Risk Performance of a Weighted Average Estimator Consisting of the Ridge Regression and OLS Estimators under Linex Loss

Ohtani, Kazuhiro

(Citation)
Kobe University Economic Review,56:1-11

(Issue Date)
2010

(Resource Type)
departmental bulletin paper

(Version)
Version of Record

(JaLCDOI)
https://doi.org/10.24546/81002754

(URL)
https://hdl.handle.net/20.500.14094/81002754

RISK PERFORMANCE OF A WEIGHTED AVERAGE ESTIMATOR CONSISTING OF THE RIDGE REGRESSION AND OLS ESTIMATORS UNDER LINEX LOSS

By **KAZUHIRO OHTANI***

In this paper, we consider a weighted average estimator consisting of the ridge regression estimator proposed by Huang (1999) and the ordinary least squares (OLS) estimator, and examine the risk performance of the weighted average estimator when the asymmetric LINEX loss function is used. It is shown that when the asymmetry of the loss function is moderate, the weighted average estimator never has the largest risk among the ridge regression estimator, the OLS estimator and the weighted average estimator. It is also shown that when the asymmetry of the loss function is severe, the ridge regression estimator has the smallest risk in a wide region of parameter space.

1. Introduction

Hoer and Kennard (1970) proposed the ridge regression estimator to avoid the problem of multicollinearity. However, since the ridge regression estimator has a smaller mean squared error (MSE) than the ordinary least squares (OLS) estimator irrespective of the problem of multicollinearity when the regression coefficient is close to zero, many researchers have studied the small sample properties of the ridge regression estimator and its variants. Some examples are Dwivedi et al. (1980), Ohtani (1986) and Firinguetti (1999).

In regression analysis, there may be a situation where our concern is to estimate a specific regression coefficient as accurately as possible. Huang (1999) showed such situations and proposed the ridge regression estimator to estimate a specific regression coefficient. Huang (1999) further examined the small sample properties of the ridge regression estimator and showed that the ridge regression estimator has a smaller MSE than the ordinary least squares (OLS) estimator when the regression coefficient is close to zero.

When the risk performances of the ridge estimator and its variants are examined, the quadratic loss function has usually been used. Since the quadratic loss function is symmetric, over-estimation and under-estimation have the same magnitude. However, as is shown in Zellner (1986), in dam construction under-estimation of the peak water level is usually much more serious than over-estimation. This indicates that the symmetric loss function may not be appropriate in some situations. From this viewpoint, Variant (1975) proposed the asymmetric LINEX loss function, and Zellner (1986) extensively discussed its properties. In several studies on the ridge regression estimator and its variants, the asymmetric LINEX loss function has been used. Some examples are Ohtani (1995), Wan (1999), Akdeniz and Namba (2003), and Namba and Ohtani (2010).

*The author is grateful to the JSPS (Japanese Society for the Promotion of Science) for partial financial support (Grant-in-Aid for Scientific Research).

In several studies on the risk performance of biased estimators, weighted average estimators consisting of the OLS estimator and any biased estimator have been used. Some examples are Stahlecker and Trenkler (1985), Tracy and Srivastava (1994) and Ohtani (1998). However, the risk performances of weighted average estimators have not been examined under LINEX loss. In this paper, we consider a weighted average estimator consisting of the ridge regression estimator proposed by Huang (1999) and the OLS estimator, and examine the risk performance of the weighted average estimator when the asymmetric LINEX loss function is used.

The organization of this paper is as follows. In section 2 the model and estimators are presented. In section 3 the LINEX loss function is shown and the exact formula of the risk function of the weighted average estimator is derived under the LINEX loss function. In section 4 the risk performance of the weighted average estimator is examined by numerical evaluations, based on the exact formula of the risk function. It is shown that when the asymmetry of the loss function is moderate, the weighted average estimator never has the largest risk among the ridge regression estimator, the OLS estimator and the weighted average estimator. It is also shown that when the asymmetry of the loss function is severe, the ridge regression estimator has the smallest risk in a wide region of parameter space.

2. Model and estimator

Consider a linear regression model,

$$y = x_1\beta_1 + X_2\beta_2 + \varepsilon, \tag{1}$$

where y is an $n \times 1$ vector of observations on a dependent variable, x_1 is an $n \times 1$ vector of observations on an explanatory variable, X_2 is an $n \times (k-1)$ matrix of observations on other explanatory variables, β_1 is a scalar coefficient for x_1, β_2 is a $k-1$ vector of coefficients for X_2, and ε is an $n \times 1$ vector of error terms. We assume that x_1 and X_2 are nonstochastic, the $n \times k$ matrix $[x_1, X_2]$ is of full column rank, and ε is distributed as $N(0, \sigma^2 I_n)$, where I_n is an $n \times n$ identity matrix. Without loss of generality, we can assume that β_1 is a specific regression coefficient which we want to estimate as accurately as possible.

Setting $X = [x_1, X_2]$ and $\beta = [\beta_1, \beta_2']'$, the ordinary least squares (OLS) estimator of β is

$$b = (X'X)^{-1}X'y, \tag{2}$$

and the OLS estimator of β_1 is

$$b_1 = (x_1'M_2x_1)^{-1}x_1'M_2y, \tag{3}$$

where

$$M_2 = I_n - X_2(X_2'X_2)^{-1}X_2'. \tag{4}$$

The distribution of b_1 is the normal distribution with mean β_1 and variance $\sigma^2/x_1'M_2x_1$:

$$b_1 \sim N(\beta_1, \sigma^2/x_1'M_2x_1). \tag{5}$$

Following Huang (1999), the feasible ridge regression estimator of β_1 is given by

$$\begin{aligned}\hat{\beta}_1 &= \left(d_1 + \frac{s^2}{b_1^2}\right)^{-1} x_1'M_2y \\ &= \left(\frac{d_1b_1^2}{d_1b_1^2 + s^2}\right)b_1,\end{aligned} \tag{6}$$

where $d_1 = x_1'M_2x_1$, $s^2 = (y - Xb)'(y - Xb)/v$ and $v = n - k$. We simply call the ridge regression estimator proposed by Huang (1999) the ridge regression estimator hereafter.

Since the ridge regression estimator has a smaller MSE than the OLS estimator when the parameter defined as $\theta_1 = \sqrt{d_1}\beta_1/\sigma$ is close to zero and vice versa, one may consider the following weighted average estimator of $\hat{\beta}_1$ and b_1:

$$\hat{\beta}_1^* = w\hat{\beta}_1 + (1 - w)b_1 \tag{7}$$

where w is a constant such that $0 \le w \le 1$. The weighted average estimator is the ridge regression estimator when $w = 1$, and it is the OLS estimator when $w = 0$. In the next section, we derive the exact formula of the risk function of the weighted average estimator under the LINEX loss function.

3. Risk under the LINEX loss function

The LINEX loss function of $\hat{\beta}_1^*$ is given by

$$\begin{aligned}L(\hat{\beta}_1^*) &= exp[a(\hat{\beta}_1^* - \beta_1)] - a(\hat{\beta}_1^* - \beta_1) - 1 \\ &= \sum_{q=2}^{\infty} \frac{a^q}{q!}(\hat{\beta}_1^* - \beta_1)^q,\end{aligned} \tag{8}$$

where $a \ne 0$ is a parameter. When $a > 0$, the loss of over-estimation is more serious than that of under-estimation and vice versa. Also, if the value of a is close to zero, the LINEX loss function is almost symmetric and close to the quadratic loss function. Then, the risk function of $\hat{\beta}_1^*$ is

$$\begin{aligned}R(\hat{\beta}_1^*) &= E[L(\hat{\beta}_1^*)] \\ &= \sum_{q=2}^{\infty} \frac{a^q}{q!} \sum_{m=0}^{q} {}_qC_m (-\beta_1)^{q-m} E[(\hat{\beta}_1^*)^m].\end{aligned} \tag{9}$$

The m-th moment of $\hat{\beta}_1^*$ is

$$E[(\hat{\beta}_1^*)^m] = E[(w\hat{\beta}_1 + (1-w)b_1)^m] = \sum_{r=0}^{m} {}_mC_r\, w^{m-r}(1-w)^r E[\hat{\beta}_1^{m-r} b_1^r]. \quad (10)$$

As is shown in the Appendix, setting $m = 2p + \delta$ ($\delta = 0$ or $\delta = 1$), $E[\hat{\beta}_1^{m-r} b_1^r]$ is given by

$$E[\hat{\beta}_1^{2p+\delta-r} b_1^r] = \frac{(\sigma/\sqrt{d_1})^{2p+\delta}}{\sqrt{\pi}\,\Gamma(v/2)} \sum_{j=0}^{\infty} \frac{\theta_1^{2j+\delta}}{(2j+\delta)!} exp(-\theta^2/2)\, 2^{j+p+\delta}\, v^{\,2p+\delta-r}$$

$$\times\, \Gamma(j+p+\delta+(v+1)/2) \int_0^1 \frac{z^{j+3p+2\delta-r-1/2}(1-z)^{v/2-1}}{[1+(v-1)z]^{2p+\delta-r}}\, dz, \quad (11)$$

where $v = n - k$ and $\theta_1 = \sqrt{d_1}\ \beta_1/\sigma$. Substituting (11) into (10), and further substituting (10) into (9), we obtain the exact formula of the risk function.

4. Numerical Analysis

Since the risk function of the weighted average estimator is very complicated, theoretical analysis of the risk function is difficult. Thus, we examine the risk performance using numerical evaluations, based on the exact formula of the risk function. The parameter values used in the numerical evaluations are as follows: $k = 3, 5, 8$; $n = 20, 30, 40$; $a = -1.0, -0.5, -0.1, 0.1, 0.5, 1.0$; $w = 0.0, 0.1, 0.3, 0.5, 0.7, 0.9, 1.0$; $\theta_1 =$ various values. The convergence tolerance of the infinite series in (11) is 10^{-12}, and the integral is calculated using Simpson's 3/8 rule with 500 equal subdivisions. Since the result for $k = 3$ and $n = 20$ is typical, we discuss the result for this case.

The risk performance of the weighted average estimator for $a = 0.1$, $k = 3$ and $n = 20$ is shown in Table 1. We see from Table 1 that when $a = 0.1$ (i.e., the asymmetry of the loss function is not severe) and the value of θ_1 is close to zero (i.e., $|\theta_1| \leq 1.0$), the risk decreases monotonically as the value of w increases. [When $\theta_1 = -1.0$, the risks for $w = 0.9$ and 1.0 are the same in Table 1. However, if we show the value of the risk down to six decimal places, the risk for $w = 0.9$ is 0.004012 and that for $w = 1.0$ is 0.003985. Although we show the value of the risk down to four decimal places to save space, the same applies hereafter.] Since the weighted average estimator with $w = 1.0$ is the ridge regression estimator, the best estimator among the weighted average estimators considered here is the ridge regression estimator. However, when $|\theta_1| \geq 2.0$, the risk increases monotonically as the value of w increases. Since the weighted average estimator with $w = 0.0$ is the OLS estimator, the best estimator is the OLS estimator. When the value of θ_1 is -1.5, the risk attains a minimum at $w = 0.3$. This indicates that the best estimator is the weighted average estimator with $w = 0.3$. Also, when the value of θ_1 is 1.5, the risk attains a minimum at $w = 0.5$. This indicates that the best estimator is the weighted average estimator with $w = 0.5$. As a whole, when the asymmetry of the loss

Table 1
Risk of the weighted average estimator under the LINEX loss function for a = 0.1, k = 3 and n = 20

	w						
θ_1	.0	.1	.3	.5	.7	.9	1.0
–10.0	.0050	.0050	.0051	.0051	.0051	.0052	.0052
–9.0	.0050	.0050	.0051	.0051	.0052	.0052	.0053
–8.0	.0050	.0050	.0051	.0051	.0052	.0053	.0053
–7.0	.0050	.0050	.0051	.0052	.0053	.0054	.0054
–6.0	.0050	.0051	.0051	.0052	.0053	.0055	.0055
–5.0	.0050	.0051	.0052	.0053	.0055	.0056	.0057
–4.5	.0050	.0051	.0052	.0054	.0056	.0058	.0059
–4.0	.0050	.0051	.0053	.0054	.0057	.0059	.0061
–3.5	.0050	.0051	.0053	.0055	.0058	.0061	.0063
–3.0	.0050	.0051	.0053	.0056	.0059	.0062	.0064
–2.5	.0050	.0051	.0053	.0055	.0059	.0062	.0064
–2.0	.0050	.0050	.0052	.0053	.0056	.0059	.0061
–1.5	.0050	.0050	.0049	.0049	.0050	.0051	.0052
–1.0	.0050	.0048	.0045	.0043	.0041	.0040	.0040
–.5	.0050	.0047	.0042	.0037	.0033	.0030	.0029
.0	.0050	.0047	.0041	.0035	.0030	.0026	.0024
.5	.0050	.0047	.0042	.0037	.0033	.0030	.0029
1.0	.0050	.0048	.0045	.0043	.0041	.0040	.0039
1.5	.0050	.0049	.0049	.0048	.0049	.0050	.0050
2.0	.0050	.0050	.0051	.0052	.0054	.0056	.0058
2.5	.0050	.0051	.0052	.0054	.0056	.0059	.0061
3.0	.0050	.0051	.0052	.0054	.0056	.0059	.0061
3.5	.0050	.0051	.0052	.0054	.0056	.0058	.0059
4.0	.0050	.0051	.0052	.0053	.0055	.0057	.0058
4.5	.0050	.0050	.0051	.0053	.0054	.0055	.0056
5.0	.0050	.0050	.0051	.0052	.0053	.0054	.0055
6.0	.0050	.0050	.0051	.0051	.0052	.0053	.0053
7.0	.0050	.0050	.0051	.0051	.0052	.0052	.0053
8.0	.0050	.0050	.0050	.0051	.0051	.0052	.0052
9.0	.0050	.0050	.0050	.0051	.0051	.0051	.0051
10.0	.0050	.0050	.0050	.0051	.0051	.0051	.0051

function is not severe, the ridge regression estimator (w = 1) or the OLS estimator (w = 0) has the smallest risk, except for some values of θ_1.

The risk performance of the weighted average estimator for a = 0.5, k = 3 and n = 20 is shown in Table 2. We see from Table 2 that when a = 0.5 (i.e., the asymmetry of the loss function is moderate) and the value of θ_1 is close to zero (i.e., $|\theta_1| \leq 1.0$), the risk decreases monotonically as the value of w increases. This indicates that the best estimator among the weighted average estimators considered here is the ridge regression estimator. However, when $\theta_1 \leq -2.0$, the risk increases monotonically as the value of w increases. This indicates that the best estimator is the OLS estimator. When $\theta_1 = -1.5$, the risk attains a minimum at w = 0.3. Also, when $1.5 \leq \theta_1 \leq 7.0$, the risk attains a minimum at w = 0.1, 0.3, 0.5 or 0.7. Comparing Tables 1 and 2, we see that when the degree of asymmetry of the loss function increases from a = 0.1 to 0.5, the region where the ridge regression estimator or the OLS estimator is best gets narrow. If we use a weight of 0.5 (i.e., w = 0.5), the risk of the weighted average

Table 2
Risk of the weighted average estimator under the LINEX loss function for a = 0.5, k = 3 and n = 20

	w						
θ_1	.0	.1	.3	.5	.7	.9	1.0
−10.0	.1331	.1341	.1362	.1384	.1408	.1433	.1446
−9.0	.1331	.1343	.1367	.1392	.1420	.1449	.1465
−8.0	.1331	.1345	.1373	.1404	.1436	.1472	.1490
−7.0	.1331	.1347	.1382	.1419	.1460	.1504	.1527
−6.0	.1331	.1351	.1394	.1442	.1494	.1550	.1580
−5.0	.1331	.1357	.1414	.1477	.1547	.1623	.1664
−4.5	.1331	.1361	.1427	.1501	.1583	.1674	.1723
−4.0	.1331	.1366	.1442	.1529	.1626	.1735	.1793
−3.5	.1331	.1370	.1457	.1557	.1670	.1798	.1867
−3.0	.1331	.1372	.1465	.1573	.1698	.1841	.1919
−2.5	.1331	.1367	.1452	.1555	.1678	.1820	.1899
−2.0	.1331	.1351	.1404	.1476	.1569	.1684	.1750
−1.5	.1331	.1321	.1314	.1326	.1359	.1413	.1448
−1.0	.1331	.1284	.1202	.1138	.1093	.1068	.1063
−.5	.1331	.1253	.1110	.0984	.0875	.0784	.0745
.0	.1331	.1242	.1078	.0931	.0800	.0687	.0636
.5	.1331	.1254	.1113	.0990	.0884	.0794	.0756
1.0	.1331	.1279	.1187	.1113	.1056	.1015	.1000
1.5	.1331	.1304	.1260	.1233	.1222	.1225	.1232
2.0	.1331	.1320	.1309	.1312	.1328	.1357	.1376
2.5	.1331	.1329	.1332	.1346	.1372	.1409	.1431
3.0	.1331	.1331	.1338	.1353	.1378	.1412	.1432
3.5	.1331	.1331	.1336	.1348	.1368	.1395	.1411
4.0	.1331	.1330	.1332	.1341	.1355	.1374	.1386
4.5	.1331	.1329	.1329	.1334	.1343	.1357	.1366
5.0	.1331	.1329	.1326	.1328	.1334	.1343	.1350
6.0	.1331	.1328	.1323	.1321	.1322	.1326	.1329
7.0	.1331	.1328	.1322	.1318	.1316	.1317	.1318
8.0	.1331	.1328	.1321	.1316	.1313	.1311	.1311
9.0	.1331	.1328	.1321	.1315	.1311	.1309	.1308
10.0	.1331	.1328	.1321	.1315	.1311	.1307	.1306

estimator is never the largest among the ridge regression estimator, the OLS estimator and the weighted average estimator with w = 0.5. This indicates that when the asymmetry of the loss function is moderate, the use of the weighted average estimator with w = 0.5 can avoid a situation where we may have the maximum risk.

The risk performance of the weighted average estimator for a = 1.0, k = 3 and n = 20 is shown in Table 3. We see from Table 3 that when a = 1.0 (i.e., the asymmetry of the loss function is severe) and $\theta_1 \leq -2.0$, the risk increases monotonically as the value of w increases. This indicates that the best estimator is the OLS estimator. Also, when $\theta_1 = -1.5$, the risk attains a minimum at w = 0.3. However, when $\theta_1 \geq -1.0$, the risk decreases monotonically as the value of w increases, except for θ_1 = 2.5. This indicates that when the asymmetry of the loss function is severe, the ridge regression estimator has the smallest risk in a wide region of θ_1.

The risk performance of the weighted average estimator for a = – 1.0, k = 3 and n = 20 is

Table 3
Risk of the weighted average estimator under the LINEX loss function for a = 1.0, k = 3 and n = 20

	w						
θ_1	.0	.1	.3	.5	.7	.9	1.0
−10.0	.6487	.6571	.6747	.6932	.7126	.7331	.7437
−9.0	.6487	.6583	.6784	.6998	.7224	.7462	.7587
−8.0	.6487	.6599	.6835	.7086	.7354	.7640	.7789
−7.0	.6487	.6621	.6904	.7209	.7537	.7889	.8074
−6.0	.6487	.6652	.7004	.7388	.7806	.8259	.8500
−5.0	.6487	.6697	.7154	.7660	.8219	.8836	.9167
−4.5	.6487	.6727	.7252	.7840	.8497	.9228	.9623
−4.0	.6487	.6759	.7359	.8039	.8807	.9669	1.0138
−3.5	.6487	.6784	.7447	.8208	.9078	1.0065	1.0606
−3.0	.6487	.6785	.7458	.8246	.9161	1.0213	1.0795
−2.5	.6487	.6734	.7309	.8008	.8842	.9823	1.0373
−2.0	.6487	.6610	.6936	.7383	.7961	.8681	.9098
−1.5	.6487	.6421	.6365	.6421	.6594	.6893	.7092
−1.0	.6487	.6218	.5755	.5393	.5135	.4984	.4949
−.5	.6487	.6074	.5324	.4671	.4114	.3649	.3451
.0	.6487	.6034	.5208	.4480	.3843	.3292	.3047
.5	.6487	.6091	.5375	.4755	.4224	.3774	.3578
1.0	.6487	.6191	.5670	.5237	.4886	.4609	.4497
1.5	.6487	.6285	.5941	.5675	.5479	.5348	.5305
2.0	.6487	.6348	.6120	.5956	.5851	.5800	.5794
2.5	.6487	.6382	.6212	.6094	.6024	.5998	.6001
3.0	.6487	.6398	.6251	.6146	.6079	.6049	.6047
3.5	.6487	.6405	.6266	.6161	.6087	.6043	.6032
4.0	.6487	.6409	.6273	.6165	.6083	.6025	.6006
4.5	.6487	.6412	.6278	.6168	.6080	.6012	.5986
5.0	.6487	.6415	.6284	.6173	.6081	.6006	.5976
6.0	.6487	.6420	.6297	.6188	.6094	.6012	.5976
7.0	.6487	.6425	.6310	.6206	.6113	.6030	.5993
8.0	.6487	.6430	.6323	.6224	.6134	.6053	.6015
9.0	.6487	.6434	.6334	.6241	.6155	.6076	.6039
10.0	.6487	.6438	.6344	.6257	.6175	.6098	.6062

shown in Table 4. We see from Tables 3 and 4 that the results for $a = -1.0$ can be obtained from the results for $a = 1.0$ in Table 3 by exchanging the sign of θ_1. For example, the risk for $\theta_1 = 1.0$ in Table 4 is the same as the risk for $\theta_1 = -1.0$ in Table 3. Thus, reversing the sign of θ_1, we can discuss the results for $a < 0$ in a parallel way to the results for $a > 0$.

Appendix

In this appendix, we derive the formula (11). Setting $u_1 = \sqrt{d_1}\, b_1/\sigma$ and $u_2 = (y - Xb)'(y - Xb)/\sigma^2$, u_1 is distributed as $N(\sqrt{d_1}\,\beta_1/\sigma, 1)$ and u_2 as the chi-square distribution with $v = n - k$ degrees of freedom. Using u_1 and u_2, b_1 and $\hat{\beta}_1$ can be written as

Table 4
Risk of the weighted average estimator under the LINEX loss function for $a = -1.0$, $k = 3$ and $n = 20$

	w						
θ_1	.0	.1	.3	.5	.7	.9	1.0
–10.0000	.6487	.6438	.6344	.6257	.6175	.6098	.6062
–9.0000	.6487	.6434	.6334	.6241	.6155	.6076	.6039
–8.0000	.6487	.6430	.6323	.6224	.6134	.6053	.6015
–7.0000	.6487	.6425	.6310	.6206	.6113	.6030	.5993
–6.0000	.6487	.6420	.6297	.6188	.6094	.6012	.5976
–5.0000	.6487	.6415	.6284	.6173	.6081	.6006	.5976
–4.5000	.6487	.6412	.6278	.6168	.6080	.6012	.5986
–4.0000	.6487	.6409	.6273	.6165	.6083	.6025	.6006
–3.5000	.6487	.6405	.6266	.6161	.6087	.6043	.6032
–3.0000	.6487	.6398	.6251	.6146	.6079	.6049	.6047
–2.5000	.6487	.6382	.6212	.6094	.6024	.5998	.6001
–2.0000	.6487	.6348	.6120	.5956	.5851	.5800	.5794
–1.5000	.6487	.6285	.5941	.5675	.5479	.5348	.5305
–1.0000	.6487	.6191	.5670	.5237	.4886	.4609	.4497
–.5000	.6487	.6091	.5375	.4755	.4224	.3774	.3578
.0000	.6487	.6034	.5208	.4480	.3843	.3292	.3047
.5000	.6487	.6074	.5324	.4671	.4114	.3649	.3451
1.0000	.6487	.6218	.5755	.5393	.5135	.4984	.4949
1.5000	.6487	.6421	.6365	.6421	.6594	.6893	.7092
2.0000	.6487	.6610	.6936	.7383	.7961	.8681	.9098
2.5000	.6487	.6734	.7309	.8008	.8842	.9823	1.0373
3.0000	.6487	.6785	.7458	.8246	.9161	1.0213	1.0795
3.5000	.6487	.6784	.7447	.8208	.9078	1.0065	1.0606
4.0000	.6487	.6759	.7359	.8039	.8807	.9669	1.0138
4.5000	.6487	.6727	.7252	.7840	.8497	.9228	.9623
5.0000	.6487	.6697	.7154	.7660	.8219	.8836	.9167
6.0000	.6487	.6652	.7004	.7388	.7806	.8259	.8500
7.0000	.6487	.6621	.6904	.7209	.7537	.7889	.8074
8.0000	.6487	.6599	.6835	.7086	.7354	.7640	.7789
9.0000	.6487	.6583	.6784	.6998	.7224	.7462	.7587
10.0000	.6487	.6571	.6747	.6932	.7126	.7331	.7437

$$b_1 = \frac{\sigma}{\sqrt{d_1}} u_1, \tag{12}$$

$$\hat{\beta}_1 = \frac{(\sigma/\sqrt{d_1})u_1^3}{u_1^2 + u_2/v} \tag{13}$$

Thus, we have

$$
\begin{aligned}
& E[\hat{\beta}_1^{2p+\delta-r} b_1^r] \\
&= E\Big[\Big(\frac{(\sigma/\sqrt{d_1})u_1^3}{u_1^2+u_2/v}\Big)^{m-r}\Big(\frac{\sigma}{\sqrt{d_1}} u_1\Big)^r\Big] \\
&= \Big(\frac{\sigma}{\sqrt{d_1}}\Big)^m \int_0^\infty \int_{-\infty}^\infty \frac{u_1^{3(m-r)+r}}{(u_1^2+u_2/v)^{m-r}} \frac{1}{\sqrt{2\pi}} exp[-(u_1-\theta_1)^2/2]
\end{aligned}
$$

$$\times \frac{1}{2^{\nu/2}\Gamma(\nu/2)} u_2^{\nu/2-1} exp(-u_2/2) du_1\, du_2$$

$$= K \int_0^\infty \int_{-\infty}^\infty \frac{u_1^{3m-2r} u_2^{\nu/2-1}}{(u_1^2+u_2/\nu)^{m-r}} exp[-(u_1-\theta_1)^2/2] exp(-u_2/2) du_1\, du_2 \quad (14)$$

where

$$K = \Big(\frac{\sigma}{\sqrt{d_1}}\Big)^m \frac{1}{2^{\nu/2+1}\sqrt{\pi}\,\Gamma(\nu/2)}.$$

Using Maclaurin's expansion, $\exp(\theta_1 u_1) = \Sigma_{i=0}^{\infty}(\theta_1 u_1)^i/i!$, (14) reduces to

$$\sum_{i=0}^{\infty} K \frac{\theta_1^i}{i!} exp(-\theta_1^2/2) \int_0^\infty \int_{-\infty}^\infty \frac{u_1^{i+3m-2r} u_2^{\nu/2-1}}{(u_1^2+u_2/\nu)^{m-r}} exp[-(u_1^2+u_2)/2] du_1 du_2. \quad (15)$$

The integral part in (15) is written as follows:

$$\int_0^\infty \int_0^\infty \frac{u_1^{i+3m-2r} u_2^{\nu/2-1}}{(u_1^2+u_2/\nu)^{m-r}} exp[-(u_1^2+u_2)/2] du_1 du_2$$
$$+\int_0^\infty \int_{-\infty}^0 \frac{u_1^{i+3m-2r} u_2^{\nu/2-1}}{(u_1^2+u_2/\nu)^{m-r}} exp[-(u_1^2+u_2)/2] du_1 du_2. \quad (16)$$

Since $u_1^2/(u_1^2+u_2/\nu) \leq 1$ and $u_1 \geq 0$ in the first term in (16), we have

$$\int_0^\infty \int_0^\infty \frac{u_1^{i+3m-2r} u_2^{\nu/2-1}}{(u_1^2+u_2/\nu)^{m-r}} exp[-(u_1^2+u_2)/2] du_1 du_2$$
$$= \int_0^\infty \int_0^\infty \frac{(u_1^2)^{m-r}(u_1^2)^{r-m} u_1^{i+3m-2r} u_2^{\nu/2-1}}{(u_1^2+u_2/\nu)^{m-r}} exp[-(u_1^2+u_2)/2] du_1 du_2$$
$$\leq \int_0^\infty \int_0^\infty u_1^{i+m} exp[-(u_1^2+u_2)/2] du_1 du_2$$
$$= \int_0^\infty u_1^{i+m} exp(-u_1^2)\, du_1 \int_0^\infty u_2^{\nu-1} exp(-u_2/2) du_2. \quad (17)$$

Since the integrals in (17) are the kernels of the moments of the normal and chi-square distributions, the first term in (16) exists. In a similar way, we can show that the second term in (16) exists.

When both m and i are even numbers, and when both m and i are odd numbers, the integrand in (15) is an even function. Since the integral exists and the integrand is an even function, (15) can be written as

$$2 \sum_{i=0}^{\infty} K \frac{\theta_1^i}{i!} exp(-\theta_1^2/2) \int_0^\infty \int_0^\infty \frac{u_1^{i+3m-2r} u_2^{\nu/2-1}}{(u_1^2+u_2/\nu)^{m-r}} exp[-(u_1^2+u_2)/2] du_1 du_2. \quad (18)$$

When m is an even number and i is an odd number, and when m is an odd number and i is an even number, the integrand in (15) is an odd function. Since the integral exists and the integrand is an odd function, the integral in (15) reduces to zero. Thus, the integral in (15) can be evaluated by calculating (18).

Putting $m = 2p + \delta$ and $i = 2j + \delta$ ($\delta = 0$ or 1), (18) is written as

$$2 \sum_{j=0}^{\infty} K \frac{\theta_1^{2j+\delta}}{(2j+\delta)!} exp(-\theta_1^2/2) \times \int_0^{\infty}\int_0^{\infty} \frac{u_1^{2j+6p+4\delta-2r} u_2^{v/2-1}}{(u_1^2+u_2/v)^{2p+\delta-r}} exp[-(u_1^2+u_2)/2]\, du_1 du_2. \tag{19}$$

Using the change of variable, $\tau_1 = u_1^2$, (19) reduces to

$$\sum_{j=0}^{\infty} K_j^* \int_0^{\infty}\int_0^{\infty} \frac{\tau_1^{j+3p+2\delta-r-1/2} u_2^{v/2-1}}{(\tau_1+u_2/v)^{2p+\delta-r}} \exp[-(\tau_1+u_2)/2]\, d\tau_1 du_2, \tag{20}$$

where

$$K_j^* = K \frac{\theta_1^{2j+\delta}}{(2j+\delta)!} exp(-\theta_1^2/2). \tag{21}$$

Further using the change of variables, $t_1 = \tau_1 + u_2$ and $t_2 = \tau_1/u_2$, and manipulating some calculations, (20) reduces to

$$\sum_{j=0}^{\infty} K_j^* \int_0^{\infty}\int_0^{\infty} \frac{t_1^{j+p+\delta+(v+1)/2-1} t_2^{j+3p+2\delta-r-1/2}}{(1+t_2)^{j+p+\delta+(v+1)/2}(t_2+1/v)^{2p+\delta-r}} \exp(-t_1/2)\, dt_2 dt_1, \tag{22}$$

Using the gamma function, (22) is expressed as

$$\sum_{j=0}^{\infty} K_j^* 2^{j+p+\delta+(v+1)/2}\Gamma(j+p+\delta+(v+1)/2) \times \int_0^{\infty} \frac{t_2^{j+3p+2\delta-r-1/2}}{(1+t_2)^{j+p+\delta+(v+1)/2}(t_2+1/v)^{2p+\delta-r}} dt_2, \tag{23}$$

Finally using the change of variable, $z = t_2/(1+t_2)$, (23) reduces to (11) in the text.

REFERENCES

Akdeniz, F. and Namba, A. (2003), "A note on new feasible generalized ridge regression estimator under the LINEX loss function," *Journal of Statistical Computation and Simulation*, 73, pp.303-310.

Dwivedi, T.D., Srivastava, V.K. and Hall, H.L. (1980), "Finite sample properties of ridge estimators," *Technometrics*, 22, pp.205-212.

Firinguetti, L. (1999), "A generalized ridge regression estimator and its finite sample properties," *Communications in Statistics-Theory and Methods*, 28, pp.1217-1229.

Hoerl, A.E. and Kennard, R.W. (1970), "Ridge regression: Biased estimation for nonorthogonal problems," *Technometrics*, 12, pp.55-82.

Huang, J.-C. (1999), "Improving the estimation precision for a selected parameter in multiple regression analysis: An algebraic approach," *Economics Letters*, pp.261-264.

Namba, A. and Ohtani, K. (2010), "Risk performance of a pre-test ridge regression estimator under the LINEX loss function when each individual regression coefficient is estimated," *Journal of Statistical Computation and Simulation*, 80, pp.255-262.

Ohtani, K. (1986), "On small sample properties of the almost unbiased generalized ridge estimator," *Communications in Statistics-Theory and Methods*, pp.1571-1578.

Ohtani, K. (1995), "Generalized ridge regression estimators under the LINEX loss function," *Statistical Papers*, 36, pp.99-110.

Ohtani, K. (1998), "The exact risk of a weighted average estimator of the OLS and Stein-rule estimators in regression under balanced loss," *Statistics and Decisions*, 16, pp.35-45.

Stahlecker, P. and Trenkler, G. (1985), "On heterogeneous version of the best linear and the ridge estimator," *Proceedings of First Tampere Seminar on Linear Models*, pp.301-322.

Tracy, D.S. and Srivastava, A.K. (1994), "Comparison of operational variants of best homogeneous and heterogeneous estimators in linear regression," *Communications in Statistics-Theory and Methods*, 23, pp.2313-2322.

Varian, H.R. (1975), "A Bayesian approach to real estate assessment," *Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage*, in S.E. Fienberg and A. Zellner (eds.), North-Holland, Amsterdam, pp.195-208.

Wan, A.T.K. (1999), "A note on almost unbiased generalized ridge regression estimators under asymmetric loss," *Journal of Statistical Computation and Simulation*, 62, pp.411-412.

Zellner, A. (1986), "Bayesian estimation and prediction using asymmetric loss functions," *Journal of the American Statistical Association*, 81, pp.446-451.