estimated k, eliminating the need for the iterative procedure.

Coniffe and Stone (1973) examined the concept of ridge regression and gave some critical comments on the usefulness of the method. They point out that Var $\underline{\hat{\beta}}_R$ and MSE($\underline{\hat{\beta}}_R$) which were derived by Hoerl and Kennard (1970a) and treating k as a constant rather than an estimated quantity, are incorrect, since, in practice, k is estimated from the data.

2.5.2 Principal Components Regression

Principal components regression is another alternative biased estimation procedure. Massy (1965) investigated the procedure and concluded that the procedure can be very useful in exploratory research. Mansfield (1975) gave a detailed discussion of the principal components regression.

Consider the standard linear regression model (2.1.3). Principal components regression transforms the columns of the X matrix into orthogonal columns of principal components through the following process:

$$Z = XV, \quad \text{so } \underline{z}_j = X\underline{V}_j \qquad j = 1, 2, \ldots, p$$

where $$Z = (\underline{z}_1, \underline{z}_2, \underline{z}_3, \ldots, \underline{z}_p) .$$

Then $\underline{z}_j$ is called the jth principal component of X; that is, principal components are an orthogonal transformation of the values of the regressor variables.

The model (2.1.3) can now be written as

$$\underline{y} = \beta_0\underline{1} + Z\underline{\gamma} + \underline{\varepsilon} \qquad \text{where } \underline{\gamma} = V'\underline{\beta} . \tag{2.5.6}$$

The ordinary least squares estimator of $\underline{\gamma}$ can be written as

$$\hat{\underline{\gamma}} = (Z'Z)^{-1}Z'\underline{y} = \Lambda^{-1}Z'\underline{y} .$$

The estimator $\hat{\underline{\beta}}$ in (2.1.3) can be written as

$$\hat{\underline{\beta}} = V\hat{\underline{\gamma}} = V\Lambda^{-1}Z'\underline{y} = V\Lambda^{-1}V'X'\underline{y} = \sum_{j=1}^{p} \ell_j^{-1}c_j\underline{V}_j \tag{2.5.7}$$

where $c_j = \underline{V}_j'X'\underline{y}$. If all components are retained in the model, the estimator of $\underline{\beta}$ will be identical to the least squares estimator.

In order to remove the damaging effects of the multicollinearity on the least squares estimator and obtain more stable estimators of the β_j, the procedure of principal components regression is to delete those components corresponding to small latent roots of X'X. Then we regress $\underline{y}$ on the retained components using ordinary least squares.

If we assume that the first s components are deleted, we can partition the matrix of latent vectors into

$$V = (V_s, V_t)$$

where V_s contains the s latent vectors corresponding to components which are deleted and V_t contains the p-s latent vectors corresponding to components which are not deleted. Partition $\underline{\gamma}$ and Λ similarly as follows:

$$\underline{\gamma} = \begin{pmatrix} \underline{\gamma}_s \\ \underline{\gamma}_t \end{pmatrix} \qquad \Lambda = \begin{pmatrix} \Lambda_s & 0 \\ 0 & \Lambda_t \end{pmatrix}$$

where $\Lambda_t = \text{diag}(\ell_{s+1}, \ldots, \ell_p)$ contains the (p-s) latent roots which are not close to zero and $\Lambda_s = (\ell_1, \ell_2, \ldots, \ell_s)$ contains the s small latent roots of X'X.

From equation (2.5.7), we have the principal components estimator, (Mansfield, 1975)

$$\hat{\underline{\beta}}_{pc} = V_t \Lambda_t^{-1} V_t' X' \underline{y} = \sum_{j=s+1}^{p} \ell_j^{-1} c_j \underline{V}_j \quad \text{where } c_j = \underline{V}_j' X' \underline{y}. \quad (2.5.8)$$

$\hat{\underline{\beta}}_{pc}$ is a linear combination of the least squares estimator $\hat{\underline{\beta}}$, since

$$\hat{\underline{\beta}}_{pc} = V_t \Lambda_t^{-1} V_t' X' X (X'X)^{-1} X' \underline{y} = V_t \Lambda_t^{-1} V_t' X' X \hat{\underline{\beta}} \ ,$$

$$E(\hat{\underline{\beta}}_{pc}) = V_t \Lambda_t^{-1} V_t' X' X \hat{\underline{\beta}} = G_2 \hat{\underline{\beta}} \qquad (2.5.9)$$

where $G_2 = V_t \Lambda_t^{-1} V_t' X' X$. Since

$$G_2 = V_t \Lambda_t^{-1} V_t' X' X = V_t \Lambda_t^{-1} V_t' X' X V \cdot V'$$

$$= V_t \Lambda_t^{-1} (0 \quad \Lambda_t) V' = (0 \quad V_t \Lambda_t^{-1} \Lambda_t) \begin{pmatrix} V_s' \\ V_t' \end{pmatrix}$$

$$= V_t V_t' = I - V_s V_s'$$

$$E(\hat{\underline{\beta}}_{pc}) = G_2\underline{\beta} = (I - V_sV_s')\underline{\beta} = \underline{\beta} - V_sV_s'\underline{\beta}$$

$$= \underline{\beta} - \sum_{j=1}^{s} (\underline{V}_j'\underline{\beta})\underline{V}_j \ . \qquad (2.5.10)$$

Thus, the principal components estimator is biased.

The variance-covariance matrix of the principal components estimator is given by (Mansfield, 1975)

$$\mathrm{Var}(\hat{\underline{\beta}}_{pc}) = \sigma^2 V_t \Lambda_t^{-1} V_t' \ . \qquad (2.5.11)$$

The mean squared error of the principal components estimator can be written as (Mansfield, 1975)

$$\mathrm{MSE}(\hat{\underline{\beta}}_{pc}) = \sigma^2 \sum_{j=s+1}^{p} \ell_j^{-1} + \sum_{j=1}^{s} (\underline{V}_j'\underline{\beta})^2 \ . \qquad (2.5.12)$$

The variance portion of (2.5.12) is much less than the variance for least squares due to the deletion of the small latent roots and if the bias portion is not too large, a great reduction in mean squared error is possible. Principal components estimators perform very well when the vector of regression coefficients, $\underline{\beta}$, and the latent vectors defining the multicollinearities, $\underline{V}_1$, $\underline{V}_2$, $\underline{V}_3$, ..., $\underline{V}_s$, are orthogonal or near-orthogonal. On the other hand, when the vector of regression coefficients and the latent vectors defining the multicollinearity are parallel or nearly parallel, the principal components procedure may perform very poorly.

Mansfield (1975) showed that the F-statistic commonly used to determine the predictive ability of a component is not reliable when the component is associated with a small latent root. He discussed the advantages and the disadvantages of this procedure and presented a technique of variable selection after some components are deleted.

The residual sum of squares of the principal components regression can be written as

$$SSE_c = (\underline{y} - X\underline{b}^c)'(\underline{y} - X\underline{b}^c) \qquad \text{where } \underline{b}^c = \hat{\underline{\beta}}_{pc} = V_t\Lambda_t^{-1}V_t'X'\underline{y}$$

$$= \underline{y}'\underline{y} - \underline{b}^{c'}X'\underline{y} - (\underline{y}'X\underline{b}^c - \underline{b}^cX'X\underline{b}c)$$

$$= \underline{y}'\underline{y} - \underline{b}^{c'}X'\underline{y}$$

since $$\underline{y}'X\underline{b}^c - \underline{b}^{c'}XX\underline{b}^c = \underline{y}'XV_t\Lambda_t^{-1}V_t'X'\underline{y} - \underline{y}'XV_t\Lambda_t^{-1}V_t'X'XV_t\Lambda_t^{-1}V_t'X'\underline{y}$$

$$= \underline{y}'XV_t\Lambda_t^{-1}V_t'X'\underline{y} - \underline{y}'XV_t\Lambda_t^{-1}\Lambda_t\Lambda_t^{-1}V_t'X'\underline{y}$$

$$= \underline{y}'XV_t\Lambda_t^{-1}V_t'X'\underline{y} - \underline{y}'XV_t\Lambda_t^{-1}V_t'X'\underline{y} = 0 \ .$$

Thus, the regression sum of squares of the principal components regression is

$$SSR_c = \underline{b}^{c'}X'\underline{y} = \underline{y}'XV_t\Lambda_t^{-1}V_t'X'\underline{y} = \sum_{j=s+1}^{p}(\underline{V}_j'X'\underline{y})^2\ell_j^{-1} \ . \quad (2.5.13)$$

It should be noted that the principal components estimator is a special case of the generalized inverse estimator discussed by Marquardt (1970). He suggests that a generalized inverse should be used according to the following two situations:

(a) If the rank of X'X is t, where t = p-s, a generalized inverse of X'X is $(X'X)^- = V_t\Lambda_t^{-1}V_t'$, and the generalized inverse estimator is then given by

$$\hat{\underline{\beta}}_G = (X'X)^- X'\underline{y} = V_t \Lambda_t^{-1} V_t' X'\underline{y} \qquad (2.5.14)$$

which is equivalent to the principal components estimator, $\hat{\underline{\beta}}_{pc}$.

(b) If $X'X$ is full rank, but some of the latent roots of $X'X$ are small, Marquardt (1970) suggests the concept of fractional rank of $X'X$. If $X'X$ has s small latent roots, he assumes the rank of $X'X$ is f, where $t < f < t + 1$ and $t = p-s$. A generalized inverse of $X'X$ is then

$$(X'X)^- = V_t \Lambda_t^{-1} V_t' + \frac{f-t}{\ell_s} \underline{V}_s \underline{V}_s' ,$$

and the generalized inverse estimator based on a fractional rank of $X'X$ is then given by

$$\begin{aligned} \hat{\underline{\beta}}_G &= (X'X)^- X'\underline{y} \\ &= \left(\frac{f-t}{\ell_s} \underline{V}_s \underline{V}_s'\right) X'\underline{y} + V_t \Lambda_t^{-1} V_t' X'\underline{y} \qquad (2.5.15) \end{aligned}$$

For more details of the derivation of the generalized inverse estimator, the reader should refer to Marquardt (1970).

2.5.3 Latent Root Regression

Latent root regression, derived by Webster, Gunst, and Mason (1974) and independently by Hawkins (1973), is another alternative to least squares estimation.

Consider again the model (2.1.3) and let $y_i^* = \frac{(y_i - \bar{y})}{\eta}$ where $\eta^2 = \sum_{i=1}^{n} (y_i - \bar{y})^2$ and let $A = (\underline{y}^* \; X)$. Thus, $A'A$ is in the form of a correlation matrix of dependent and independent variables.

Let $\lambda_0 \leq \lambda_1 \leq \ldots \leq \lambda_p$ be the ordered latent roots of A'A, and let $\underline{\gamma}_0, \underline{\gamma}_1, \ldots, \underline{\gamma}_p$ be the corresponding latent vectors. The elements of the jth latent vector are denoted by

$$\underline{\gamma}_j' = (\gamma_{0j}, \gamma_{1j}, \ldots, \gamma_{pj}) = (\gamma_{0j}, \underline{\delta}_j') .$$

Finally, let

$$\Gamma = (\underline{\gamma}_0, \underline{\gamma}_1, \ldots, \underline{\gamma}_p) \qquad \Lambda^* = \mathrm{diag}(\lambda_0, \lambda_1, \ldots, \lambda_p)$$

then $$\Gamma'(A'A)\Gamma = \Lambda^* .$$

From this we obtain

$$\underline{\gamma}_j'A'A\underline{\gamma}_j = \lambda_j \qquad j = 1, 2, \ldots, p . \tag{2.5.16}$$

When multicollinearity is present, some of the latent roots of A'A will be near zero since $\lambda_0 \leq \ell_1$(Gunst, Webster, and Mason, 1976). Suppose $\lambda_0 \doteq 0$. Hence $\underline{\gamma}_0'A'A\underline{\gamma}_0 \doteq 0$, and $(A\underline{\gamma}_0)'(A\underline{\gamma}_0) \doteq 0$.

This implies $$A\underline{\gamma}_0 \doteq 0,$$

and $$A\underline{\gamma}_0 = \underline{y}\gamma_{00} + X\underline{\delta}_0 \doteq 0. \tag{2.5.17}$$

Thus elements of the latent vector $\underline{\gamma}_0$ are coefficients of a linear combination involving the response and regressor variables. Equation (2.5.17) can be written as

$$\begin{bmatrix} y_1\gamma_{00} + \sum_{r=1}^{p} X_{1r}\gamma_{r0} \\ \vdots \\ y_n\gamma_{00} + \sum_{r=1}^{p} X_{nr}\gamma_{r0} \end{bmatrix} \doteq 0 \, . \qquad (2.5.18)$$

If $\lambda_0 \doteq 0$ and $\gamma_{00} \doteq 0$, we can see from (2.5.18), that the linear dependence exists among the columns of X and does not involve $\underline{y}$. This is termed a nonpredictive multicollinearity by Webster, et al. (1974). If $\lambda_0 \doteq 0$, but γ_{00} is not small, the response variable is involved in the relationship. This is termed a predictive multicollinearity.

The least squares estimator of $\underline{\beta}$ can be written in terms of the latent roots and latent vectors of A'A as

$$\hat{\underline{\beta}} = -\eta \sum_{j=0}^{p} a_j\underline{\delta}_j \quad \text{where} \quad \eta^2 = \sum_{i=1}^{n} (y_i-\bar{y})^2, \; a_j = \gamma_{0j}\lambda_j^{-1}\Big(\sum_{\ell=0}^{p} \frac{\gamma_{0\ell}^2}{\lambda_\ell}\Big)^{-1}. \qquad (2.5.19)$$

If we assume that the latent vectors $\underline{\gamma}_0, \underline{\gamma}_1, \ldots, \underline{\gamma}_{q-1}$ correspond to nonpredictive multicollinearities, the latent root estimator of regression coefficients can be written as

$$\hat{\underline{\beta}}_{LR} = -\eta \sum_{j=q}^{p} a_j\underline{\delta}_j \quad \text{where} \quad a_j = \gamma_{0j}\lambda_j^{-1}\Big(\sum_{\ell=q}^{p} \frac{\gamma_{0\ell}^2}{\lambda_\ell}\Big)^{-1}, \; j = q, q+1, \ldots, p \, .$$

(Webster, Gunst, and Mason, 1974) (2.5.20)

Thus, we remove the latent vectors corresponding to the nonpredictive multicollinearities in order to improve the estimation. The actual improvement is strongly dependent upon the orientation of coefficient vector $\underline{\beta}$, to latent vectors which are removed.

The advantage of this method is that by deleting the effect of the nonpredictive multicollinearities, the true influence of the regressor variables on the response variable are more clearly represented. The disadvantage of this method is that distributional properties, including expectation, variance, and mean squared error of the estimator are unknown.

CHAPTER III

MODIFIED PRINCIPAL COMPONENTS REGRESSION

3.1 The Estimation Procedures

In this chapter, we propose a compromise estimation procedure for regression coefficients when multicollinearities exist among the regressor variables. Certain theoretical properties of the resulting estimator will be derived and discussed.

Consider again the classical linear regression model (2.1.3). This model can be written in terms of the principal components as in equation (2.5.6). The dependent variable is now regressed on the principal components, $\underline{z}_j$, rather than on the original variables, x_j. As mentioned in Chapter II, the ordinary least squares estimator of $\underline{\beta}$ for the model (2.1.3) can be written in terms of latent roots and latent vectors of X'X as

$$\hat{\underline{\beta}} = \sum_{j=1}^{s} \ell_j^{-1} c_j \underline{V}_j + \sum_{j=s+1}^{p} \ell_j^{-1} c_j \underline{V}_j \ . \tag{3.1.1}$$

The variance of the least squares estimator can be written as

$$\text{Var } \hat{\underline{\beta}} = \sigma^2 \sum_{j=1}^{s} \ell_j^{-1} \underline{V}_j \underline{V}_j' + \sigma^2 \sum_{j=s+1}^{p} \ell_j^{-1} \underline{V}_j \underline{V}_j' \ . \tag{3.1.2}$$

Thus, the least squares estimator tends to place large weights on the latent vectors corresponding to the small latent roots of X'X.

The elements of $\hat{\underline{\beta}}$ corresponding to regressors involved in multicollinearities tend to be dominated by the first few terms of (3.1.1). From equations (3.1.1) and (3.1.2), we can see that the least squares estimator is unstable primarily due to the presence of small latent roots in equation (3.1.1). Therefore, the least squares estimator is severely affected by the linear dependence among regressors and could provide poor estimates of the true parameters.

As we have seen from Chapter II, the disadvantage of the principal components regression is that the components are deleted without regard to their predictive ability. If a component has predictive ability and we delete the component, we will lose information on the linear combination of x's represented by this component. As mentioned in Chapter II, the measure commonly used to determine the predictive ability of a component is not reliable when the component is associated with a small latent root.

The modified principal components regression is designed to remove the dominant effects of the multicollinearity, and at the same time, to keep as much important information as possible. Instead of deleting the components associated with small latent roots, the modified principal components regression reduces the weights of latent vectors corresponding to small latent roots of X'X.

The modified principal components estimator can be written as

$$\hat{\underline{\beta}}_{mpc} = \sum_{j=1}^{s} (\ell_j + k_j)^{-1} c_j \underline{V}_j + \sum_{j=s+1}^{p} \ell_j^{-1} c_j \underline{V}_j \qquad (3.1.3)$$

where $c_j = \underline{V}_j' X' \underline{y}$.

The k_j are determined so that the variance inflation factors of x_j are less than ten but close to ten, since a V.I.F(x_j) of ten or greater is considered an indication that the variable x_j is involved in multicolleariy (Marquardt, 1970). The reason for this has been explained in Chapter II (section 2.4).

Equation (3.1.3) can also be written as

$$\hat{\underline{\beta}}_{mpc} = V_s(\Lambda_s + K_s)^{-1}V_s'X'Y + V_t\Lambda_t^{-1}V_t'X'\underline{Y} = VD^{-1}V'X'\underline{Y} \quad (3.1.4)$$

where $$D = \text{diag}((\ell_1+k_1),(\ell_2+k_2), \ldots, (\ell_s+k_s), \ell_{s+1}, \ldots, \ell_p)$$

$$= \begin{pmatrix} \Lambda_s+K_s & 0 \\ 0 & \Lambda_t \end{pmatrix}$$

and $K_s = \text{diag}(k_1, k_2, \ldots, k_s)$.

3.2 The Properties of Modified Principal Components Regression

We now consider certain theoretical properties of this estimator. In the derivation which follows, we assume k_j to be nonstochastic.

The estimator, $\hat{\underline{\beta}}_{mpc}$, is a linear combination of the least squares estimator:

$$\hat{\underline{\beta}}_{mpc} = VD^{-1}V'X'\underline{Y} = VD^{-1}V'(X'X) \cdot (X'X)^{-1}X'\underline{Y}$$

$$= VD^{-1}V'(X'X)V \cdot V'\hat{\underline{\beta}}$$

$$= VD^{-1}\Lambda V'\hat{\underline{\beta}}$$

$$= G\hat{\underline{\beta}} \quad (3.2.1)$$

where $$G = VD^{-1}\Lambda V' = [I_p + (X'X)^{-1}V_sK_sV_s']^{-1} .$$

Since

$$D^{-1}\Lambda = (\Lambda + K)^{-1}\Lambda = V'(X'X + V_s K_s V_s')^{-1}V \cdot V'(X'X)V$$

where $K = \mathrm{diag}(k_1, k_2, \ldots, k_s, 0, \ldots, 0)$ is a p×p matrix, then

$$VD^{-1}\Lambda V' = (X'X + V_s K_s V_s)^{-1}(X'X) = [I_p + (X'X)^{-1}V_s K_s V_s']^{-1}.$$

It follows immediately that $\hat{\underline{\beta}}_{mpc}$ is a biased estimator of $\underline{\beta}$.

$$E(\hat{\underline{\beta}}_{mpc}) = G \cdot E(\hat{\underline{\beta}}) = G\underline{\beta} \qquad \text{where} \quad G = VD^{-1}\Lambda V'.$$

Since $V'GV = D^{-1}\Lambda = D^{-1}(D - K) = I_p - D^{-1}K$, then we have

$$\begin{aligned} E(\hat{\underline{\beta}}_{mpc}) &= G\underline{\beta} = V \cdot V'GV \cdot V'\underline{\beta} \\ &= V(I_p - D^{-1}K)V'\underline{\beta} \\ &= VV'\underline{\beta} - VD^{-1}KV'\underline{\beta} \\ &= \underline{\beta} - \sum_{j=1}^{s} k_j(\ell_j + k_j)^{-1}(\underline{V}_j\underline{\beta})\underline{V}_j . \end{aligned} \tag{3.2.2}$$

Thus, the bias of the modified principal components estimator depends on the orientation of the vector of regression coefficients to the latent vectors corresponding to small latent roots. The bias should be relatively small if k_j is small and $\underline{\beta}$ is not parallel to $\underline{V}_j$ for $j = 1, 2, \ldots, s$.

The variance-covariance matrix of the modified principal components estimator is

$$\begin{aligned} \mathrm{Var}\ \hat{\underline{\beta}}_{mpc} &= \mathrm{Var}\ G\hat{\underline{\beta}} = G \cdot \mathrm{Var}\ \hat{\underline{\beta}} \cdot G' \\ &= \sigma^2 VD^{-1}\Lambda V'(X'X)^{-1}V\Lambda D^{-1}V' \end{aligned}$$

$$= \sigma^2 V D^{-1} \Lambda D^{-1} V'$$

$$= \sigma^2 \left[\sum_{j=1}^{s} \ell_j (\ell_j + k_j)^{-2} \underline{V}_j \underline{V}_j' + \sum_{j=s+1}^{p} \ell_j^{-1} \underline{V}_j \underline{V}_j' \right] . \qquad (3.2.3)$$

The variance inflation factors for the modified principal components estimator are the diagonal elements of $VD^{-1}\Lambda D^{-1}V'$.

Comparing equation (2.2.4) with equation (3.2.3), we can see that the variances and pairwise covariances of the modified principal components estimators of individual coefficients are smaller than the corresponding ones of the least squares estimators for any $k_j > 0$.

The mean squared error of the modified principal components estimator can be expressed as:

$$MSE(\hat{\underline{\beta}}_{mpc}) = \sigma^2 \sum_{j=1}^{s} \ell_j (\ell_j + k_j)^{-2} + \sigma^2 \sum_{j=s+1}^{p} \ell_j^{-1} + \sum_{j=1}^{s} k_j^2 (\ell_j + k_j)^{-2} (\underline{V}_j' \underline{\beta})^2 .$$

(3.2.4)

Since,

$$MSE(\hat{\underline{\beta}}_{mpc}) = E(\underline{b}^m - \underline{\beta})'(\underline{b}^m - \underline{\beta}) \qquad \text{where } \underline{b}^m = \hat{\underline{\beta}}_{mpc}$$

$$= E(\underline{b}^m - E\underline{b}^m)'(\underline{b}^m - E\underline{b}^m) + (E\underline{b}^m - \underline{\beta})'(E\underline{b}^m - \underline{\beta})$$

$$= E \sum_{j=1}^{p} [b_j^m - E(b_j^m)]^2 + \sum_{j=1}^{p} [Eb_j^m - \beta_j]^2$$

$$= \sum_{j=1}^{p} \text{Var } b_j^m + \sum_{j=1}^{p} (\text{bias of } b_j^m)^2 = V + B$$

where

$$V = \sum_{j=1}^{p} \text{Var } b_j^m = \text{tr Var } \underline{b}^m = \sigma^2 \text{ tr } VD^{-1}\Lambda D^{-1}V'$$

$$= \sigma^2 \; tr[V_s(\Lambda_s+k_s)^{-2}\Lambda_s V_s' + V_t \Lambda_t^{-1} V_t']$$

$$= \sigma^2 \sum_{j=1}^{s} \ell_j(\ell_j+k_j)^{-2} \; tr \; \underline{V}_j\underline{V}_j' + \sigma^2 \sum_{j=s+1}^{p} \ell_j^{-1} \; tr \; \underline{V}_j\underline{V}_j'$$

$$= \sigma^2 \sum_{j=1}^{s} \ell_j(\ell_j+k_j)^{-2} + \sigma^2 \sum_{j=s+1}^{p} \ell_j^{-1}$$

$$B = (E\underline{b}^m-\underline{\beta})'(E\underline{b}^m-\underline{\beta}) = (G\underline{\beta}-\underline{\beta})'(G\underline{\beta}-\underline{\beta})$$

$$= \underline{\beta}'(G-I)'(G-I)\underline{\beta} \; .$$

Since $G = VD^{-1}\Lambda V'$,

$$VGV' = D^{-1}\Lambda = D^{-1}(D-K) = I_p - D^{-1}K$$

$$G = V(I_p-D^{-1}K)V' = I_p - VD^{-1}KV', \text{ then } G - I = (-VD^{-1}KV')$$

$$B = \underline{\beta}'(-VD^{-1}KV')'(-VD^{-1}KV')\underline{\beta}$$

$$= \underline{\beta}'VKD^{-1}V' \cdot VD^{-1}KV'\underline{\beta}$$

$$= \underline{\beta}'VKD^{-2}KV'\underline{\beta}$$

$$= \underline{\beta}'(V_s V_t)\begin{pmatrix} K_s & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} (\Lambda_s+K_s)^{-2} & 0 \\ 0 & \Lambda_t^{-2} \end{pmatrix}\begin{pmatrix} K_s & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} V_s'\underline{\beta} \\ V_t'\underline{\beta} \end{pmatrix}$$

$$= \underline{\beta}'V_s K_s(\Lambda_s+K_s)^{-2}K_s V_s'\underline{\beta}$$

$$= \sum_{j=1}^{s} k_j^2(\ell_j+k_j)^{-2}(\underline{V}_j'\underline{\beta})^2$$

and the result follows.

Since the determination of k_j, using the V.I.F.(x) as a criterion, depends only on the X matrix, which is assumed to be nonstochastic, k_j's for the modified principal components estimator in equation (3.1.3) are nonstochastic in practice. Therefore, the theoretical properties, derived above under the assumption of nonrandom k_j, are all valid.

From equations (3.2.4) and (2.2.5), we can see that the variance portion of MSE($\hat{\underline{\beta}}_{mpc}$) is smaller than the corresponding one of MSE($\hat{\underline{\beta}}$) for any $k_j > 0$. From equation (3.2.4) and (2.5.12), it is obvious that the bias portion of the MSE($\hat{\underline{\beta}}_{mpc}$) is always smaller than that of the MSE($\hat{\underline{\beta}}_{pc}$) for any $k_j > 0$. Each term of the bias part of MSE($\hat{\underline{\beta}}_{mpc}$) is weighted by $(\frac{k_j}{\ell_j+k_j})^2$ which is less than one and will be close to zero if k_j is near zero. Thus, the bias of the modified principal components regression will be much smaller than that of the principal components regression when the vector of regression coefficients and the latent vector defining the multicollinearity are parallel or nearly parallel. This will be illustrated in our simulation results.

From equations (3.2.4) and (2.2.5), we have

$$\text{MSE}(\hat{\underline{\beta}}_{mpc}) = \sigma^2 \sum_{j=1}^{s} \ell_j(\ell_j+k_j)^{-2} + \sigma^2 \sum_{j=s+1}^{p} \ell_j^{-1} + \sum_{j=1}^{s} k_j^2(\ell_j+k_j)^{-2}(\underline{V}_j'\underline{\beta})^2$$

$$\text{MSE}(\hat{\underline{\beta}}) = \sigma^2 \sum_{j=1}^{s} \ell_j^{-1} + \sigma^2 \sum_{j=s+1}^{p} \ell_j^{-1} .$$

Comparing MSE($\hat{\underline{\beta}}_{mpc}$) with MSE($\hat{\underline{\beta}}$), we see that MSE($\hat{\underline{\beta}}_{mpc}$) < MSE($\hat{\underline{\beta}}$), if and only if

$$\sigma^2 \sum_{j=1}^{s} \ell_j(\ell_j+k_j)^{-2} + \sum_{j=1}^{s} (\frac{k_j}{\ell_j+k_j})^2 (\underline{V}_j'\underline{\beta})^2 < \sigma^2 \sum_{j=1}^{s} \ell_j^{-1}$$

or $$\sum_{j=1}^{s} (\frac{k_j}{\ell_j+k_j})^2 (\underline{V}_j'\underline{\beta})^2 < \sigma^2 \sum_{j=1}^{s} \ell_j^{-1} - \sigma^2 \sum_{j=1}^{s} \ell_j(\ell_j+k_j)^{-2} .$$

After some simplification, we obtain that a necessary and sufficient condition for $MSE(\hat{\underline{\beta}}_{mpc}) < MSE(\hat{\underline{\beta}})$ is

$$\sum_{j=1}^{s} k_j^2(\ell_j+k_j)^{-2}(\underline{V}_j'\underline{\beta})^2 < \sigma^2 \sum_{j=1}^{s} k_j(\ell_j+k_j)^{-2}\ell_j^{-1}(k_j+2\ell_j)$$

or $$\sum_{j=1}^{s} k_j(\underline{V}_j'\underline{\beta})^2 < \sigma^2 \sum_{j=1}^{s} \ell_j^{-1}(k_j+2\ell_j) . \qquad (3.2.5)$$

From equation (2.2.5) and (2.5.12), we have

$$MSE(\hat{\underline{\beta}}) = \sigma^2 \sum_{j=1}^{s} \ell_j^{-1} + \sigma^2 \sum_{j=s+1}^{p} \ell_j^{-1}$$

$$MSE(\hat{\beta}_{pc}) = \sigma^2 \sum_{j=s+1}^{p} \ell_j^{-1} + \sum_{j=1}^{s} (\underline{V}_j'\underline{\beta})^2 .$$

Comparing $MSE(\hat{\underline{\beta}})$ with $MSE(\hat{\underline{\beta}}_{pc})$, we see that a necessary and sufficient condition for $MSE(\hat{\underline{\beta}}_{pc})$ to be less than $MSE(\hat{\underline{\beta}})$ is

$$\sigma^2 \sum_{j=1}^{s} \ell_j^{-1} > \sum_{j=1}^{s} (\underline{V}_j'\underline{\beta})^2 . \qquad (3.2.6)$$

From equation (3.2.5), it follows that $MSE(\hat{\underline{\beta}}_{mpc})$ is less than $MSE(\hat{\underline{\beta}})$, if $\sigma^2 \ell_j^{-1} > (\frac{k_j}{k_j+2\ell_j})(\underline{V}_j'\underline{\beta})^2$, $j = 1, 2, \ldots, s.$ (3.2.7)

From equation (3.2.6), it follows that $MSE(\hat{\underline{\beta}}_{pc})$ is less than $MSE(\hat{\underline{\beta}})$

$$\text{if } \sigma^2 \ell_j^{-1} > (\underline{V}_j'\underline{\beta})^2, \; j = 1, 2, \ldots, s. \quad (3.2.8)$$

The two equations (3.2.7) and (3.2.8) imply that the range of orientations for which $MSE(\hat{\underline{\beta}}_{mpc}) < MSE(\hat{\underline{\beta}})$ is larger than the range of orientations for which $MSE(\hat{\underline{\beta}}_{pc}) < MSE(\hat{\underline{\beta}})$ since $(\frac{k_j}{k_j+2\ell_j})$ in equation (3.2.7) is less than one.

From equations (3.2.4) and (2.5.12), we have

$$MSE(\hat{\underline{\beta}}_{mpc}) = \sigma^2 \sum_{j=1}^{s} \ell_j(\ell_j+k_j)^{-2} + \sigma^2 \sum_{j=s+1}^{p} \ell_j^{-1} + \sum_{j=1}^{s} k_j^2(\ell_j+k_j)^{-2}(\underline{V}_j'\underline{\beta})^2$$

$$MSE(\hat{\underline{\beta}}_{pc}) = \sigma^2 \sum_{j=s+1}^{p} \ell_j^{-1} + \sum_{j=1}^{s} (\underline{V}'\underline{\beta})^2 .$$

If we compare $MSE(\hat{\underline{\beta}}_{mpc})$ with $MSE(\hat{\underline{\beta}}_{pc})$, it is clear that $MSE(\hat{\underline{\beta}}_{mpc}) <$ $MSE(\hat{\underline{\beta}}_{pc})$ if and only if

$$\sigma^2 \sum_{j=1}^{s} \ell_j(\ell_j+k_j)^{-2} + \sum_{j=1}^{s} (\frac{k_j}{\ell_j+k_j})^2(\underline{V}_j'\underline{\beta})^2 < \sum_{j=1}^{s} (\underline{V}_j'\underline{\beta})^2$$

or $$\sigma^2 \sum_{j=1}^{s} \ell_j(\ell_j+k_j)^{-2} < \sum_{j=1}^{s} (\underline{V}_j'\underline{\beta})^2 - \sum_{j=1}^{s} (\frac{k_j}{\ell_j+k_j})^2(\underline{V}'\underline{\beta})^2 .$$

After some simplification, we obtain that a necessary and sufficient condition for $MSE(\hat{\underline{\beta}}_{mpc}) < MSE(\hat{\underline{\beta}}_{pc})$ is

$$\sigma^2 \sum_{j=1}^{s} \ell_j(\ell_j+k_j)^{-2} < \sum_{j=1}^{s} (\underline{V}_j'\underline{\beta})^2 \ell_j(\ell_j+k_j)^{-2}(\ell_j+2k_j)$$

or

$$s\sigma^2 < \sum_{j=1}^{s} (\ell_j+2k_j)(\underline{V}_j'\underline{\beta})^2 \ . \tag{3.2.9}$$

From equation (3.2.9), it follows that $\text{MSE}(\hat{\underline{\beta}}_{mpc})$ is less than $\text{MSE}(\hat{\underline{\beta}}_{pc})$ if

$$(\underline{V}_j'\underline{\beta})^2 > \left(\frac{\ell_j}{\ell_j+2k_j}\right)\sigma^2\ell_j^{-1} \qquad j = 1, 2, \ldots, s. \tag{3.2.10}$$

Using equations (3.2.7) and (3.2.10), we see that a sufficient condition for both $\text{MSE}(\hat{\underline{\beta}}_{mpc}) < \text{MSE}(\hat{\underline{\beta}})$ and $\text{MSE}(\hat{\underline{\beta}}_{mpc}) < \text{MSE}(\hat{\underline{\beta}}_{pc})$ is

$$\left(\frac{\ell_j}{\ell_j+2k_j}\right)\sigma^2\ell_j^{-1} < (\underline{V}_j'\underline{\beta})^2 < \left(\frac{k_j+2\ell_j}{k_j}\right)\sigma^2\ell_j^{-1} \quad j = 1, 2, \ldots, s. \tag{3.2.11}$$

Thus, for given ℓ_j and k_j, and estimated σ^2, we can obtain an estimated range of $(\underline{V}_j'\underline{\beta})^2$, $j = 1, 2, \ldots, s$, where the modified principal components estimator will always perform better in terms of mean squared error criterion than both the least squares estimator and the principal components estimator.

If we compare equation (2.5.10) with equation (3.2.2), we can observe that the biases of both modified principal components estimator and principal components estimator are functions of $(\underline{V}_j'\underline{\beta})\underline{V}_j$ and because of the nature of the weights, $\left(\frac{k_j}{\ell_j+k_j}\right)$, in equation (3.2.2), the bias of the modified principal components estimator will always be smaller than that of the principal components estimator.

The residual sum of squares of the modified principal components estimator can be written as

$$\begin{aligned}
SSE_m &= (\underline{y}-X\underline{b}^m)'(\underline{y}-X\underline{b}^m) \\
&= \underline{y}'\underline{y} - \underline{b}^{m'}X'\underline{y} - (\underline{y}'X\underline{b}^m - \underline{b}^{m'}X'X\underline{b}^m) \\
&= \underline{y}'\underline{y} - \underline{b}^{m'}X'\underline{y} - \underline{b}^{m'}V_s K_s V_s'\underline{b}^m
\end{aligned}$$

since

$$\begin{aligned}
\underline{y}'X\underline{b}^m - \underline{b}^{m'}X'X\underline{b}^m &= \underline{y}'X\underline{b}^m - \underline{y}'XVD^{-1}V'X'X\underline{b}^m \\
&= \underline{y}'XVD^{-1}V'(VDV'-X'X)\underline{b}^m \\
&= \underline{y}'XVD^{-1}V'(VDV'-V\Lambda V')\underline{b}^m \\
&= \underline{b}^{m'}V(D-\Lambda)V'\underline{b}^m = \underline{b}^{m'}VKV'\underline{b}^m \\
&= \underline{b}^{m'}V_s K_s V_s'\underline{b}^m \ .
\end{aligned}$$

The regression sum of squares of the modified principal components regression is

$$\begin{aligned}
SSR_m = \underline{y}'\underline{y} - SSE_m &= \underline{b}^{m'}X'\underline{y} + \underline{b}^{m'}V_s K_s V_s'\underline{b}^m \\
&= (\underline{b}'X'\underline{y} - \underline{b}'VKD^{-1}V'X'\underline{y}) + \underline{b}^{m'}V_s K_s V_s'\underline{b}^m \\
&= \sum_{j=1}^{p} (\underline{V}_j'X'\underline{y})^2 \ell_j^{-1} - \sum_{j=1}^{s} (\underline{V}_j'X'\underline{y})^2 k_j(\ell_j+k_j)^{-1}[\ell_j^{-1} - (\ell_j+k_j)^{-1}] \\
&= \sum_{j=1}^{p} (\underline{V}_j'X'\underline{y})^2 \ell_j^{-1} - \sum_{j=1}^{s} (\underline{V}_j'X'\underline{y})^2 k_j(\ell_j+k_j)^{-1}\left[\frac{k_j}{\ell_j(\ell_j+k_j)}\right]
\end{aligned}$$

$$= \sum_{j=1}^{p} (\underline{V}_j'X'\underline{y})^2 \ell_j^{-1} - \sum_{j=1}^{s} (\underline{V}_j'X'\underline{y})^2 \ell_j^{-1} \left(\frac{k_j}{\ell_j+k_j}\right)^2$$

since, $\underline{b}^m = G\underline{b} = V \cdot V'GV \cdot V'\underline{b}$

$$= V(I-D^{-1}K)V'\underline{b} = \underline{b} - VD^{-1}KV'\underline{b}$$

$$\underline{b}^{m'}X'\ \underline{y} = (\underline{b}-VD^{-1}KV'\underline{b})'X'\underline{y} = \underline{b}'X'\underline{y} - \underline{b}'VKD^{-1}V'X'\underline{y}$$

$$= \underline{y}'X(X'X)^{-1}X'\underline{y} - \underline{y}'X(X'X)^{-1}VKD^{-1}V'X\underline{y}$$

$$= \underline{y}'XV \cdot V'(X'X)^{-1}V \cdot V'X'\underline{y} - \underline{y}'XV \cdot V'(X'X)^{-1}VKD^{-1}V'X'\underline{y}$$

$$= \underline{y}'XV\Lambda^{-1}V'X'\underline{y} - \underline{y}'XV\Lambda^{-1}KD^{-1}V'X'\underline{y}$$

$$= \sum_{j=1}^{p} (\underline{V}_j'X'\underline{y})^2 \ell_j^{-1} - \sum_{j=1}^{s} (\underline{V}_j'X'\underline{y})^2 \ell_j^{-1} k_j (\ell_j+k_j)^{-1}$$

$$\underline{b}^{m'} V_s K_s V_s' \underline{b}^m = \underline{b}^m VKV'\underline{b}^m = \underline{y}'XVD^{-1}V'VKV'VD^{-1}V'X'\underline{y}$$

$$= \underline{y}'XVD^{-1}KD^{-1}V'X'\underline{y}$$

$$= \sum_{j=1}^{s} (\underline{V}_j'X'\underline{y})^2 (\ell_j+k_j)^{-2} k_j \ .$$

Thus, the regression sum of squares of the modified principal components regression can be written as

$$SSR_m = \sum_{j=1}^{p} (\underline{V}_j'X'\underline{y})^2 \ell_j^{-1} - \sum_{j=1}^{s} (\underline{V}_j'X'\underline{y}) \ell_j^{-1} \left(\frac{k_j}{\ell_j+k_j}\right)^2 \ . \qquad (3.2.12)$$

If all the k_j are zero, the regression sum of squares for the modified principal components method will be identical to that of the least squares estimator, that is,

$$SSR_m = \sum_{j=1}^{p} (\underline{V}_j'X'\underline{y})^2 \ell_j^{-1} = SSR_{LS} \quad \text{if } k_j = 0,$$

$$j = 1, 2, \ldots, s.$$

If we compare equation (3.2.12) with equation (2.5.13), we obtain

$$SSR_m - SSR_c = \sum_{j=1}^{s} (\underline{V}_j'X'\underline{y})^2 \ell_j^{-1} - \sum_{j=1}^{s} (\underline{V}_j'X'\underline{y})^2 \ell_j^{-1} \left(\frac{k_j}{\ell_j + k_j}\right)^2$$

$$= \sum_{j=1}^{s} (\underline{V}_j'X'\underline{y})^2 \ell_j^{-1} \left[1 - \frac{k_j^2}{(\ell_j + k_j)^2}\right]$$

$$= \sum_{j=1}^{s} (\underline{V}_j'X'\underline{y})^2 \ell_j^{-1} \left[\frac{\ell_j^2 + k_j^2 + 2\ell_j k_j - k_j^2}{(\ell_j + k_j)^2}\right]$$

$$= \sum_{j=1}^{s} (\underline{V}_j'X'\underline{y})^2 \ell_j^{-1} \left[\frac{\ell_j(\ell_j + 2k_j)}{(\ell_j + k_j)^2}\right]$$

$$= \sum_{j=1}^{s} (\underline{V}_j'X'\underline{y})^2 (\ell_j + 2k_j)(\ell_j + k_j)^{-2} > 0 .$$

Thus, the regression sum of squares for the modified principal components method is always greater than that for the principal components method. Alternatively, the residual sum of squares of the modified principal components regression will be always smaller than that of the principal components regression.

3.3 The Relationship Between the Modified Principal Components Estimator and Other Estimators

1. The Least Squares Estimator and the Modified Principal Components Estimator.

From equation (3.2.1), we can see that $\hat{\underline{\beta}}_{mpc}$ is a linear combination of the least squares estimator, $\hat{\underline{\beta}}$.

$$\hat{\underline{\beta}}_{mpc} = G\hat{\underline{\beta}} \qquad \text{where} \qquad G = VD^{-1}\Lambda V' = [I + (X'X)^{-1}V_s k_s V_s']^{-1} \ . \tag{3.3.1}$$

2. The Principal Components Estimator and the Modified Principal Components Estimator.

From equation (2.5.8), the principal components estimator can be written as

$$\hat{\underline{\beta}}_{pc} = V_t \Lambda_t^{-1} V_t' X' y$$

From equation (3.1.4), the modified principal components estimator can be written as

$$\hat{\underline{\beta}}_{mpc} = V_s(\Lambda_s + k_s)^{-1} V_s' X' \underline{y} + V_t \Lambda_t^{-1} V_t' X' \underline{y}$$

Therefore, the relationship between $\hat{\underline{\beta}}_{mpc}$ and $\hat{\underline{\beta}}_{pc}$ is

$$\hat{\underline{\beta}}_{mpc} = V_s(\Lambda_s + k_s)^{-1} V_s' X' \underline{y} + \hat{\underline{\beta}}_{pc} \ . \tag{3.3.2}$$

3. The Generalized Inverse Estimator and the Modified Principal Components Estimator.

The relationship between the generalized inverse estimator discussed by Marquardt (1970) and the modified principal components estimator can be outlined according to the following two situations:

(a) If the rank of $X'X$ is t, where $t = p-s$, the generalized inverse estimator can be written as equation (2.5.14), which is equivalent to the principal components estimator, $\hat{\underline{\beta}}_{pc}$. In this case, the relationship between the modified principal components estimator and the generalized inverse estimator will be the same as the relationship given in equation (3.3.2).

(b) If the rank of X'X is p, but some of the latent roots of X'X are small, the generalized inverse estimator, based on a fractional rank of X'X, can be written as equation (2.5.15).

From equation (3.1.4), the modified principal components estimator can be written as

$$\begin{aligned} \hat{\underline{\beta}}_{mpc} &= V_s(\Lambda_s+K_s)^{-1}V_s'X'\underline{y} + V_t\Lambda_t^{-1}V_t'X'\underline{y} \\ &= [\sum_{j=1}^{s}(\ell_j+k_j)^{-1}\underline{V}_j\underline{V}_j']X'\underline{y} + V_t\Lambda_t^{-1}V_t'X'\underline{y} \ . \qquad (3.3.3) \end{aligned}$$

Comparing equation (2.5.15) with equation (3.3.3), we can see that the modified principal components estimator uses information from all the latent vectors, $\underline{V}_1$, $\underline{V}_2$, $\underline{V}_3$, ..., $\underline{V}_p$; whereas, the generalized inverse estimator based on a fractional rank of X'X, contains information only from the latent vectors $\underline{V}_s$, $\underline{V}_{s+1}$, $\underline{V}_{s+2}$, ..., $\underline{V}_p$. In addition to this, $\underline{V}_s$ of $\hat{\underline{\beta}}_{mpc}$ is weighted by $(\ell_s+k_s)^{-1}c_s$; whereas, $\underline{V}_s$ of $\hat{\underline{\beta}}_G$ is weighted by $(\frac{f-t}{\ell_s})c_s$ where $c_s = \underline{V}_s'X'\underline{y}$.

For s = 1, $\hat{\underline{\beta}}_{mpc}$ in equation (3.3.3) can be written as

$$\hat{\underline{\beta}}_{mpc} = (\ell_1+k_1)^{-1}c_1\underline{V}_1 + V_t\Lambda_t^{-1}V_t'X'\underline{y} \quad \text{where } c_1 = \underline{V}_1'X'\underline{y}; \qquad (3.3.4)$$

whereas, $\hat{\underline{\beta}}_G$ in equation (2.5.15) can be written as

$$\underline{\hat{\beta}}_G = (\frac{f-t}{\ell_1})c_1\underline{v}_1 + V_t\Lambda_t^{-1}V_t'X'\underline{y} . \tag{3.3.5}$$

Thus, f in equation (3.3.5) can be chosen so that $\underline{\hat{\beta}}_G$ is equal to $\underline{\hat{\beta}}_{mpc}$; this requires $f = t + \ell_1(\ell_1+k_1)^{-1}$. However, f is not generally chosen in this fashion, so that, in practice, $\underline{\hat{\beta}}_{mpc}$ and $\underline{\hat{\beta}}_G$ will be different estimators.

4. The Partitioned Ridge Regression and the Modified Principal Components Regression.

From equation (2.5.5), the partitioned ridge estimator can be written as

$$\underline{\hat{\beta}}_R = (X'X + kA)^{-1}X'\underline{y} \qquad k > 0$$

$$\text{where } A = \begin{pmatrix} 0 & 0 \\ 0 & A_{22} \end{pmatrix}$$

and A_{22} is a positive semidefinite matrix.

From equation (3.1.4), the modified principal components estimator can be written in another form as

$$\underline{\hat{\beta}}_{mpc} = VD^{-1}V'X'\underline{y} = (X'X+V_sK_sV_s')^{-1}X'\underline{y} . \tag{3.3.6}$$

$$\text{Since } V'(X'X+V_sK_sV_s')^{-1}V = [V'(X'X+V_sK_sV_s')V]^{-1}$$

$$= \begin{pmatrix} \Lambda_s+K_s & 0 \\ 0 & \Lambda_t \end{pmatrix}^{-1} = D^{-1}$$

thus, $(X'X+V_sK_sV_s')^{-1} = VD^{-1}V'$.

The matrix $V_sK_sV_s'$ in the equation (3.3.6) can be expanded and written as follows:

$$V_sK_sV_s' = (\underline{V}_1, \underline{V}_2, \ldots, \underline{V}_s)\begin{pmatrix} k_1 & & & 0 \\ & k_2 & & \\ & & \ddots & \\ 0 & & & k_s \end{pmatrix}\begin{pmatrix} \underline{V}_1' \\ \underline{V}_2' \\ \vdots \\ \underline{V}_s' \end{pmatrix}$$

$$= \begin{pmatrix} V_{11} & V_{21} & \cdots & V_{s1} \\ V_{12} & V_{22} & \cdots & V_{s2} \\ \vdots & \vdots & & \vdots \\ V_{1p} & V_{2p} & \cdots & V_{sp} \end{pmatrix}\begin{pmatrix} k_1 & & & 0 \\ & k_2 & & \\ & & \ddots & \\ 0 & & & k_s \end{pmatrix}\begin{pmatrix} V_{11} & V_{12} & \cdots & V_{1p} \\ V_{21} & V_{ss} & \cdots & V_{2p} \\ \vdots & \vdots & & \vdots \\ V_{s1} & V_{s2} & \cdots & V_{sp} \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{j=1}^{s} k_jV_{j1}^2 & \sum_{j=1}^{s} k_jV_{j1}V_{j2} & \cdots & \sum_{j=1}^{s} k_jV_{j1}V_{js} \\ \sum_{j=1}^{s} k_jV_{j2}V_{j1} & \sum_{j=1}^{s} k_jV_{j2}^2 & \cdots & \sum_{j=1}^{s} k_jV_{j2}V_{js} \\ \vdots & \vdots & & \vdots \\ \sum_{j=1}^{s} k_jV_{js}V_{j1} & \sum_{j=1}^{s} k_jV_{js}V_{j2} & \cdots & \sum_{j=1}^{s} k_jV_{js}^2 \end{pmatrix}$$

Therefore, no element of the matrix $V_sK_sV_s'$ is zero.

If we compare the equation (2.5.5) with the equation (3.3.6), we can see that in the partitioned ridge estimator, some of the elements of the matrix kA, which is added to the matrix X'X, are zero. However, in the modified principal components estimator, the elements in the matrix $V_sK_sV_s'$, added to the matrix X'X, are nonzero. In addition to this, the criteria of the determination of k or k_j are different under these two procedures.

CHAPTER IV

FURTHER RESULTS OF STATISTICAL INFERENCE

4.1 Introduction

In Chapter III, the modified principal components (MPC) estimator of the regression coefficients was established and some of its theoretical properties were developed. We shall now consider the use of the regression model for testing hypotheses about the regression coefficients and for constructing confidence intervals for the individual coefficients based on estimates obtained through the use of the MPC estimator.

4.2 The t-statistic for the Tests of Significance and a 100(1-α)% Confidence Interval for β_j

We will now consider testing the hypothesis, H_0: $\beta_j = 0$.

Define

$$t^* = \frac{(b_j^m - e_j) - m_j\beta_j}{\sqrt{MS_E \cdot h_{jj}}} \tag{4.2.1}$$

where $b_j^m = \underline{u}_j' D^{-1}V'X'\underline{y}$ is the jth element of $\underline{b}^m$,

$\underline{u}_j'$ = is the jth row vector of V,

$e_j = \widehat{E(b_j^m)} = \underline{u}_j'D^{-1}\Lambda V'\underline{\hat{\beta}}_W$,

$\underline{\hat{\beta}}_W = \underline{b} - V\Lambda^{-1}\underline{u}_j\ c_{jj}^{-1}b_j$, and b_j is the jth element of $\underline{b}$,

$h_{jj} = m_j^2 c_{jj}$ and $m_j = \underline{u}_j'D^{-1}\underline{u}_j c_{jj}^{-1}$.

We can now state the following theorem.

Theorem 4.2.1. For the model (2.1.3), with $\underline{\varepsilon} \sim MVN(0, \sigma^2 I)$, the t* statistic in (4.2.1) has central t distribution with n-p-1 degree of freedom.

Proof: (Adapted from Obenchain, 1977)

From equation (3.2.1), we have $\underline{b}^m = VD^{-1}\Lambda V'\underline{b}$. So, $b_j^m = \underline{u}_j' D^{-1}\Lambda V'\underline{b}$ and $E(b_j^m) = \underline{u}_j' D^{-1}\Lambda V'\underline{\beta}$. We can then let $e_j = \widehat{E(b_j^m)} = \underline{u}_j' D^{-1}\Lambda V'\underline{\beta}_W$, where

$$\hat{\underline{\beta}}_W = \underline{b} - H(\underline{a}_j'\underline{b} - \rho), \quad H = (X'X)^{-1}\underline{a}_j(\underline{a}_j'(X'X)^{-1}\underline{a}_j)^{-1} .$$

For the null hypothesis of $\beta_j = 0$ we have $\rho = 0$ and $\underline{a}_j' = (0,\ldots,0,1,0,\ldots,0)$ is a row vector with the value of one in the jth position and zero elsewhere. Thus, $\hat{\underline{\beta}}_W$ is the restricted least squares estimator of $\underline{\beta}$ under H_0: $\beta_j = 0$.

Hence, $H = VV'(X'X)^{-1}VV'\underline{a}_j[\underline{a}_j'(X'X)^{-1}\underline{a}_j]^{-1} = V\Lambda^{-1}\underline{u}_j c_{jj}^{-1}$ and $\hat{\underline{\beta}}_W = \underline{b} - V\Lambda^{-1}\underline{u}_j c_{jj}^{-1} b_j$. Therefore, $E(\hat{\underline{\beta}}_W) = \underline{\beta}$ under H_0. Thus, we have

$$\begin{aligned} b_j^m - e_j &= \underline{u}_j' D^{-1}\Lambda V'\underline{b} - \underline{u}_j' D^{-1}\Lambda V'\hat{\underline{\beta}}_W = \underline{u}_j' D^{-1}\Lambda V'(\underline{b} - \hat{\underline{\beta}}_W) \\ &= \underline{u}_j' D^{-1}\Lambda V'(V\Lambda^{-1}\underline{u}_j c_{jj}^{-1} b_j) \\ &= \underline{u}_j' D^{-1}\underline{u}_j c_{jj}^{-1} b_j = m_j b_j \qquad (4.2.2) \\ &\text{where } m_j = \underline{u}_j' D^{-1}\underline{u}_j c_{jj}^{-1}. \end{aligned}$$

This implies that $E(b_j^m - e_j) = m_j\beta_j$ and $Var(b_j^m - e_j) = \sigma^2 h_{jj}$, where $h_{jj} = m_j^2 c_{jj}$. Therefore, on the basis of the normality assumption, we have

$$(b_j^m - e_j) \sim N(m_j\beta_j, \sigma^2 h_{jj}), \text{ and}$$

$\dfrac{(b_j^m - e_j) - m_j\beta_j}{\sqrt{\sigma^2 h_{jj}}}$ is distributed N(0,1). Also, $\dfrac{SSE}{\sigma^2} = \dfrac{\underline{y}'MY}{\sigma^2}$, where M = [1-

$X(X'X)^{-1}X']$, is distributed as χ^2_{n-p-1}. (4.2.3)

Furthermore,

$$\underline{b}^m - \underline{e} = VD^{-1}\Lambda V'\underline{b} - VD^{-1}\Lambda V'\underline{\hat{\beta}}_W = VD^{-1}\Lambda V'(\underline{b}-\underline{\hat{\beta}}_W)$$

$$= VD^{-1}\Lambda V'H(\underline{a}_j'\underline{b}-\rho) = VD^{-1}\Lambda V'H[\underline{a}_j'(X'X)^{-1}X'\underline{y}-\rho]$$

$$= C^*\underline{y}-VD^{-1}\Lambda V'H\rho$$

where $C^* = VD^{-1}\Lambda V'H\underline{a}_j'(X'X)^{-1}X'$

and $C^*M = VD^{-1}\Lambda V'H\underline{a}_j'(X'X)^{-1}X'[1-X(X'X)^{-1}X'] = 0$. (4.2.4)

Since $\underline{y} \sim N(X\underline{\beta}, \sigma^2 I)$, $C^*M = 0$, and this implies that $C^*\underline{y}$ and $\underline{y}'M\underline{y}$ are independent. Therefore, $(\underline{b}^m - \underline{e})$ and SS_E are independent. It follows that

$$t^* = \frac{\dfrac{(b_j^m-e_j)-m_j\beta_j}{\sqrt{\sigma^2 h_{jj}}}}{\sqrt{\dfrac{SS_E}{\sigma^2(n-p-1)}}} = \frac{(b_j^m-e_j)-m_j\beta_j}{\sqrt{MS_E \cdot h_{jj}}}$$

is distributed as t_{n-p-1} and the proof is completed.

Thus, H_0: $\beta_j = 0$ can be tested using

$$t_0^* = \frac{(b_j^m-e_j)}{\sqrt{MS_E \cdot h_{jj}}}$$

as the test statistic.

Corollary 4.2.1. A $100(1-\alpha)\%$ confidence interval for β_j is given by

$$\frac{1}{m_j}(b_j^m-e_j)-\frac{1}{m_j}\,t_{\frac{\alpha}{2},n-p-1}\sqrt{MS_E\cdot h_{jj}} < \beta_j < \frac{1}{m_j}(b_j^m-e_j)+\frac{1}{m_j}\,t_{\frac{\alpha}{2},n-p-1}\sqrt{MS_E\cdot h_{jj}}. \quad (4.2.5)$$

4.3 The General Linear Hypothesis

Suppose we wish to test the general linear hypothesis, H_0: $A\underline{\beta} = \underline{\rho}$ where A is a known r×p matrix of rank r, and $\underline{\rho}$ is a known r×1 vector. Define

$$F_0^* = \frac{(A\underline{b}^m-\underline{e}^*)'[TA(X'X)^{-1}A'T']^{-1}(A\underline{b}^m-\underline{e}^*)}{r\ MS_E} \quad (4.3.1)$$

where $\underline{b}^m = VD^{-1}V'X'\underline{y}$,

$\underline{e}^* = \widehat{E(A\underline{b}^m)} = AVD^{-1}\Lambda V'\hat{\underline{\beta}}_W$, and $\hat{\underline{\beta}}_W = \underline{b} - H(A\underline{b}-\underline{\rho})$,

$T = AVD^{-1}\Lambda V'H$ and

$H = (X'X)^{-1}A'[A(X'X)^{-1}A']^{-1}$.

The distribution of F* is stated in the following theorem.

Theorem 4.3.1. Given the model (2.1.3), with $\underline{\varepsilon} \sim MVN(0, \sigma^2 I)$, the F_0^* statistic in (4.3.1) is distributed as $F_{r,n-p-1}$ under H_0: $A\underline{\beta} = \underline{\rho}$.

Proof: (Adapted from Obenchain, 1977)

$$\begin{aligned}\underline{u}^* &= A\underline{b}^m - \underline{e}^* = AVD^{-1}\Lambda V'\underline{b} - AVD^{-1}\Lambda V'\hat{\underline{\beta}}_W\\ &= AVD^{-1}\Lambda V'(\underline{b}-\hat{\underline{\beta}}_W) = AVD^{-1}\Lambda V'H(A\underline{b}-\underline{\rho})\\ &= T(A\underline{b}-\underline{\rho}) \end{aligned} \quad (4.3.2)$$

where $T = AVD^{-1}\Lambda V'H$ and $H = (X'X)^{-1}A'[A(X'X)^{-1}A']^{-1}$.

Thus, $E(\underline{u}^*) = E(A\underline{b}^m - \underline{e}^*) = T(A\underline{\beta} - \underline{\rho}) = 0$ under H_0, (4.3.3)

and $\mathrm{Var}(\underline{u}^*) = \mathrm{Var}(A\underline{b}^m - \underline{e}^*) = \mathrm{Var}\ T(A\underline{b} - \underline{\rho})$

$$= T\ \mathrm{Var}\ A\underline{b}\ T' = T \cdot \sigma^2 A(X'X)^{-1}A'T' = \sigma^2 TA(X'X)^{-1}A'T', \tag{4.3.4}$$

and $\widehat{\mathrm{Var}(\underline{u}^*)} = MS_E TA(X'X)^{-1}A'T'$ is the unbiased estimator of $\mathrm{Var}(\underline{u}^*)$. Therefore, $\underline{u}^*$ is distributed $N[T(A\underline{\beta} - \underline{\rho}),\ \sigma^2 TA(X'X)^{-1}A'T']$ under H_1, and $\underline{u}^*$ is distributed $N[0,\ \sigma^2 TA(X'X)^{-1}A'T']$ under H_0. Consequently,

$$\underline{u}^{*\prime}(\mathrm{Var}\ \underline{u}^*)^{-1}\underline{u}^* = (A\underline{b}^m - \underline{e}^*)'[\sigma^2 TA(X'X)^{-1}A'T'](A\underline{b}^m - \underline{e}^*)$$

$$= \frac{(A\underline{b}^m - \underline{e}^*)[TA(X'X)^{-1}A'T']^{-1}(A\underline{b}^m - \underline{e}^*)}{\sigma^2}$$

is distributed as $\chi^2_{r,\lambda}$, where $\lambda = \frac{1}{2}[T(A\underline{\beta} - \underline{\rho})'[\sigma^2 TA(X'X)^{-1}A'T']^{-1}[T(A\underline{\beta} - \underline{\rho})]$ under H_1 and $\lambda = 0$, under H_0.

From equation (4.2.3), we have $\frac{SS_E}{\sigma^2} = \frac{\underline{y}'M\underline{y}}{\sigma^2}$, which is distributed as χ^2_{n-p-1}, and $SS_E = \underline{\varepsilon}'M\underline{\varepsilon}$ where $M = [I - X(X'X)^{-1}X']$. (4.3.5)

Also, $\underline{u}^*\ (\mathrm{Var}\ \underline{u}^*)^{-1}\underline{u}^* = (A\underline{b} - \underline{\rho})'T'[\sigma^2 TA(X'X)^{-1}A'T']^{-1}T(A\underline{b} - \underline{\rho})$

$$= \frac{(A\underline{b} - A\underline{\beta}_0)'T'T'^{-1}A'^{-1}(X'X)A^{-1}T^{-1}T(A\underline{b} - A\underline{\beta}_0)}{\sigma^2}$$

$$= \frac{\underline{\varepsilon}'X(X'X)^{-1}A'A'^{-1}(X'X)A^{-1}A(X'X)^{-1}X'\underline{\varepsilon}}{\sigma^2} = \frac{\underline{\varepsilon}'X(X'X)^{-1}X'\underline{\varepsilon}}{\sigma^2}$$

$$= \frac{\underline{\varepsilon}'M^*\underline{\varepsilon}}{\sigma^2}\ , \text{ where } M^* = X(X'X)^{-1}X'.$$

We have $\underline{\varepsilon}$ distributed as $N(0,\ \sigma^2 I)$, and $M^*M = X(X'X)^{-1}X'[I - X(X'X)^{-1}X'] = 0$. This implies that $\underline{\varepsilon}'M^*\underline{\varepsilon}$ and $\underline{\varepsilon}'M\underline{\varepsilon}$ are independent. Therefore,

$\underline{u}^*$ (Var $\underline{u}^*)^{-1}\underline{u}^*$ and SS_E are independent. (Graybill, 1976)

It follows that

$$F_0^* = \frac{\dfrac{(A\underline{b}^m-e^*)'[TA(X'X)^{-1}A'T']^{-1}(A\underline{b}^m-\underline{e}^*)}{\sigma^2 r}}{\dfrac{SS_E}{\sigma^2(n-p-1)}} = \frac{(A\underline{b}^m-\underline{e}^*)'[TA(X'X)^{-1}A'T']^{-1}(A\underline{b}^m-\underline{e}^*)}{r\ MS_E}$$

is distributed as $F_{r,n-p-1}$ under H_0 and the proof is complete.

Thus, the F_0^* statistic in (4.3.1) can be used as the test statistic for testing H_0: $A\underline{\beta} = \underline{\rho}$.

4.4 Confidence Region for $A\underline{\beta}$

We now may construct a confidence region for $A\underline{\beta}$.

Theorem 4.4.1. With the model (2.1.3), and with $\underline{\varepsilon} \sim MVN(0, \sigma^2 I)$, we have that

$$P\{\frac{[(A\underline{b}^m-\underline{e}^*)-T(A\underline{b}-\underline{\rho})]'[TA(X'X)^{-1}A'T']^{-1}[(A\underline{b}^m-\underline{e}^*)-T(A\underline{\beta}-\underline{\rho})]}{r\ MS_E} < F_{\alpha,r,n-p-1}\} = 1-\alpha, \quad (4.4.1)$$

Proof: (Adapted from Obenchain, 1977)

From equations (4.3.2) and (4.3.3), and by setting

$$R^* = \underline{u}^* - E(\underline{u}^*), \text{ we have } \underline{R}^* = \underline{u}^* - E(\underline{u}^*) = T(Ab - A\underline{\beta}) . \quad (4.4.2)$$

Thus, $E(\underline{R}^*) = T\ E(A\underline{b} - A\underline{\beta}) = 0$, and

$$\text{Var}(\underline{R}^*) = \text{Var}(\underline{u}^* - E\ \underline{u}^*) = \text{Var}\ u^* = \sigma^2 TA(X'X)^{-1}A'T'. \quad (4.4.3)$$

Then, $\underline{R}^*$ is distributed $N[\underline{0}, \sigma^2 TA(X'X)^{-1}A'T']$. Therefore,

$$\underline{R}^*(\text{Var } \underline{R}^*)^{-1}\underline{R}^* = (\underline{u}^* - E\underline{u}^*)'[\text{Var}(\underline{u}^* - E\underline{u}^*)]^{-1}(\underline{u}^* - E\underline{u}^*)$$

$$= \frac{[(A\underline{b}^m - \underline{e}^*) - T(A\underline{\beta} - \underline{\rho})]'[TA(X'X)^{-1}A'T']^{-1}[(A\underline{b}^m - \underline{e}^*) - T(A\underline{\beta} - \underline{\rho})]}{\sigma^2}$$

is distributed as $\chi^2(r, \lambda)$, where $\lambda = \frac{1}{2}(\underline{0})'[\sigma^2 TA(X'X)^{-1}A'T']^{-1}(\underline{0}) = 0$.

From equation (4.2.3), we have $\frac{SS_E}{\sigma^2} = \frac{\underline{y}'M\underline{y}}{\sigma^2}$, which is distributed as χ^2_{n-p-1}. Also, from equation (4.3.5), we have $SS_E = \underline{\varepsilon}'M\underline{\varepsilon}$, where $M = [I - X(X'X)^{-1}X']$.

Therefore, referring to equations (4.4.2) and (4.4.3), we have

$$\underline{R}^*\ (\text{Var } \underline{R}^*)^{-1}\ \underline{R}^* = (A\underline{b} - A\underline{\beta})'T'[\sigma^2 TA(X'X)^{-1}A'T']^{-1}T(A\underline{b} - A\underline{\beta})$$

$$= \frac{(A\underline{b} - A\underline{\beta})'T'T'^{-1}A'^{-1}(X'X)A^{-1}T^{-1}T(A\underline{b} - A\underline{\beta})}{\sigma^2} = \frac{\underline{\varepsilon}'X(X'X)^{-1}X'\underline{\varepsilon}}{\sigma^2}$$

$$= \frac{\underline{\varepsilon}'M^*\underline{\varepsilon}}{\sigma^2}\ , \text{ where } M^* = X(X'X)^{-1}X'\ ,$$

since $\underline{b} = (X'X)^{-1}X'\underline{y} = (X'X)^{-1}X'(X\underline{\beta}+\underline{\varepsilon}) = \underline{\beta} + (X'X)^{-1}X'\underline{\varepsilon}$, and $A(\underline{b}-\underline{\beta}) = A(X'X)^{-1}X'\underline{\varepsilon}$. Then, $M^*M = X(X'X)^{-1}[I - X(X'X)^{-1}X'] = 0$. This implies that $\underline{\varepsilon}'M^*\underline{\varepsilon}$ and $\underline{\varepsilon}'M\underline{\varepsilon}$ are independent and hence $\underline{R}^{*\prime}(\text{Var } \underline{R}^*)^{-1}\ \underline{R}^*$ and SS_E are independent. Therefore,

$$F^* = \frac{\dfrac{[(A\underline{b}^m - \underline{e}^*) - T(A\underline{\beta} - \underline{\rho})]'[TA(X'X)^{-1}A'T']^{-1}[(A\underline{b}^m - \underline{e}^*) - T(A\underline{\beta} - \underline{\rho})]}{r\sigma^2}}{\dfrac{SS_E}{\sigma^2(n-p-1)}}$$

$$= \frac{[(A\underline{b}^m - \underline{e}^*) - T(A\underline{\beta} - \underline{\rho})]'[TA(X'X)^{-1}A'T']^{-1}[(A\underline{b}^m - \underline{e}^*) - T(A\underline{\beta} - \underline{\rho})]}{r\ MS_E}$$

is distributed as $F_{r,\, n-p-1}$.

It follows that $100(1-\alpha)\%$ confidence region for $A\underline{\beta}$ is given by the set of all $\underline{\beta}$ for which

$$\left\{\frac{[(A\underline{b}^{m}-e^{*})-T(A\underline{\beta}-\underline{\rho})]'[TA(X'X)^{-1}A'T']^{-1}[(A\underline{b}^{m}-\underline{e}^{*})-T(A\underline{\beta}-\underline{\rho})]}{r\ MS_{E}} <\right.$$

$$\left. F_{\alpha,r,n-p-1}\right\}.$$

CHAPTER V

SIMULATIONS

5.1 Introduction

In this chapter, Monte Carlo techniques are employed to evaluate the effectiveness of the modified principal components estimator in reducing the damaging effects of multicollinearity. The estimator will be compared with the least squares estimator and the principal components estimator.

The simulations were conducted on an IBM-370 computing using a program written by the author, which incorporated subroutines from the IBM Scientific Subroutine Package (1970).

5.2 Construction

From equations (3.2.7), (3.2.8), and (3.2.10), we see that for constant k_j, the magnitude of improvement of MSE($\hat{\underline{\beta}}_{mpc}$) over MSE($\hat{\underline{\beta}}$) or MSE($\hat{\underline{\beta}}_{pc}$) depends on the magnitude of $(\frac{V_j'\beta}{\sigma})^2$ relative to ℓ_j for i = 1, 2, ..., s. For this simulation, we controlled the following parameters:

n = number of observations,

p = number of regressor variables,

ℓ_j = smallest latent roots of X'X, j = 1, 2, ..., s, s = 1, or 2,

$$\phi_j = \text{orientation} = \frac{\underline{V}_j'\underline{\beta}}{\sqrt{\underline{\beta}'\underline{\beta}}}, \; j = 1, 2, \ldots, s, \; s = 1, \text{ or } 2,$$

$$\rho = \text{signal-to-noise ratio} = \frac{\underline{\beta}'\underline{\beta}}{\sigma^2}.$$

Six X matrices were generated for this study. They are as follows:

first X matrix: with $n = 30$, $p = 10$, $\ell_1 = 0.0105992$,

second X matrix: with $n = 30$, $p = 10$, $\ell_1 = 0.005009$,

third X matrix: with $n = 30$, $p = 10$, $\ell_1 = 0.0490105$,

fourth X matrix: with $n = 30$, $p = 5$, $\ell_1 = 0.013574$,

fifth X matrix: with $n = 30$, $p = 15$, $\ell_1 = 0.005374$,

sixth X matrix: with $n = 30$, $p = 10$, $\ell_1 = 0.0178895$

$\ell_2 = 0.0303788$, with two multicollinearities.

The strength of the multicollinearities is indicated by the size of the smallest latent roots of X'X. For convenience, the X matrices are identified by ℓ_1, the smallest latent root. The complete set of the latent roots of X'X for the six X matrices is given in Table 5.1,where the latent roots have been arranged in ascending order of magnitude. Each matrix represents a different degree of multicollinearity. The first X matrix through the fifth X matrix contains a single multicollinearity involving the first three regressor variables. The sixth X matrix contains two multicollinearities. It consists of eight columns of random observations, and the first column is a linear combination of the third column and the fourth column plus a random error vector, and the second column is a linear combination of the fifth column and the sixth column plus a random error vector.

The latent vectors associated with small latent roots of X'X for each matrix are given in Table 5.2. In this table, the latent vectors corresponding to small latent roots reveal that the multicollinearity involves the first three regressor variables for the first X matrix through the fifth X matrix and that the multicollinearity involves the first six regressor variables for the sixth X matrix.

The values of $\underline{\beta}$ used in the simulation are given in the Tables 5.3a and 5.3b. They were selected in order to obtain various orientations of $\underline{\beta}$ with $\underline{V}_1$. The degree of parallelism is used as the measure of orientation. The orientations of $\underline{\beta}$ and $\underline{V}_j$ are given below:

$$\cos\theta = \phi_j = \frac{\underline{V}_j'\underline{\beta}}{\sqrt{\underline{V}'\underline{V}_j}\sqrt{\underline{\beta}'\underline{\beta}}} = \frac{\underline{V}_j'\underline{\beta}}{\sqrt{\underline{\beta}'\underline{\beta}}} \quad j = 1, 2, \ldots, s$$

where θ is angle between $\underline{\beta}$ and $\underline{V}_j$.

The choices of $\underline{\beta}$ in Tables 5.3a and 5.3b were made in order to compare the performance of the estimators for various orientations of $\underline{\beta}$ ranging from parallel to $\underline{V}_1$ to orthogonal to $\underline{V}_1$. A value of $|\phi_1|$ near one indicates that $\underline{\beta}$ is nearly parallel to $\underline{V}_1$. On the other hand, a value of $|\phi_1|$ near zero indicates that $\underline{\beta}$ is nearly orthogonal to $\underline{V}_1$. The measure is scaled so that it is not affected by the length of $\underline{\beta}$. The signal-to-noise ratios, $\rho = \frac{\underline{\beta}'\underline{\beta}}{\sigma^2}$, are given in the Table 5.3a.

One hundred samples of size 30 of response variables were generated for each X matrix, using each of nine choices of orientation and five choices of signal-to-noise ratios according to the model:

$$\underline{y} = X\underline{\beta} + \underline{\varepsilon} \;, \quad \underline{\varepsilon} \sim NVM(0, \sigma^2 I_n) \;.$$

The random error terms were generated by GAUSS, a subroutine from the IBM Scientific Subroutine Package.

5.3 Evaluation

For each sample so generated, we obtain the estimates of $\underline{\beta}$ for the least squares estimator, the principal components estimator, and the modified principal components estimator. In order to evaluate the performance of the modified principal components estimator and to compare its performance with that of the least squares estimator and that of the principal components estimator, the total estimated squared error was calculated for each sample using

$$m_{i1} = SE(\hat{\underline{\beta}}) = \sum_{j=1}^{p} (\hat{\beta}_j - \beta_j)^2, \; i = 1, 2, \ldots, 100$$

$$m_{i2} = SE(\hat{\underline{\beta}}_{pc}) = \sum_{j=1}^{p} (\hat{\beta}_{jpc} - \beta_j)^2, \; i = 1, 2, \ldots, 100$$

$$m_{i3} = SE(\hat{\underline{\beta}}_{mpc}) = \sum_{j=1}^{p} (\hat{\beta}_{jmpc} - \beta_j)^2, \; i = 1, 2, \ldots, 100$$

for the least squares estimator, the principal components estimator and the modified principal components estimator respectively.

We also determined the average value, the standard deviation, the maximum and the minimum of the 100 total estimated squared errors for each combination of orientation and signal-to-noise ratio. We refer to these average estimated total squared errors as estimated mean squared errors. These summary statistics for the 100 total estimated squared errors for the least squares estimates, the principal components estimates, and the modified principal components estimates are presented in Tables 5.4 through 5.9. Within each table, various values of the orientation $|\phi|$ are included, ranging from zero to one.

For the principal components procedure used here, we always delete the components associated with the small latent roots of X'X, since Mansfield (1975) has demonstrated that the use of the response variables as an aid to determine which terms to delete does not consistently indicate the proper terms to delete.

Table 5.16 through Table 5.21 contain the number of times (out of 100) that the total estimated squared error of the modified principal components estimator is less than that of the least squares estimator and the principal components estimator respectively for the first X matrix through the sixth X matrix.

Table 5.22 (a through f) contains the variance inflation factors for both the least squares estimator and the modified principal components estimator.

5.4 Conclusions

In this section, the performance of each estimator will be discussed and the situations where each does the best will be mentioned.

From the simulation results presented in Table 5.4 through Table 5.9, we can make the following comments.

The least squares estimator is very unstable as can be seen from the maximum values and the minimum values of the 100 total estimated squared errors in each table. Among each set of 100 runs there is considerable variation of the total estimated squared errors, with some being extremely large and some very small. For example, for the X matrix with a minimum latent root of 0.005009 and the signal-to-noise ratio of 100, the minimum total estimated squared error is 10.35 and the maximum total estimated squared error is 2248.02. This indicates that the least squares estimator is very unreliable when multicollinearity is present. Hence, predictions obtained from least squares estimates will also be very unreliable.

The least squares estimator has larger estimated mean squared error than either of the other two estimators for most of the orientations when the signal-to-noise ratio is small. For example, when ℓ_1 = 0.0105992 and ρ = 100, the MSE($\underline{\hat{\beta}}$) is 110.73, which is larger than either MSE($\underline{\hat{\beta}}_{pc}$) or MSE($\underline{\hat{\beta}}_{mpc}$) in most of the orientations.

For the least squares estimates, the average of 100 runs is affected a great deal by the size of the smallest latent root. This is because the least squares estimator in equation (2.2.2) still contains the smallest latent root. For example, when ℓ_1 = 0.0490105, and ρ = 100, the average of 100 total estimated squared error is 36.77; when ℓ_1 = 0.005009, and ρ = 100, the average of 100 total estimated squared error is 254.18. However, the average of 100 runs for the least squares estimates is not affected by the orientation. This is because the least squares estimator is unbiased and its mean squared error in equation (2.2.5) does not contain a ($\underline{V}_1'\underline{\beta}$) term. This can be seen from Tables 5.4 through 5.9 which show that as the orientation increases, the average values of the 100 total estimated squared errors for the 100 samples remain nearly constant across orientations.

For the principal components estimates, the average values of the total estimated squared errors for the 100 samples increase as the orientation increases. This is caused by the addition of the bias in the mean squared error function. However, these averages are not affected by the changes in the magnitude of the smallest latent root. This is due to the fact that the smallest latent root has been removed from the estimator, as seen in equation (2.5.8).

The principal components procedure performs well in those cases where $\underline{\beta}$ is orthogonal or nearly orthogonal to $\underline{V}_1$. This is to be expected since the deletion of the first component reduces the variance part of $MSE(\hat{\underline{\beta}}_{pc})$ in equation (2.5.12) by an amount equal to $\sigma^2 \ell_1^{-1}$, without adding any bias. However, it performs very poorly when $\underline{\beta}$ is parallel or nearly parallel to $\underline{V}_1$. This can be seen by examining the estimated mean squared error for the cases where $\phi_1 > 0.625$ in Tables 5.4 through 5.9.

The modified principal components estimator has a smaller estimated mean squared error than the other two estimators over a wide range of middle values of orientations as indicated by the notation "*" in Tables 5.4 through 5.9. For those cases for which modified principal components regression does not perform the best, it is always the second best and is always close to the best. For example, when $\ell_1 = 0.005009$, $\rho = 100$, and $\phi_1 = 0$, the estimated mean squared error of the modified principal components estimate is 33.96, and it is much closer to that of the principal components estimate (10.706) than that of the least squares estimate (254.182).

When $\underline{\beta}$ is parallel or nearly parallel to $\underline{V}_1$, the modified principal components estimator produces substantially smaller estimated mean squared error than the principal components estimator does. For example, when $\ell_1 = 0.0105992$, $\rho = 100$, and $\phi_1 = 1.0$, the MSE($\hat{\underline{\beta}}_{mpc}$) is 59.47 while MSE($\hat{\underline{\beta}}_{pc}$) is 111.02. This due to the fact that the bias in MSE($\hat{\underline{\beta}}_{mpc}$) has been tremendously reduced since the weight in each term of the bias part of MSE($\hat{\underline{\beta}}_{mpc}$) is $(\frac{k_j}{\ell_j+k_j})^2$ which is less than one and will be close to zero when k_j is small.

For the X matrix with the smallest latent root of 0.0490105, principal components regression performs noticeably worse than the least squares estimator or the modified principal components estimator for all the signal-to-noise ratios when ϕ_1 is greater than 0.625. This is due to the addition of the very large bias in the mean squared error function. The modified principal components procedure effectively reduces the bias and hence effectively reduces the mean squared error. This phenomenon can be seen by examining the estimated mean squared error for the cases where $\phi_1 > 0.625$ in the Table 5.6 (a through f). For example, when $\ell_1 = 0.0490105$, $\rho = 1000$, and $\phi_1 = 1.0$, the MSE($\hat{\underline{\beta}}_{mpc}$) is 3.47, while the MSE($\hat{\underline{\beta}}_{pc}$) is 101.05. The use of the principal components estimator in this case could be very misleading because the estimates are very biased.

The standard deviations of the 100 total estimated squared errors for the least squares estimator are very large relative to those of the other two estimators. This indicates the large variability of the

the 100 total estimated squared errors of the least squares estimator with the multicollinear data; that is, least squares estimates of regression coefficients fluctuate from sample to sample and are very unstable. Thus, least squares estimates of regression coefficients can be very unreliable when multicollinearity exists in the matrix of regressor variables.

For the X matrix with 5 variables in Table 5.7 (a through f) and the X matrix with 15 variables in Table 5.8 (a through f), we obtain similar conclusions as for the X matrix with 10 regressor variables.

The simulation is also extended to the case of two small latent roots as displayed in Table 5.9 (a through f). For this table, all possible combinations of ϕ_j are used where ϕ_j was chosen to be 0, 0.25, 0.5, 0.75, and 1.0 for j = 1, 2. The simulation results of this X matrix with two small latent roots are similar to those of the X matrix with one small latent root. Thus, the comments, made above for the summary statistics, also apply to the X matrix with two multicollinearities. For example, in this matrix, when $\rho = 200$ or 300, the modified principal components estimator produces more satisfactory results than either of the other two estimators in 12 out of 17 cases. (See Table 5.9.)

Figures 5.1 through 5.6 show the relationships between the estimated mean squared error of the three estimators (the least squares estimator, the principal components estimator, the modified principal components estimator) and the orientation, ϕ_1, of $\underline{\beta}$ to $\underline{V}_1$ for the first X matrix ($\ell_1 = 0.0105992$). These figures, which are plotted by using information from Table 5.4 (a through f) provide a clear picture of the performance of the three estimators in the first X matrix. The comments made above concerning the estimated mean squared errors in Table 5.4 can also apply to these figures. They can be outlined briefly as follows. The estimated MSE($\hat{\underline{\beta}}_{pc}$) and the estimated MSE($\hat{\underline{\beta}}_{mpc}$) increases as ϕ_1 increases, while the estimated MSE($\hat{\underline{\beta}}$) remains constant as ϕ_1 increases. The estimated MSE($\hat{\underline{\beta}}_{pc}$) is smaller than either the estimated MSE($\hat{\underline{\beta}}$) or the estimated MSE($\hat{\underline{\beta}}_{mpc}$) when ϕ_1 is zero or near zero. However, it increases rapidly as ϕ_1 increases and it becomes larger than the estimated MSE($\hat{\underline{\beta}}_{mpc}$). (This change occurs when the value of ϕ_1 is between 0.2 and 0.5.) The estimated MSE($\hat{\underline{\beta}}_{mpc}$) is much smaller than the estimated MSE($\hat{\underline{\beta}}_{pc}$) when ϕ_1 is equal to one or near one. Overall, these figures indicate that the estimated MSE($\hat{\underline{\beta}}_{mpc}$) is smaller than the estimated MSE($\hat{\underline{\beta}}$) and the estimated MSE($\hat{\underline{\beta}}_{pc}$) over a wide range values of ϕ_1 and ρ.

Table 5.10 through Table 5.15 contain the theoretical mean squared errors of the least squares estimator, the principal components estimator, and the modified principal components estimator. These theoretical mean squared errors are obtained by using the equations (2.2.5), (2.5.12), and (3.2.4) for the least squared estimator, the principal components estimator, and the modified principal components estimator respectively. They provide a further check on the conclusions stated above. The fact that the estimated mean squared errors in the Table 5.4 through the Table 5.9 are very close to the corresponding theoretical mean squared errors in Table 5.10 through

Table 5.15 indicates that our sample size of 100 is large enough to compute the estimated mean squared error. The conclusions drawn above for estimated mean squared errors (the average values of the total estimated squared errors for the 100 samples) of least squares estimates, principal components estimates, and modified principal components estimates, are unaltered by these corresponding theoretical mean squared errors.

In Tables 5.16 through 5.21, for each set of 100 samples, counts are made of the number of samples out of 100 for which the total estimated squared error of the modified principal components estimator is smaller than that of the principal components estimator and the least squares estimator respectively. These counts presented in Tables 5.16 through 5.21, can be used to support the conclusions made above. For example, in Table 5.18, when $\rho = 100$, $MSE(\hat{\underline{\beta}}_{mpc})$ is less than $MSE(\hat{\underline{\beta}})$ more than 55% of the time for all the orientation, and the $MSE(\hat{\underline{\beta}}_{mpc})$ is less than the $MSE(\hat{\underline{\beta}}_{pc})$ more than 53% of the time for $\phi_1 > 0.375$; when $\rho = 200$, $MSE(\hat{\underline{\beta}}_{mpc})$ is less than $MSE(\hat{\underline{\beta}})$ more than 52% of the time for all the orientations and the $MSE(\hat{\underline{\beta}}_{mpc})$ is less than the $MSE(\hat{\underline{\beta}}_{pc})$ more than 50% of the time for $\phi_1 \geq 0.25$. Principal components regression performs better only for the cases when $\underline{\beta}$ is orthogonal or nearly orthogonal to $\underline{V}_1$.

Further results of the simulation can be seen by examining the variance inflation factors corresponding to each of regressor variables for least squares and for the modified principal components procedure respectively. As can be seen in the Table 5.22 (a through f), the modified principal components procedure effectively reduces the variance inflation factors for those regressor variables involved in the multicollinearity, while its effect on the remaining variance inflation factors is negligible.

For each X matrix, the k_j's, determined by the variance inflation factor criterion, are given below:

first X matrix: $k_1 = 0.014$,

second X matrix: $k_1 = 0.0112$,

third X matrix: $k_1 = 0.003$,

fourth X matrix: $k_1 = 0.0126$,

fifth X matrix: $k_1 = 0.0118$,

sixth X matrix: $k_1 = 0.0122$,

$k_2 = 0.0113$.

5.5 Tables and Figures

Table 5.1 Latent Roots of X'X

	1st X Matrix	2nd X Matrix	3rd X Matrix	4th X Matrix	5th X Matrix	6th X Matrix
ℓ_1	0.010599	0.0050099	0.049010	0.013574	0.005374	0.017889
ℓ_2	0.269799	0.373841	0.379009	0.720962	0.189588	0.030378
ℓ_3	0.474545	0.586275	0.584896	1.024013	0.337236	0.484500
ℓ_4	0.606893	0.780317	0.775352	1.168488	0.403558	0.606272
ℓ_5	0.741168	0.902358	0.902387	2.072935	0.482485	0.897948
ℓ_6	1.054171	0.945521	0.943341		0.515671	1.029147
ℓ_7	1.168946	1.079109	1.077696		0.778568	1.160969
ℓ_8	1.475028	1.445634	1.449540		0.936636	1.299992
ℓ_9	1.647505	1.530352	1.536986		1.017790	2.118703
ℓ_{10}	2.551177	2.351444	2.301644		1.100485	2.354039
ℓ_{11}					1.378034	
ℓ_{12}					1.471312	
ℓ_{13}					1.592427	
ℓ_{14}					2.052575	
ℓ_{15}					2.737864	

1st X Matrix: $X_{30\times 10}$

2nd X Matrix: $X_{30\times 10}$

3rd X Matrix: $X_{30\times 10}$

4th X Matrix: $X_{30\times 5}$

5th X Matrix: $X_{30\times 15}$

6th X Matrix: $X_{30\times 10}$

Table 5.2 The Latent Vectors Associated With the Smallest Latent Root(s) of Each Matrix

1st X Matrix:

$\underline{v}_1' = (+0.735991, -0.628152, +0.246981, -0.015178, -0.009951,$
$+0.012783, +0.001666, +0.007821, +0.040456, -0.023332)$

2nd X Matrix:

$\underline{v}_1' = (+0.704443, -0.531947, -0.467219, +0.012128, +0.022515,$
$-0.018542, -0.024426, -0.009148, -0.016042, +0.023637)$

3rd X Matrix:

$\underline{v}_1' = (+0.698568, -0.538589, -0.439688, +0.045079, +0.077228,$
$-0.065633, -0.078594, -0.021569, -0.052244, +0.083175)$

4th X Matrix:

$\underline{v}_1' = (+0.699432, -0.415915, -0.580736, +0.001846, +0.023435)$

5th X Matrix:

$\underline{v}_1' = (-0.711466, +0.559907, +0.412157, +0.043568, +0.030222,$
$+0.046455, +0.026068, +0.003482, +0.013782, +0.033798,$
$-0.036655, +0.009884, -0.025930, -0.007693, +0.035666)$

6th X Matrix:

$\underline{v}_1' = (-0.212935, -0.669694, +0.103216, +0.148562, +0.559638,$
$+0.397542, +0.008405, +0.017728, +0.024430, -0.034971)$

$\underline{v}_2' = (+0.700238, -0.184906, -0.426562, -0.486455, +0.214913,$
$+0.061895, -0.054415, -0.014938, -0.033233, +0.050689)$

Table 5.3a The Orientations of $\underline{\beta}$ to $\underline{V}_1$ (for One Small Latent Root) and Signal-to-Noise Ratio

Orientation

ϕ_1	$\underline{\beta}$
0	$\underline{\beta} = 10\underline{V}_{10}$
0.125	$\underline{\beta} = 1.25\underline{V}_1 + 9.9215\underline{V}_{10}$
0.25	$\underline{\beta} = 2.5\underline{V}_1 + 9.6824\underline{V}_{10}$
0.375	$\underline{\beta} = 3.75\underline{V}_1 + 9.2702\underline{V}_{10}$
0.5	$\underline{\beta} = 5(\underline{V}_1 + \underline{V}_2 + \underline{V}_9 + \underline{V}_{10})$
0.625	$\underline{\beta} = 6.25\underline{V}_1 + 7.8062\underline{V}_{10}$
0.75	$\underline{\beta} = 7.5\underline{V}_1 + 6.6143\underline{V}_{10}$
0.875	$\underline{\beta} = 8.75\underline{V}_1 + 4.8412\underline{V}_{10}$
1.0	$\underline{\beta} = 10\underline{V}_1$

Signal-to-Noise Ratio

ρ	σ^2
100	$\sigma^2 = 1^2 = 1.000$
200	$\sigma^2 = (0.7071)^2 = 0.4999$
300	$\sigma^2 = (0.5774)^2 = 0.3334$
500	$\sigma^2 = (0.4472)^2 = 0.1999$
750	$\sigma^2 = (0.3652)^2 = 0.1334$
1000	$\sigma^2 = (0.3162)^2 = 0.0999$

Table 5.3b The Orientations of $\underline{\beta}$ to $\underline{V}_1$ and $\underline{\beta}$ to $\underline{V}_2$

(for Two Small Latent Roots)

ϕ_1	ϕ_2	$\underline{\beta}$
0	0	$\underline{\beta} = 10\underline{V}_{10}$
0	0.25	$\underline{\beta} = 2.5\underline{V}_2 + 9.6824\underline{V}_{10}$
0	0.5	$\underline{\beta} = 5\underline{V}_2 + 8.6602\underline{V}_{10}$
0	0.75	$\underline{\beta} = 7.5\underline{V}_2 + 6.6143\underline{V}_{10}$
0	1	$\underline{\beta} = 10\underline{V}_2$
0.25	0	$\underline{\beta} = 2.5\underline{V}_1 + 9.6824\underline{V}_{10}$
0.25	0.25	$\underline{\beta} = 2.5\underline{V}_1 + 2.5\underline{V}_2 + 9.3541\underline{V}_{10}$
0.25	0.5	$\underline{\beta} = 2.5\underline{V}_1 + 5\underline{V}_2 + 8.2916\underline{V}_{10}$
0.25	0.75	$\underline{\beta} = 2.5\underline{V}_1 + 7.5\underline{V}_2 + 6.1237\underline{V}_{10}$
0.5	0	$\underline{\beta} = 5\underline{V}_1 + 8.6602\underline{V}_{10}$
0.5	0.25	$\underline{\beta} = 5\underline{V}_1 + 2.5\underline{V}_2 + 8.2916\underline{V}_{10}$
0.5	0.5	$\underline{\beta} = 5\underline{V}_1 + 5\underline{V}_2 + 7.0711\underline{V}_{10}$
0.75	0	$\underline{\beta} = 7.5\underline{V}_1 + 6.6144\underline{V}_{10}$
0.75	0.25	$\underline{\beta} = 7.5\underline{V}_1 + 2.5\underline{V}_2 + 6.1237\underline{V}_{10}$
0.75	0.5	$\underline{\beta} = 7.5\underline{V}_1 + 5\underline{V}_2 + 4.3301\underline{V}_{10}$
1.0	0	$\beta = 10\underline{V}_1$

Table 5.4 Summary Statistics of 100 Total Estimated Squared Errors for LS, PC, MPC; 1st X Matrix ($X_{30\times10}$, ℓ_1 = 0.0105992)

(a) ρ = 100

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0.0	110.7346	11.0276	29.5388	Ave.	0.625	110.7346	50.0894	40.6542*	Ave.
	160.9329	6.8412	31.0220	S.D.		160.8346	6.8431	39.5269	S.D.
	1143.5200	35.6419	216.3083	Max.		1143.5510	74.7040	260.2416	Max.
	3.5738	2.5996	2.9992	Min.		3.5734	41.6618	4.4743	Min.
0.125	110.7353	12.5901	29.7378	Ave.	0.75	110.7349	67.2766	45.9139*	Ave.
	160.9357	6.8413	30.5184	S.D.		160.9356	6.8463	43.6914	S.D.
	1143.5670	37.2044	196.1363	Max.		1143.5650	91.8913	282.1979	Max.
	3.5738	4.1621	4.0203	Min.		3.5735	58.8492	4.2386	Min.
0.25	110.7351	17.2775	30.9487	Ave.	0.875	110.7347	87.5888	52.1856*	Ave.
	160.9351	6.8413	31.2387	S.D.		160.9351	6.8510	48.2805	S.D.
	1143.5520	41.8917	200.4584	Max.		1143.5580	112.2035	305.1604	Max.
	3.5737	8.8495	3.8198	Min.		3.5737	79.1616	4.9964	Min.
0.375	110.7345	25.0900	33.1716	Ave.	1.0	110.7337	111.0259	59.4708*	Ave.
	160.9346	6.8414	33.1035	S.D.		160.9311	6.8519	53.1855	S.D.
	1143.5640	49.7042	219.3726	Max.		1143.5060	135.6406	329.1418	Max.
	3.5736	16.6620	3.7353	Min.		3.5740	102.5989	3.7508	Min.
0.5	110.7348	36.0274	36.4071	Ave.					
	160.9349	6.8406	35.9358	S.D.					
	1143.5450	60.6419	239.3078	Max.					
	3.5736	27.5995	4.0679	Min.					

Table 5.4 (Continued) 1st X Matrix ($X_{30\times10}$, $\ell_1 = 0.0105992$)

(b) $\rho = 200$

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	55.3663	5.5137	14.7690	Ave.	0.625	55.3670	44.5757	26.3377*	Ave.
	80.4666	3.4206	15.5110	S.D.		80.4679	3.4206	24.3087	S.D.
	571.7119	17.8207	108.1563	Max.		571.7871	56.8828	153.4104	Max.
	1.7867	1.2990	1.4995	Min.		1.7862	40.3620	2.5442	Min.
0.125	55.3662	7.0761	15.0577	Ave.	0.75	55.3666	61.7627	31.6832*	Ave.
	80.4652	3.4205	15.3343	S.D.		80.4673	3.4296	27.8252	S.D.
	571.7568	19.3831	94.9181	Max.		571.7812	74.0702	170.5645	Max.
	1.7869	2.8623	2.1698	Min.		1.7865	57.5494	1.6629	Min.
0.25	55.3665	11.7636	16.3586	Ave.	0.875	55.3665	82.0747	38.0449*	Ave.
	80.4670	3.4206	16.3543	S.D.		80.4668	3.4415	31.5487	S.D.
	571.7770	24.0705	108.0252	Max.		571.7797	94.3825	188.7281	Max.
	1.7867	7.5498	1.8389	Min.		1.7864	77.8618	1.7228	Min.
0.375	55.3666	19.5760	18.6714*	Ave.	1.0	55.3660	105.5119	45.4202*	Ave.
	80.4676	3.4209	18.3719	S.D.		80.4653	3.4510	35.4151	S.D.
	571.7846	31.8830	122.1408	Max.		571.7573	117.8196	207.9100	Max.
	1.7861	15.3622	2.0367	Min.		1.7867	101.2991	2.7951	Min.
0.5	55.3664	30.5134	21.9966*	Ave.					
	80.4665	3.4201	21.1035	S.D.					
	571.7697	42.8207	137.2696	Max.					
	1.7861	26.2997	2.1059	Min.					

Table 5.4 (Continued) 1st X Matrix ($X_{30\times10}$, ℓ_1 = 0.0105992)

(c) ρ = 300

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	36.9184	3.6764	9.8480	Ave.	0.625	36.9185	42.7385	21.6114*	Ave.
	53.6548	2.2808	10.3427	S.D.		53.6554	2.2818	18.8527	S.D.
	381.2587	11.8828	72.1189	Max.		381.2717	50.9449	115.1833	Max.
	1.1913	0.8667	0.9999	Min.		1.1909	39.9290	1.0773	Min.
0.125	36.9185	5.2390	10.1764	Ave.	0.75	36.9185	59.9255	27.0006*	Ave.
	53.6550	2.2809	10.3108	S.D.		53.6558	2.2945	21.9233	S.D.
	381.2587	13.4453	65.1948	Max.		381.2758	68.1324	130.2124	Max.
	1.1911	2.4292	1.3386	Min.		1.1910	57.1163	1.3206	Min.
0.25	36.9185	9.9264	11.5169	Ave.	0.875	36.9186	80.2376	33.4021*	Ave.
	53.6551	2.2809	11.4380	S.D.		53.6554	2.3102	25.1202	S.D.
	381.2678	18.1326	76.1721	Max.		381.2700	88.4447	146.2530	Max.
	1.1921	7.1167	1.3467	Min.		1.1910	77.4287	2.5760	Min.
0.375	36.9185	17.7388	13.8695*	Ave.	1.0	36.9182	103.6747	40.8173	Ave.
	53.6546	2.2812	13.4357	S.D.		53.6541	2.2390	28.4012	S.D.
	381.2595	25.9451	88.1637	Max.		381.2521	111.8818	163.3084	Max.
	1.1909	14.9291	1.4548	Min.		1.1913	100.8661	2.5529	Min.
0.5	36.9184	28.6762	17.2346*	Ave.					
	53.6552	2.2804	15.9811	S.D.					
	381.2658	36.8828	101.1704	Max.					
	1.1910	25.8665	1.6354	Min.					

Table 5.4 (Continued) 1st X Matrix ($X_{30\times10}$, ℓ_1 = 0.0105992)

(d) ρ = 500

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	22.1262	2.2034	5.9021	Ave.	0.625	22.1461	41.2675	17.8705*	Ave.
	32.1569	1.3669	6.1987	S.D.		32.1865	1.3656	14.0171	S.D.
	228.5037	7.1216	43.2234	Max.		228.7199	46.1902	82.4170	Max.
	0.7138	0.5194	0.5992	Min.		0.7143	39.5821	1.0599	Min.
0.125	22.1460	3.7678	6.2757	Ave.	0.75	22.1263	58.4524	23.2947	Ave.
	32.1856	1.3682	6.3156	S.D.		32.1572	1.3853	16.5263	S.D.
	228.7094	8.6905	40.9612	Max.		228.5112	63.3712	95.2613	Max.
	0.7145	2.0824	0.7417	Min.		0.7136	56.7691	1.4768	Min.
0.25	22.1262	8.4533	7.6511*	Ave.	0.875	22.1461	78.7664	29.7411	Ave.
	32.1569	1.3671	7.5036	S.D.		32.1860	1.4158	19.1150	S.D.
	228.5036	13.3717	49.7718	Max.		228.7156	83.6899	109.2183	Max.
	0.7137	6.7694	0.8260	Min.		0.7142	77.0820	2.1582	Min.
0.375	22.1460	16.2677	10.0488*	Ave.	1	22.1259	102.2016	37.1914	Ave.
	32.1854	1.3687	9.3797	S.D.		32.1562	1.4475	21.7270	S.D.
	228.7066	21.1903	59.6639	Max.		228.4985	107.1208	124.0838	Max.
	0.7142	14.5823	0.9339	Min.		0.7138	100.5188	1.6361	Min.
0.5	22.1259	27.2031	13.4487*	Ave.					
	32.1565	1.3669	11.5990	S.D.					
	228.5025	32.1216	70.4922	Max.					
	0.7137	25.5193	0.6278	Min.					

Table 5.4 (Continued) 1st X Matrix ($X_{30\times10}$, $\ell_1 = 0.0105992$)

(e) $\rho = 750$

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	14.7692	1.4708	3.9396	Ave.	0.625	14.7690	40.5329	16.0285	Ave.
	21.4650	0.9124	4.1376	S.D.		21.4653	0.9056	11.2356	S.D.
	152.5319	4.7537	28.8527	Max.		152.6361	43.8159	64.6628	Max.
	0.4765	0.3467	0.3999	Min.		0.4761	39.4089	1.0047	Min.
0.125	14.7690	3.0332	4.3331	Ave.	0.75	14.7691	57.7197	21.4826	Ave.
	21.4648	0.9124	4.3361	S.D.		21.4649	0.9428	13.3520	S.D.
	152.6309	6.3162	28.5823	Max.		152.5327	61.0033	76.2120	Max.
	0.4764	1.9092	0.4880	Min.		0.4763	56.5964	1.7413	Min.
0.25	14.7692	7.7207	5.7387*	Ave.	0.875	14.7692	78.0318	27.9493	Ave.
	21.4650	0.9127	5.5277	S.D.		21.4652	0.9847	15.5043	S.D.
	152.5312	11.0036	36.0837	Max.		152.5333	81.3156	88.7767	Max.
	0.4764	6.5967	0.5662	Min.		0.4761	76.9088	0.7654	Min.
0.375	14.7691	15.5331	8.1565*	Ave.	1	14.7688	101.4889	35.4296	Ave.
	21.4647	0.9133	7.2373	S.D.		21.4643	1.0283	17.6796	S.D.
	152.5294	18.8160	44.5982	Max.		152.5234	104.7528	102.3541	Max.
	0.4762	14.4092	0.4715	Min.		0.4764	100.3461	0.8018	Min.
0.5	14.7691	26.4705	11.5865*	Ave.					
	21.4648	0.9136	9.1801	S.D.					
	152.5302	29.7536	54.1243	Max.					
	0.4061	25.3466	0.6466	Min.					

Table 5.4 (Continued) 1st X Matrix ($X_{30\times10}$, $\ell_1 = 0.0105992$)

(f) $\rho = 1000$

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	11.0577	1.0112	2.9496	Ave.	0.625	11.0717	40.1647	15.1173	Ave.
	16.0709	0.6831	3.0978	S.D.		16.0914	0.6779	9.6492	S.D.
	114.1986	3.5591	21.6016	Max.		114.3490	42.6259	55.1158	Max.
	0.3567	0.2596	0.2994	Min.		0.3570	39.3222	1.1224	Min.
0.125	11.0717	2.6650	3.3619	Ave.	0.75	11.0578	57.3502	20.5832	Ave.
	16.0913	0.6840	3.3495	S.D.		16.0714	0.7169	11.5014	S.D.
	114.3461	5.1261	22.2491	Max.		114.2075	59.8087	65.8239	Max.
	0.3569	1.8224	0.3816	Min.		0.3564	56.5093	0.5979	Min.
0.25	11.0578	7.3511	4.7789*	Ave.	0.875	11.0719	77.6638	27.0682	Ave.
	16.0714	0.6835	4.5091	S.D.		16.0918	0.7571	13.3922	S.D.
	114.2032	9.8091	28.9203	Max.		114.3525	80.1256	77.6245	Max.
	0.3565	6.5096	0.4369	Min.		0.3568	76.8220	0.6380	Min.
0.375	11.0718	15.1649	7.2152*	Ave.	1.0	11.0575	101.0993	34.5603	Ave.
	16.0917	0.6852	6.0880	S.D.		16.0707	0.8318	15.2806	S.D.
	114.3483	17.6260	36.6596	Max.		114.1981	103.5583	90.3523	Max.
	0.3570	14.3224	0.3200	Min.		0.3566	100.2590	1.6940	Min.
0.5	11.0577	26.1008	10.6568*	Ave.					
	16.0711	0.6853	7.8247	S.D.					
	114.2042	28.5590	45.3477	Max.					
	0.3565	25.2594	0.7373	Min.					

Table 5.5 Summary Statistics of 100 Total Estimated Squared Errors for LS, PC, MPC; 2nd X Matrix ($X_{30\times10}$, $\ell_1 = 0.005009$)

(a) $\rho = 100$

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	254.1852	10.7062	33.9633	Ave.	0.625	254.1862	49.7681	46.8301*	Ave.
	370.8283	4.8712	35.5742	S.D.		370.8269	4.8730	44.7353	S.D.
	2248.0710	25.6825	225.2420	Max.		2247.9950	64.7446	280.6794	Max.
	10.3517	2.8960	5.2209	Min.		10.3494	41.9582	4.6268	Min.
0.125	254.1849	12.2687	33.5528	Ave.	0.75	254.1821	66.9552	53.8793*	Ave.
	370.8266	4.8712	33.7293	S.D.		370.8193	4.8786	50.6054	S.D.
	2248.0200	27.2449	200.7366	Max.		2247.9890	81.9320	309.5944	Max.
	10.3500	4.4585	6.9470	Min.		10.3463	59.1456	5.4627	Min.
0.25	254.1856	16.9562	34.6342	Ave.	0.875	254.1810	87.2673	62.4420*	Ave.
	370.8298	4.8713	33.8722	S.D.		370.8178	4.8870	57.0804	S.D.
	2248.0580	31.9324	202.8755	Max.		2247.9320	102.2443	340.0119	Max.
	10.3507	9.1460	4.1978	Min.		10.3459	79.4581	5.8164	Min.
0.375	254.1875	24.7686	37.2076	Ave.	1.0	254.1718	110.7044	72.4617*	Ave.
	370.8271	4.8715	35.9784	S.D.		370.7968	4.8918	63.9792	S.D.
	2248.0040	39.7448	227.3128	Max.		2247.7750	125.6815	371.9338	Max.
	10.3500	16.9584	2.9728	Min.		10.3406	102.8953	6.4277	Min.
0.5	254.1779	35.7092	41.2758	Ave.					
	370.8122	4.8704	39.7401	S.D.					
	2247.9330	50.6855	253.2615	Max.					
	10.3457	27.8991	3.2404	Min.					

Table 5.5 . (Continued) 2nd X Matrix ($X_{30\times10}$, ℓ_1 = 0.005009)

(b) ρ = 200

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	127.0921	5.3530	16.9815	Ave.	0.625	127.0926	44.4151	31.5389*	Ave.
	185.4174	2.4356	17.7875	S.D.		185.4178	2.4354	28.7755	S.D.
	1124.0670	12.8410	112.6238	Max.		1124.0710	51.9032	171.1008	Max.
	5.1776	1.4480	2.6096	Min.		5.1773	40.5102	2.9052	Min.
0.125	127.0928	6.9155	16.9092	Ave.	0.75	127.0912	61.6021	38.9269*	Ave.
	185.4190	2.4356	16.7703	S.D.		185.4149	2.4463	33.7137	S.D.
	1124.0740	14.4036	95.5181	Max.		1124.0350	69.0906	193.9522	Max.
	5.1781	3.0105	2.8047	Min.		5.1766	57.6976	3.5337	Min.
0.25	127.0922	11.6029	18.3286	Ave.	0.875	127.0897	81.9142	47.8076*	Ave.
	185.4178	2.4357	17.7430	S.D.		185.4106	2.4605	38.9101	S.D.
	1124.0770	19.0909	111.4922	Max.		1123.9930	89.4030	218.2993	Max.
	5.1773	7.6979	1.5385	Min.		5.1754	78.0101	3.4676	Min.
0.375	127.0930	19.4153	21.2400	Ave.	1.0	127.0838	105.3513	58.1851*	Ave.
	185.4184	2.4359	20.4238	S.D.		185.3973	2.4782	44.2765	S.D.
	1124.0570	26.9034	129.8720	Max.		1123.8870	112.8402	244.1523	Max.
	5.1777	15.5104	1.7645	Min.		5.1711	101.4473	4.1324	Min.
0.5	127.0887	30.3560	25.6468*	Ave.					
	185.4094	2.4349	24.2558	S.D.					
	1123.9790	37.8440	149.7542	Max.					
	5.1741	26.4510	2.5040	Min.					

Table 5.5 (Continued) 2nd X Matrix ($X_{30\times10}$, ℓ_1 = 0.005009)

(c) ρ = 300

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	84.7458	3.5693	11.3233	Ave.	0.625	84.7467	42.6314	26.6293*	Ave.
	123.6397	1.6241	11.8610	S.D.		123.6394	1.6231	22.8977	S.D.
	749.5825	8.5624	75.1027	Max.		749.5527	47.6246	131.2657	Max.
	3.4545	0.9655	1.7396	Min.		3.4533	40.0277	2.3817	Min.
0.125	84.7469	5.1318	11.4007	Ave.	0.75	84.7450	59.8183	34.1670*	Ave.
	123.6419	1.6241	11.2132	S.D.		123.6369	1.6430	27.1763	S.D.
	749.6086	10.1249	65.5291	Max.		749.5417	64.8120	151.4319	Max.
	3.4541	2.5280	1.5942	Min.		3.4520	57.2151	2.4209	Min.
0.25	84.7462	9.8193	12.9699	Ave.	0.875	84.7442	80.1305	43.1978*	Ave.
	123.6414	1.6242	12.5147	S.D.		123.6340	1.6629	31.6022	S.D.
	749.5778	14.8123	79.7220	Max.		749.5007	85.1243	173.0979	Max.
	3.4544	7.2155	0.9705	Min.		3.4511	77.5276	3.3683	Min.
0.375	84.7461	17.6317	16.0309*	Ave.	1.0	84.7393	103.5676	53.7248*	Ave.
	123.6415	1.6246	15.2753	S.D.		123.6629	1.6877	36.1226	S.D.
	749.5776	22.6248	95.4101	Max.		749.4138	108.5616	196.2552	Max.
	3.4544	15.0280	1.5577	Min.		3.4484	100.9649	3.2838	Min.
0.5	84.7438	28.5723	20.5876*	Ave.					
	123.6345	1.6233	18.8679	S.D.					
	749.5183	33.5654	112.6060	Max.					
	3.4511	25.9686	1.9437	Min.					

Table 5.5 (Continued) 2nd X Matrix ($X_{30\times10}$, $\ell_1 = 0.005009$)

(d) $\rho = 500$

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	50.8382	2.1411	6.7926	Ave.	0.625	50.8376	41.2032	22.8498*	Ave.
	74.1718	0.9742	7.1154	S.D.		74.1706	0.9702	17.5017	S.D.
	449.6809	5.1363	45.0545	Max.		449.6708	44.1984	96.6388	Max.
	2.0733	0.5792	1.0433	Min.		2.0728	39.6414	1.6102	Min.
0.125	50.8382	3.7036	7.0202	Ave.	0.75	50.8373	58.3901	30.5381*	Ave.
	74.1729	0.9743	6.8392	S.D.		74.1688	1.0035	20.9751	S.D.
	449.6887	6.6987	41.6909	Max.		449.6286	61.3858	114.1122	Max.
	2.0730	2.1417	0.7551	Min.		2.0718	56.8288	2.1166	Min.
0.25	50.8383	8.3910	8.7397	Ave.	0.875	50.8352	78.7020	39.7191*	Ave.
	74.1722	0.9746	8.3913	S.D.		74.1657	1.0526	24.5174	S.D.
	449.6752	11.3862	53.1918	Max.		449.6154	81.6983	133.0742	Max.
	2.0735	6.8292	0.7763	Min.		2.0712	77.1412	1.9711	Min.
0.375	50.8376	16.2034	11.9510*	Ave.	1.0	50.8322	102.1392	50.3965*	Ave.
	74.1709	0.9749	11.0249	S.D.		74.1583	1.0964	28.1049	S.D.
	449.6682	19.1986	66.1815	Max.		449.5549	105.1355	153.5370	Max.
	2.0733	14.6416	1.1840	Min.		2.0690	100.5786	2.4624	Min.
0.5	50.8351	27.1441	16.6580*	Ave.					
	74.1646	0.9744	14.1506	S.D.					
	449.6184	30.1393	80.6731	Max.					
	2.0712	25.5822	1.4663	Min.					

Table 5.5 (Continued) 2nd X Matrix ($X_{30\times10}$, $\ell_1 = 0.005009$)

(e) $\rho = 750$

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	33.9052	1.4279	4.5300	Ave.	0.625	33.9049	40.4900	21.0607*	Ave.
	49.4668	0.6497	4.7455	S.D.		49.4663	0.6421	14.2752	S.D.
	299.9060	3.4254	30.0482	Max.		299.9038	42.4875	77.5918	Max.
	1.3832	0.3863	0.6957	Min.		1.3836	39.4485	1.3462	Min.
0.125	33.9048	2.9904	4.8523	Ave.	0.75	33.9034	57.6770	28.8434*	Ave.
	49.4667	0.6497	4.7007	S.D.		49.4633	0.6789	17.1765	S.D.
	299.9064	4.9879	29.4366	Max.		299.8835	59.6750	93.3588	Max.
	1.3837	1.9488	0.4202	Min.		1.3825	56.6359	1.5133	Min.
0.25	33.9044	7.6778	6.6663*	Ave.	0.875	33.9025	77.9891	38.1193	Ave.
	49.4667	0.6502	6.3253	S.D.		49.4616	0.7247	20.1161	S.D.
	229.9150	9.6753	39.2375	Max.		299.8605	79.4873	110.6257	Max.
	1.3835	6.6362	0.6451	Min.		1.3820	76.9483	2.0500	Min.
0.375	33.9056	15.4902	9.9725*	Ave.	1.0	33.9005	101.4262	48.8914	Ave.
	49.4686	0.6507	8.7328	S.D.		49.4564	0.7853	23.0806	S.D.
	229.9150	17.4878	50.5297	Max.		299.8146	103.4246	129.3907	Max.
	1.3837	14.4487	0.8998	Min.		1.3800	100.3856	4.0195	Min.
0.5	33.9025	26.4309	14.7741*	Ave.					
	49.4619	0.6513	11.4406	S.D.					
	299.8664	28.4284	63.3228	Max.					
	1.3814	25.3893	1.0345	Min.					

Table 5.5 (Continued) 2nd X Matrix ($X_{30\times10}$, $\ell_1 = 0.05009$)

(f) $\rho = 100$

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	25.4179	1.0704	3.3960	Ave.	0.625	25.4173	40.1326	20.2095*	Ave.
	37.0849	0.4870	3.5576	S.D.		37.0835	0.4734	12.3835	S.D.
	224.8509	2.5679	22.5281	Max.		224.8445	41.6301	67.2268	Max.
	1.0378	0.2896	0.5213	Min.		1.0380	39.3518	1.0085	Min.
0.125	25.4176	2.6329	3.7749	Ave.	0.75	25.4171	57.3194	28.0488	Ave.
	37.0849	0.4870	3.6471	S.D.		37.0832	0.5360	14.9242	S.D.
	224.8593	4.1303	23.1292	Max.		224.8270	58.8175	81.9835	Max.
	1.0379	1.8521	0.2915	Min.		1.0372	56.5392	1.4748	Min.
0.25	25.4179	7.3203	5.6456*	Ave.	0.875	25.4158	77.6316	37.3812	Ave.
	37.0858	0.4876	5.2625	S.D.		37.0797	0.5935	17.4890	S.D.
	224.8471	8.8178	31.9181	Max.		224.7895	79.1299	98.2315	Max.
	1.0380	6.5395	0.6047	Min.		1.0363	76.8516	3.1889	Min.
0.375	25.4176	15.1328	9.0082*	Ave.	1.0	25.4138	101.0687	48.2099	Ave.
	37.0840	0.4885	7.4720	S.D.		37.0754	0.6662	20.0705	S.D.
	224.8307	16.6303	42.1945	Max.		224.7597	102.5672	115.9790	Max.
	1.0379	14.3520	0.6939	Min.		1.0347	100.2890	6.3957	Min.
0.5	25.4160	26.0734	13.8663*	Ave.					
	37.0813	0.4898	9.8867	S.D.					
	224.8057	27.5709	53.9731	Max.					
	1.0359	25.2926	1.1823	Min.					

Table 5.6 Summary Statistics of 100 Total Estimated Squared Errors for LS, PC, MPC; 3rd X Matrix ($X_{30 \times 10}$, $\ell_1 = 0.0490105$)

(a) $\rho = 100$

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	36.7739	10.6063	33.8423	Ave.	0.625	36.7738	49.6680	33.4954*	Ave.
	39.3859	4.7238	35.0305	S.D.		39.3857	4.7270	34.0002	S.D.
	263.5954	25.7160	235.3538	Max.		263.5944	64.7782	224.6950	Max.
	4.5975	2.9365	4.5051	Min.		4.5976	41.9986	5.2522	Min.
0.125	36.7740	12.1688	33.7521	Ave.	0.75	36.7737	66.8550	33.4572*	Ave.
	39.3860	4.7238	34.7995	S.D.		39.3853	4.7335	33.8328	S.D.
	263.5976	27.2785	233.2051	Max.		263.5917	81.9655	222.5925	Max.
	4.5975	4.4989	4.6337	Min.		4.5978	59.1860	5.4330	Min.
0.25	36.7740	16.8562	33.6724	Ave.	0.875	36.7736	87.1672	33.4294*	Ave.
	39.3859	4.7239	34.5807	S.D.		39.3851	4.7422	33.6788	S.D.
	263.5961	31.9660	231.0602	Max.		263.5908	102.2778	220.5020	Max.
	4.5975	9.1864	4.7727	Min.		4.5969	79.4984	5.6242	Min.
0.375	36.7739	24.6686	33.6030	Ave.	1.0	36.7732	110.6043	33.4119*	Ave.
	39.3860	4.7241	34.3745	S.D.		39.3844	4.7450	33.5380	S.D.
	263.5983	39.7784	228.9307	Max.		263.5847	125.7151	218.4170	Max.
	4.5976	16.9989	4.9221	Min.		4.5982	102.9357	5.8260	Min.
0.5	36.7734	35.6061	33.5437*	Ave.					
	39.3844	4.7228	34.1799	S.D.					
	263.5834	50.7158	226.7963	Max.					
	4.5982	27.9364	5.0826	Min.					

Table 5.6 (Continued) 3rd X Matrix ($X_{30\times10}$, $\ell_1 = 0.0490105$)

(b) $\rho = 200$

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	18.3867	5.3030	16.9208	Ave.	0.625	18.3865	44.3650	16.7135*	Ave.
	19.6930	2.3619	17.5154	S.D.		19.6927	2.3610	16.8342	S.D.
	131.7991	12.8578	117.6774	Max.		131.7963	51.9199	110.1773	Max.
	2.2985	1.4682	2.2522	Min.		2.2987	40.5304	2.8187	Min.
0.125	18.3867	6.8655	16.8586	Ave.	0.75	18.3865	61.5519	16.7033*	Ave.
	19.6930	2.3619	17.3538	S.D.		19.6925	2.3745	16.7377	S.D.
	131.7994	14.4203	116.1573	Max.		131.7946	69.1073	108.7074	Max.
	2.2986	3.0307	2.3448	Min.		2.2987	57.7178	2.9632	Min.
0.25	18.3866	11.5529	16.8067	Ave.	0.875	18.3864	81.8640	16.7034*	Ave.
	19.6929	2.3619	17.2044	S.D.		19.6924	2.3934	16.6549	S.D.
	131.7976	19.1078	114.6458	Max.		131.7939	89.4197	107.2488	Max.
	2.2986	7.7182	2.4477	Min.		2.2988	78.0301	3.0382	Min.
0.375	18.3867	19.3653	16.7653*	Ave.	1.0	18.3862	105.3012	16.7140*	Ave.
	19.6929	2.3623	17.0680	S.D.		19.6919	2.4037	16.5860	S.D.
	131.7978	26.9201	113.1463	Max.		131.7902	112.8569	105.7983	Max.
	2.2986	15.5306	2.5609	Min.		2.2991	101.4675	2.8624	Min.
0.5	18.3863	30.3028	16.7341*	Ave.					
	19.6919	2.3610	16.9439	S.D.					
	131.7900	37.8576	111.6504	Max.					
	2.2991	26.4681	2.6852	Min.					

Table 5.6 (Continued) 3rd X Matrix ($X_{30\times10}$, $\ell_1 = 0.0490105$)

(c) $\rho = 300$

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	12.2601 13.1313 87.8847 1.5326	3.5360 1.5749 8.5735 0.9789	11.2827 11.6792 78.4682 1.5018	Ave. S.D. Max. Min.	0.625	12.2601 13.1311 87.8814 1.5327	42.5980 1.5731 47.6357 40.0412	11.1373* 11.1533 72.3666 1.9880	Ave. S.D. Max. Min.
0.125	12.2601 13.1313 87.8833 1.5415	5.0985 1.5750 10.1360 2.5415	11.2328 11.5483 77.2264 1.5781	Ave. S.D. Max. Min.	0.75	12.2000 13.1310 87.8812 1.5328	59.7848 1.5972 64.8231 57.2286	11.1394* 11.0884 71.1780 1.9846	Ave. S.D. Max. Min.
0.25	12.2601 13.1312 87.8834 1.5327	9.7859 1.5751 14.8235 7.2289	11.1933 11.4299 75.9963 1.6651	Ave. S.D. Max. Min.	0.875	12.2600 13.1309 87.8798 1.5329	80.0969 1.6206 85.1354 77.5409	11.1519* 11.0376 69.9987 1.8450	Ave. S.D. Max. Min.
0.375	12.2601 13.1314 87.8844 1.5326	17.5983 1.5755 22.6359 15.0414	11.1643* 11.3245 74.7775 1.7623	Ave. S.D. Max. Min.	1.0	12.2598 13.1305 87.8767 1.5330	103.5340 1.6484 108.5727 100.9783	11.1749* 11.0011 68.8284 1.7158	Ave. S.D. Max. Min.
0.5	12.2598 13.1305 87.8768 1.5331	28.5357 1.5745 33.5733 25.9789	11.1455* 11.2317 73.5615 1.8705	Ave. S.D. Max. Min.					

Table 5.6 (Continued) 3rd X Matrix ($X_{30\times10}$, ℓ_1 = 0.0490105)

(d) ρ = 500

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	7.3544	2.1211	6.7681	Ave.	0.625	7.3543	41.1832	6.6847*	Ave.
	7.8770	0.9447	7.0060	S.D.		7.8769	0.9401	6.6366	S.D.
	52.7191	5.1429	47.0705	Max.		52.7180	44.2051	42.3752	Max.
	0.9193	0.5872	0.9008	Min.		0.9194	39.6494	1.1517	Min.
0.125	7.3544	3.6836	6.7306	Ave.	0.75	7.3543	58.3701	6.6992*	Ave.
	7.8770	0.9448	6.9060	S.D.		7.8768	0.9666	6.6039	S.D.
	52.7193	6.7054	46.1109	Max.		52.7171	61.3925	41.4668	Max.
	0.9193	2.1497	0.9612	Min.		0.9194	56.8368	1.0485	Min.
0.25	7.3544	8.3710	6.7035*	Ave.	0.875	7.3543	78.6821	6.7242*	Ave.
	7.8770	0.9450	6.8186	S.D.		7.8767	1.0104	6.5856	S.D.
	52.7185	11.3929	45.1610	Max.		52.7162	81.7048	40.5688	Max.
	0.9193	6.8372	1.0319	Min.		0.9195	77.1492	0.9557	Min.
0.375	7.3544	16.1834	6.6869*	Ave.	1.0	7.3541	102.1191	6.7596*	Ave.
	7.8770	0.9455	6.7443	S.D.		7.8764	1.0660	6.5818	S.D.
	52.7193	19.2053	44.2229	Max.		52.7137	105.1421	39.6799	Max.
	0.9193	14.6497	1.1131	Min.		0.9196	100.5865	0.8732	Min.
0.5	7.3542	27.1208	6.6806*	Ave.					
	7.8764	0.9448	6.6832	S.D.					
	52.7140	30.1427	43.2899	Max.					
	0.9196	25.5871	1.2051	Min.					

Table 5.6 (Continued) 3rd X Matrix ($X_{30\times 10}$, ℓ_1 = 0.0490105)

(e) ρ = 750

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	4.9046	1.4146	4.5136	Ave.	0.625	4.9046	40.4767	4.4693*	Ave.
	5.2533	0.6300	4.6723	S.D.		5.2532	0.6201	4.4021	S.D.
	35.1584	3.4298	31.3913	Max.		35.1582	42.4920	27.5814	Max.
	0.6130	0.3916	0.6007	Min.		0.6131	39.4538	0.6914	Min.
0.125	4.9046	2.9770	4.4839	Ave.	0.75	4.9046	57.6634	4.4917*	Ave.
	5.2532	0.6301	4.5918	S.D.		5.2530	0.6704	4.3899	S.D.
	35.1589	4.9923	30.6089	Max.		35.1573	59.6794	26.8499	Max.
	0.6130	1.9541	0.6509	Min.		0.6131	56.6412	0.6175	Min.
0.25	4.9047	7.6644	4.4647*	Ave.	0.875	4.9045	77.9755	4.5245*	Ave.
	5.2532	0.6305	4.5240	S.D.		5.2530	0.7291	4.3923	S.D.
	35.1588	9.6797	29.8366	Max.		35.1561	79.9917	26.1286	Max.
	0.6130	6.6416	0.7116	Min.		0.6132	76.9536	0.5542	Min.
0.375	4.9046	15.4769	4.4558*	Ave.	1.0	4.9044	101.4125	4.5677*	Ave.
	5.2532	0.6313	4.4695	S.D.		5.2528	0.8017	4.4093	S.D.
	35.1585	17.4922	29.0743	Max.		35.1544	103.4290	25.4173	Max.
	0.6131	14.4540	0.7826	Min.		0.6133	100.3909	0.5011	Min.
0.5	4.9045	26.4142	4.4574*	Ave.					
	5.2528	0.6317	4.4286	S.D.					
	35.1548	28.4296	28.3197	Max.					
	0.6133	25.3915	0.7752	Min.					

Table 5.6 (Continued) 3rd X Matrix ($X_{30\times10}$, $\ell_1 = 0.0490105$)

(f) $\rho = 100$

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	3.6768	1.0604	3.3836	Ave.	0.625	3.6768	40.1225	3.3627*	Ave.
	3.9382	0.4723	3.5027	S.D.		3.9381	0.4633	3.2921	S.D.
	26.3573	2.5712	23.5332	Max.		26.3570	41.6334	20.2518	Max.
	0.4595	0.2936	0.4502	Min.		0.4596	39.3558	0.4748	Min.
0.125	3.6768	2.6629	3.3587	Ave.	0.75	3.6768	57.3094	3.3898*	Ave.
	3.9383	0.4722	3.4339	S.D.		3.9380	0.5246	3.2922	S.D.
	26.3579	4.1337	22.8567	Max.		26.3560	58.8207	19.6259	Max.
	0.4595	1.8561	0.4944	Min.		0.4596	56.5431	0.4186	Min.
0.25	3.6768	7.3103	3.3441*	Ave.	0.875	3.6767	77.6214	3.4272*	Ave.
	3.9382	0.4732	3.3777	S.D.		3.9380	0.5946	3.3069	S.D.
	26.3569	8.8211	22.1892	Max.		26.3553	79.1331	19.0108	Max.
	0.4596	6.5435	0.5491	Min.		0.4597	76.8555	0.3727	Min.
0.375	3.6768	15.1228	3.3399*	Ave.	1.0	3.6766	101.0585	3.4751*	Ave.
	3.9382	0.4739	3.3351	S.D.		3.9378	0.6624	3.3361	S.D.
	26.3573	16.6336	21.5333	Max.		26.3537	102.5704	19.3772	Max.
	0.4595	14.3560	0.6141	Min.		0.4598	100.2929	0.3373	Min.
0.5	3.6766	26.0601	3.3462*	Ave.					
	3.9377	0.4754	3.3063	S.D.					
	26.3535	27.5710	20.8843	Max.					
	0.4598	25.2934	0.5411	Min.					

Table 5.7 Summary Statistics of 100 Total Estimated Squared Errors for LS,

PC, MPC; 4th X Matrix ($X_{30\times5}$, ℓ_1 = 0.0135742)

(a) ρ = 100

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	86.0009	4.1534	26.1670	Ave.	0.625	86.0012	43.2157	39.7433*	Ave.
	140.4362	3.0823	38.2120	S.D.		140.4373	3.0811	57.1233	S.D.
	883.3913	14.1217	238.8975	Max.		883.4086	53.1840	340.6132	Max.
	0.6433	0.2835	0.5947	Min.		0.6434	39.3458	0.7186	Min.
0.125	86.0011	5.7159	27.4339	Ave.	0.75	86.0013	60.4025	44.6311*	Ave.
	140.4376	3.0824	41.1734	S.D.		140.4376	3.0936	61.7411	S.D.
	883.4400	15.6842	257.7973	Max.		883.4072	70.3715	363.1271	Max.
	0.6433	1.8460	1.1176	Min.		0.6433	56.5333	0.5952	Min.
0.25	86.0009	10.4034	29.4249	Ave.	0.875	86.0015	80.7146	50.2428*	Ave.
	140.4372	3.0824	44.6437	S.D.		140.4386	3.1089	66.5106	S.D.
	883.4079	20.3717	277.4135	Max.		883.4223	90.6840	386.3718	Max.
	0.6433	6.5335	1.1313	Min.		0.6433	76.8458	1.0326	Min.
0.375	86.0013	18.2158	32.1402	Ave.	1.0	86.0014	104.1518	56.5797*	Ave.
	140.4370	3.0826	48.5143	S.D.		140.4388	3.1205	71.4008	S.D.
	883.3996	28.1841	297.7534	Max.		883.4187	114.1212	410.3337	Max.
	0.6432	14.3460	0.7335	Min.		0.6434	100.2832	1.9971	Min.
0.5	86.0016	29.1533	35.5800	Ave.					
	140.4385	3.0819	52.6976	S.D.					
	883.4104	39.1215	318.8229	Max.					
	0.6433	25.2834	0.3639	Min.					

Table 5.7 (Continued) 4th X Matrix ($X_{30\times 5}$, $\ell_1 = 0.0135742$)

(b) $\rho = 200$

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	43.0003	2.0767	13.0832	Ave.	0.625	43.0000	41.1389	25.3341*	Ave.
	70.2170	1.5411	19.1057	S.D.		70.2170	1.5380	33.4259	S.D.
	441.6882	7.0607	119.4466	Max.		441.6906	46.1230	194.0184	Max.
	0.3216	0.1417	0.2973	Min.		0.3216	39.2041	0.5455	Min.
0.125	42.9999	3.6392	14.0850	Ave.	0.75	43.0000	58.3258	29.9570*	Ave.
	70.2168	1.5412	21.2774	S.D.		70.2176	1.5598	36.9038	S.D.
	441.6933	8.6232	132.9140	Max.		441.7001	63.3105	211.1085	Max.
	0.3216	1.7042	0.6319	Min.		0.3216	56.3916	1.2197	Min.
0.25	43.0000	8.3266	15.8110	Ave.	0.875	43.0002	78.6379	35.3040*	Ave.
	70.2169	1.5413	23.9122	S.D.		70.2178	1.5871	40.4706	S.D.
	441.6826	13.3107	147.1007	Max.		441.6953	83.6230	228.9179	Max.
	0.3217	6.3917	0.4349	Min.		0.3217	76.7040	0.9042	Min.
0.375	42.9997	16.1391	18.2611	Ave.	1.0	43.0003	102.0751	41.3759*	Ave.
	70.2165	1.5416	26.8745	S.D.		70.2182	1.6141	44.1050	S.D.
	441.6896	21.1232	162.0171	Max.		441.7004	107.0603	247.4551	Max.
	0.3216	14.2042	0.1917	Min.		0.3217	100.1414	0.8356	Min.
0.5	43.0002	27.0765	21.4357*	Ave.					
	70.2177	1.5415	30.0681	S.D.					
	441.6987	32.0606	177.6584	Max.					
	0.3217	25.1416	0.3060	Min.					

Table 5.7 (Continued) 4th X Matrix ($X_{30\times5}$, ℓ_1 = 0.0135742)

(c) ρ = 300

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	28.6720	1.3847	8.7237	Ave.	0.625	28.6721	40.4470	20.3879*	Ave.
	46.8198	1.0276	12.7394	S.D.		46.8205	1.0225	24.8971	S.D.
	294.5153	4.7080	79.6463	Max.		294.5214	43.7704	142.2017	Max.
	0.2145	0.0945	0.1982	Min.		0.2144	39.1569	0.7904	Min.
0.125	28.6721	2.9472	9.6082	Ave.	0.75	28.6723	57.6339	24.8933*	Ave.
	46.8202	1.0276	14.5590	S.D.		46.8202	1.0499	27.8234	S.D.
	294.5175	6.2706	90.7092	Max.		294.5178	60.9579	156.8835	Max.
	0.2145	1.6570	0.3871	Min.		0.2145	56.3443	0.5322	Min.
0.25	28.6721	7.6346	11.2168	Ave.	0.875	28.6722	77.9460	30.1230	Ave.
	46.8203	1.0279	16.8094	S.D.		46.8203	1.0868	30.8116	S.D.
	294.5197	10.9580	102.4963	Max.		294.5180	81.2703	172.2912	Max.
	0.2145	6.3445	0.1573	Min.		0.2145	76.6568	0.7482	Min.
0.375	28.6723	15.4471	13.5497*	Ave.	1.0	28.6723	101.3833	36.0774	Ave.
	46.8202	1.0283	19.3410	S.D.		44.8210	1.1237	33.8453	S.D.
	294.5195	18.7705	115.0070	Max.		294.5241	104.7077	188.4253	Max.
	0.2144	14.1570	0.2564	Min.		0.2145	100.0942	1.4432	Min.
0.5	28.6724	26.3846	16.6069*	Ave.					
	46.8208	1.0286	22.0571	S.D.					
	294.5234	29.7080	128.2433	Max.					
	0.2144	25.0944	0.3255	Min.					

Table 5.7 (Continued) 4th X Matrix ($X_{30\times5}$, $\ell_1 = 0.0135742$)

(d) $\rho = 500$

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	17.1991	0.8306	5.2330	Ave.	0.625	17.1990	39.8929	16.3079*	Ave.
	28.0852	0.6164	7.6418	S.D.		28.0848	0.6077	17.5029	S.D.
	176.6643	2.8242	47.7755	Max.		176.6612	41.8865	98.2630	Max.
	0.1286	0.0567	0.1189	Min.		0.1287	39.1190	0.3244	Min.
0.125	17.1992	2.3932	5.9996	Ave.	0.75	17.1992	57.0798	20.6955	Ave.
	28.0852	0.6164	9.1043	S.D.		28.0853	0.6552	19.8392	S.D.
	176.6669	4.3867	56.4257	Max.		176.6666	59.0740	110.5354	Max.
	0.1287	1.6192	0.2295	Min.		0.1286	56.3065	0.8238	Min.
0.25	17.1992	7.0805	7.4904	Ave.	0.875	17.1992	77.3919	25.8074	Ave.
	28.0853	0.6170	10.9513	S.D.		28.0855	0.7204	22.2126	S.D.
	176.6679	9.0742	65.7997	Max.		175.6684	79.3864	123.5307	Max.
	0.1286	6.3067	0.0882	Min.		0.1287	76.6190	0.8347	Min.
0.375	17.1993	14.8930	9.7054*	Ave.	1.0	17.1993	100.8291	31.6441	Ave.
	28.0854	0.6176	13.0201	S.D.		28.0859	0.7793	24.6119	S.D.
	176.6666	16.8866	75.8970	Max.		176.6734	102.8238	137.2517	Max.
	0.1287	14.1192	0.1660	Min.		0.1287	100.0564	1.0731	Min.
0.5	17.1993	25.8304	12.6448*	Ave.					
	28.0857	0.6195	15.2207	S.D.					
	176.6712	27.8241	86.7206	Max.					
	0.1286	25.0566	0.4612	Min.					

Table 5.7 (Continued) 4th X Matrix ($X_{30\times5}$, $\ell_1 = 0.0135742$)

(e) $\rho = 750$

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	11.4701	0.5540	3.4898	Ave.	0.625	11.4701	39.6162	14.1938	Ave.
	18.7300	0.4111	5.0963	S.D.		18.7304	0.3965	13.4239	S.D.
	117.8173	1.8834	31.8615	Max.		117.8225	40.9458	74.7545	Max.
	0.0857	0.0378	0.0793	Min.		0.0858	39.1002	0.5670	Min.
0.125	11.4699	2.1164	4.1822	Ave.	0.75	11.4700	56.8032	18.5072	Ave.
	18.7298	0.4110	6.3306	S.D.		18.7301	0.4592	15.3673	S.D.
	117.8182	3.4459	38.9915	Max.		117.8206	58.1333	85.5045	Max.
	0.0858	1.6003	0.1266	Min.		0.0858	56.2876	0.5571	Min.
0.25	11.4701	6.8039	5.5989*	Ave.	0.875	11.4700	77.1153	23.5448	Ave.
	18.7301	0.4120	7.9090	S.D.		18.7301	0.5288	17.3341	S.D.
	117.8186	8.1334	46.8457	Max.		117.8185	78.4457	96.9785	Max.
	0.0858	6.2878	0.0874	Min.		0.0858	76.6001	1.0235	Min.
0.375	11.4700	14.6163	7.7396*	Ave.	1.0	11.4702	100.5525	29.3072	Ave.
	18.7302	0.4134	9.6642	S.D.		18.7305	0.6082	19.3178	S.D.
	117.8220	15.9459	55.4251	Max.		117.8229	101.8831	109.1797	Max.
	0.0858	14.1003	0.2978	Min.		0.0858	100.0375	2.2141	Min.
0.5	11.4701	25.5538	10.6048*	Ave.					
	18.7303	0.4158	11.5159	S.D.					
	117.8207	26.8834	64.7271	Max.					
	0.0858	25.0377	0.2092	Min.					

Table 5.7 (Continued) 4th X Matrix ($X_{30\times5}$, ℓ_1 = 0.0135742)

(f) ρ = 1000

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	8.5986	0.4153	2.6162	Ave.	0.625	8.5985	39.4775	13.0984	Ave.
	14.0410	0.3082	3.8204	S.D.		14.0411	0.2930	11.1903	S.D.
	88.3207	1.4119	23.8846	Max.		88.3253	40.4743	62.2369	Max.
	0.0643	0.0283	0.0595	Min.		0.0643	39.0907	0.4138	Min.
0.125	8.5985	1.9778	3.2642	Ave.	0.75	8.5986	56.6648	17.3675	Ave.
	14.0408	0.3081	4.9164	S.D.		14.0411	0.3606	12.8908	S.D.
	88.3201	2.9744	30.1062	Max.		88.3236	57.6618	72.0789	Max.
	0.0643	1.5908	0.0624	Min.		0.0643	56.2782	0.7249	Min.
0.25	8.5985	6.6652	4.6365*	Ave.	0.875	8.5986	76.9766	22.3608	Ave.
	14.0411	0.3096	6.3263	S.D.		14.0410	0.4599	14.6081	S.D.
	88.3220	7.6619	37.0527	Max.		88.3235	77.9742	82.6457	Max.
	0.0643	6.2783	0.0657	Min.		0.0643	76.5906	1.7603	Min.
0.375	8.5985	14.4776	6.7329*	Ave.	1.0	8.5986	100.4138	28.0788	Ave.
	14.0410	0.3105	7.8834	S.D.		14.0415	0.5488	16.3369	S.D.
	88.3247	15.4744	44.7238	Max.		88.3266	101.4116	93.9384	Max.
	0.0643	14.0908	0.2121	Min.		0.0643	100.0281	3.5200	Min.
0.5	8.5986	25.4151	9.5538	Ave.					
	14.0413	0.3148	9.5157	S.D.					
	88.3252	26.4118	53.1184	Max.					
	0.0643	25.0282	0.2871	Min.					

Table 5.8 Summary Statistics of 100 Total Estimated Squared Errors for LS, PC, MPC; 5th X Matrix ($X_{30\times 15}$, ℓ_1 = 0.005374)

(a) ρ = 100

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	117.3953	21.9613	37.1810	Ave.	0.625	177.3920	61.0229	58.9331*	Ave.
	218.2216	10.5184	24.3795	S.D.		218.2173	10.5025	44.4605	S.D.
	1472.9170	64.5528	160.1553	Max.		1472.8810	103.6150	281.1118	Max.
	13.8719	6.9582	8.6970	Min.		13.8729	46.0202	7.9298	Min.
0.125	177.3951	23.5237	38.5815	Ave.	0.75	177.3903	78.2100	67.7080*	Ave.
	218.2206	10.5185	26.2904	S.D.		218.2155	10.5235	50.2259	S.D.
	1472.8840	66.1153	181.3939	Max.		1472.8380	120.8025	309.7233	Max.
	13.8720	8.5206	7.3048	Min.		13.8734	63.2076	9.0095	Min.
0.25	177.3936	28.1112	41.4570	Ave.	0.875	177.3899	98.5221	77.9569*	Ave.
	218.2189	10.5181	29.6283	S.D.		218.2121	10.5277	56.1991	S.D.
	1472.8740	70.8029	204.1097	Max.		1472.7910	141.1153	339.8090	Max.
	13.8720	13.2081	7.0180	Min.		13.8737	83.5199	11.5639	Min.
0.375	177.3942	36.0236	45.8075	Ave.	1.0	177.3852	121.9592	89.6768*	Ave.
	218.2171	10.5178	33.9754	S.D.		218.2032	10.5241	62.3177	S.D.
	1472.8500	78.6151	228.2992	Max.		1472.6540	164.5530	371.3559	Max.
	13.8723	21.0205	8.1755	Min.		13.8751	106.9571	12.1537	Min.
0.5	177.3907	46.9608	51.6308	Ave.					
	218.2137	10.5189	38.9936	S.D.					
	1472.8020	89.5534	253.9605	Max.					
	13.8729	31.9580	8.3253	Min.					

Table 5.8 (Continued) 5th X Matrix ($X_{30\times15}$, ℓ_1 = 0.005374)

(b) ρ = 200

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	88.6968	10.9804	18.5902	Ave.	0.625	88.6946	50.0421	39.3740*	Ave.
	109.1106	5.2591	12.1899	S.D.		109.1065	2.2607	28.3156	S.D.
	736.4868	32.2757	80.0804	Max.		736.4379	71.3379	171.0091	Max.
	6.9346	3.4791	4.3485	Min.		6.9357	42.5411	5.8031	Min.
0.125	88.6965	12.5428	19.7969	Ave.	0.75	88.6941	67.2291	47.9552*	Ave.
	109.1095	5.2591	13.7628	S.D.		109.1069	5.2671	32.6674	S.D.
	736.4570	33.8379	95.3138	Max.		736.4262	88.5253	193.6196	Max.
	6.9351	5.0415	3.5124	Min.		6.9364	59.7285	6.4738	Min.
0.25	88.6962	17.2303	22.4787	Ave.	0.875	88.6931	87.5412	58.0101*	Ave.
	109.1076	5.2592	16.5872	S.D.		109.1035	5.2749	37.1134	S.D.
	736.4528	38.5253	112.0251	Max.		736.3825	108.8378	217.6999	Max.
	6.9350	9.7290	3.9363	Min.		6.9368	80.0408	7.8347	Min.
0.375	88.6953	25.0427	26.6356	Ave.	1.0	88.6902	110.9782	69.5365*	Ave.
	109.1087	5.2594	20.1438	S.D.		109.0986	5.2762	41.6225	S.D.
	736.4633	46.3380	130.2134	Max.		736.3063	132.2754	243.2486	Max.
	6.9352	17.5414	4.0466	Min.		6.9377	103.4780	9.8164	Min.
0.5	88.6942	35.9801	32.2654*	Ave.					
	109.1048	5.2583	24.1094	S.D.					
	736.3913	57.2758	149.8654	Max.					
	6.9361	28.4788	4.2361	Min.					

Table 5.8 (Continued) 5th X Matrix ($X_{30\times15}$, ℓ_1 = 0.005374)

(c) ρ = 300

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	59.1441	7.3216	12.3959	Ave.	0.625	59.1420	46.3835	32.7508*	Ave.
	72.7571	3.5068	8.1283	S.D.		72.7536	3.5073	22.1449	S.D.
	491.1147	21.5211	53.4001	Max.		491.0761	60.5832	131.0342	Max.
	4.6237	2.3198	2.8990	Min.		4.6246	41.3819	4.4400	Min.
0.125	59.1433	8.8841	13.5168	Ave.	0.75	59.1412	63.5705	41.2462*	Ave.
	72.7563	3.5068	9.5420	S.D.		72.7516	3.5165	25.7891	S.D.
	491.1000	23.0835	65.9760	Max.		491.0441	77.7707	150.9821	Max.
	4.6241	3.8823	2.3199	Min.		4.6249	58.5692	5.5728	Min.
0.25	59.1428	13.5715	16.1128	Ave.	0.875	59.1401	83.8824	51.2153*	Ave.
	72.7557	3.5068	12.0819	S.D.		72.7512	3.5290	29.4913	S.D.
	491.1052	27.7709	80.0292	Max.		491.0344	98.0832	172.4084	Max.
	4.6238	8.5698	3.0285	Min.		4.6253	78.8816	7.7463	Min.
0.375	59.1425	21.3839	20.1839*	Ave.	1.0	59.1380	107.3195	62.6559	Ave.
	72.7540	3.5071	15.1935	S.D.		72.7470	3.5362	33.2302	S.D.
	491.0830	35.5836	95.5542	Max.		490.9614	121.5207	195.2974	Max.
	4.6240	16.3822	2.6757	Min.		4.6261	102.3188	11.3929	Min.
0.5	59.1411	32.3213	25.7278*	Ave.					
	72.7522	3.5058	18.5904	S.D.					
	491.0444	46.5213	112.5514	Max.					
	4.6249	27.3196	3.7773	Min.					

Table 5.8 (Continued) 5th X Matrix ($X_{30\times15}$, ℓ_1 = 0.005374)

(d) ρ = 500

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	35.4785	4.3920	7.4358	Ave.	0.625	35.4779	43.4539	27.3600*	Ave.
	43.6436	2.1036	4.8758	S.D.		43.6417	2.1034	16.4792	S.D.
	294.6162	12.9096	32.0342	Max.		294.5712	51.9718	96.3187	Max.
	2.7734	1.3915	1.7389	Min.		2.7739	40.4536	3.8279	Min.
0.125	35.4787	5.9544	8.4706	Ave.	0.75	35.4772	60.6408	35.7694	Ave.
	43.6446	2.1037	6.1163	S.D.		43.6416	2.1185	19.3648	S.D.
	294.6088	14.4719	41.9413	Max.		294.5703	69.1592	113.6021	Max.
	2.7733	2.9541	1.4317	Min.		2.7741	57.6410	6.1983	Min.
0.25	35.4781	10.6419	10.9805	Ave.	0.875	35.4761	80.9529	45.6526	Ave.
	43.6437	2.1038	8.3069	S.D.		43.6409	2.1345	22.2801	S.D.
	294.6091	19.1594	53.3245	Max.		294.5541	89.4717	132.3574	Max.
	2.7736	7.6415	1.5907	Min.		2.7743	77.9533	10.0438	Min.
0.375	35.4783	18.4542	14.9655*	Ave.	1.0	35.4747	104.3899	57.0070	Ave.
	43.6426	2.1040	10.8889	S.D.		43.6373	2.1485	25.2138	S.D.
	294.5910	26.9718	66.1802	Max.		294.5053	112.9091	152.5801	Max.
	2.7733	15.4540	2.1336	Min.		2.7751	101.3906	15.3623	Min.
0.5	35.4770	29.3916	20.4233*	Ave.					
	43.6417	2.1028	13.6405	S.D.					
	294.5708	37.9096	80.5103	Max.					
	2.7741	26.3913	2.8033	Min.					

Table 5.8 (Continued) 5th X Matrix ($X_{30\times15}$, ℓ_1 = 0.005374)

(e) ρ = 750

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	23.6604	2.9290	4.9589	Ave.	0.625	23.6604	41.9909	24.6121	Ave.
	29.1060	1.4028	3.2517	S.D.		29.1058	1.4012	13.1527	S.D.
	196.4820	8.6093	21.3636	Max.		196.4698	47.6714	77.2465	Max.
	1.8495	0.9281	1.1595	Min.		1.8496	39.9901	4.3967	Min.
0.125	23.6608	4.4914	5.9395	Ave.	0.75	23.6595	59.1779	32.9670	Ave.
	29.1063	1.4030	4.3703	S.D.		29.1043	1.4191	15.5370	S.D.
	196.4818	10.1717	29.5902	Max.		196.4526	64.8589	92.8453	Max.
	1.8494	2.4905	1.0312	Min.		1.8498	57.1775	7.8203	Min.
0.25	23.6607	9.1789	8.3951*	Ave.	0.875	23.6591	79.4898	42.7959	Ave.
	29.1071	1.4030	6.2933	S.D.		29.1035	1.4519	17.9386	S.D.
	196.4904	14.8592	39.2931	Max.		196.4343	85.1713	109.9185	Max.
	1.8493	7.1780	1.1019	Min.		1.8502	77.4898	12.7180	Min.
0.375	23.6606	16.9913	12.3258*	Ave.	1.0	23.7687	102.9269	54.0962	Ave.
	29.1071	1.4035	8.4907	S.D.		29.1011	1.4771	20.3512	S.D.
	196.4973	22.6716	50.4714	Max.		196.4005	108.6087	128.4623	Max.
	1.8496	14.9904	1.6624	Min.		1.8508	100.9271	19.0891	Min.
0.5	23.7697	27.9286	17.7294*	Ave.					
	29.1052	1.4026	10.7953	S.D.					
	196.4614	33.6093	63.1177	Max.					
	1.8500	25.9277	2.4467	Min.					

Table 5.8 (Continued) 5th X Matrix ($X_{30\times15}$, $\ell_1 = 0.005374$)

(f) $\rho = 1000$

ϕ_1	LS	PC	MPC		ϕ_1	LS	PC	MPC	
0	17.7375	2.1957	3.7175	Ave.	0.625	17.7368	41.2577	23.2083	Ave.
	21.8204	1.0516	2.4377	S.D.		21.8193	1.0487	11.2442	S.D.
	147.3075	6.4539	16.0168	Max.		147.2883	45.5161	66.8753	Max.
	1.3862	0.6957	0.8691	Min.		1.3866	39.7578	5.1886	Min.
0.125	17.7376	3.7582	4.6656	Ave.	0.75	17.7365	58.4446	31.5310	Ave.
	21.8202	1.0517	3.4756	S.D.		21.8186	1.0722	13.3212	S.D.
	147.3009	8.0164	23.2379	Max.		147.2763	62.7036	81.4704	Max.
	1.3862	2.2582	0.8505	Min.		1.3867	56.9452	9.2418	Min.
0.25	17.7372	8.4456	7.0888*	Ave.	0.875	17.7364	78.7566	41.3278	Ave.
	21.8198	1.0520	5.2152	S.D.		21.8177	1.1058	15.4100	S.D.
	147.3007	12.7039	31.9350	Max.		147.2622	83.0160	97.5396	Max.
	1.3865	6.9457	0.9683	Min.		1.3870	77.2575	14.7690	Min.
0.375	17.7375	16.2580	10.9872*	Ave.	1.0	17.7352	102.1935	52.5956	Ave.
	21.8207	1.0526	7.1623	S.D.		21.8159	1.1626	17.5063	S.D.
	147.3078	20.5163	42.1086	Max.		147.2321	106.4533	115.0790	Max.
	1.3863	14.7581	1.5070	Min.		1.3875	100.6947	21.7688	Min.
0.5	17.7366	27.1954	16.3583*	Ave.					
	21.8185	1.0523	9.1849	S.D.					
	147.2733	31.4539	53.7499	Max.					
	1.3868	25.6955	2.6097	Min.					

Table 5.9 Summary Statistics of 100 Total Estimated Squared Errors for LS, PC, MPC; 6th X Matrix ($X_{30\times 10}$, ℓ_1 = 0.0178895, ℓ_2 = 0.0303788)

(a) ρ = 100

ϕ_1	ϕ_2	LS	PC	MPC		ϕ_1	ϕ_2	LS	PC	MPC	
0	0	98.0262	8.3331	47.1313	Ave.	0.25	0	98.0258	14.5831	49.5784	Ave.
		89.1022	4.0444	39.7381	S.D.			89.1019	4.0445	41.8414	S.D.
		453.2504	20.8391	213.2895	Max.			453.2531	27.0890	220.8649	Max.
		6.9641	2.2510	5.5062	Min.			6.9640	8.5009	4.4749	Min.
0	0.25	98.0256	14.5831	46.8143	Ave.	0.25	0.25	98.0255	20.8330	49.2612	Ave.
		89.1023	4.0445	38.2049	S.D.			89.1012	4.0447	40.3729	S.D.
		453.2497	27.0890	194.5937	Max.			453.2460	33.3390	207.4274	Max.
		6.9643	8.5009	4.9907	Min.			6.9637	14.7509	5.6571	Min.
0	0.5	98.0255	33.3329	47.4161	Ave.	0.25	0.5	98.0254	39.5829	49.8631	Ave.
		89.1008	4.0432	37.6508	S.D.			89.1012	4.0428	39.8338	S.D.
		453.2419	45.8388	187.5017	Max.			453.2429	52.0888	194.9108	Max.
		6.9635	27.2507	5.3937	Min.			6.9633	33.5007	6.3495	Min.
0	0.75	98.0249	64.5819	48.9369*	Ave.	0.25	0.75	98.0244	70.8318	51.3834*	Ave.
		89.1004	4.0540	38.1189	S.D.			89.0993	4.0577	40.2611	S.D.
		453.2292	77.0885	205.8253	Max.			453.2246	83.3385	206.6871	Max.
		6.9630	58.5005	6.4960	Min.			6.9626	64.7505	6.9905	Min.
0	1	98.0231	108.3311	51.3769*	Ave.	0.5	0	98.0252	33.3329	54.0798	Ave.
		89.0977	4.0701	39.5731	S.D.			89.1011	4.0423	45.4347	S.D.
		453.2065	120.8379	225.0697	Max.			453.2365	45.8388	242.8817	Max.
		6.9613	102.2502	6.4048	Min.			6.9636	27.2508	2.8236	Min.

Table 5.9 (Continued) 6th X Matrix ($X_{30\times10}$, $\ell_1 = 0.0178895$, $\ell_2 = 0.0303788$)

(a) $\rho = 100$

ϕ_1	ϕ_2	LS	PC	MPC		ϕ_1	ϕ_2	LS	PC	MPC	
0.5	0.25	98.0253	39.5829	53.7628	Ave.	0.75	0.25	98.0242	70.8318	60.3184*	Ave.
		89.1011	4.0427	44.0725	S.D.			89.0998	4.0584	48.9568	S.D.
		453.2373	52.0888	229.4479	Max.			453.2253	83.3386	253.5214	Max.
		6.9634	33.5007	4.0060	Min.			6.9624	64.7506	3.6278	Min.
0.5	0.5	98.0248	58.3321	54.3645*	Ave.	0.75	0.5	98.0240	89.5816	60.9197*	Ave.
		89.1001	4.0517	43.5646	S.D.			89.0991	4.0668	48.4880	S.D.
		453.2309	70.8389	216.9301	Max.			453.2172	102.0883	241.0026	Max.
		6.9626	52.2506	6.1076	Min.			6.9619	83.5004	5.5512	Min.
0.5	0.75	98.0240	89.5816	55.8845*	Ave.	1.0	0	98.0227	108.3312	69.2435*	Ave.
		89.0992	4.0659	43.9419	S.D.			89.0980	4.0701	55.8341	S.D.
		453.2197	102.0883	209.6044	Max.			453.2062	120.8380	293.0786	Max.
		6.9618	83.5004	8.1849	Min.			6.9611	102.2502	3.2611	Min.
0.75	0	98.0246	64.5820	60.6355*	Ave.						
		89.1003	4.0546	50.1990	S.D.						
		453.2304	77.0886	266.9575	Max.						
		6.9627	58.5006	2.6235	Min.						

Table 5.9 (Continued) 6th X Matrix ($X_{30\times10}$, $\ell_1 = 0.0178895$, $\ell_2 = 0.0303788$)

(b) ρ = 200

ϕ_1	ϕ_2	LS	PC	MPC		ϕ_1	ϕ_2	LS	PC	MPC	
0	0	49.0131	4.1665	23.5655	Ave.	0.25	0	49.0127	10.4164	25.5971	Ave.
		44.5511	2.0222	19.8689	S.D.			44.5511	2.0224	21.5834	S.D.
		226.6331	10.4194	106.6450	Max.			226.6300	16.6693	114.8702	Max.
		3.4824	1.1255	2.7532	Min.			3.4823	7.3755	1.7705	Min.
0	0.25	49.1027	10.4164	23.4757	Ave.	0.25	0.25	49.0126	16.6663	25.5073	Ave.
		44.5508	2.0223	18.9259	S.D.			44.5505	2.0226	20.7038	S.D.
		226.6271	16.6693	93.5587	Max.			226.6261	22.9193	105.5037	Max.
		3.4824	7.3754	2.5231	Min.			3.4822	13.6254	2.7410	Min.
0	0.5	49.0126	29.1662	24.3050*	Ave.	0.25	0.5	49.0123	35.4162	26.3363*	Ave.
		44.5505	2.0215	18.9827	S.D.			44.5503	2.0204	20.7410	S.D.
		226.6255	35.4191	101.3041	Max.			226.6242	41.6691	102.2148	Max.
		3.4822	26.1253	3.2117	Min.			3.4821	32.3752	3.7874	Min.
0	0.75	49.0119	60.4152	26.0531*	Ave.	0.25	0.75	49.0116	66.6651	28.0842*	Ave.
		44.5497	2.0406	20.0310	S.D.			44.5496	2.0470	21.6903	S.D.
		226.6139	66.6689	114.9326	Max.			226.6164	72.9188	115.8422	Max.
		3.4818	57.3750	3.1670	Min.			3.4813	63.6250	4.4756	Min.
0	1.0	49.0105	104.1645	28.7208*	Ave.	0.5	0	49.0123	29.1662	29.6831	Ave.
		44.5480	2.0695	21.9295	S.D.			44.5504	2.0214	24.6560	S.D.
		226.5985	110.4184	129.4830	Max.			226.6218	35.4191	131.3430	Max.
		3.4804	101.1247	3.4419	Min.			3.4821	26.1253	1.3600	Min.

Tabel 5.9 (Continued) 6th X Matrix ($X_{30\times10}$, ℓ_1 = 0.0178895, ℓ_2 = 0.0303788)

(b) ρ = 200

ϕ_1	ϕ_2	LS	PC	MPC		ϕ_1	ϕ_2	LS	PC	MPC	
0.5	0.25	49.0121	35.4162	29.5932*	Ave.	0.75	0.25	49.0116	66.6652	35.7333*	Ave.
		44.5590	2.0206	23.8771	S.D.			44.5493	2.0470	27.9748	S.D.
		226.6189	41.6691	121.9768	Max.			226.6135	72.9189	140.5052	Max.
		3.4818	32.3753	2.2047	Min.			3.4814	63.6251	2.7077	Min.
0.5	0.5	49.0117	54.1655	30.4221*	Ave.	0.75	0.5	49.0112	85.4149	36.5621*	Ave.
		44.5495	2.0332	23.8966	S.D.			44.5490	2.0648	27.9804	S.D.
		226.6140	60.4190	113.5287	Max.			226.6063	91.6687	132.0537	Max.
		3.4814	51.1251	3.1893	Min.			3.4810	82.3750	3.3053	Min.
0.5	0.75	49.0112	85.4148	32.1698*	Ave.	1.0	0	49.0103	104.1646	44.0158*	Ave.
		44.5488	2.0652	24.7126	S.D.			44.5480	2.0719	33.2421	S.D.
		226.6086	91.6687	118.8079	Max.			226.5966	110.4185	170.4464	Max.
		3.4810	82.3749	3.7271	Min.			3.4802	101.1248	4.4199	Min.
0.75	0	49.0118	60.4153	35.8232*	Ave.						
		44.5497	2.0411	28.6531	S.D.						
		226.6156	66.6689	149.8711	Max.						
		3.4814	57.3751	1.8631	Min.						

Table 5.9 (Continued) 6th X Matrix ($X_{30\times10}$, $\ell_1 = 0.0178895$, $\ell_2 = 0.0303788$)

(c) $\rho = 300$

ϕ_1	ϕ_2	LS	PC	MPC		ϕ_1	ϕ_2	LS	PC	MPC	
0	0	32.6819	2.7782	15.7133	Ave.	0.25	0	32.6818	9.0281	17.5609	Ave.
		29.7065	1.3483	13.2487	S.D.			29.7065	1.3486	14.7842	S.D.
		151.1203	6.9476	71.1116	Max.			151.1194	13.1976	78.9450	Max.
		2.3225	0.7505	1.8360	Min.			2.3223	7.0004	1.0215	Min.
0	0.25	32.6818	9.0281	15.7244	Ave.	0.25	0.25	32.6815	15.2780	17.5718	Ave.
		29.7063	1.3486	12.5689	S.D.			29.7062	1.3490	14.1638	S.D.
		151.1180	13.1976	60.9242	Max.			151.1164	19.4474	71.3808	Max.
		2.3222	7.0004	1.7322	Min.			2.3220	13.2504	1.8983	Min.
0	0.5	32.6816	27.7779	16.6542*	Ave.	0.25	0.5	32.6814	34.0279	18.5016*	Ave.
		29.7060	1.3482	12.8941	S.D.			29.7059	1.3456	14.4391	S.D.
		151.1152	31.9473	71.5554	Max.			151.1120	38.1974	72.4876	Max.
		2.3221	25.7503	2.1740	Min.			2.3218	32.0003	3.0610	Min.
0	0.75	32.6812	59.0271	18.5031*	Ave.	0.25	0.75	32.6809	65.2769	20.3501*	Ave.
		29.7057	1.3693	14.1553	S.D.			29.7052	1.3787	15.5624	S.D.
		151.1106	63.1971	83.1053	Max.			151.1056	69.4471	84.0373	Max.
		2.3217	57.0000	2.1919	Min.			2.3217	63.2500	2.9089	Min.
0	1.0	32.6801	102.7763	21.2715*	Ave.	0.5	0	32.6814	27.7779	21.4629*	Ave.
		29.7040	1.4249	16.1347	S.D.			29.7059	1.3484	17.5774	S.D.
		151.0942	106.9467	95.5762	Max.			151.1114	31.9474	92.9612	Max.
		2.3207	100.7497	2.1508	Min.			2.3219	25.7503	0.8881	Min.

Table 5.9 (Continued) 6th X Matrix ($\ell_1 = 0.0178895$, $\ell_2 = 0.0303788$)

(c) $\rho = 300$

ϕ_1	ϕ_2	LS	PC	MPC		ϕ_1	ϕ_2	LS	PC	MPC	
0.5	0.25	32.6814	34.0279	21.4738*	Ave.	0.75	0.25	32.6808	65.2770	27.4299*	Ave.
		29.7059	1.3454	17.0471	S.D.			29.7053	1.3803	20.6865	S.D.
		151.1122	38.1973	85.3987	Max.			151.1073	69.4471	101.4705	Max.
		2.3219	32.0003	1.6620	Min.			2.3214	63.2501	2.3841	Min.
0.5	0.5	32.6811	52.7773	22.4033*	Ave.	0.75	0.5	32.6805	84.0267	28.3593*	Ave.
		29.7055	1.3638	17.2645	S.D.			29.7048	1.4016	20.8562	S.D.
		151.1080	56.9472	78.7527	Max.			151.1013	88.1970	94.8235	Max.
		2.3215	50.7502	2.0771	Min.			2.3211	81.9999	2.9147	Min.
0.5	0.75	32.6806	84.0266	24.2517*	Ave.	1.0	0	32.6799	102.7763	35.4277	Ave.
		29.7046	1.4041	18.2030	S.D.			29.7041	1.4284	25.1345	S.D.
		151.0996	88.1970	87.0231	Max.			151.0929	106.9468	127.1535	Max.
		2.3211	81.9999	2.9370	Min.			2.3205	100.7498	2.2917	Min.
0.75	0	32.6811	59.0271	27.4191*	Ave.						
		29.7054	1.3718	21.1352	S.D.						
		151.1068	63.1971	109.0331	Max.						
		2.3216	57.0000	2.2411	Min.						

Table 5.9 (Continued) 6th X Matrix ($\ell_1 = 0.0178895$, $\ell_2 = 0.0303788$)

(d) $\rho = 500$

ϕ_1	ϕ_2	LS	PC	MPC		ϕ_1	ϕ_2	LS	PC	MPC	
0	0	19.6048	1.6665	9.4259	Ave.	0.25	0	19.6046	7.9164	11.0886	Ave.
		17.8203	0.8088	7.9475	S.D.			17.8201	0.8091	9.2933	S.D.
		90.6562	4.1676	42.6576	Max.			90.6538	10.4176	49.6788	Max.
		1.3935	0.4502	1.1015	Min.			1.3932	6.7001	0.5465	Min.
0	0.25	19.6047	7.9164	9.5379	Ave.	0.25	0.25	19.6045	14.1664	11.2006*	Ave.
		17.8199	0.8092	7.5335	S.D.			17.8199	0.8096	8.9282	S.D.
		90.6506	10.4176	38.3453	Max.			90.6499	16.6675	43.9231	Max.
		1.3934	6.7001	1.1246	Min.			1.3930	12.9501	1.3291	Min.
0	0.5	19.6045	26.6662	10.5688*	Ave.	0.25	0.5	19.6044	32.9162	12.2314*	Ave.
		17.8199	0.8107	8.1142	S.D.			17.8197	0.8044	9.4104	S.D.
		90.6495	29.1674	46.8892	Max.			90.6487	35.4174	47.8428	Max.
		1.3932	25.4499	1.2633	Min.			1.3931	31.6999	1.7847	Min.
0	0.75	19.6042	57.9154	12.5188*	Ave.	0.25	0.75	19.6040	64.1653	14.1811*	Ave.
		17.8194	0.8439	9.5092	S.D.			17.8193	0.8547	10.6251	S.D.
		90.6466	60.4171	56.3525	Max.			90.6443	66.6671	57.3064	Max.
		1.3929	56.6997	1.2842	Min.			1.3927	62.9497	1.9778	Min.
0	1.0	19.6034	101.6646	15.3883*	Ave.	0.5	0	19.6044	26.6663	14.8059*	Ave.
		17.8183	0.9273	11.4243	S.D.			17.8199	0.8099	11.7463	S.D.
		90.6342	104.6667	66.7372	Max.			90.6486	29.1674	61.2299	Max.
		1.3920	100.4494	1.7622	Min.			1.3930	25.4500	0.8557	Min.

Table 5.9 (Continued) 6th X Matrix (ℓ_1 = 0.0178895, ℓ_2 = 0.0303788)

(d) ρ = 500

ϕ_1	ϕ_2	LS	PC	MPC		ϕ_1	ϕ_2	LS	PC	MPC	
0.5	0.5	19.6043	32.9162	14.9178*	Ave.	0.75	0.25	19.6040	64.1654	20.6893	Ave.
		17.8197	0.8036	11.4490	S.D.			17.8193	0.8539	14.5206	S.D.
		90.6483	35.4173	55.4758	Max.			90.6433	66.6672	69.0812	Max.
		1.3929	31.7000	1.3811	Min.			1.3927	62.9498	0.7506	Min.
0.5	0.5	19.6041	51.6657	15.9485*	Ave.	0.75	0.5	19.6037	82.9151	21.7197	Ave.
		17.8194	0.8288	11.8184	S.D.			17.8189	0.8851	14.8053	S.D.
		90.6449	54.1672	50.8519	Max.			90.6408	85.4170	64.2449	Max.
		1.3927	50.4499	1.6917	Min.			1.3923	81.6997	0.9663	Min.
0.5	0.75	19.6038	82.9151	17.8979*	Ave.	1.0	0	19.6032	101.6648	28.4014	Ave.
		17.8190	0.8819	12.7971	S.D.			17.8183	0.9218	18.0656	S.D.
		90.6418	85.4170	60.3141	Max.			90.6332	104.1668	90.4916	Max.
		1.3924	81.6997	1.6110	Min.			1.3919	100.4495	1.5719	Min.
0.75	0	19.6041	57.9155	20.5774	Ave.						
		17.8194	0.8409	14.7644	S.D.						
		90.6434	60.4172	74.8352	Max.						
		1.3928	56.6998	1.4541	Min.						

Table 5.9 (Continued) 6th X Matrix (ℓ_1 = 0.0178895, ℓ_2 = 0.0303788)

(e) ρ = 750

ϕ_1	ϕ_2	LS	PC	MPC		ϕ_1	ϕ_2	LS	PC	MPC	
0	0	13.0744	1.1114	6.2861	Ave.	0.25	0	13.0744	7.3613	7.8325	Ave.
		11.8844	0.5394	5.3002	S.D.			11.8843	0.5400	6.5167	S.D.
		60.4591	2.7794	28.4479	Max.			60.4565	9.0294	34.7413	Max.
		0.9294	0.3002	0.7346	Min.			0.9293	6.5502	0.3727	Min.
0	0.25	13.0744	7.3613	6.4618*	Ave.	0.25	0.25	13.0743	13.6112	8.0081*	Ave.
		11.8845	0.5399	5.0546	S.D.			11.8843	0.5406	6.3057	S.D.
		60.4579	9.0293	26.7974	Max.			60.4578	15.2793	30.1271	Max.
		0.9292	6.5502	0.8367	Min.			0.9293	12.8002	0.8664	Min.
0	0.5	13.0743	26.1111	7.5563*	Ave.	0.25	0.5	13.0742	32.3611	9.1026*	Ave.
		11.8842	0.5424	5.7740	S.D.			11.8841	0.5348	6.8839	S.D.
		60.4553	27.7791	34.0277	Max.			60.4546	34.2091	34.9955	Max.
		0.9291	25.2999	0.9098	Min.			0.9291	31.5500	1.3357	Min.
0	0.75	13.0740	57.3603	9.5700*	Ave.	0.25	0.75	13.0738	63.6101	11.1159*	Ave.
		11.8838	0.5871	7.1739	S.D.			11.8837	0.6025	8.0839	S.D.
		60.4527	59.0289	42.1773	Max.			60.4511	65.2789	43.1442	Max.
		0.9290	56.5498	1.0038	Min.			0.9289	62.7998	0.9881	Min.
0	1.0	13.0734	101.1096	12.5032*	Ave.	0.5	0	13.0742	26.1111	11.4334*	Ave.
		11.8829	0.6918	8.9405	S.D.			11.8842	0.5424	8.7055	S.D.
		60.4429	102.7785	51.2475	Max.			60.4542	27.7792	44.7405	Max.
		0.9283	100.2994	0.7164	Min.			0.9290	25.3000	1.0620	Min.

Table 5.9 (Continued) 6th X Matrix ($\ell_1 = 0.0178895$, $\ell_2 = 0.0303788$)

(e) $\rho = 750$

ϕ_1	ϕ_2	LS	PC	MPC		ϕ_1	ϕ_2	LS	PC	MPC	
0.5	0.25	13.0741	32.3611	11.6090*	Ave.	0.75	0.25	13.0738	63.6103	17.2640	Ave.
		11.8841	0.5342	8.5392	S.D.			11.8837	0.6020	11.1810	S.D.
		60.4537	34.0291	40.1255	Max.			60.4501	65.2789	52.1783	Max.
		0.9291	31.5500	1.0211	Min.			0.9288	62.7998	0.3821	Min.
0.5	0.5	13.0740	51.1106	12.7033*	Ave.	0.75	0.5	13.0736	82.3599	18.3582	Ave.
		11.8840	0.5590	8.9656	S.D.			11.8834	0.6513	11.5028	S.D.
		60.4521	52.7790	38.0175	Max.			60.4472	84.0288	49.3865	Max.
		0.9290	50.2999	1.2622	Min.			0.9286	81.5497	0.8109	Min.
0.5	0.75	13.0737	82.3599	14.7164	Ave.	1.0	0	13.0732	101.1096	24.7963	Ave.
		11.8834	0.6538	9.9084	S.D.			11.8829	0.7053	14.1149	S.D.
		60.4478	84.0288	46.1664	Max.			60.4421	102.7786	70.8971	Max.
		0.9286	81.5497	0.7221	Min.			0.9282	100.2995	2.2869	Min.
0.75	0	13.0739	57.3604	17.0885	Ave.						
		11.8839	0.5839	11.3158	S.D.						
		60.4508	59.0289	56.7935	Max.						
		0.9288	56.5499	0.8724	Min.						

Table 5.9 (Continued) 6th X Matrix ($X_{30\times10}$, $\ell_1 = 0.0178895$, $\ell_2 = 0.0303788$)

(f) $\rho = 1000$

ϕ_1	ϕ_2	LS	PC	MPC		ϕ_1	ϕ_2	LS	PC	MPC	
0	0	9.8014	0.8332	4.7124	Ave.	0.25	0	9.8014	7.0831	6.1893*	Ave.
		8.9093	0.4044	3.9734	S.D.			8.9094	0.4052	5.1061	S.D.
		45.3239	2.0836	21.3261	Max.			45.3254	8.3336	27.1068	Max.
		0.6968	0.2251	0.5507	Min.			0.6967	6.4750	0.2587	Min.
0	0.25	9.8014	7.0830	4.9261*	Ave.	0.25	0.25	9.8014	13.3330	6.4030*	Ave.
		8.9094	0.4051	3.8284	S.D.			8.9094	0.4058	4.9819	S.D.
		45.3238	8.3335	20.8831	Max.			45.3240	14.5835	23.1721	Max.
		0.6967	6.4750	0.6486	Min.			0.6967	12.7250	0.6149	Min.
0	0.5	9.8013	25.8328	6.0587*	Ave.	0.25	0.5	9.8012	32.0828	7.5354*	Ave.
		8.9089	0.4680	5.9695	S.D.			8.9091	0.3989	5.5956	S.D.
		45.3213	27.0834	27.3289	Max.			45.3223	33.3333	28.3035	Max.
		0.6967	25.2249	0.6473	Min.			0.6966	31.4748	1.1354	Min.
0	0.75	9.8011	57.0820	8.1104*	Ave.	0.25	0.75	9.8010	63.3319	9.5869*	Ave.
		8.9089	0.4680	5.9695	S.D.			8.9087	0.4859	6.7492	S.D.
		45.3200	58.3331	34.6930	Max.			45.3174	64.5831	35.6686	Max.
		0.6964	56.4746	0.8987	Min.			0.6963	62.7246	0.7340	Min.
0	1.0	9.8005	100.8311	11.0817	Ave.	0.5	0	9.8012	25.8329	9.7207*	Ave.
		8.9080	0.6332	7.5993	S.D.			8.9091	0.4075	7.1104	S.D.
		45.3115	102.0828	42.9776	Max.			45.3211	27.0834	36.1768	Max.
		0.6959	100.2243	0.3649	Min.			0.6967	25.2249	0.8198	Min.

Table 5.9 (Continued) 6th X Matrix ($X_{30\times10}$, $\ell_1 = 0.0178895$, $\ell_2 = 0.0303788$)

(f) $\rho = 100$

ϕ_1	ϕ_2	LS	PC	MPC		ϕ_1	ϕ_2	LS	PC	MPC	
0.5	0.25	9.8011	32.0829	9.9343	Ave.	0.75	0.25	9.8008	63.3319	15.5198	Ave.
		8.9090	0.3981	7.0130	S.D.			8.9087	0.4936	9.3672	S.D.
		45.3205	33.3333	32.2425	Max.			45.3168	64.5832	43.3671	Max.
		0.6965	31.4749	0.4570	Min.			0.6964	62.7247	0.5929	Min.
0.5	0.5	9.8009	50.8323	11.0666	Ave.	0.75	0.5	9.8007	82.0816	16.6520	Ave.
		8.9088	0.4516	7.4534	S.D.			8.9085	0.5696	9.6950	S.D.
		45.3185	52.0832	31.3343	Max.			45.3150	83.3301	42.4204	Max.
		0.6964	50.2248	0.8229	Min.			0.6962	81.4746	1.1490	Min.
0.5	0.75	9.8008	82.0815	13.1178	Ave.	1.0	0	9.8004	100.8312	22.9445	Ave.
		8.9085	0.5719	8.3468	S.D.			8.9080	0.6291	11.9212	S.D.
		45.3157	83.3330	38.6983	Max.			45.3103	102.0828	60.4777	Max.
		0.6962	81.4746	0.6359	Min.			0.6958	100.2244	3.1453	Min.
0.75	0	9.8009	57.0821	15.3063	Ave.						
		8.9088	0.4701	9.4469	S.D.						
		45.3186	58.3332	47.3023	Max.						
		0.6964	56.4747	0.9558	Min.						

Table 5.10 Theoretical Mean Squared Errors of LS, PC, MPC (1st X Matrix)

ϕ_1	ρ = 100			ρ = 200			ρ = 300		
	LS	PC	MPC	LS	PC	MPC	LS	PC	MPC
0	106.632	12.285	29.800	53.315	6.142	14.900	35.550	4.096	9.935
0.125	106.632	13.848	30.306	53.315	7.705	15.406	35.550	5.658	10.441
0.25	106.632	18.535	31.825	53.315	12.392	16.924	35.550	10.346	11.959
0.375	106.632	26.348	34.355	53.315	20.205	19.455*	35.550	18.158	14.490*
0.5	106.632	37.285	37.898	53.315	31.142	22.997*	35.550	29.095	18.032*
0.625	106.632	51.348	42.453*	53.315	45.205	27.552*	35.550	43.158	22.587*
0.75	106.632	68.535	48.020*	53.315	62.392	33.119*	35.550	60.346	28.154*
0.875	106.632	88.848	54.599*	53.315	82.705	39.698*	35.550	80.658	34.733*
1.0	106.632	112.285	62.190*	53.315	106.142	47.290*	35.550	104.096	42.325*

ϕ_1	ρ = 500			ρ = 750			ρ = 1000		
	LS	PC	MPC	LS	PC	MPC	LS	PC	MPC
0	21.325	2.457	5.959	14.222	1.638	3.975	10.661	1.228	2.979
0.125	21.325	4.019	6.465	14.222	3.201	4.481	10.661	2.791	3.485
0.25	21.325	8.707	7.984*	14.222	7.889	5.998*	10.661	7.478	5.003*
0.375	21.325	16.519	10.514*	14.222	15.701	8.529*	10.661	15.291	7.534*
0.5	21.325	27.457	14.057*	14.222	26.639	12.072*	10.661	26.228	11.077
0.625	21.325	41.519	18.612*	14.222	40.701	16.627	10.661	40.291	15.631
0.75	21.325	58.707	24.179	14.222	57.888	22.193	10.661	57.478	21.198
0.875	21.325	79.019	30.758	14.222	78.201	28.773	10.661	77.791	27.778
1.0	21.325	102.457	38.349	14.222	101.639	36.364	10.661	101.228	35.369

*MSE($\hat{\underline{\beta}}_{mpc}$) is lower than either MSE($\hat{\underline{\beta}}$) or MSE($\hat{\underline{\beta}}_{pc}$).

Table 5.11 Theoretical Mean Squared Errors of LS, PC, MPC (2nd X Matrix)

ϕ_1	ρ = 100			ρ = 200			ρ = 300		
	LS	PC	MPC	LS	PC	MPC	LS	PC	MPC
0	210.130	10.525	29.591	105.063	5.263	14.795	70.055	3.509	9.865
0.125	210.130	12.088	30.337	105.063	6.825	15.541	70.055	5.072	10.611
0.25	210.130	16.775	32.574	105.063	11.513	17.778	70.055	9.759	12.848
0.375	210.130	24.588	36.303	105.063	19.325	21.507	70.055	17.572	16.577*
0.5	210.130	35.525	41.524	105.063	30.263	26.727*	70.055	28.509	21.798*
0.625	210.130	49.588	48.236*	105.063	44.325	33.439*	70.055	42.572	28.510*
0.75	210.130	66.775	56.439*	105.063	61.512	41.643*	70.055	59.759	36.713*
0.875	210.130	87.088	66.134*	105.063	81.825	51.338*	70.055	80.072	46.408*
1.0	210.130	110.525	77.321*	105.063	105.263	62.525*	70.055	100.509	57.595*

ϕ_1	ρ = 500			ρ = 750			ρ = 1000		
	LS	PC	MPC	LS	PC	MPC	LS	PC	MPC
0	42.023	2.105	5.917	28.025	1.404	3.946	21.009	1.052	2.958
0.125	42.023	3.667	6.663	28.025	2.966	4.692	21.009	2.615	3.704
0.25	42.023	8.355	8.901	28.025	7.654	6.929	21.009	7.302	5.941*
0.375	42.023	16.167	12.629*	28.025	15.466	10.658*	21.009	15.115	9.670*
0.5	42.023	27.105	17.850*	28.025	26.404	15.879*	21.009	26.052	14.891*
0.625	42.023	41.167	24.562*	28.025	40.466	22.591*	21.009	40.115	21.603
0.75	42.023	58.355	32.766*	28.025	57.654	30.794	21.009	57.302	29.806
0.875	42.023	78.667	42.461	28.025	77.966	40.489	21.009	77.615	39.501
1.0	42.023	102.105	53.647	28.025	101.404	51.676	21.009	101.052	50.688

*MSE($\hat{\underline{\beta}}_{mpc}$) is lower than either MSE($\hat{\underline{\beta}}$) or MSE($\hat{\underline{\beta}}_{pc}$).

Table 5.12 Theoretical Mean Squared Errors of LS, PC, MPC (3rd X Matrix)

ϕ_1	$\rho = 100$			$\rho = 200$			$\rho = 300$		
	LS	PC	MPC	LS	PC	MPC	LS	PC	MPC
0	30.913	10.509	28.626	15.456	5.254	14.313	10.306	3.504	9.543
0.125	30.913	12.072	28.631	15.456	6.817	14.318	10.306	5.066	9.549
0.25	30.913	16.759	28.647	15.456	11.504	14.333	10.306	9.754	9.564*
0.375	30.913	24.572	28.673	15.456	19.317	14.359*	10.306	17.566	9.590*
0.5	30.913	35.509	28.709*	15.456	30.254	14.396*	10.306	28.504	9.627*
0.625	30.913	49.572	28.756*	15.456	44.317	14.443*	10.306	42.566	9.673*
0.75	30.913	66.759	28.813*	15.456	61.504	14.500*	10.306	59.754	9.731*
0.875	30.913	87.072	28.881*	15.456	81.817	14.567*	10.306	80.066	9.798*
1.0	30.913	110.509	28.959*	15.456	105.254	14.645*	10.306	103.504	9.876*

ϕ_1	$\rho = 500$			$\rho = 750$			$\rho = 1000$		
	LS	PC	MPC	LS	PC	MPC	LS	PC	MPC
0	6.182	2.102	5.725	4.123	1.402	3.817	3.091	1.051	2.862
0.125	6.182	3.664	5.730	4.123	2.964	3.823	3.091	2.613	2.867
0.25	6.182	8.352	5.745*	4.123	7.652	3.838*	3.091	7.301	2.882*
0.375	6.182	16.164	5.771*	4.123	15.464	3.864*	3.091	15.113	2.908*
0.5	6.182	27.102	5.808*	4.123	26.402	3.901*	3.091	26.051	2.945*
0.625	6.182	41.164	5.854*	4.123	40.464	3.947*	3.091	40.113	2.992*
0.75	6.182	58.352	5.912*	4.123	57.652	4.005*	3.091	57.301	3.049*
0.875	6.182	78.664	5.979*	4.123	77.964	4.072*	3.091	77.613	3.116
1.0	6.182	102.102	6.057*	4.123	101.402	4.150	3.091	101.051	3.194

*MSE($\hat{\underline{\beta}}_{mpc}$) is lower than either MSE($\hat{\underline{\beta}}$) or MSE($\hat{\underline{\beta}}_{pc}$).

Table 5.13 Theoretical Mean Squared Errors of LS, PC, MPC (4th X Matrix)

ϕ_1	ρ = 100			ρ = 200			ρ = 300		
	LS	PC	MPC	LS	PC	MPC	LS	PC	MPC
0	77.371	3.702	23.515	38.685	1.851	11.757	25.795	1.234	7.839
0.125	77.371	5.264	23.877	38.685	3.413	12.119	25.795	2.797	8.201
0.25	77.371	9.952	24.963	38.685	8.101	13.205	25.795	7.484	9.287
0.375	77.371	17.764	26.773	38.685	15.913	15.015*	25.795	15.297	11.098*
0.5	77.371	28.702	29.308	38.685	26.851	17.550*	25.795	26.234	13.632*
0.625	77.371	42.764	32.566*	38.685	40.913	20.808*	25.795	40.297	16.890*
0.75	77.371	59.952	36.548*	38.685	58.101	24.790*	25.795	57.484	20.872*
0.875	77.371	80.264	41.255*	38.685	78.413	29.497*	25.795	77.796	25.579*
1.0	77.371	103.702	46.685*	38.685	101.851	34.927*	25.795	101.234	31.009

ϕ_1	ρ = 500			ρ = 750			ρ = 1000		
	LS	PC	MPC	LS	PC	MPC	LS	PC	MPC
0	15.473	0.740	4.702	10.319	0.494	3.136	7.736	0.370	2.351
0.125	15.473	2.303	5.064	10.319	2.056	3.498	7.736	1.933	2.713
0.25	15.473	6.990	6.151*	10.319	6.744	4.584*	7.736	6.620	3.799*
0.375	15.473	14.803	7.961*	10.319	14.556	6.394*	7.736	14.433	5.609*
0.5	15.473	25.740	10.495*	10.319	25.494	8.928*	7.736	25.370	8.143
0.625	15.473	39.803	13.753*	10.319	39.556	12.186	7.736	39.433	11.401
0.75	15.473	56.990	17.735	10.319	56.744	16.169	7.736	56.620	15.384
0.875	15.473	77.303	22.442	10.319	77.056	20.895	7.736	76.933	20.090
1.0	15.473	100.740	27.872	10.319	100.494	26.306	7.736	100.370	25.521

*MSE($\hat{\underline{\beta}}_{mpc}$) is lower than either MSE($\hat{\underline{\beta}}$) or MSE($\hat{\underline{\beta}}_{pc}$).

Table 5.14 Theoretical Mean Squared Errors of LS, PC, MPC (5th X Matrix)

ϕ_1	$\rho = 100$			$\rho = 200$			$\rho = 300$		
	LS	PC	MPC	LS	PC	MPC	LS	PC	MPC
0	207.940	21.859	40.078	103.968	10.929	20.036	69.325	7.288	13.361
0.125	207.940	23.421	40.816	103.968	12.492	20.776	69.325	8.850	14.099
0.25	207.940	28.109	43.028	103.968	17.179	22.989	69.325	13.538	16.311
0.375	207.940	35.921	46.716	103.968	24.992	26.676	69.325	21.350	19.999*
0.5	207.940	46.859	51.878	103.968	35.929	31.839*	69.325	32.288	25.161*
0.625	207.940	60.921	58.516*	103.968	49.992	38.476*	69.325	46.350	31.799*
0.75	207.940	78.109	66.628*	103.968	67.179	46.589*	69.325	63.538	39.911*
0.875	207.940	98.421	76.216*	103.968	87.492	56.176*	69.325	83.850	49.499*
1.0	207.940	121.859	87.278*	103.968	110.929	67.239*	69.325	107.288	60.561*

ϕ_1	$\rho = 500$			$\rho = 750$			$\rho = 1000$		
	LS	PC	MPC	LS	PC	MPC	LS	PC	MPC
0	41.585	4.372	8.015	27.733	2.915	5.345	20.790	2.186	4.007
0.125	41.585	5.934	8.752	27.733	4.478	6.082*	20.790	3.748	4.744
0.25	41.585	10.622	10.965	27.733	9.165	8.295*	20.790	8.436	6.957*
0.375	41.585	18.434	14.652*	27.733	16.978	11.982*	20.790	16.248	10.644*
0.5	41.585	29.372	19.815*	27.733	27.915	17.145*	20.790	27.186	15.807*
0.625	41.585	43.434	26.452*	27.733	41.978	23.782*	20.790	41.248	22.444
0.75	41.585	60.622	34.565*	27.733	59.165	31.895	20.790	58.436	30.557
0.875	41.585	80.934	44.152	27.733	79.478	41.482	20.790	78.748	40.144
1.0	41.585	104.372	55.215	27.733	102.915	52.545	20.790	102.186	51.207

*MSE($\hat{\underline{\beta}}_{mpc}$) is lower than either MSE($\hat{\underline{\beta}}$) or MSE($\hat{\underline{\beta}}_{pc}$).

Table 5.15 Theoretical Mean Squared Errors of LS, PC, MPC (6th X Matrix)

ϕ_1	ϕ_1	$\rho = 100$			$\rho = 200$			$\rho = 300$		
		LS	PC	MPC	LS	PC	MPC	LS	PC	MPC
0	0	97.143	8.326	45.573	48.570	4.163	22.786	32.386	2.776	15.193
0	0.25	97.143	14.576	46.032	48.570	10.413	23.245	32.386	9.026	15.653
0	0.5	97.143	33.326	47.410	48.570	29.163	24.623*	32.386	27.776	17.031*
0	0.75	97.143	64.576	49.707*	48.570	60.413	26.920*	32.386	59.026	19.328*
0	1	97.143	108.326	52.923*	48.570	104.163	30.136*	32.386	102.776	22.543*
0.25	0	97.143	14.576	46.600	48.570	10.413	23.813	32.386	9.026	16.221
0.25	0.25	97.143	20.826	47.060	48.570	16.663	24.272	32.386	15.276	16.680
0.25	0.5	97.143	39.576	48.438	48.570	35.413	25.651*	32.386	34.026	18.058*
0.25	0.75	97.143	70.826	50.735*	48.570	66.663	27.947*	32.386	65.276	20.355*
0.5	0	97.143	33.326	49.683	48.570	29.163	26.895*	32.386	27.776	19.303*
0.5	0.25	97.143	39.576	50.142	48.570	35.413	27.355*	32.386	34.026	19.762*
0.5	0.5	97.143	58.326	51.520*	48.570	54.163	28.733*	32.386	52.776	21.140*
0.5	0.75	97.143	89.576	53.817*	48.570	85.413	31.030*	32.386	84.026	23.437*
0.75	0	97.143	64.576	54.820*	48.570	60.413	32.033*	32.386	59.026	19.303*
0.75	0.25	97.143	70.826	55.279*	48.570	66.663	32.492*	32.386	65.276	24.900*
0.75	0.5	97.143	89.576	56.657*	48.570	85.413	33.870*	32.386	84.026	26.278*
1.0	0	97.143	108.326	62.012*	48.570	104.163	39.225*	32.386	102.776	31.632*

*MSE($\hat{\underline{\beta}}_{mpc}$) is lower than either MSE($\hat{\underline{\beta}}$) or MSE($\hat{\underline{\beta}}_{pc}$).

Table 5.15 (Continued) (6th X Matrix)

ϕ_1	ϕ_2	ρ = 500			ρ = 750			ρ = 1000		
		LS	PC	MPC	LS	PC	MPC	LS	PC	MPC
0	0	19.427	1.665	9.114	12.956	1.111	6.078	9.713	0.833	4.556
0	0.25	19.427	7.915	9.573	12.956	7.361	6.537*	9.713	7.083	5.015*
0	0.5	19.427	26.665	10.951*	12.956	26.111	6.915*	9.713	25.833	6.394*
0	0.75	19.427	57.915	13.248*	12.956	57.361	10.212*	9.713	57.083	8.690*
0	1.0	19.427	101.665	16.464*	12.956	101.111	13.428	9.713	100.833	11.906
0.25	0	19.427	7.915	10.141	12.956	7.361	7.105*	9.713	7.083	5.583*
0.25	0.25	19.427	14.155	10.600*	12.956	13.611	7.564*	9.713	13.333	6.043*
0.25	0.5	19.427	32.915	11.979*	12.956	32.361	8.943*	9.713	32.083	7.421*
0.25	0.75	19.427	64.165	14.275*	12.956	63.611	11.239*	9.713	63.333	9.718
0.5	0	19.427	26.665	13.223*	12.956	26.111	10.187*	9.713	25.833	8.666*
0.5	0.25	19.427	32.915	13.683*	12.956	32.361	10.647*	9.713	32.083	9.125*
0.5	0.5	19.427	51.665	15.061*	12.956	51.110	12.025*	9.713	50.833	10.503
0.5	0.75	19.427	82.915	17.358*	12.956	82.361	14.322	9.713	82.083	12.800
0.75	0	19.427	57.915	18.361*	12.956	57.361	15.325	9.713	57.083	13.803
0.75	0.25	19.427	64.165	18.820*	12.956	63.611	15.784	9.713	63.333	14.262
0.75	0.5	19.427	82.915	20.198	12.956	82.361	17.162	9.713	82.083	15.640
1.0	0	19.427	101.665	25.553	12.956	101.111	22.517	9.713	100.833	20.995

*MSE($\hat{\underline{\beta}}_{mpc}$) is lower than either MSE($\hat{\underline{\beta}}$) or MSE($\hat{\underline{\beta}}_{pc}$).

Table 5.16 Number of Times (of 100) that $SE(\hat{\underline{\beta}}_{mpc}) < SE(\hat{\underline{\beta}})$, or $SE(\hat{\underline{\beta}}_{mpc}) < SE(\hat{\underline{\beta}}_{pc})$; 1st X Matrix

ϕ_1	ρ = 100		ρ = 200		ρ = 300	
	LS	PC	LS	PC	LS	PC
0	100	0	100	0	100	0
0.125	94	20	94	27	90	40
0.25	88	43	80	56	74	62
0.375	80	56	70	69	64	76
0.5	72	67	61	77	56	81
0.625	66	75	56	81	48	88
0.75	60	78	50	87	40	93
0.875	56	81	42	92	35	96
1.0	54	84	40	93	28	97

ϕ_1	ρ = 500		ρ = 750		ρ = 100	
	LS	PC	LS	PC	LS	PC
0	100	0	100	0	100	0
0.125	88	45	82	55	78	56
0.25	69	70	62	77	59	79
0.375	56	81	52	86	44	90
0.5	47	88	40	93	32	97
0.625	40	93	29	97	22	98
0.75	30	98	20	99	17	99
0.875	23	98	16	99	14	100
1.0	19	99	14	100	9	100

Table 5.17 Number of Times (of 100) that $SE(\hat{\underline{\beta}}_{mpc}) < SE(\hat{\underline{\beta}})$, or $SE(\hat{\underline{\beta}}_{mpc}) < SE(\hat{\underline{\beta}}_{pc})$; 2nd X Matrix

ϕ_1	ρ = 100		ρ = 200		ρ = 300	
	LS	PC	LS	PC	LS	PC
0	100	0	100	0	100	0
0.125	95	21	94	29	90	32
0.25	88	35	86	48	80	54
0.375	86	50	72	66	70	69
0.5	75	60	69	70	64	76
0.625	72	68	63	77	60	79
0.75	68	71	62	78	52	84
0.875	64	76	52	84	46	88
1.0	62	78	48	86	39	92

ϕ_1	ρ = 500		ρ = 750		ρ = 1000	
	LS	PC	LS	PC	LS	PC
0	100	0	100	0	100	0
0.125	88	42	86	47	83	52
0.25	72	68	70	69	66	74
0.375	64	75	62	78	54	82
0.5	57	81	49	85	45	89
0.625	48	86	40	92	35	94
0.75	40	92	33	96	30	97
0.875	36	94	30	97	24	97
1.0	31	96	25	97	17	99

Table 5.18 Number of Times (of 100) that $SE(\underline{\hat{\beta}}_{mpc}) < SE(\underline{\hat{\beta}})$ or $SE(\underline{\hat{\beta}}_{mpc}) < SE(\underline{\hat{\beta}}_{pc})$; 3rd X Matrix

ϕ_1	ρ = 100		ρ = 200		ρ = 300	
	LS	PC	LS	PC	LS	PC
0	100	0	100	0	100	0
0.125	91	23	86	33	81	41
0.25	79	43	75	50	71	61
0.375	73	53	63	76	59	83
0.5	66	71	57	86	56	93
0.625	60	82	56	93	53	97
0.75	56	88	54	97	52	99
0.875	56	93	52	99	52	100
1.0	55	96	52	99	51	100

ϕ_1	ρ = 500		ρ = 750		ρ - 1000	
	LS	PC	LS	PC	LS	PC
0	100	0	100	0	100	0
0.125	77	47	76	49	72	57
0.25	61	78	57	86	56	91
0.375	56	93	54	97	53	98
0.5	53	97	52	99	52	100
0.625	52	99	52	100	50	100
0.75	52	100	50	100	50	100
0.875	50	100	50	100	48	100
1.0	50	100	48	100	47	100

Table 5.19 Number of Times (of 100) that $SE(\hat{\underline{\beta}}_{mpc}) < SE(\hat{\underline{\beta}})$ or $SE(\hat{\underline{\beta}}_{mpc}) < SE(\hat{\underline{\beta}}_{pc})$ (4th X Matrix)

ϕ_1	$\rho = 100$		$\rho = 200$		$\rho = 300$	
	LS	PC	LS	PC	LS	PC
0	100	0	100	0	100	0
0.125	89	25	86	34	84	41
0.25	83	44	75	54	71	59
0.375	74	58	68	67	60	75
0.5	68	66	57	78	51	83
0.625	61	74	50	84	41	89
0.75	55	79	42	88	34	91
0.875	50	84	36	90	31	91
1.0	45	87	32	91	27	94

ϕ_1	$\rho = 500$		$\rho = 750$		$\rho = 1000$	
	LS	PC	LS	PC	LS	PC
0	100	0	100	0	100	0
0.125	81	47	77	51	73	58
0.25	66	70	58	77	54	80
0.375	51	83	44	87	37	90
0.5	40	89	33	91	31	91
0.625	32	91	27	94	20	96
0.75	27	94	20	96	15	98
0.875	20	96	15	98	12	99
1.0	16	98	12	99	10	100

Table 5.20 Number of Times (of 100) that $SE(\hat{\underline{\beta}}_{mpc}) < SE(\hat{\underline{\beta}})$ or $SE(\hat{\underline{\beta}}_{mpc}) < SE(\hat{\underline{\beta}}_{pc})$ (5th X Matrix)

ϕ_1	ρ = 100		ρ = 200		ρ = 300	
	LS	PC	LS	PC	LS	PC
0	100	0	100	0	100	0
0.125	93	20	90	32	88	37
0.25	87	42	82	52	82	54
0.375	82	54	79	61	74	64
0.5	80	60	73	76	63	69
0.725	76	63	61	71	51	79
0.75	71	66	51	79	43	85
0.875	63	69	45	83	38	88
1.0	51	79	40	86	29	91

ϕ_1	ρ = 500		ρ = 750		ρ = 1000	
	LS	PC	LS	PC	LS	PC
0	100	0	100	0	100	0
0.125	87	43	82	50	82	54
0.25	78	61	73	65	66	69
0.375	63	69	51	79	46	82
0.5	49	80	41	86	34	89
0.625	41	86	29	91	24	93
0.75	30	90	23	93	17	98
0.875	27	92	16	98	11	99
1.0	19	96	11	99	9	99

Table 5.21 Number of Times (of 100) that $SE(\hat{\underline{\beta}}_{mpc}) < SE(\hat{\underline{\beta}})$ or $SE(\hat{\underline{\beta}}_{mpc}) < SE(\hat{\underline{\beta}}_{pc})$ (6th X Matrix)

ϕ_1	ϕ_2	$\rho = 100$		$\rho = 200$		$\rho = 300$	
		LS	PC	LS	PC	LS	PC
0	0	100	0	100	0	100	0
0	0.25	100	12	99	24	99	36
0	0.5	93	44	86	73	81	85
0	0.75	85	75	75	93	68	99
0	1	77	89	67	99	59	100
0.25	0	98	12	93	25	90	35
0.25	0.25	96	26	91	45	85	59
0.25	0.5	87	51	81	75	70	87
0.25	0.75	81	77	69	93	61	99
0.5	0	86	42	77	64	66	76
0.5	0.25	85	49	72	69	59	83
0.5	0.5	80	64	62	86	52	92
0.5	0.75	74	83	57	97	49	99
0.75	0	74	66	53	85	44	92
0.75	0	71	69	54	87	44	93
0.75	0.5	63	79	48	92	43	98
1	0	57	84	41	93	32	99

Table 5.21 (Continued) (6th X Matrix)

ϕ_1	ϕ_2	ρ = 500		ρ = 750		ρ = 1000	
		LS	PC	LS	PC	LS	PC
0	0	100	0	100	0	100	0
0	0.25	89	52	86	70	83	78
0	0.5	71	97	67	99	63	99
0	0.75	61	100	50	100	48	100
0	1	49	100	37	100	35	100
0.25	0	83	45	78	60	74	66
0.25	0.25	75	73	64	84	59	89
0.25	0.5	59	97	51	99	48	100
0.25	0.75	48	100	44	100	42	100
0.5	0	51	85	42	93	37	96
0.5	0.25	50	91	41	98	35	99
0.5	0.5	45	100	40	100	33	100
0.5	0.75	42	100	36	100	31	100
0.75	0	36	97	26	100	20	100
0.75	0.25	32	99	27	100	24	100
0.75	0.5	32	100	27	100	23	100
1	0	22	100	18	100	12	100

Table 5.22 The Variance Inflation Factors for LSE and PCE Respectively

(a) 1st X Matrix

Variable	(LSE) VIF	(MPC) VIF
x_1	51.4717	9.8580
x_2	38.0357	7.7230
x_3	7.0567	2.3705
x_4	1.4551	1.4374
x_5	1.4419	1.4248
x_6	1.4683	1.4557
x_7	1.4793	1.4791
x_8	1.4190	1.4142
x_9	1.4864	1.3606
x_{10}	1.3156	1.2737

(b) 2nd X Matrix

Variable	(LSE) VIF	(MPC) VIF
x_1	99.3487	9.7706
x_2	57.2463	6.1666
x_3	44.5093	5.1041
x_4	1.1506	1.1240
x_5	1.4377	1.3461
x_6	1.1737	1.1116
x_7	1.5051	1.3973
x_8	1.2912	1.2760
x_9	1.1108	1.0643
x_{10}	1.3354	1.2345

(c) 3rd X Matrix

Variable	(LSE) VIF	(MPC) VIF
x_1	10.2891	9.1736
x_2	6.6161	5.9530
x_3	5.0027	4.5607
x_4	1.1506	1.1459
x_5	1.4377	1.4240
x_6	1.1737	1.1638
x_7	1.5051	1.4909
x_8	1.2912	1.2901
x_9	1.1108	1.1046
x_{10}	1.3355	1.3196

(d) 4th X Matrix

Variable	(LSE) VIF	(MPC) VIF
x_1	36.3284	9.8806
x_2	13.4827	4.1654
x_3	25.3506	7.1869
x_4	1.0630	1.0625
x_5	1.1489	1.1189

Table 5.22 (Continued) The Variance Inflation Factors for LSE and PCE Respectively

(c) 5th X Matrix

Variable	(LSE) VIF	(MPC) VIF
x_1	94.7532	9.7912
x_2	59.0034	6.3837
x_3	33.0418	4.5288
x_4	1.9253	1.6066
x_5	1.7639	1.6105
x_6	2.6438	2.2815
x_7	1.4955	1.3814
x_8	1.7559	1.7544
x_9	1.3322	1.3003
x_{10}	1.8254	1.6336
x_{11}	1.9023	1.6765
x_{12}	1.6792	1.5934
x_{13}	1.4769	1.3641
x_{14}	1.4988	1.4888
x_{15}	1.8268	1.6132

(f) 6th X Matrix

Variable	(LSE) VIF	(MPC) VIF
x_1	18.9445	9.7401
x_2	26.5060	9.7908
x_3	7.5085	4.3159
x_4	9.6702	5.2275
x_5	19.6083	7.5777
x_6	9.7421	3.9719
x_7	1.4935	1.4452
x_8	1.2634	1.2485
x_9	1.1358	1.0971
x_{10}	1.2677	1.1838

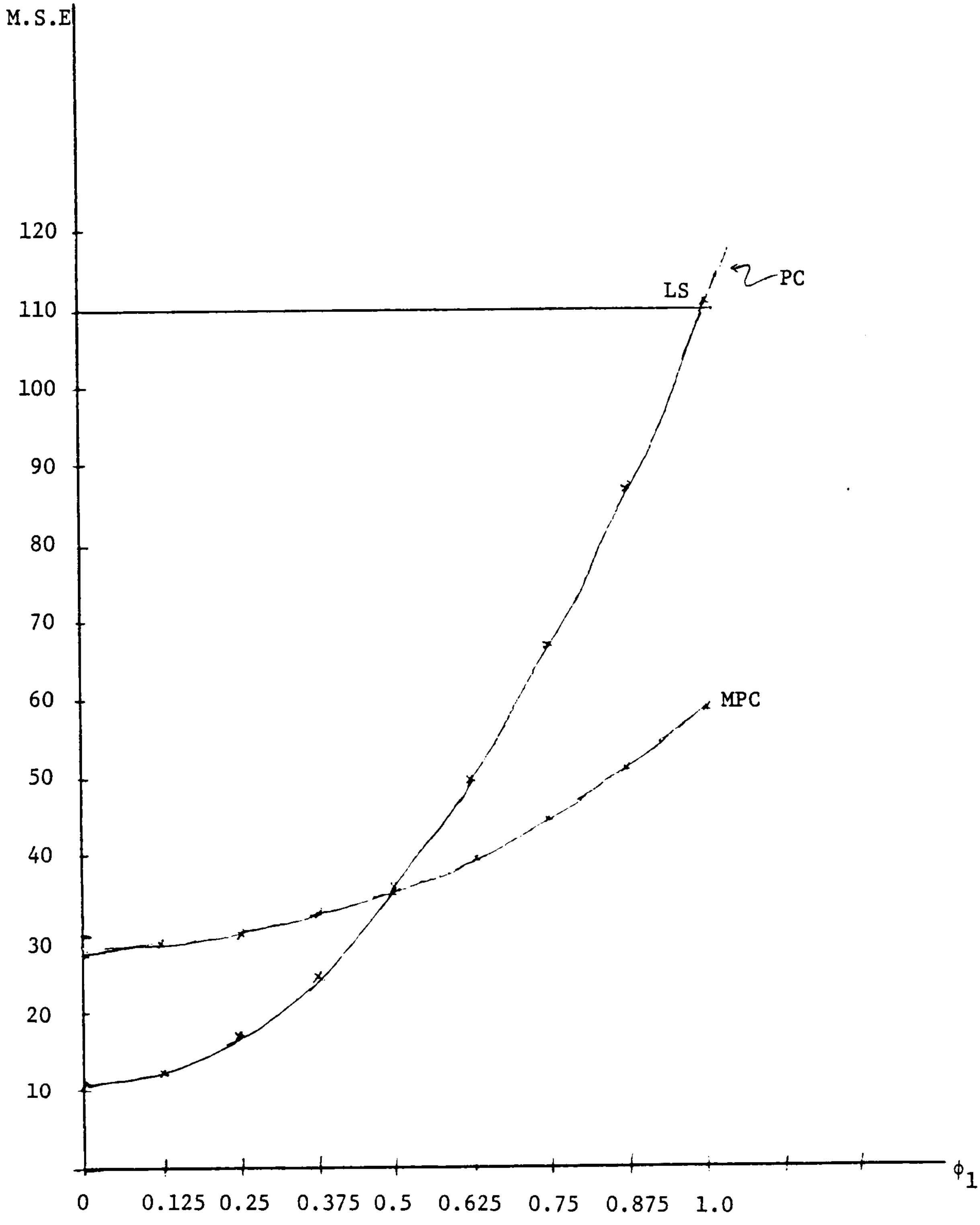

Fig. 5.1 The Relationships Between the Estimated Mean Squared Errors of LS, PC, MPC and the Orientation ϕ_1 (ℓ_1 = 0.0105992, ρ = 100)

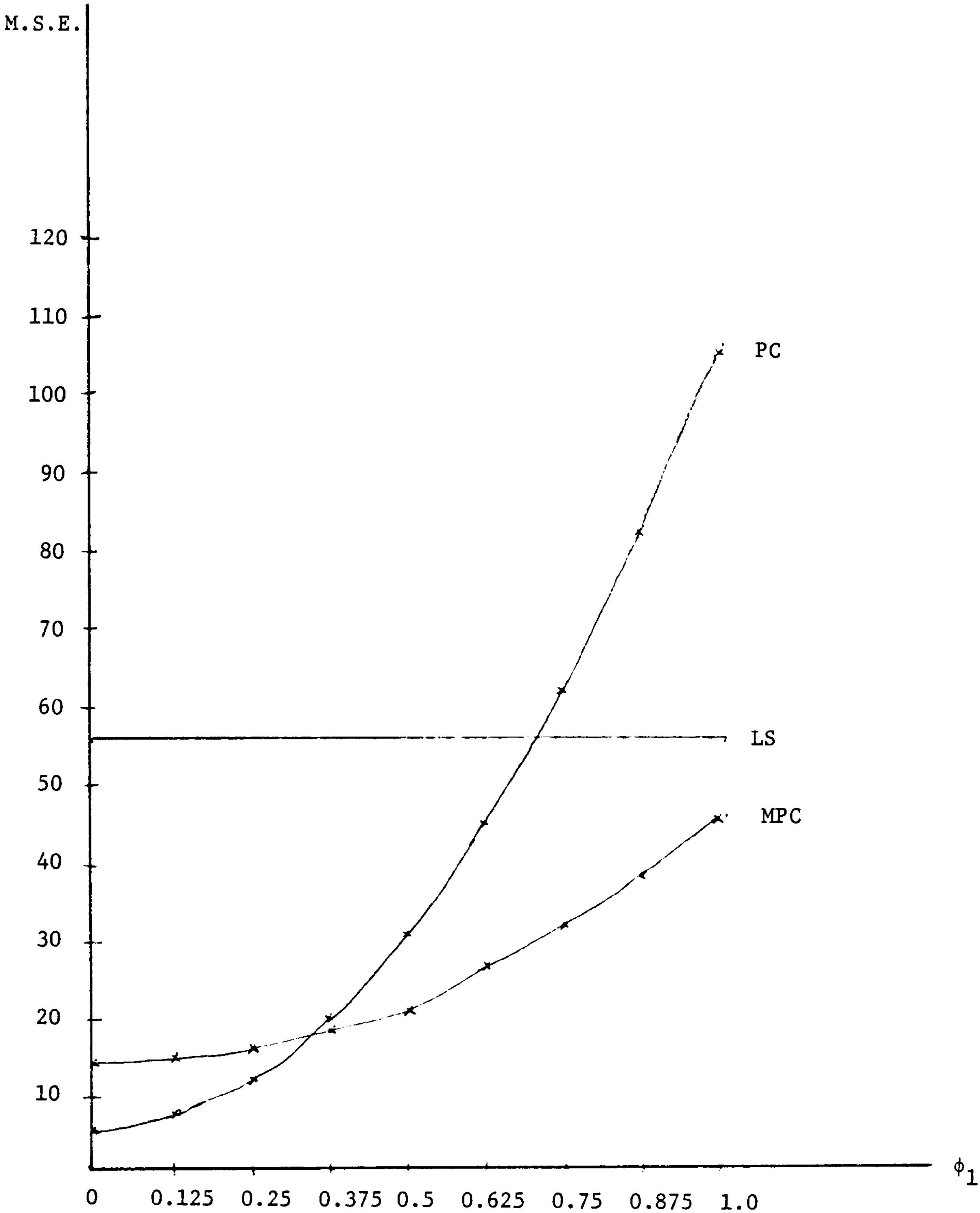

Fig. 5.2 The Relationships Between the Estimated Mean Squared Errors of LS, PC, MPC and the Orientation ϕ_1 (ℓ_1 = 0.0105992, ρ = 200)

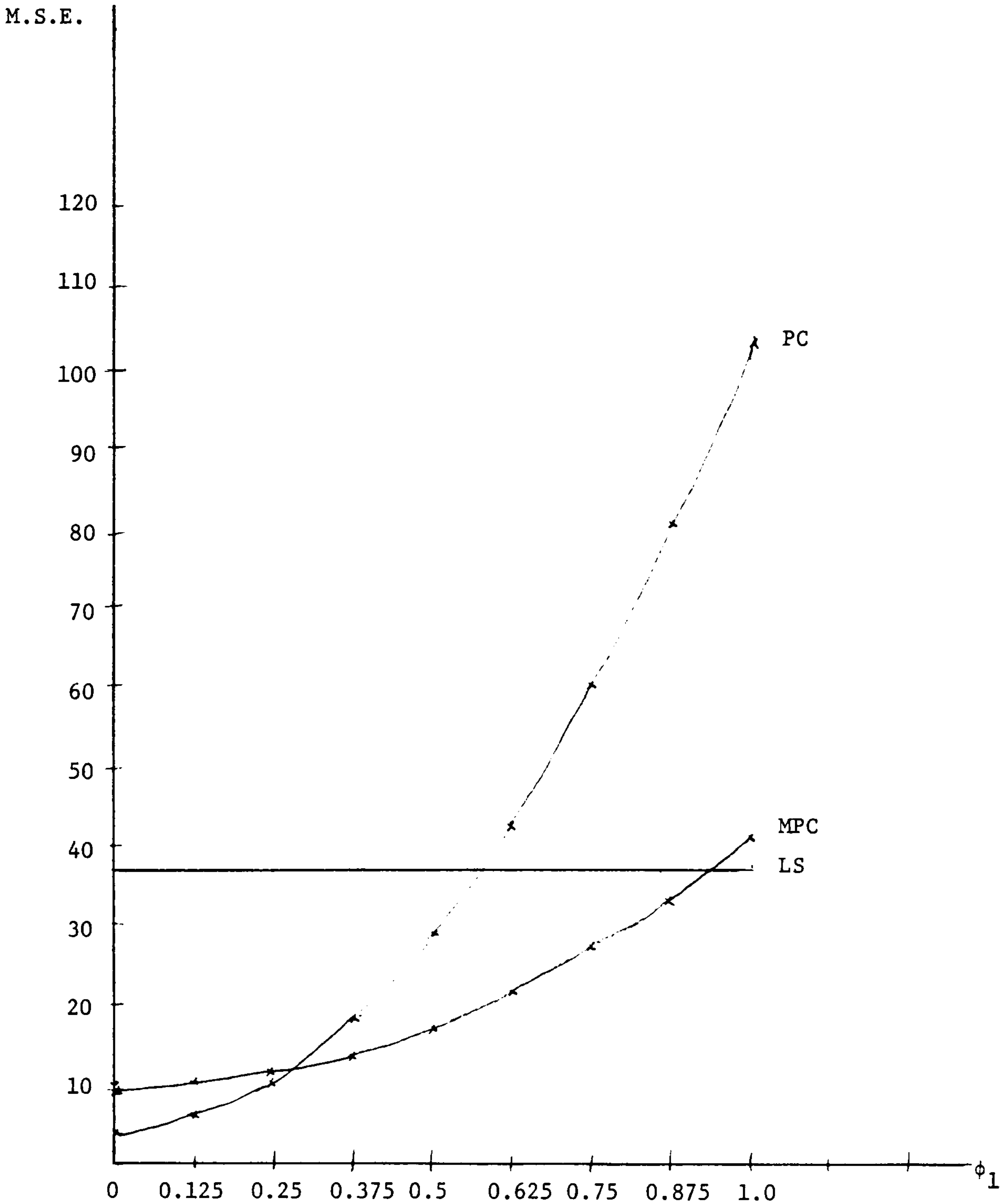

Fig. 5.3 The Relationships Between the Estimated Mean Squared Errors of LS, PC, MPC and the Orientation ϕ_1 ($\ell_1 = 0.0105992$, $\rho = 300$)

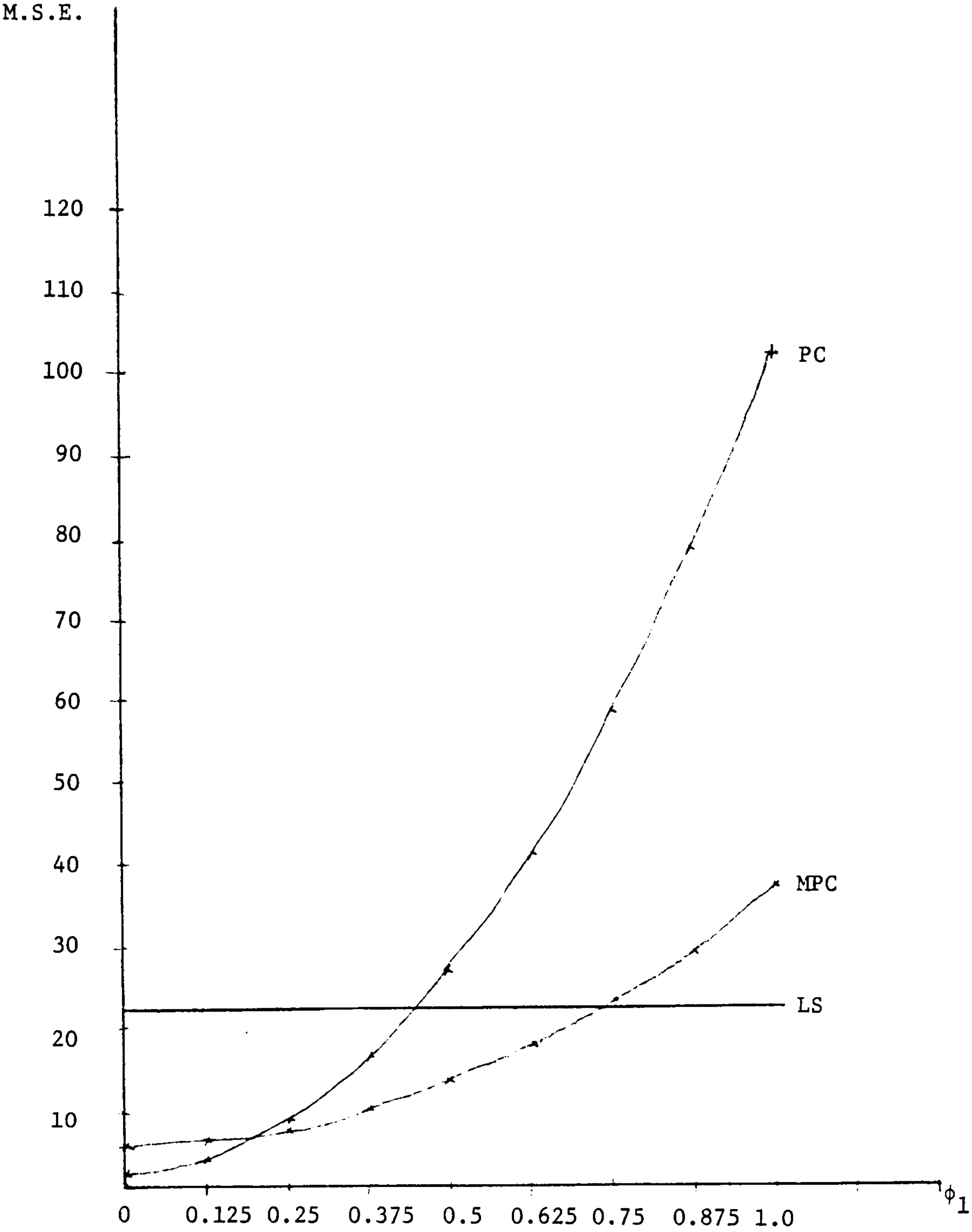

Fig. 5.4 The Relationships Between the Estimated Mean Squared Errors of LS, PC, MPC and the Orientation ϕ_1 (ℓ_1 = 0.0105992, ρ = 500)

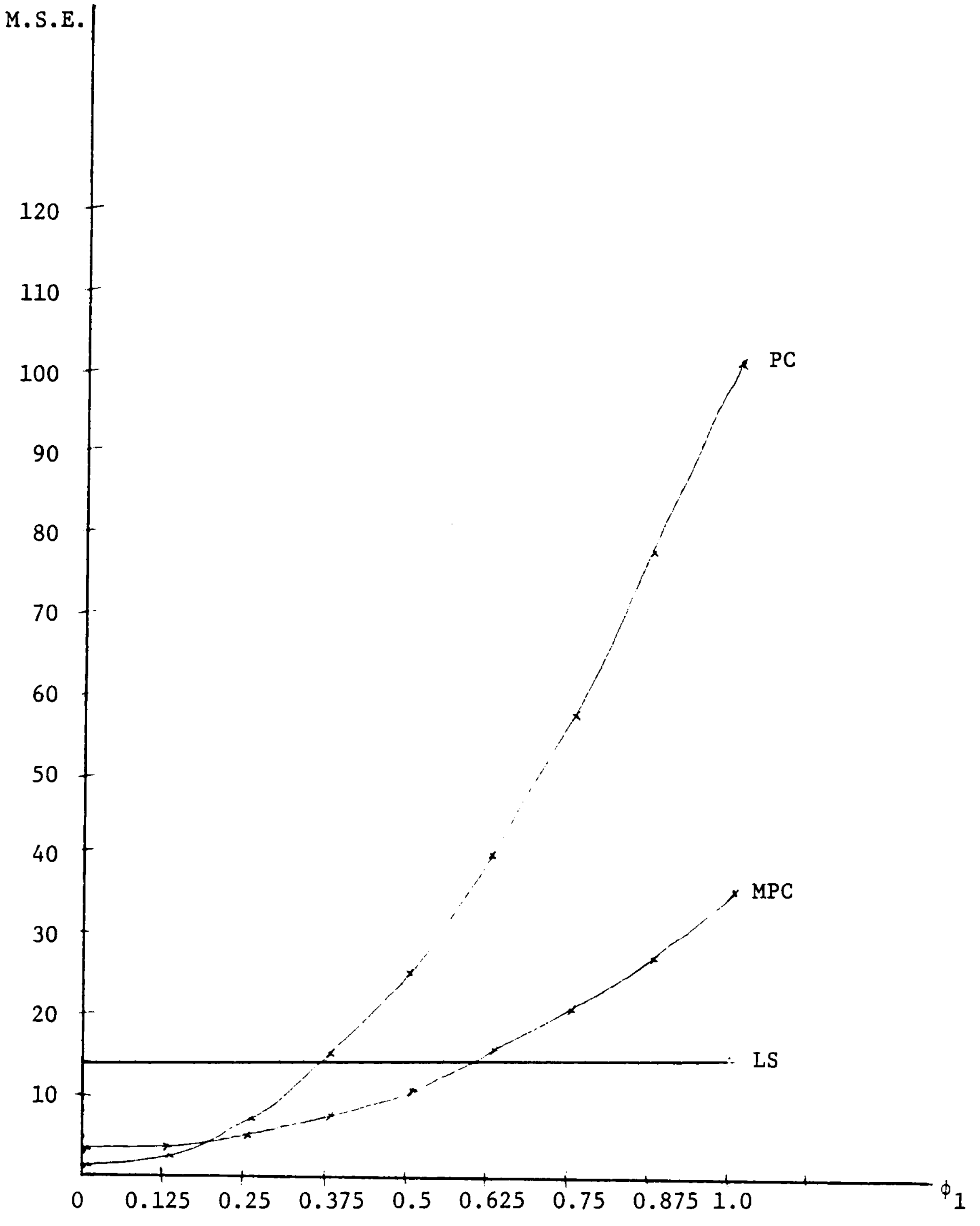

Fig. 5.5 The Relationships Between the Estimated Mean Squared Errors of LS, PC, MPC and the Orientation ϕ_1 (ℓ_1 = 0.0105992, ρ = 750)

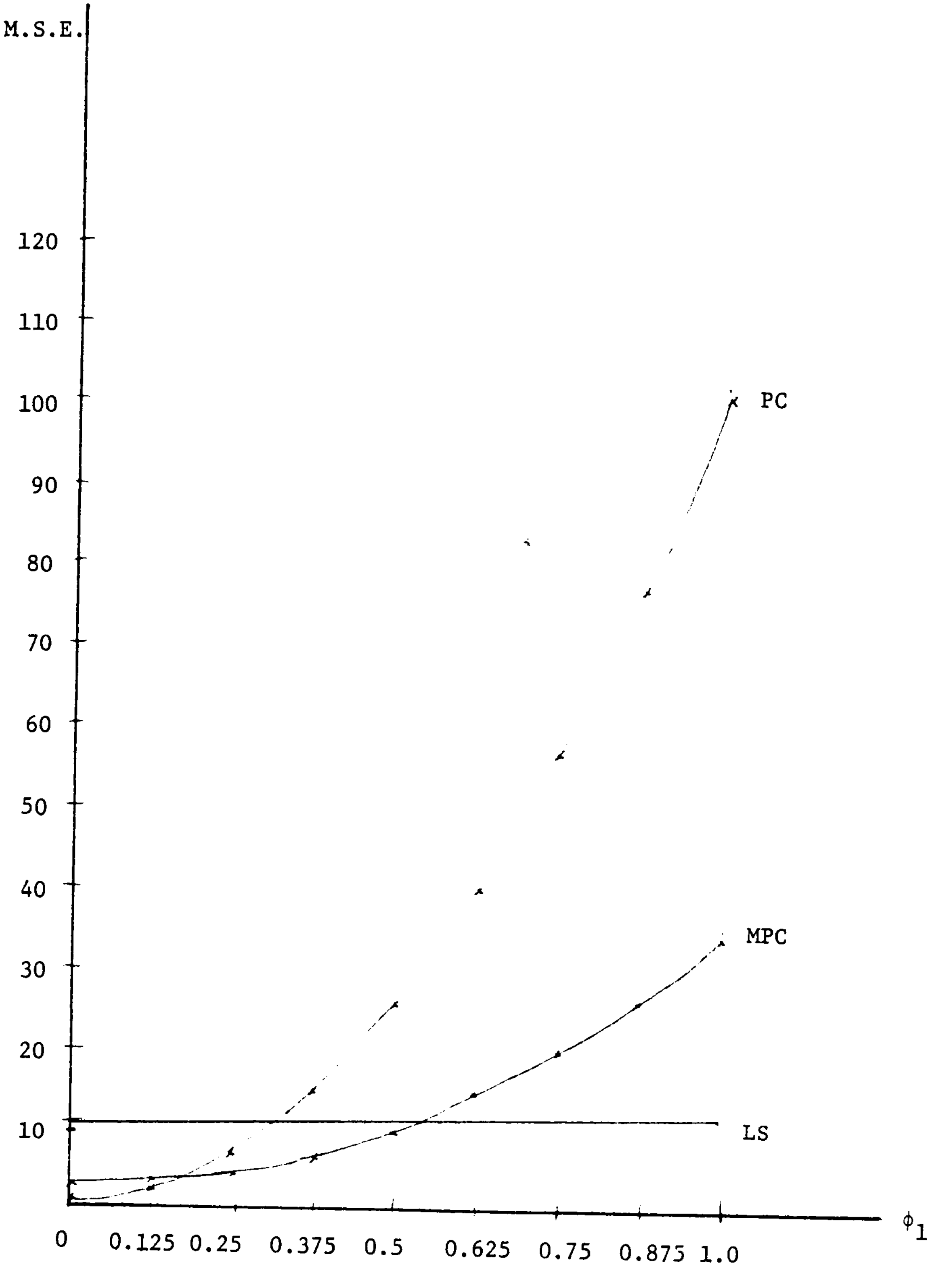

Fig. 5.6 The Relationship Between the Estimated Mean Squared Errors of LS, PC, MPC and the Orientation ϕ_1 (ℓ_1 = 0.0105992, ρ = 1000)

CHAPTER VI

CONCLUSIONS

When near linear relationships exist in the matrix of independent variables of a multiple linear regression model, the least squares estimators of the coefficients of the variables involved in the linear relationships are very unreliable. As seen in Chapter II, the least squares estimator of the regression coefficients tends to place considerable weight on the latent vectors corresponding to small latent roots of X'X. Therefore, the elements of the $\underline{\hat{\beta}}$ corresponding to the regressors involved in multicollinearities tend to be dominated by the multicollinearities. Thus, the least squares estimators are severely affected by the linear dependence among the regressors, and they could estimate the true parameters poorly.

Three biased estimation procedures which have been proposed as alternative procedures to least squares estimation, are ridge regression, principal components regression, and latent root regression.

Principal components regression is designed to remove the ill-effects of the multicollinearity by an orthogonal transformation of the data and by deleting the components associated with small latent roots of X'X in order to obtain more stable estimates of the β's (Mansfield, 1975). Several simulation studies in the statistical literature have indicated that the principal components estimator performs very well, in terms of mean squared error, relative to the least squares estimator and other estimators. (Gunst and Mason, 1977) (Mansfield, 1975) However, the performance of the principal components estimator depends very strongly on the orientation of $\underline{\beta}$ to $\underline{V}_1$, the latent vector defining the multicollinearity. The principal components estimate tends to be dominated by the bias when the $\underline{\beta}$ is parallel or nearly parallel to $\underline{V}_1$. Since the measure commonly used to test the predictive ability of a component is not reliable when the component is associated with a small latent root, principal components regression, which deletes a component associated with a small latent root without regard to its predictive ability, may result in loss of information when the deleted component has predictive value.

This paper proposes a modified principal components regression, which is a compromise procedure between the least squares regression and the principal components regression. For the modified principal components estimator, the weights are reduced on the latent vector corresponding to the small latent root from the least squares estimator while retaining the same weights on the other latent vectors, but these terms are not completely deleted as in principal components estimator.

Certain theoretical properties of the resulting estimator are investigated and are compared to those of the least squares estimator and those of the principal component estimator.

A major feature of the modified principal components estimator is that k_j's are nonstochastic. This is important because all the theoretical properties, including the expectation, the variance and the mean squared error, are derived under the assumption that k_j's are constant. This allows some measure of the precision of the estimates of the regression coefficients. Certain hypotheses about the β_j can be tested and confidence intervals for the β_j can be constructed. For this purpose, the t-statistic for the test of significance of the individual coefficient and the F-statistic for testing the general linear hypothesis have been constructed for the resulting estimator.

Since the mean squared error of the modified principal components estimator depends upon the unknown parameters, a Monte Carlo study was undertaken to investigate the performance of this estimator by using various orientations. signal-to-noise ratios, three different numbers of regressor variables, and several degrees of multicollinearity. The general results of these simulation can be summarized as follows:

1. The modified principal components estimator performs better than either of the other two estimators over a wide range of orientations and signal-to-noise ratios. It provides a reasonable compromise choice between the least squares estimator and the principal components estimator when the orientation is unknown.

2. For those orientations where the modified principal components estimator does not perform the best, it is always the second best and is always close to the best.

3. The least squares estimator is unstable for all the orientations with multicollinear data.

4. Principal components regression performs very well in those cases where $\underline{\beta}$ is orthogonal to or nearly orthogonal to $\underline{V}_1$. However,

it performs very poorly for $\phi_1 \geq 0.625$; that is, when $\underline{\beta}$ is parallel or nearly parallel to $\underline{V}_1$. For these orientations, the mean squared errors of the modified principal components estimator are much smaller than those of the principal components estimator.

Suggestion for Further Research:

In this paper, the k_js of the modified principal components estimator are determined under the situation where the orientation of $\underline{\beta}$ and $\underline{V}_1$ is unknown, since the measure commonly used to determine the predictive ability of a component is not reliable when the component is associated with a small latent root. If the orientation between $\underline{\beta}$ and $\underline{V}_1$ can be estimated accurately, the k_js of the modified principal components estimator can be determined so that the maximum variance inflation factor of x_j is equal to 1, 2.5, 5, 7.5, and 9.9 for ϕ_1 = 0, 0.25, 0.5, 0.75, and 1.0 respectively. Thus, the modified principal components regression has potential for improvement, but information is needed about the orientation of the vector of regression coefficients and the latent vector defining the multicollinearity before this potential can be realized. Therefore, further studies for determining the orientations of $\underline{\beta}$ and $\underline{V}_1$ need to be done. Also needed are investigations to determine other methods of choosing k_j in order to reduce the estimated mean squared error of this estimator.

In addition, other measures of performance, such as componentwise comparisons, can be used as criteria to evaluate estimators instead of using mean squared error. Componentwise comparison is a criterion based on estimated mean squared error of the individual regression coefficients used in the various procedures.

BIBLIOGRAPHY

Conniffe, O. and Stone, J. (1973). "A Critical View of Ridge Regression." The American Statistician, Vol. 22, No. 2:181-187.

Farebrother, R. W. (1978). "Partitioned Ridge Regression." Technometrics, Vol. 22, No. 2:121-122.

Farrar, D. E. and Glauber, R. R. (1967). "Multicollinearity in Regression Analysis: the Problem Revisited." Review of Economics and Statistics, 49:92-107.

Graybill, F. A. (1976). Theory and Application of the Linear Model, Duxbury Press, Wadsworth Publishing Company, Inc., Belmont, California.

Gunst, R. F. and Manson, R. L. (1977). "Biased Estimation in Regression: An Evaluation Using Mean Squared Error." Journal of the American Statistical Association, 72:616-628.

Gunst, R. F., Webster, J. T., and Manson, R. L. (1976). "A Comparison of Least Squares and Latent Root Regression Estimators." Technometrics, 18:75-83.

Hawkins, D. M. (1973). "On the Investigation of Alternative Regressions by Principal Component Analysis." Applied Statistics, 22:257-286.

Hemmerle, W. J. (1975). "An Explicit Solution for Generalized Ridge Regression." Technometrics, Vol. 17, No. 3:309-314.

Hemmerle, W. J. and Brantle, T. F. (1978). "Explicit and Constrained Generalized Ridge Estimation." Technometrics, Vol. 20, No. 2:109-119.

Hocking, R. R. (1976). "The Analysis and Selection of Variables in Linear Regression." Biometrics, 32:1-50.

Hoerl, A. E. and Kennard, R. W. (1970a). "Ridge Regression: Biased Estimation for Nonorthogonal Problems." Technometrics, 12:55-68.

Hoerl, A. E. and Kennard, R. W. (1970b). "Ridge Regression: Applications to Nonorthogonal Problems." Technometrics, 12:69-82.

Hoerl, A. E., Kennard, R. W., and Baldwin, K. F. (1975). "Ridge Regression: Some Simulations." Communications in Statistics, 4:105-123.

Hoerl, A. E. and Kennard, R. W. (1976). "Ridge Regression Iterative Estimation of the Biasing Parameter." Communications in Statistics: Theory and Methods, A5(1):77-88.

Jeffers, J. N. R. (1967). "Two Case Studies in the Application of Principal Component Analysis." Applied Statistics, 3:225-236.

Johnston, J. (1972). Econometric Methods, 2nd Ed., McGraw-Hill Book Co., New York. N.Y.

Lott, W. F. (1973). "The Optimal Set of Principal Component Restrictions on a Least-Squares Regression." Communications in Statistics, 2(5):449-464.

Mansfield, E. R. (1975). "Principal Component Approach to Handling Multicollinearity in Regression Analysis." Ph.D. Dissertation, Department of Statistics, Southern Methodist University.

Marquardt, D. W. (1970). "Generalized Inverses, Ridge Regression, Biased Linear Estimation and Nonlinear Estimation." Technometrics, 12:591-612.

Marquardt, D. W. and Snee, R. D. (1975). "Ridge Regression in Practice." The American Statistician, Vol. 29, No. 1:3-20.

Mason, R. L., Gunst, R. F., and Webster, J. T. (1975). "Regression Analysis and Problems of Multicollinearity." Communications in Statistics, 4(3):277-292.

Massy, W. F. (1965). "Principal Components Regression in Exploratory Statistical Research." Journal of the American Statistical Association, 60, 234-257.

Obenchain, R. L. (1977). "Classical F-Tests and Confidence Regions for Ridge Regression." Technometrics, Vol. 19, No. 4:429-439.

Silvey, S. D. (1969). "Multicollinearity and Imprecise Estimation." J. R. Statistical Society B, 31:539-552.

Webster, J. T., Gunst, R. F., and Mason, R. L. (1974). "Latent Root Regression Analysis." Technometrics, 16:513-522.

White, J. W. (1976). "Inference Procedures for Latent Root Regression Analysis." Ph.D. Dissertation, Department of Statistics, Southern Methodist University.

The vita has been removed from the scanned document

MODIFIED PRINCIPAL COMPONENTS REGRESSION

by

Huan-ter Wu

(ABSTRACT)

When near linear relationships exist among the columns of regressor variables, the variances of the least squares estimators of the regression coefficients become very large. The least squares estimator of the vector of the regression coefficients, which can be written in terms of latent roots and latent vectors of X'X, tends to place heavy weights on the latent vectors corresponding to small latent roots of X'X. Thus, the estimates of regression coefficients corresponding to the regressors involved in multicollinearities tend to be dominated by the multicollinearities. Therefore, the least squares estimators could estimate the true parameters poorly and could be very unreliable.

In order to overcome the ill-effects of multicollinearities on the least squares estimator, the procedure of principal components regression deletes those components corresponding to the small latent roots of X'X. Then we regress $\underline{y}$ on the retained components using ordinary least squares. When principal components regression is used as an alternative to the least squares in the presence of a near singular X'X matrix, its performance depends strongly on the

orientation of the deleted components to the vector of regression coefficients. In this paper, we present a modification of the principal components procedure in which components associated with near singularities are dampened but are not completely deleted.

The resulting estimator was compared in a Monte Carlo study with the least squares estimator and the principal component estimator using mean squared error as the basis of comparison. The results indicate that the modified principal components estimator will perform better than either of the other two estimators over a wide range of orientations and signal-to-noise ratios and that is provides a reasonable compromise choice when the orientation is unknown.