Hacettepe Journal of Mathematics and Statistics
Volume 43 (5) (2014), 827 – 841

Modified ridge regression parameters: A comparative Monte Carlo study

Yasin Asar* , Adnan Karaibrahimoğlu† and Aşır Genç‡

Received 25 : 07 : 2013 : Accepted 16 : 12 : 2013

Abstract

In multiple regression analysis, the independent variables should be uncorrelated within each other. If they are highly intercorrelated, this serious problem is called multicollinearity. There are several methods to get rid of this problem and one of the most famous one is the ridge regression. In this paper, we will propose some modified ridge parameters. We will compare our estimators with some estimators proposed earlier according to mean squared error (MSE) criterion. All results are calculated by a Monte Carlo simulation. According to simulation study, our estimators perform better than the others in most of the situations in the sense of MSE.

Keywords: Multicollinearity, multiple linear regression, ridge regression, ridge estimator, Monte Carlo Simulation.

1. Introduction

Multiple linear regression is one of the most widely used statistical method. Linear modeling is often used in regression analysis. But in some cases, if the correlation between dependent and independent variables is not linear, the exponential or quadratic models could be preferred. For linear regression the generalized model is as follows

$$Y = X\beta + \epsilon \tag{1.1}$$

where Y is an $n \times 1$ vector of dependent (response) variables, X is a design matrix of order $n \times p$ where p is the number of independent (explanatory) variables, β is a $p \times 1$ vector of coefficients and ϵ is an error vector of order $n \times 1$ which is distributed as $N(0, \sigma^2 I_n)$. The most common method of estimating β is to use the ordinary least squared (OLS)

*Department of Mathematics–Computer Sciences, Necmettin Erbakan University, Konya 42090, Turkey,
Email: yasar@konya.edu.tr
†Necmettin Erbakan University, Meram Faculty of Medicine,Medical Education and Informatics Department, Biostatistics Unit, Konya 42080, Turkey,
Email:akara@konya.edu.tr
‡Department of Statistics, Selcuk University, Konya, 42250, Turkey,
Email:agenc@selcuk.edu.tr

estimator where the residual sum of squares is minimized. The OLS estimator of β is written as

$$\hat{\beta} = (X'X)^{-1}X'Y. \tag{1.2}$$

Therefore Equation (1.2) is an unbiased estimator of β. However, this equation is valid under several assumptions of multiple linear regression analysis. Regression assumptions clarify the conditions under which multiple regression works well, ideally with unbiased and efficient estimates. One of the assumptions is that explanatory variables are uncorrelated to one another. In many situations, this assumption is not the case since the variables are highly intercorrelated. Thus the design matrix X becomes linearly dependent and not having full rank. This leads that the matrix becomes close to singular. This problem is called multicollinearity. Multicollinearity is perfect or sometimes called exact if the predictors are highly correlated. Then the regression coefficients are indeterminate and their standard errors are infinite. If multicollinearity is moderate, then the regression coefficients are determinate but possess large standard errors which mean that the coefficients cannot be estimated with great accuracy [7].

It is an especially common problem in time series regressions, that is, where the data consists of a series of observations on the variables over a number of time periods. If two or more of the explanatory variables have a strong time trend, they will be highly correlated and this condition may give rise to multicollinearity. It should be noted that the presence of multicollinearity does not mean that the model is misspecified. Accordingly, the regression coefficients remain unbiased and the standard errors remain valid [6].

When the matrix $X'X$ is close to be singular, the numerical reliability of the calculations is reduced. In extreme cases it is possible that the reported calculations will be wrong. A more relevant implication of near multicollinearity is that individual coefficient estimates will be imprecise. We can see this most simply in homoscedastic linear regression models. As the correlation coefficient ρ approaches to 1, the matrix $X'X$ becomes singular. Therefore the collinearity is indexed by ρ. In this case we can also observe that the variance of a coefficient estimate

$$\sigma^2[n(1-\rho^2)]^{-1} \tag{1.3}$$

tends to infinity. It can be easily seen that the expression (1.3) depends on the correlation ρ and the sample size n [8].

Collinearity is also an important phenomenon in sampling. Unbiasedness is a repeated sampling property. The explanatory variables in the population may not be linearly related but they may be related in the particular sample. Therefore one should pay attention in dealing with the sample values and computing the OLS estimators for each of these samples [7].

There are many methods to overcome multicollinearity. One of the most commonly used method is the ridge regression (RR) firstly suggested by Hoerl and Kennard [9, 10]. If the matrix $X'X$ is ill-conditioned, then the OLS estimate $\hat{\beta}$ becomes instable. The main idea of RR is to add a positive constant k to the diagonals of $X'X$ before computing $\hat{\beta}$. Then the solution vector becomes

$$\hat{\beta}_{ridge} = (X'X + kI_p)^{-1}X'Y \tag{1.4}$$

where $k > 0$ is called a biased (ridge) estimator and I_p is the $p \times p$ identity matrix. The question is how to choose k. In equation (1.4), if $k = 0$, $\hat{\beta}_{ridge}$ becomes the unbiased OLS estimator whereas if $k \neq 0$, $\hat{\beta}_{ridge}$ is the biased ridge estimator of β. There are various numbers of proposed ridge estimatorss. In papers, we see that the researchers compared their estimators with the one proposed by Hoerl et al. [11]. Some examples of the researchers studying in this area are the followings: Lawless and Wang [16], McDonald and Galarneau [19], Saleh and Kibria [26], Kibria [15], Khalaf and Shukur [14], Alkhamisi et al. [2], Norliza et al. [23], Akhamisi and Shukur [3], Sakallioglu and Kaciranlar [25], Muniz and Kibria [20], Liu and Gao [17], Mansson et al. [18], Dorugade and Kashid [5], Al-Hassan [1], Muniz et al. [21], Dorugade [4] and Karaibrahimoğlu et al.[13].

The aim of this study is to propose some new modified ridge estimators and compare some of them to early proposed estimators from the literature. Multicollinearity is a serious problem and it must be overcomed to get an accurate result in regression models. Therefore we will choose the best k parameter after comparison according to mean squared error (MSE) criterion. In section 2, we will explain the theoretical background of the ridge regression and define the new estimators. In section 3, we will give the details of the Monte Carlo simulation. We will present the results and discussions in section 4. Conclusively, we will make some comments on the results and choose the best estimator of k. All the tables of the simulation results and the graphs will be present in the appendix.

2. Ridge Regression and Ridge Estimators

2.1. Detection of Multicollinearity. There are several methods to detect multicollinearity problem. However some of them do not tell anything about the degree of multicollinearity. Two of the most commonly used methods are the followings

(1) Condition number: Let $\lambda_1, \lambda_2, ..., \lambda_p$ are eigenvalues of the matrix $X'X$. Let λ_{max} and λ_{min} denote the maximum and minimum of these eigenvalues, respectively. The condition number κ is defined as

$$\kappa = \frac{\lambda_{max}}{\lambda_{min}}. \tag{2.1}$$

If $10 < \sqrt{\kappa} < 30$, then there is an intermediate multicollinearity and if $\sqrt{\kappa} > 30$ there is a severe multicollinearity in the model. One can also see that if λ_{min} is equal to 0 or very close to 0 then the ratio is infinite. Equivalently, if λ_{min} and λ_{max} are close to each other, then the value of $\kappa = 1$ or close to 1, meaning that the predictors are said to be orthogonal. That is, there is no collinearity problem.

(2) Variance Inflation Factor (VIF): VIF value is computed as

$$VIF = \frac{1}{1 - R_j^2} \tag{2.2}$$

where R_j^2 is the coefficient of determination in the regression of explanatory variable X_j on the remaining explanatory variables of the model. Generally, when $VIF > 10$, it is assumed that there exists highly multicollinearity.

2.2. Theoretical Background of Ridge Regression. Two American statisticians, Arthur Hoerl and Robert Kennard, published a paper in 1970 on ridge regression, a method for solving badly conditioned linear regression problems. Bad conditioning means numerical difficulties in performing the inverse of a matrix which is necessary to obtain the variance matrix. Meanwhile, the Russian theoretician Andre Tikhonov (1977) was working on the solution of ill-posed problems for which no unique solution exists because, in effect, there is not enough information specified in the problem. Hoerl and Kennard called the method as "ridge regression" whereas Tikhonov developed a method known as regularization. Hoerl and Kennard's method was in fact a crude form of regularization [12].

One of the main methods to solve the multicollinearity problem is principle component regression. This method can be considered as the basis of the ridge regression idea.

Principal component regression (PCR) is a regression analysis that uses principal component analysis when estimating regression coefficients. It is a procedure used to overcome problems which arise when the multicollinearity exists. Often the principal components with the highest variance are selected. However, the low-variance principal components may also be important in some cases even more important [24].

The idea of ridge regression was developed based on PCR in the same manner. After writing the general regression equation in canonical form, the orthogonalization process is applied. At this point, ridge regression differs from the principal component regression. The OLS estimator of β is obtained via using different components in latter whereas the biased estimator of β is calculated adding a parameter k to the diagonal elements in former. The theoretical procedure of ridge regression is as follows: Consider the general model given in (1.1). First, write this equation in canonical form. Suppose that there exits an orthogonal matrix Q such that

$$Q'X'XQ = \Lambda = diag(\lambda_1, \lambda_2, ..., \lambda_p) \tag{2.3}$$

where Q is a $p \times p$ orthogonal matrix and Λ is a $p \times p$ diagonal matrix whose diagonal elements are the eigenvalues $\lambda_1, \lambda_2, ..., \lambda_p$ of $X'X$. Thus, we obtain the equivalent model

$$Y = Z\alpha + \epsilon \tag{2.4}$$

where $Z = XQ$, $Z'Z = \Lambda$ and $\alpha = Q'\beta$. Therefore the OLS estimator of α is defined as

$$\hat{\alpha} = \Lambda^{-1}Z'Y. \tag{2.5}$$

Also, the OLS estimator of β is given by

$$\hat{\beta} = Q\hat{\alpha}. \tag{2.6}$$

We can write the ridge estimator of $\hat{\alpha}$ as

$$\hat{\alpha}_{ridge} = (Z'Z + kI_p)^{-1}Z'Y \tag{2.7}$$

and the ordinary ridge estimator of β is given as

$$\hat{\beta}_{ridge} = Q(I - kA_k^{-1})\hat{\alpha} \tag{2.8}$$

where $A_k = (\Lambda + kI_p)$.

It is known since Hoerl and Kennard [9] that the value of k which minimizes the $MSE(\hat{\alpha}_{ridge})$ is

$$k_i = \frac{\sigma^2}{\alpha_i^2} \tag{2.9}$$

where

$$MSE(\hat{\alpha}_{ridge}) = \sigma^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k)^2} + k^2 \sum_{i=1}^{p} \frac{\alpha_i^2}{(\lambda_i + k)^2}. \tag{2.10}$$

The first part of the above function is the variance function and the second part is the squared bias function. We know that the variance function is a continuous and monotonically decreasing function and the squared bias function is a continuous and monotonically increasing function of k [9].

We obtain $MSE(\hat{\alpha}_{OLS})$ if we put $k = 0$ in equation (2.10). Thus

$$MSE(\hat{\alpha}_{OLS}) = \sigma^2 \sum_{i=1}^{p} \frac{1}{\lambda_i}. \tag{2.11}$$

It is obvious from equation (2.10) that k depends on σ^2 and α. But we don't know σ^2 and α in practice. Therefore we use the estimators $\hat{\sigma}^2$ and $\hat{\alpha}$ instead of the unknown parameters σ^2 and α respectively such that $k_i = \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}$ where $\hat{\sigma}^2 = (Y - X\hat{\beta})'(Y - X\hat{\beta})/(n-p)$ [9, 10].

2.3. Proposed Ridge Estimators. It is proved that a sufficient condition that $MSE(\hat{\alpha}_{Ridge}) < MSE(\hat{\alpha}_{OLS})$ is the following $k < k_{HK} = \frac{\hat{\sigma}^2}{\hat{\alpha}_{\max}^2}$ [9]. In the literature, some of the parameters are smaller than k_{HK}. However, it is possible to get good estimators (in the sense of MSE) bigger than k_{HK}. As it can be seen from the figure given in [9], k_{HK} value makes the first derivative of the function $MSE(\hat{\alpha}_{Ridge})$ smaller than zero. Thus, any estimator k satisfying the sufficient condition given in [9] such that $0 < k < k_{HK}$ makes the derivative smaller than zero as well. However, the intersection point of the graphs of the variance and the squared bias functions is definitely greater than k_{HK}. One can also find estimators at which point the first derivative of the $MSE(\hat{\alpha}_{Ridge})$ may be greater than zero. There are some estimators greater than k_{HK}, for example see [3] for the estimators $k_{NAS} = \max(\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2} + \frac{1}{\lambda_i})$ and $k_{AS} = k_{HK} + \frac{1}{\lambda_i}, i = 1, ..., p$ which are clearly greater than k_{HK}.

Now, we suggest some estimators which are modifications of $k_K = \frac{1}{p}\sum_{i=1}^{p}\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}$ proposed in [15] and $k_{AD} = \frac{2p}{\lambda_{\max}}\frac{\hat{\sigma}^2}{\sum_{i=1}^{p}\hat{\alpha}_i^2}$ proposed in [4]. We apply some transformations namely, taking i^{th} root of the parameter or taking squares or cubes of the parameters, the number of predictors p and the eigenvalues λ_i following [3] and [14] in order to obtain some new estimators greater than k_{HK} and have better performance.

Now, we give some ridge estimators proposed earlier and our new proposed estimators. We compare the performances of the estimators given below according to the average MSE values. The descriptions of earlier estimators are presented with indices belonging to their author's name.

(1) $k_{HK} = \frac{\hat{\sigma}^2}{\hat{\alpha}_{\max}^2}$ where $\hat{\alpha}_{\max}$ is the maximum element of $\hat{\alpha}_{OLS}$. (Hoerl and Kennard, 1970a)
(2) $k_{NAS} = \max(\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2} + \frac{1}{\lambda_i})$ which is the maximum element of $k_i = \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2} + \frac{1}{\lambda_i}$ being greater than k_{HK}. (Alkhamisi and Shukur, $pp.543, 2007$)
(3) $k_K = \frac{1}{p}\sum_{i=1}^{p}\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}$ which is the arithmetic mean of $k_i = \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}$. (Kibria, $pp.423, 2003$)
(4) $k_{AD} = \frac{2p}{\lambda_{\max}}\frac{\hat{\sigma}^2}{\sum_{i=1}^{p}\hat{\alpha}_i^2}$ which is the harmonic mean of $k_i = \frac{2\hat{\sigma}^2}{\lambda_{\max}\hat{\alpha}_i^2}$. (Dorugade, $pp.3, 2013$)
(5) $k_{KM8} = \max\left(\frac{1}{\frac{\lambda_{\max}\hat{\alpha}_i^2}{(n-p)\hat{\sigma}^2+\lambda_{\max}\hat{\alpha}_i^2}}\right)$ which is the maximum element of $k_i = \frac{1}{\frac{\lambda_{\max}\hat{\alpha}_i^2}{(n-p)\hat{\sigma}^2+\lambda_{\max}\hat{\alpha}_i^2}}$. (Mansson et. al, $pp.5, 2010$)
(6) $k_{KM12} = median\left(\frac{1}{\frac{\lambda_{\max}\hat{\alpha}_i^2}{(n-p)\hat{\sigma}^2+\lambda_{\max}\hat{\alpha}_i^2}}\right)$ which is the median of $k_i = \frac{1}{\frac{\lambda_{\max}\hat{\alpha}_i^2}{(n-p)\hat{\sigma}^2+\lambda_{\max}\hat{\alpha}_i^2}}$. (Mansson et. al, $pp.5, 2010$)

Our proposed estimators are as follows:

(1) $k_{AY1} = \frac{p^2}{\lambda_{\max}^2}\frac{\hat{\sigma}^2}{\sum_{i=1}^{p}\hat{\alpha}_i^2}$
(2) $k_{AY2} = \frac{p^3}{\lambda_{max}^3}\frac{\hat{\sigma}^2}{\sum_{i=1}^{p}\hat{\alpha}_i^2}$
(3) $k_{AY3} = \frac{p}{\lambda_{max}^{1/3}}\frac{\hat{\sigma}^2}{\sum_{i=1}^{p}\hat{\alpha}_i^2}$
(4) $k_{AY4} = \frac{p}{(\sum_{i=1}^{p}\sqrt{\lambda_i})^{1/3}}\frac{\hat{\sigma}^2}{\sum_{i=1}^{p}\hat{\alpha}_i^2}$
(5) $k_{AY5} = \frac{2p}{\sqrt{\lambda_{\max}}}\frac{\hat{\sigma}^2}{\sum_{i=1}^{p}\hat{\alpha}_i^2}$

Since the matrix $X'X$ is in the correlation form, we have $\sum_{i=1}^{p} \lambda_i = p$. So $p > \lambda_{\max}$. Thus, all of the new proposed estimators are clearly greater than k_{HK}. We will show in the next section that our estimators have better performance than the given early proposed estimators especially when there is severe multicollinearity.

3. Application: A Monte Carlo Simulation

This section is related to the Monte Carlo simulation. In conducting the simulation we compare the performances of the estimators. For a valuable Monte Carlo simulation two criteria are used in design. One criterion is to determine the effective factors affecting the properties of the estimators. The other one is to specify the criteria of judgment. We choose the sample size n, the number of predictors p, the correlation coefficient ρ and the variances between the error terms σ^2 as effective factors. Also the mean squared error (MSE) is chosen to be the criteria for comparison of the performances. However, in literature, we realized that the average MSE is used to compare the performances of the estimators. Thus we computed the average MSE (AMSE) values of all estimators with respect to different effective factors. There are many ridge estimators proposed in papers and we have simulated many of them, but we give six of them that are $k_{HK}, k_{NAS}, k_K, k_{AD}, k_{KM8}$ and k_{KM12} in this study. Additionally, we suggest five new ridge parameters that are $k_{AY1}, k_{AY2}, k_{AY3}, k_{AY4}$ and k_{AY5}.

The general regression model (1.1) is considered with independent error terms, that is $IID - \epsilon \sim N(0, \sigma^2 I)$. If β is chosen to be the eigenvector of the largest eigenvalue of the matrix $X'X$ such that $\beta'\beta = 1$, then the minimized value of MSE is obtained [22].

In order to generate the explanatory variables, the following common device is used:

$$x_{ij} = (1 - \rho^2)^{1/2} z_{ij} + \rho z_{ip} \tag{3.1}$$

where $i = 1, 2, ..., n$, $j = 1, 2, ..., p$, ρ^2 represents the correlation between the explanatory variables and z_{ij}'s are independent random numbers obtained from standard normal distribution. We can generate the dependent variable Y with the following equation:

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_p x_{ip}, i = 1, 2, ..., n \tag{3.2}$$

where ϵ_i's are independent and identically normal distributed pseudorandom numbers with zero mean and variance σ^2 and $\beta_0 = 0$.

In the simulation, we consider different cases of effective factors on estimators: $n = 50, 100, 150$, $\rho^2 = 0.95, 0.99, 0.999$, $p = 4, 6$ and $\sigma^2 = 0.1, 0.5, 1.0$. For the given values of $\rho^2 = 0.95, 0.99, 0.999$, we consider the following condition numbers of the generated data sets, respectively, $\kappa \approx 15$,$\kappa \approx 30$,$\kappa \approx 90$ for $p = 4$, and $\kappa \approx 25$, $\kappa \approx 45$, $\kappa \approx 120$ for $p = 6$.

First we generated the matrix of explanatory variables X and the vector of dependent variable Y, and then we standardized both X and Y in such a way that $X'X$ and $X'Y$ are in the correlation form. For different values of n, p, ρ and σ^2, the iteration was performed 5.000 times by generating the error terms of the general linear regression equation (1.1).

We computed average mean squared errors of the estimators via the following equation, $AMSE(\hat{\alpha}) = \frac{1}{5000} \sum_{r=1}^{5000} (\hat{\alpha} - \alpha)'(\hat{\alpha} - \alpha)$ where $\hat{\alpha}$ is $\hat{\alpha}_{OLS}$ and $\hat{\alpha}_{Ridge}$.

The biasedness plays secondary important role in comparing the performances of the estimators and thus we created the squared bias tables. All the results are given in tables. To comprehend the tables, we illustrated the results in graphs. All tables and figures are given in Appendix.

4. Results and Discussion

In Table. 1, AMSE values of the estimators are presented for fixed n, p, ρ and different σ^2's. There are 18 sub-tables in the arrangement of n, p, ρ trio. All of the proposed estimators have better performance than k_{HK} and OLS estimator having the largest AMSE values. One can see from tables that when the error variance σ^2 increases, AMSE values increases for all estimators. For the case $p = 4, \rho = 0.95$, k_{AY2} has the least AMSE value and other new proposed estimators except for k_{AY3} have better performance than the ones chosen from the literature. When we increase the correlation, k_{AY5} and k_{AD} have less AMSE values than other estimators for $\rho = 0.99$ and $\rho = 0.999$. If we compare the estimators for the case $p = 6$, one can see that all of the new proposed estimators except for k_{AY3} are quite better than the others, especially k_{AY4} is the best among them for $\rho = 0.95$ and $\rho = 0.99$. However, k_{AY5} and k_{AD} perform almost equally for $\rho = 0.999$. These results can easily be seen from Figure 7.1 as well.

All comparison graphs have been plotted using earlier estimators k_{NAS}, k_K, k_{AD} and new estimators k_{AY2}, k_{AY3} selected randomly. Since the values of OLS and k_{HK} estimators are larger than the others as a scale, they are not included in the graphs. We know that multicollinearity becomes severe when the correlation increases. However, there is an interesting result that AMSE values of new proposed estimators decrease, when the correlation increases. In other words, new proposed estimators are robust to the correlation. This feature is also observed for k_{AD} and presented in Figure. 7.3. Although MSE is used as a comparison criterion, the bias of an estimator is another indicator of good performance. Thus, we have provided the squared bias values of the estimators in Table. 2. In most of the cases, k_{HK} has the least bias value (Figure. 7.4(a)). If we increase the correlation, our new estimators have less bias as it is observed from Figure. 7.5. When $\rho = 0.999, k_K$ becomes the estimator having least bias. Also k_{AY1}, k_{AY2} and k_{AY4} have quite less biases for this situation as well. If the biasses are compared to the error variance, one can see that when the error variance increases, AMSE values of all estimators except for k_{KM8} and k_{KM12} increases monotonically when $\rho = 0.95$ and $p = 4$. Figure. 7.5, obviously, shows us the performances of the bias values between earlier and new parameters.

5. Summary and Conclusion

In this paper, we studied ridge regression and ridge estimators. We reviewed six old estimators and proposed five new estimators. We compared all of the estimators according to mean squared error and the squared bias criteria. We have conducted a Monte Carlo simulation to compare the results by generating random numbers for dependent and independent variables and pseudo-random numbers for the error terms from the standard normal distribution. We created tables consisting of AMSE values according to different values of the sample size n, the correlation coefficient between the explanatory variables ρ, the number of predictors p and the variance of error terms σ^2. We plotted some graphs for selected situations. According to tables and figures, we may say that our new suggested ridge estimators are better than the older ones proposed in [3, 4, 9, 10, 15, 20, 21]. We concluded that our new estimators are functional for solving multicollinearity problem. Finally, among our estimators, k_{AY1} and k_{AY2} are the best ones as a performance of AMSE and the bias. The superiority of new estimators changes according to the situation. Although k_{AY2} has the best performance in the sense of AMSE, k_{HK} gives better results in the sense of bias. However, in applying a ridge estimator to a real data, just one estimator is not enough to get rid of the collinearity problem. One should notice that every data set is different and each has its own statistical characteristics, and each of the proposed estimators has its own superiority too. Therefore, we advise researchers that one should apply many estimators until getting a better solution to encounter this problem.

6. Acknowledgment

The authors express their gratitude to the editor and the referees for quick process, valuable suggestions and helpful comments.

7. Appendix

Table 1. AMSE values of the estimators fixed n, p, ρ and different σ^2's

A: $n = 50, p = 4, \rho = 0.95$

σ^2	k_{HK}	k_{NAS}	k_K	k_{AD}	k_{KM8}	k_{KM12}	k_{AY1}	k_{AY2}	k_{AY3}	k_{AY4}	k_{AY5}	OLS
0.1	2.4842	0.8896	0.6768	0.3321	0.8602	0.7360	0.3106	0.3097	0.3509	0.3136	0.3303	8.7159
0.5	12.0972	0.8991	1.1215	0.4413	0.8547	0.7054	0.4194	0.4176	0.4631	0.4248	0.4392	45.0968
1.0	20.5269	0.8892	1.3803	0.5346	0.8480	0.7002	0.5234	0.5203	0.5531	0.5326	0.5324	78.8936

B: $n = 50, p = 4, \rho = 0.99$

σ^2	k_{HK}	k_{NAS}	k_K	k_{AD}	k_{KM8}	k_{KM12}	k_{AY1}	k_{AY2}	k_{AY3}	k_{AY4}	k_{AY5}	OLS
0.1	9.0286	0.9681	0.9821	0.2140	0.9329	0.8075	0.2372	0.2359	0.2236	0.2398	0.2139	33.8276
0.5	46.3548	0.9649	0.9760	0.3110	0.9369	0.8055	0.2989	0.2984	0.3300	0.2999	0.3107	171.7628
1.0	80.5670	0.9653	1.0929	0.4210	0.9293	0.7887	0.4105	0.4099	0.4391	0.4117	0.4207	311.1999

C: $n = 50, p = 4, \rho = 0.999$

σ^2	k_{HK}	k_{NAS}	k_K	k_{AD}	k_{KM8}	k_{KM12}	k_{AY1}	k_{AY2}	k_{AY3}	k_{AY4}	k_{AY5}	OLS
0.1	84.0152	0.9968	0.6063	0.0905	0.9887	0.9358	0.1022	0.1022	0.0951	0.1024	0.0905	331.5741
0.5	446.4061	0.9967	0.3491	0.2279	0.9900	0.9372	0.1981	0.1981	0.2453	0.1980	0.2279	1729.6702
1.0	996.2741	0.9967	0.4311	0.3490	0.9904	0.9418	0.3189	0.3190	0.3659	0.3189	0.3490	3715.5852

D: $n = 50, p = 6, \rho = 0.95$

σ^2	k_{HK}	k_{NAS}	k_K	k_{AD}	k_{KM8}	k_{KM12}	k_{AY1}	k_{AY2}	k_{AY3}	k_{AY4}	k_{AY5}	OLS
0.1	2.5995	0.8731	0.4942	0.4906	0.8005	0.5462	0.4058	0.4161	0.5640	0.3873	0.4839	10.4329
0.5	15.7874	0.8617	1.3840	0.4689	0.8200	0.5287	0.4107	0.4130	0.5393	0.4082	0.4653	62.2329
1.0	25.5628	0.8650	1.6620	0.5650	0.7982	0.5071	0.5097	0.5133	0.6300	0.5070	0.5600	106.6135

E: $n = 50, p = 6, \rho = 0.99$

σ^2	k_{HK}	k_{NAS}	k_K	k_{AD}	k_{KM8}	k_{KM12}	k_{AY1}	k_{AY2}	k_{AY3}	k_{AY4}	k_{AY5}	OLS
0.1	11.0756	0.9673	1.0008	0.2799	0.9226	0.6596	0.2508	0.2506	0.3328	0.2514	0.2790	49.2345
0.5	72.1513	0.9660	1.6474	0.3038	0.9385	0.7002	0.2788	0.2784	0.3600	0.2796	0.3033	293.2311
1.0	145.7039	0.9662	1.7144	0.3841	0.9388	0.7014	0.3548	0.3545	0.4421	0.3554	0.3835	594.9463

F: $n = 50, p = 6, \rho = 0.999$

σ^2	k_{HK}	k_{NAS}	k_K	k_{AD}	k_{KM8}	k_{KM12}	k_{AY1}	k_{AY2}	k_{AY3}	k_{AY4}	k_{AY5}	OLS
0.1	114.8290	0.9956	1.3833	0.1018	0.9905	0.9121	0.1139	0.1138	0.1176	0.1141	0.1018	474.1472
0.5	481.0077	0.9956	0.8418	0.1966	0.9887	0.8942	0.1727	0.1727	0.2382	0.1727	0.1966	2076.1783
1.0	1388.6341	0.9955	0.5608	0.2817	0.9922	0.9291	0.2458	0.2458	0.3286	0.2457	0.2816	5357.3082

G: $n = 100, p = 4, \rho = 0.95$

σ^2	k_{HK}	k_{NAS}	k_K	k_{AD}	k_{KM8}	k_{KM12}	k_{AY1}	k_{AY2}	k_{AY3}	k_{AY4}	k_{AY5}	OLS
0.1	2.1653	0.8868	0.5850	0.3674	0.8885	0.8160	0.3296	0.3305	0.3889	0.3295	0.3647	7.8563
0.5	10.4590	0.8815	1.0933	0.4407	0.8850	0.8041	0.4159	0.4145	0.4628	0.4207	0.4385	39.3817
1.0	18.6198	0.8803	1.3710	0.5459	0.8804	0.8001	0.5348	0.5314	0.5636	0.5456	0.5434	71.2297

H: $n = 100, p = 4, \rho = 0.99$

σ^2	k_{HK}	k_{NAS}	k_K	k_{AD}	k_{KM8}	k_{KM12}	k_{AY1}	k_{AY2}	k_{AY3}	k_{AY4}	k_{AY5}	OLS
0.1	9.0049	0.9644	0.9785	0.1939	0.9446	0.8615	0.2189	0.2177	0.2028	0.2214	0.1938	33.3718
0.5	45.8357	0.9651	0.9566	0.3079	0.9459	0.8598	0.2934	0.2929	0.3274	0.2943	0.3076	170.3914
1.0	86.2032	0.9675	1.0202	0.4195	0.9430	0.8512	0.4021	0.4017	0.4390	0.4029	0.4191	332.3339

I: $n = 100, p = 4, \rho = 0.999$

σ^2	k_{HK}	k_{NAS}	k_K	k_{AD}	k_{KM8}	k_{KM12}	k_{AY1}	k_{AY2}	k_{AY3}	k_{AY4}	k_{AY5}	OLS
0.1	94.3318	0.9968	0.4769	0.0794	0.9904	0.9516	0.0852	0.0851	0.0844	0.0853	0.0794	361.6911
0.5	515.7422	0.9966	0.3070	0.2212	0.9914	0.9557	0.1922	0.1922	0.2376	0.1921	0.2212	1909.1391
1.0	882.1237	0.9966	0.4494	0.3515	0.9899	0.9475	0.3215	0.3215	0.3686	0.3214	0.3514	3402.4834

J: $n = 100, p = 6, \rho = 0.95$

σ^2	k_{HK}	k_{NAS}	k_K	k_{AD}	k_{KM8}	k_{KM12}	k_{AY1}	k_{AY2}	k_{AY3}	k_{AY4}	k_{AY5}	OLS
0.1	2.8249	0.8647	0.5239	0.4654	0.8385	0.6705	0.3832	0.3909	0.5390	0.3692	0.4601	11.1193
0.5	14.2895	0.8642	1.2808	0.4982	0.8394	0.6493	0.4326	0.4364	0.5700	0.4273	0.4938	57.8058
1.0	27.6500	0.8619	1.5735	0.5491	0.8376	0.6499	0.4880	0.4912	0.6176	0.4842	0.5450	113.1446

K: $n = 100, p = 6, \rho = 0.99$

σ^2	k_{HK}	k_{NAS}	k_K	k_{AD}	k_{KM8}	k_{KM12}	k_{AY1}	k_{AY2}	k_{AY3}	k_{AY4}	k_{AY5}	OLS
0.1	12.6430	0.9665	1.0644	0.2600	0.9363	0.7490	0.2352	0.2350	0.3094	0.2359	0.2594	54.2543
0.5	69.3508	0.9657	1.5355	0.3113	0.9418	0.7584	0.2810	0.2808	0.3681	0.2814	0.3107	284.3854
1.0	133.7324	0.9650	1.6167	0.3825	0.9397	0.7572	0.3481	0.3480	0.4423	0.3484	0.3819	551.2684

L: $n = 100, p = 6, \rho = 0.999$

σ^2	k_{HK}	k_{NAS}	k_K	k_{AD}	k_{KM8}	k_{KM12}	k_{AY1}	k_{AY2}	k_{AY3}	k_{AY4}	k_{AY5}	OLS
0.1	106.5513	0.9955	1.1515	0.1038	0.9905	0.9214	0.1073	0.1072	0.1225	0.1074	0.1038	447.8673
0.5	552.2347	0.9954	0.7148	0.1864	0.9907	0.9224	0.1640	0.1640	0.2246	0.1640	0.1863	2268.7474
1.0	1108.7436	0.9956	0.6246	0.2960	0.9906	0.9245	0.2567	0.2567	0.3462	0.2566	0.2959	4628.2175

M: $n = 150, p = 4, \rho = 0.95$

σ^2	k_{HK}	k_{NAS}	k_K	k_{AD}	k_{KM8}	k_{KM12}	k_{AY1}	k_{AY2}	k_{AY3}	k_{AY4}	k_{AY5}	OLS
0.1	2.1744	0.8751	0.6332	0.3407	0.9082	0.8613	0.3105	0.3104	0.3615	0.3120	0.3387	7.7457
0.5	8.7842	0.8620	1.0892	0.4620	0.9025	0.8540	0.4332	0.4320	0.4844	0.4382	0.4592	32.7942
1.0	18.7026	0.8720	1.3276	0.5320	0.9050	0.8534	0.5155	0.5131	0.5512	0.5233	0.5297	71.8151

N: $n = 150, p = 4, \rho = 0.99$

σ^2	k_{HK}	k_{NAS}	k_K	k_{AD}	k_{KM8}	k_{KM12}	k_{AY1}	k_{AY2}	k_{AY3}	k_{AY4}	k_{AY5}	OLS
0.1	9.0232	0.9622	1.0105	0.1907	0.9521	0.8913	0.2164	0.2153	0.1990	0.2186	0.1906	32.8662
0.5	41.1732	0.9637	0.9941	0.3143	0.9499	0.8815	0.3012	0.3006	0.3340	0.3025	0.3138	156.7651
1.0	75.3489	0.9643	1.0904	0.4411	0.9474	0.8782	0.4257	0.4251	0.4603	0.4270	0.4406	291.9112

O: $n = 150, p = 4, \rho = 0.999$

σ^2	k_{HK}	k_{NAS}	k_K	k_{AD}	k_{KM8}	k_{KM12}	k_{AY1}	k_{AY2}	k_{AY3}	k_{AY4}	k_{AY5}	OLS
0.1	86.9109	0.9967	0.5220	0.0886	0.9906	0.9547	0.0958	0.0957	0.0936	0.0959	0.0886	338.6129
0.5	478.9288	0.9966	0.3001	0.2150	0.9915	0.9594	0.1842	0.1843	0.2323	0.1842	0.2149	1831.0742
1.0	896.9235	0.9966	0.4209	0.3458	0.9909	0.9561	0.3157	0.3157	0.3628	0.3156	0.3457	3471.8863

P: $n = 150, p = 6, \rho = 0.95$

σ^2	k_{HK}	k_{NAS}	k_K	k_{AD}	k_{KM8}	k_{KM12}	k_{AY1}	k_{AY2}	k_{AY3}	k_{AY4}	k_{AY5}	OLS
0.1	2.8426	0.8588	0.5603	0.4486	0.8590	0.7396	0.3698	0.3763	0.5215	0.3584	0.4439	10.9617
0.5	14.3225	0.8524	1.3054	0.4855	0.8606	0.7260	0.4214	0.4247	0.5570	0.4168	0.4815	57.1206
1.0	24.6180	0.8573	1.5599	0.5806	0.8555	0.7277	0.5204	0.5247	0.6460	0.5154	0.5757	101.3096

Q: $n = 150, p = 6, \rho = 0.99$

σ^2	k_{HK}	k_{NAS}	k_K	k_{AD}	k_{KM8}	k_{KM12}	k_{AY1}	k_{AY2}	k_{AY3}	k_{AY4}	k_{AY5}	OLS
0.1	12.2021	0.9651	1.1453	0.2550	0.9396	0.7921	0.2346	0.2343	0.3014	0.2355	0.2545	51.3831
0.5	64.4840	0.9645	1.5561	0.2996	0.9428	0.7912	0.2710	0.2708	0.3567	0.2716	0.2990	272.0730
1.0	132.2975	0.9648	1.6828	0.3974	0.9422	0.7895	0.3639	0.3638	0.4566	0.3643	0.3967	545.3588

R: $n = 150, p = 6, \rho = 0.999$

σ^2	k_{HK}	k_{NAS}	k_K	k_{AD}	k_{KM8}	k_{KM12}	k_{AY1}	k_{AY2}	k_{AY3}	k_{AY4}	k_{AY5}	OLS
0.1	106.2218	0.9954	1.1887	0.1045	0.9904	0.9269	0.1086	0.1085	0.1235	0.1087	0.1045	444.5699
0.5	527.4235	0.9954	0.6478	0.1903	0.9905	0.9297	0.1638	0.1638	0.2312	0.1637	0.1903	2227.5910
1.0	1061.1549	0.9954	0.6507	0.2898	0.9903	0.9275	0.2544	0.2545	0.3377	0.2544	0.2898	4423.6314

Table 2. Bias values of the estimators for fixed n, p, ρ and different σ^2's

A: $n = 50, p = 4, \rho = 0.95$

σ^2	k_{HK}	k_{NAS}	k_K	k_{AD}	k_{KM8}	k_{KM12}	k_{AY1}	k_{AY2}	k_{AY3}	k_{AY4}	k_{AY5}
0.1	0.0007	0.8893	0.0367	0.2041	0.8583	0.7345	0.1258	0.1302	0.2346	0.1174	0.2009
0.5	0.0060	0.8986	0.0518	0.2933	0.8508	0.7001	0.1836	0.1901	0.3339	0.1710	0.2889
1.0	0.0036	0.8881	0.0651	0.3592	0.8426	0.6905	0.2328	0.2421	0.4030	0.2149	0.3534

B: $n = 150, p = 4, \rho = 0.95$

σ^2	k_{HK}	k_{NAS}	k_K	k_{AD}	k_{KM8}	k_{KM12}	k_{AY1}	k_{AY2}	k_{AY3}	k_{AY4}	k_{AY5}
0.1	0.0005	0.8748	0.0396	0.2177	0.9078	0.8611	0.1354	0.1401	0.2495	0.1264	0.2144
0.5	0.0072	0.8610	0.0591	0.3215	0.9017	0.8532	0.2063	0.2149	0.3626	0.1898	0.3159
1.0	0.0076	0.8704	0.0647	0.3603	0.9038	0.8516	0.2334	0.2429	0.4042	0.2151	0.3543

C: $n = 100, p = 6, \rho = 0.99$

σ^2	k_{HK}	k_{NAS}	k_K	k_{AD}	k_{KM8}	k_{KM12}	k_{AY1}	k_{AY2}	k_{AY3}	k_{AY4}	k_{AY5}
0.1	0.0172	0.9955	0.0022	0.0274	0.9904	0.9198	0.0133	0.0134	0.0457	0.0133	0.0273
0.5	0.1614	0.9954	0.0024	0.0596	0.9907	0.9206	0.0254	0.0254	0.1056	0.0253	0.0596
1.0	0.1643	0.9956	0.0068	0.1173	0.9905	0.9224	0.0555	0.0556	0.1902	0.0553	0.1172

D: $n = 100, p = 6, \rho = 0.999$

σ^2	k_{HK}	k_{NAS}	k_K	k_{AD}	k_{KM8}	k_{KM12}	k_{AY1}	k_{AY2}	k_{AY3}	k_{AY4}	k_{AY5}
0.1	0.0017	0.9665	0.0169	0.1443	0.9354	0.7465	0.0840	0.0851	0.2070	0.0818	0.1434
0.5	0.0136	0.9657	0.0186	0.1759	0.9409	0.7534	0.0988	0.0999	0.2566	0.0968	0.1750
1.0	0.0242	0.9650	0.0190	0.2165	0.9386	0.7503	0.1201	0.1214	0.3129	0.1173	0.2154

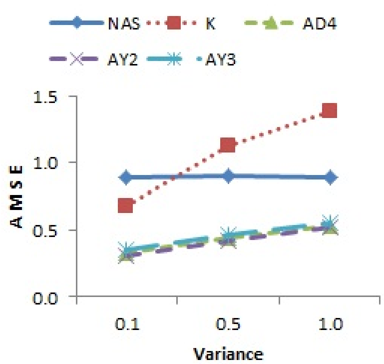
Table 3. Bias values of the estimators for fixed p, σ^2 and different ρ

A: $n = 50, p = 4, \sigma^2 = 0.1$

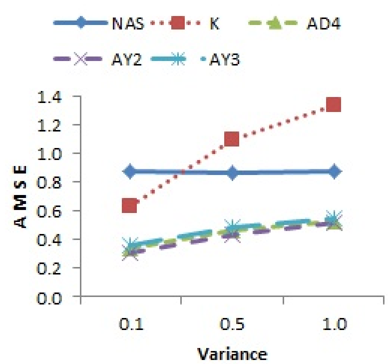
ρ^2	k_{HK}	k_{NAS}	k_K	k_{AD}	k_{KM8}	k_{KM12}	k_{AY1}	k_{AY2}	k_{AY3}	k_{AY4}	k_{AY5}
0.95	0.0007	0.8893	0.0367	0.2041	0.8583	0.7345	0.1258	0.1302	0.2346	0.1174	0.2009
0.99	0.0023	0.9681	0.0128	0.0875	0.9316	0.8048	0.0475	0.0480	0.1061	0.0465	0.0871
0.999	0.0004	0.9968	0.0016	0.0188	0.9886	0.9343	0.0083	0.0083	0.0247	0.0082	0.0188

B: $n = 50, p = 4, \sigma^2 = 1.0$

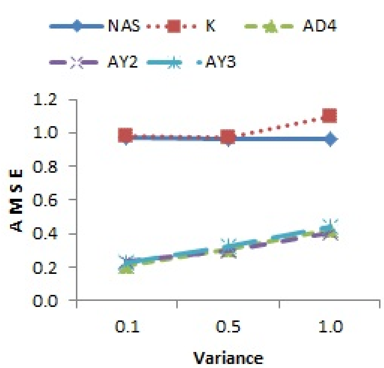
ρ^2	k_{HK}	k_{NAS}	k_K	k_{AD}	k_{KM8}	k_{KM12}	k_{AY1}	k_{AY2}	k_{AY3}	k_{AY4}	k_{AY5}
0.95	0.0036	0.8881	0.0651	0.3592	0.8426	0.6905	0.2328	0.2421	0.4030	0.2149	0.3534
0.99	0.0094	0.9652	0.0221	0.2003	0.9270	0.7805	0.1059	0.1071	0.2410	0.1035	0.1993
0.999	0.9092	0.9967	0.0061	0.1109	0.9902	0.9391	0.0477	0.0478	0.1423	0.0476	0.1109



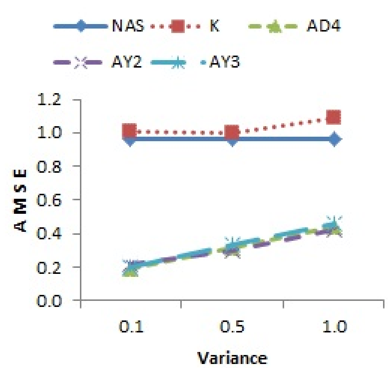
(a) n=50,ρ=0.95, p=4



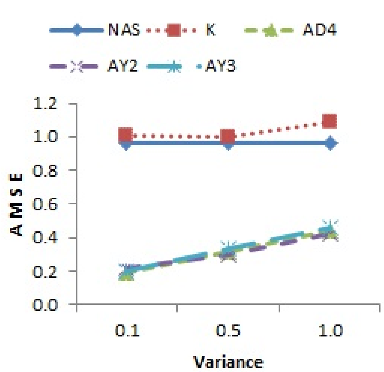
(b) n=150,ρ=0.95, p=4



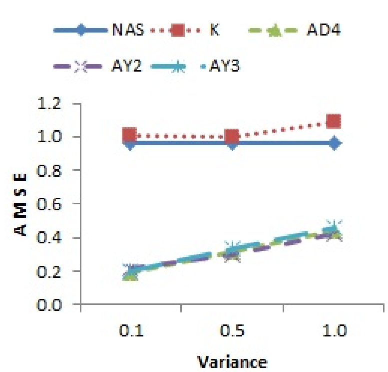
(c) n=50,ρ=0.99, p=4



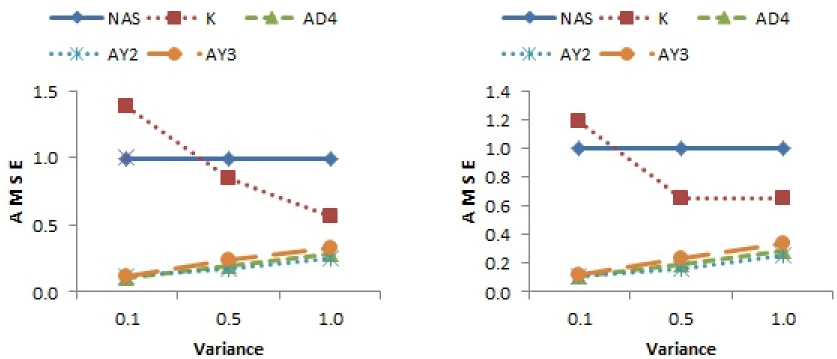
(d) n=150,ρ=0.99, p=4



(e) n=50,ρ=0.95, p=6



(f) n=150,ρ=0.95, p=6



(g) n=50,ρ=0.999, p=6 (h) n=100,ρ=0.999, p=6

Figure 7.1. Comparison graphs of AMSE Values for selected k's

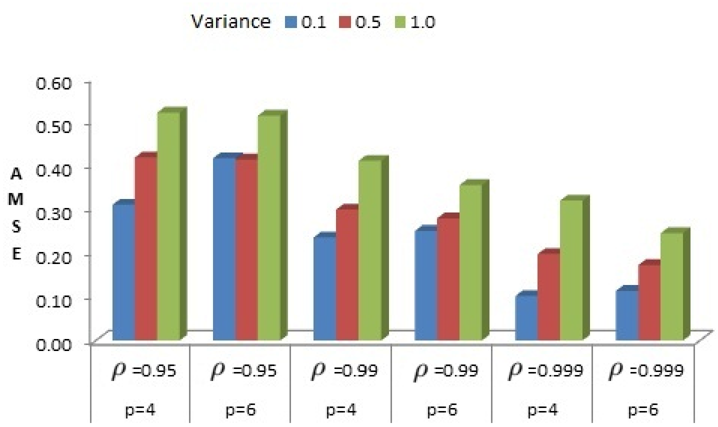
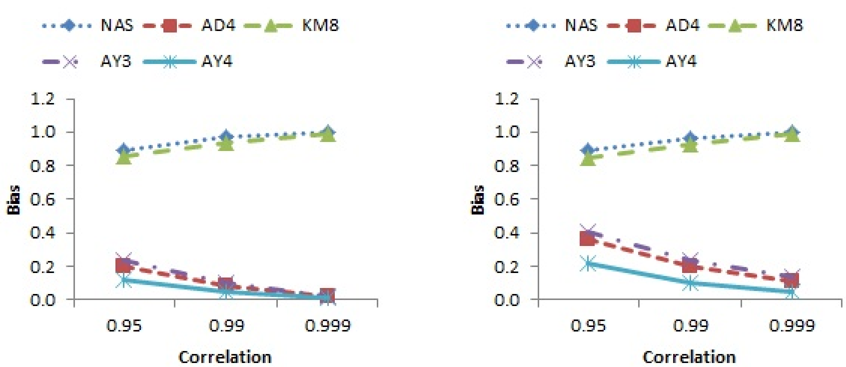
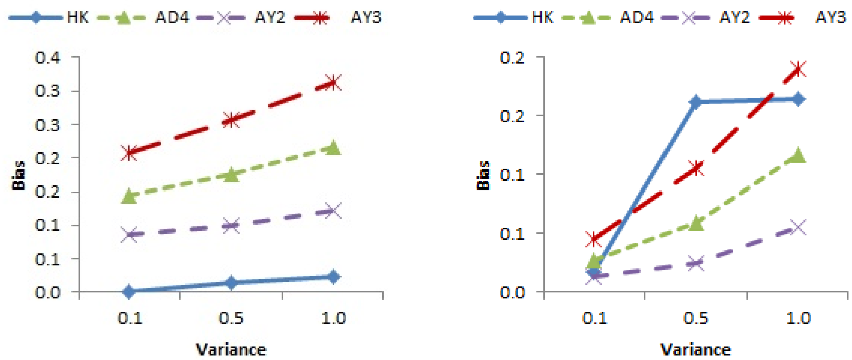


Figure 7.2. Bar graphs of different variance values of k_{AY2} for different p and ρ



(a) n=50, σ^2=0.1, p=4 (b) n=50, σ^2=1.0, p=4

Figure 7.3. Comparison graphs of bias values with respect to different correlations for selected k's



(a) n=100, ρ=0.99, p=6 (b) n=100, ρ=0.999, p=6

Figure 7.4. Comparison graphs of bias values with respect to different variances for selected k's

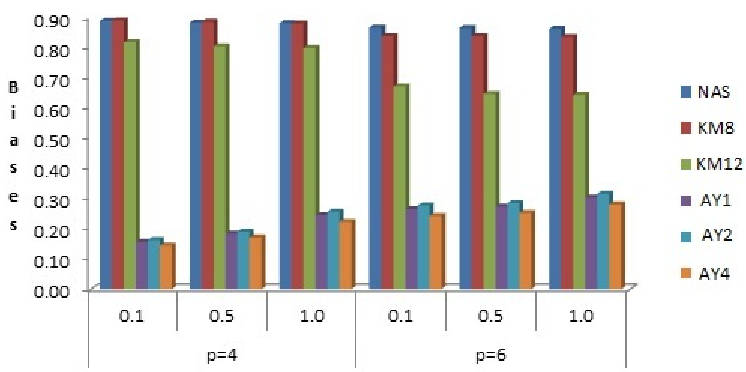


Figure 7.5. Bar graphs of bias values of selected k's for different p's and variances

References

[1] Al-Hassan, Y., Performance of new ridge regression estimators, Journal of the Association of Arab Universities for Basic and Applied Science 9, 23-26, 2010

[2] Alkhamisi, M., Khalaf, G. and Shukur, G., Some modifications for choosing ridge parameters, Communications in Statistics-Theory and Methods, 35, 2005-2020, 2006

[3] Alkhamisi, M. and Shukur, G., A Monte Carlo study of recent ridge parameters, Communications in Statistics - Simulation and Computation 36 (3), 535-547, 2007

[4] Dorugade, A.V., New ridge parameters for ridge regression, Journal of the Association of Arab Universities for Basic and Applied Science http://dx.doi.org/10.1016/j.jaubas.2013.03.005, 2013

[5] Dorugade, A.V. and Kashid, D.N., Alternative method for choosing ridge parameter for regression, International Journal of Applied Mathematical Sciences 4 (9), 447-456, 2010

[6] Dougherty, C., Introduction to Econometrics, (Oxford University Press, 2011)

[7] Gujarati, D., Basic Econometrics, (McGraw Hill Publications, 2003).

[8] Hansen, B., Econometrics (Wisconsin University, 2012) www.ssc.wisc.edu/ bhansen.

[9] Hoerl, A. E. and Kennard, R. W. Ridge regression: biased estimation for non-orthogonal problems, Technometrics, 12, 55-67, 1970a.

[10] Hoerl, A. E. and Kennard, R. W. Ridge regression: application to non-orthogonal problems, Technometrics, 12, 69-82, 1970b.

[11] Hoerl, A. E., Kennard, R. W. and Baldwin, K. F. Ridge regression: some simulation, Communications in Statistics, 4, 105-123, 1975.

[12] Jolliffe, I. T., A note on the Use of Principal Components in Regression , Journal of the Royal Statistical Society, Series C Vol. 31(3), 300-303, 1982.

[13] Karaibrahimoğlu A., Asar Y., Genç A., Some new modifications of Kibria's and Dorugade's methods: An application to Turkish GDP data, Journal of the Association of Arab Universities for Basic and Applied Sciences, Available online 7 October 2014, ISSN 1815-3852, http://dx.doi.org/10.1016/j.jaubas.2014.08.005.

[14] Khalaf, G.and Shukur, G., Choosing ridge parameter for regression problem, Communications in Statistics–Theory and Methods, 34, 1177-1182, 2005.

[15] Kibria, G., Performance of some new ridge regression estimators, Communications in Statistics-Theory and Methods, 32, 419-435, 2003.

[16] Lawless, J. F. and Wang, P. A, Simulation study of ridge and other regression estimators, Communications in Statistics A, 5, 307-323, 1976.

[17] Lıu, Xu-Qing and Gao, F., Linearized ridge regression estimator in linear regression, Communications in Statistics-Theory and Methods, 40, 2182-2192, 2011.

[18] Mansson K., Shukur G. and Kibria B. G., On Some Ridge Regression Estimators: A Monte Carlo Simulation Study Under Different Error Variances, Journal of Statistics, 17, 1-22, 2010.

[19] McDonald, G. C. and Galarneau, D. I., A Monte Carlo evaluation of some ridge-type estimators, Journal of the American Statistical Association, 70, 407-416, 1975.

[20] Muniz, G. and Kibria, G., On some ridge regression estimators: an empirical comparisons, Communications in Statistics-Simulation and Computation, 38(3), 621-630, 2009

[21] Muniz, G., Kibria, G, Mansson K. and Shukur G., On developing ridge regression parameters: a graphical investigation, SORT, 36 (2), 115-138, 2012.

[22] Newhouse, J. P., and Oman, S. D., An evaluation of ridge estimators, Rand Corporation, P-716-PR,1-28, 1971.

[23] Norliza, A., Maizah, H.A.and Robin, A., A comparative study on some methods for handling multicollinearity problems, Mathematika 22(2), 109-119, 2006.

[24] Paul, R.K., Multicollinearity: Causes, Effects And Remedies, 2006 http://iasri.res.in/seminar/as299/ebooks

[25] Sakallioglu S. and Kaciranlar S., A new biased estimator based on ridge estimation, Stat Papers, 49, 669-689, 2008.

[26] Saleh, A.K.and Kibria, B.M., Performances of some new preliminary test ridge regression estimators and their properties, Communications in Statistics - Theory and Methods, 22, 2747-2764, 1993.