The Exact MSE-Efficiency of the

General Ridge Estimator Relative to OLS*

by

Rudolf Teekens and Paul de Boer

ABSTRACT

In this paper it has been pointed out that the merits of the Ridge procedure as proposed by Hoerl and Kennard tend to be overvalued due to an incorrect analysis of the associated Mean Square Error. For the case of the so-called General Ridge Estimator it is then shown how the exact MSE can be derived and finally it is seen that General Ridge Estimator dominates the OLS estimator only in a limited interval of the parameter space.

Contents

* The authors are indebted to Mr A.S. Louter for performing the calculations for appendix C.

I. Introduction

Since Hoerl and Kennard first published their so-called Ridge Regression method in 1970, a considerable amount of research has been devoted to this subject. Some Econometricians as well seem to be taken with the Ridge Method, witness the publications of Vinod (1976 a, 1976 b) and Moulaert (1976).

In our opinion, however, the Ridge Regression Method

a) is based on a dubious method, which consists of optimizing an unknown loss-function, and

b) dominates the OLS estimator in MSE only in a limited range of parameter values.

The subsequent analysis is more or less analogous to the one carried out by Feldstein (1973) who studied the mean square error efficiency of COV (conditional omitted variable) and WTD (weighted average) estimators relative to the OLS estimator.

As Hoerl and Kennard (1970 a) we consider the standard linear model

(1.1) $\quad y = X\beta + \varepsilon$

where y is an observable random vector of n elements

X is an observable fixed matrix of order $n \times p$ with rank p

β is a vector of p unknown parameters

ε is a non-observable random vector of n elements which has a multivariate normal distribution: $\varepsilon \sim n(0,\sigma^2 I)$.

We also follow Hoerl and Kennard in reducing the above model to a canonical form in which the X'X matrix is diagonal. This may be achieved by applying the following orthogonal transformation to X and β. Let

(1.2) $\quad \alpha = P'\beta$

and

(1.3) $\quad X^* = XP$

where the columns of P are the eigenvectors of X'X and

(1.4) $$P' = P^{-1}$$

so that $X^{*\prime}X^* = P'X'XP = \Lambda$

with Λ the diagonal matrix of (positive) eigenvalues of X'X.

Then model (1.1) may be rewritten as

(1.5) $$y = X^*\alpha + \varepsilon$$

where it should be noted that

$$\alpha'\alpha = \beta'PP'\beta = \beta'\beta$$

The OLS estimator of α is

$$\hat{\alpha} = (X^{*\prime}X^*)^{-1}X^{*\prime}y = \Lambda^{-1}X^{*\prime}y$$

If we define the vector c as

(1.6) $$c = X^{*\prime}y = P'X'y$$

$\hat{\alpha}$ may be written as

(1.7) $$\hat{\alpha} = \Lambda^{-1}c$$

Moreover, it is easily verified that

(1.8) $$\hat{\beta} = P\hat{\alpha}$$

The general ridge procedure is defined from

(1.9) $$\hat{\alpha}^* = [X^{*\prime}X^* + K]^{-1}X^{*\prime}y = [\Lambda + K]^{-1}c$$

with K a diagonal matrix with non-negative elements.

The corresponding Ridge estimator in model (1.1) is then defined as

$$\hat{\beta}^* = P\hat{\alpha}^* \tag{1.10}$$

It should be noted here that $\hat{\beta}^*$ equals

$$\hat{\beta}^* = [X'X + K]^{-1}X'y$$

<u>only if</u> $K = kI$. If K is not a scalar matrix it follows from (1.10) that

$$\hat{\beta}^* = [X'X + PKP']^{-1}X'y$$

In order to determine the matrix K, we minimize the mean square error (MSE) of $\hat{\alpha}^*$ relative to α; this function will be denoted as $\pi(\hat{\alpha}^*;\alpha)$

$$\pi(\hat{\alpha}^*;\alpha) = E[(\hat{\alpha}^* - \alpha)'(\hat{\alpha}^* - \alpha)] = \sum_{i=1}^{p} E(\hat{\alpha}^*_i - \alpha_i)^2 \tag{1.11}$$

Writing (1.9) in scalar terms, we obtain

$$\hat{\alpha}^*_i = \frac{c_i}{\lambda_i + k_i} \qquad i = 1, \ldots, p$$

Realizing furthermore that $[c_i] = c = X^{*\prime}y = \Lambda\alpha + X^{*\prime}u$ has a multivariate normal distribution:

$$c \sim n(\Lambda\alpha, \sigma^2\Lambda)$$

or

$$c_i \sim n(\lambda_i\alpha_i, \lambda_i\sigma^2)$$

we can write (1.11) as

$$(1.12) \qquad \pi(\hat{\alpha}^*;\alpha) = \sum_{i=1}^{p} E\left[\frac{c_i}{\lambda_i + k_i} - \alpha_i\right]^2$$

$$= \sum_{i=1}^{p} \frac{\sigma^2\lambda_i}{(\lambda_i + k_i)^2} E\left[\frac{c_i - \lambda_i\alpha_i}{\sigma\sqrt{\lambda_i}} - \frac{\alpha_i k_i}{\sigma\sqrt{\lambda_i}}\right]^2$$

$$= \sum_{i=1}^{p} \frac{\sigma^2\lambda_i}{(\lambda_i + k_i)^2}\left[1 + \frac{\alpha_i^2 k_i^2}{\sigma^2\lambda_i}\right]$$

$$= \sum_{i=1}^{p} \frac{\sigma^2\lambda_i + \alpha_i^2 k_i^2}{(\lambda_i + k_i)^2}$$

Minimizing $\pi(\hat{\alpha}^*;\alpha)$ with respect to the k_i's yields the following optimal values of k_i:

$$(1.13) \qquad k_i = \frac{\sigma^2}{\alpha_i^2} \qquad i = 1, \ldots, p$$

Inspection of the second order conditions shows that (1.13) indeed constitutes minimum for (1.12). Obviously this solution for k_i, $i = 1, \ldots, p$ is useless for estimation purposes, since k_i depends on the unknown parameters α_i, $i = 1, \ldots, p$ and σ^2. Therefore Hoerl and Kennard propose the following method for approximating the theoretical optimal values of k_i:

(i) determine the OLS estimates $\hat{\alpha}_i$ and $\hat{\sigma}^2$

(ii) determine $\hat{k}_{i(0)}$ from $\hat{k}_{i(0)} = \hat{\sigma}^2/\hat{\alpha}_i^2$

(iii) continue the process as follows

$$\hat{\alpha}^*_{i(l)} = \frac{c_i}{\lambda_i + \hat{k}_{i(l-1)}}, \quad \hat{k}_{i(l)} = \frac{\hat{\sigma}^2}{\hat{\alpha}^2_{i(l)}}, \quad l = 1, 2, 3, \ldots$$

until stability is achieved in $\hat{\alpha}^{*\prime}_{(l)}\hat{\alpha}^*_{(l)}$.

In the remainder of this paper we will limit ourselves to this "general form of the ridge regression", i.e. K a non-scalar diagonal matrix. But we will conclude this introduction with a number of remarks about the case of a scalar K-matrix, to which Hoerl and Kennard devote the main part of their paper. First, it is observed that in case of a unique k the MSE-function of $\hat{\alpha}^*$ becomes

$$(1.14) \qquad \pi'(\hat{\alpha}^*;\alpha) = \sum_{i=1}^{p} \frac{\sigma^2\lambda_i + \alpha_i^2 k}{(\lambda_i + k)^2}$$

so that in this case the first order condition for a minimum of $\pi'(\hat{\alpha}^*;\alpha)$ does not yield an explicit value of k. This first order condition reads as:

$$(1.15) \qquad \sum_{i=1}^{p} \lambda_i \frac{\alpha_i^2 k - \sigma^2}{(\lambda_i + k)^3} = 0$$

Unlike the case of different k_i, $i = 1, \ldots, p$, Hoerl and Kennard do not consider the possibility of applying an iterative method for approximating the theoretical optimal value of k. Such a method could consist of substituting for the unknown parameter α_i the value of its ridge estimate into (1.15),

$$\hat{\alpha}_i^* = \frac{c_i}{\lambda_i + k}$$

and then solving by a numerical method the resulting education

$$(1.16) \qquad \sum_{i=1}^{p} \lambda_i \frac{k^2 + (2\lambda_i - c_i^2/\sigma^2)k + \lambda_i^2}{(k + \lambda_i)^5} = 0$$

The alternative method as presented by Hoerl and Kennard, viz. the use of the so-called ridge trace, has already been critizised by Conniffe and Stone (1973) and Farebrother (1975). Their stability criterion leads them to choose a too high value of k.

As indicated by Conniffe and Stone (1973), and Conniffe, Stone and O'Neill (1976) and Newhouse and Oman (1971), the proof given by Hoerl and Kennard, that for some <u>fixed</u> k $\hat{\alpha}^*$ has lower MSE than the OLS estimator of α, is inapplicable for the practical case where k is a <u>random variable</u>. However, for a unique k it seems impossible to determine analitically the true MSE of $\hat{\alpha}^*$, since k cannot be determined explicitely from (1.16). Fortunately, the case of general ridge regression as presented earlier in the introduction lends itself much better to an analysis of the MSE of $\hat{\alpha}^*$ and this will be the subject of the rest of this paper.

II. The explicit general ridge estimator

The iterative method as proposed by Hoerl and Kennard to solve for $\hat{k}_i$ and $\hat{\alpha}_i^*$ which has been described in the previous section makes things unnecessarily complicated and unclear. This method provides with a solution of the following two equations

$$\left.\begin{aligned} &(2.1) \qquad \alpha_i^* = \frac{c_i}{\lambda_i + k_i} \\ &(2.2) \qquad k_i = \hat{\sigma}^2/\alpha_i^{*2} \end{aligned}\right\} \qquad i = 1, \ldots, p$$

But these equations may easily be solved analitically; substitution of (2.2) into (2.1) yields the following quadratic equation in α_i^*:

$$(2.3) \qquad \lambda_i\alpha_i^{*2} - c_i\alpha_i^* + \hat{\sigma}^2 = 0$$

and the roots of α_i^* are [1)]

$$(2.4) \qquad \hat{\alpha}_i^* = \frac{1}{2\lambda_i}\left[c_i \pm \sqrt{c_i^2 - 4\lambda_i\hat{\sigma}^2}\right]$$

<u>provided that</u> $c_i^2 \geq 4\lambda_i\hat{\sigma}^2$. The corresponding roots of k_i follow immediately from substitution of (2.4) into (2.2).

----

1) Hemmerle (1975) found independently the same solution.

Given this analytical solution a number of remarks can be made

(i) The general ridge estimator does not always exist; it is defined only if $c_i^2 \geq 4\lambda_i\hat{\sigma}^2$

(ii) The general ridge estimator is not unique, since for each i there exist two pairs $(\hat{\alpha}_i^*, \hat{k}_i)$ which satisfy (2.1) and (2.2) simultaneously, provided that $c_i^2 > 4\lambda_i\hat{\sigma}^2$.

(iii) The iteration procedure as proposed by Hoerl and Kennard selects by definition the set of roots $(\hat{\alpha}_i^*, \hat{k}_i)$ with the smallest k_i (or the highest $|\alpha^*|$). This may easily be seen from figure 1 in which the two functions $\alpha_i^* = \dfrac{c_i}{\lambda_i + k_i}$ and $\alpha_i^* = \sqrt{\hat{\sigma}^2}/\sqrt{k_i}$ are plotted for the case where $c_i > 0$.

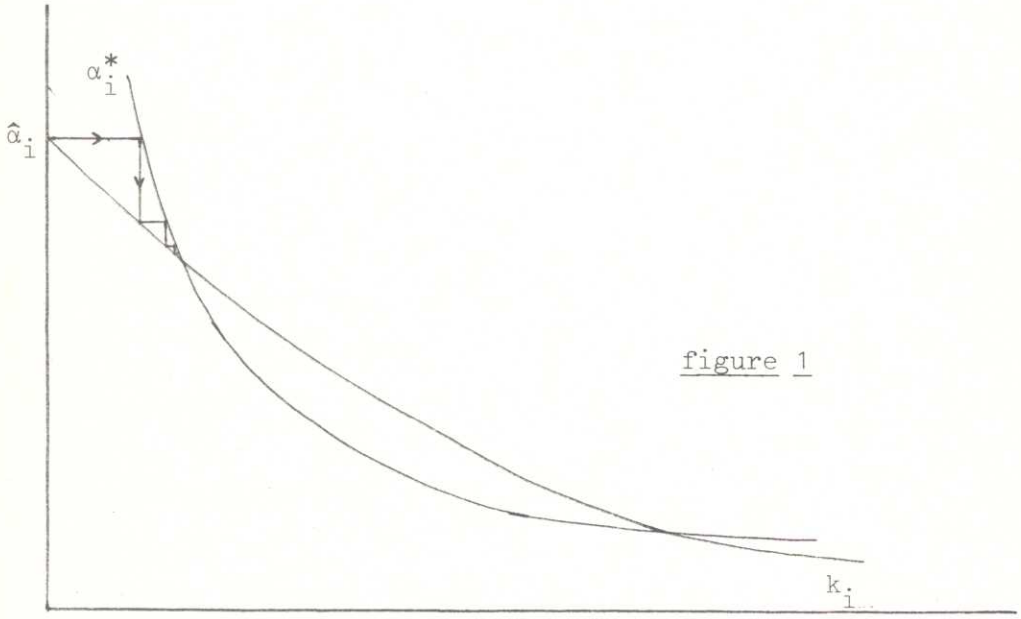


figure 1

From the graph it is clear that, provided the two curves intersect (i.e. provided that $c_i^2 > 4\lambda_i\hat{\sigma}^2$), starting the iteration with $k_i = 0$ (i.e. chosing for $\alpha_{i(0)}^* = \hat{\alpha}_i$) one always moves to the point of intersection with the smallest k_i.

In the sequel we shall concentrate on the general ridge estimator as proposed by Hoerl and Kennard, i.e. the $\hat{\alpha}^*$ corresponding to the smallest k_i. From the previous paragraphs it is clear that this estimator is not defined if $c_i^2 < 4\lambda_i\hat{\sigma}^2$. Since it is our intention to compare the MSE-performance of $\hat{\alpha}_i^*$ with that of the OLS estimator of α_i, it seems fair to complete the definition of $\hat{\alpha}_i^*$ by defining it for $c_i^2 < 4\lambda_i\hat{\sigma}^2$ to be identical to the OLS estimator of α_i.

The general ridge estimator is then defined for the entire sample space. If, moreover, we realize that the OLS estimator $\hat{\alpha}_i^*$ is defined as c_i/λ_i, we may write $\hat{\alpha}_i^*$ as

$$(2.5)\quad \begin{cases} \hat{\alpha}_i^* = \hat{\alpha}_i & \text{for } |\hat{\alpha}_i| < 2\hat{\sigma}_i \\ \hat{\alpha}_i^* = \tfrac{1}{2}\hat{\alpha}_i \left(1 + \sqrt{1 - 4\hat{\sigma}_i/\hat{\alpha}_i^2}\right) & \text{for } |\hat{\alpha}_i| > 2\hat{\sigma}_i \end{cases}$$

where $\hat{\sigma}_i^2 = \hat{\sigma}^2/\lambda_i$ is the estimated variance of $\hat{\alpha}_i$.

III. The exact mean square error efficiency

We define the MSE-efficiency of $\hat{\alpha}_i^*$ with respect to $\hat{\alpha}$ (the OLS estimator of α) as the ratio of the Mean Square Errors of the two estimators

$$(3.1)\quad \varepsilon_\alpha(\hat{\alpha}_i^*;\hat{\alpha}_i) = E[\hat{\alpha}_i^* - \alpha_i]^2 \;/\; E[\hat{\alpha}_i - \alpha_i]^2$$

First we consider the simple case, where the variance of the system, σ^2, is known. In that case the general ridge estimator $\hat{\alpha}_i^*$, as defined in (2.5) becomes

$$(3.2)\quad \begin{cases} \hat{\alpha}_i^* = \hat{\alpha}_i & \text{for } |\hat{\alpha}_i| < 2\sigma_i \\ \hat{\alpha}_i^* = \tfrac{1}{2}\,\hat{\alpha}_i \left(1 + \sqrt{1 - 4\sigma_i^2/\hat{\alpha}_i^2}\right) & \text{for } |\hat{\alpha}_i| > 2\sigma_i \end{cases}$$

Since $\hat{\alpha}_i$ is Normally distributed with mean α_i and variance $\sigma_i^2 = \sigma^2/\lambda_i$

we may apply the result of appendix A and conclude that

$$\varepsilon_\alpha(\hat{\alpha}_i^*;\hat{\alpha}_i) = \varphi_1\left(\frac{\alpha_i}{\sigma_i}\right) = \varphi_1\left(\frac{\alpha_i\sqrt{\lambda_i}}{\sigma}\right)$$

where $\varphi_1(.)$ is defined in appendix A and tabulated in appendix C. In figure 2 we have given $\varphi_1(.)$ for positive arguments only, since $\varphi_1(.)$ is symmetric about the origin.

From figure 2 it can be seen that the general ridge estimator of α_i dominates[1) the OLS estimator for

$$\left|\frac{\alpha_i\sqrt{\lambda_i}}{\sigma}\right| < 2.59668 \qquad \text{or } |\alpha_i| < 2.59668\,\sigma/\sqrt{\lambda_i}$$

provided that the variance of the system is known.

Next, we consider the realistic case where the variance σ^2 is unknown. In that case the general ridge estimator as defined in (2.5) applies. Since $\hat{\alpha}_i$ is normally distributed and $\frac{m\hat{\sigma}^2}{\sigma_i^2}$ has a χ^2-distribution with m degrees of freedom, we may apply the result of appendix B and conclude that the MSE-efficiency of $\hat{\alpha}_i^*$ with respect to $\hat{\alpha}_i$ equals

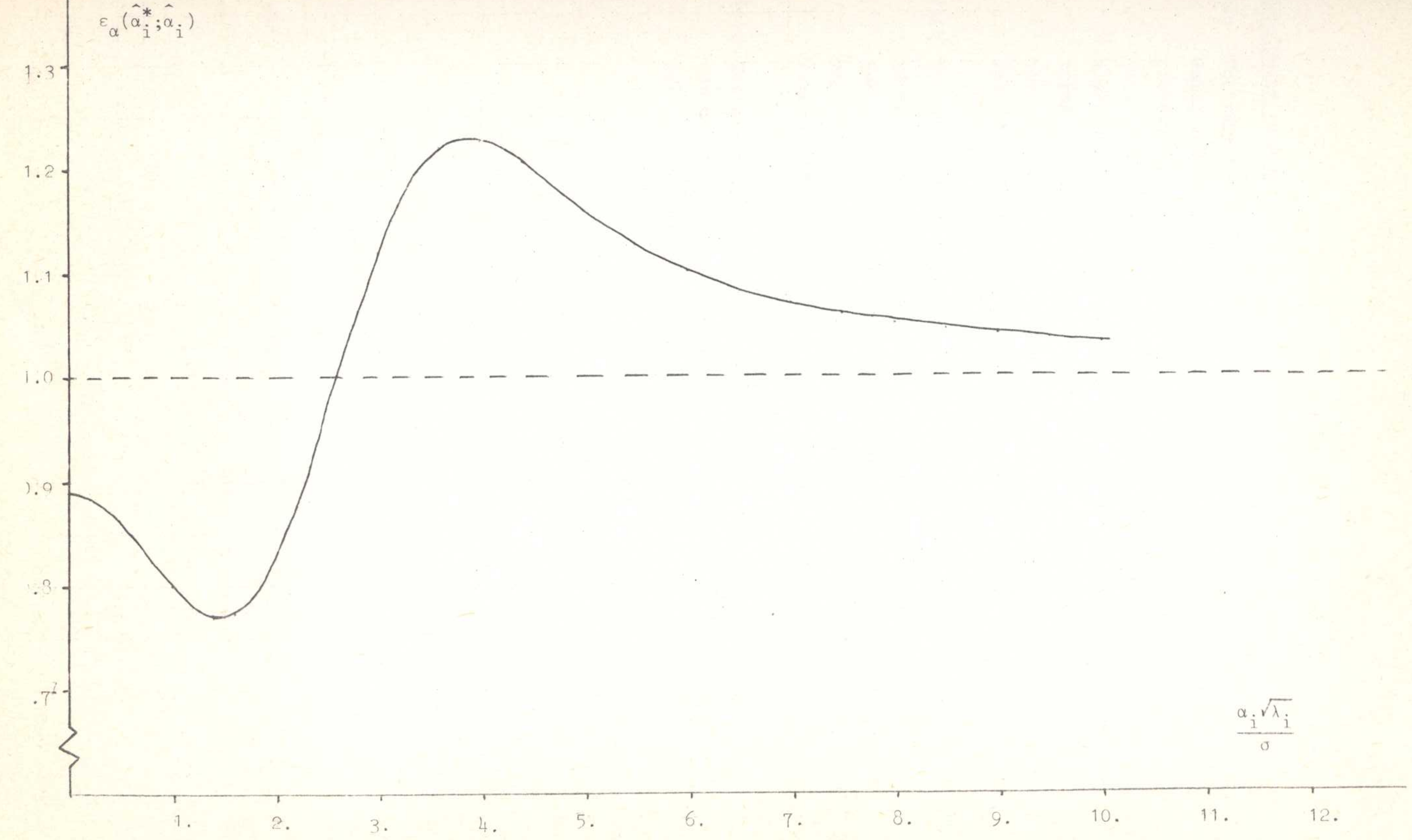
$$\varepsilon'_\alpha(\hat{\alpha}_i^*;\hat{\alpha}_i) = \varphi_2\left(\frac{\alpha_i}{\sigma_i}, m\right) = \varphi_2\left(\frac{\alpha_i\sqrt{\lambda_i}}{\sigma}, m\right)$$

where $\varphi_2(.,.)$ is defined in appendix B and tabulated in appendix C. From this table it can be seen that $\varphi_2(.,.)$ is very close to $\varphi_1(.)$ for different values of m. Therefore we may consider $\varphi_1(.)$ as a good approximation to the mean square error efficiency of the general ridge estimator for unknown variance.

----

1) in the sense of having a lower mean square error

figure 2


$\varepsilon_\alpha(\hat{\alpha}_i^*;\hat{\alpha}_i)$
1.3
1.2
1.1
1.0
0.9
.8
.7
1.
2.
3.
4.
5.
6.
7.
8.
9.
10.
11.
12.
$\frac{\alpha_i\sqrt{\lambda_i}}{\sigma}$

IV. Concluding remarks

(i) Contrary to the Stein-rule estimation procedures (see Baranchik (1973), James and Stein (1961) and Stein (1960)), the Hoerl and Kennard procedure is based on the minimization of an objective function with unknown parameters. Therefore, any resulting optimal value of k depends on the unknown parameters and should be redefined in terms of estimated parameters in order to obtain an estimator. But then the original MSE properties of $\hat{\alpha}^*$ are no longer valid and the redefined $\hat{\alpha}^*$ should be reconsidered with respect to its MSE.

(ii) The exact MSE has been calculated for the case of the so-called general ridge estimator, with a non-scalar diagonal matrix K. Here the function $\varphi_1(.)$ as tabulated in appendix C and pictured in figure 2 turns out to be a good approximation to the mean square error of $\hat{\alpha}^*$. This MSE is a function of $(\alpha_i\sqrt{\lambda_i})/\sigma$ only.

(iii) The general ridge estimator $\hat{\alpha}_i^*$ dominates the ordinary least squares estimator $\hat{\alpha}_i$ only if $\alpha_i\sqrt{\lambda_i}/\sigma < 2.59668$, a condition which can be made the subject of a pre-test.

References

Baranchik, A.J. (1973). Inadmissibility of Maximum Likelihood Estimators in Some Multiple Regression Problems with Three or More Independent Variables. The Annals of Statistics, Vol. 1, No. 2, p. 312-321.

Conniffe, D. and J. Stone (1973). A Critical View of Ridge Regression, The Statistician, 22, p. 181-187.

Conniffe, D. and J. Stone (1975). A Reply to Smith and Goldstein, The Statistician, 24, No. 1, p. 67,68.

Conniffe, D., Stone, J. and F. O'Neill (1976). Is Ridge Regression a Useful Technique in Economic Analysis . Unpublished Research Memorandum.

Farebrother, R.W. (1975). The Minimum Mean Square Error Linear Estimator and Ridge Regression, Technometrics, Vol. 17, No. 1, p. 127,128.

Feldstein, M.S. (1973). Multicollenearity and the Mean Square Error of Alternative Estimators, Econometrica, Vol. 41, No. 2.

Hemmerle, W.J. (1975). An Explicite Solution for Generalized Ridge Regression, Technometrics, 17, p. 309-314.

Hoerl, A.E. and R.W. Kennard (1970-a). Ridge Regression: Biased Estimation for Non-orthogonal Problems, Technometrics, Vol. 12, No. 1, p. 55-67.

Hoerl, A.E. and R.W. Kennard (1970-b). Ridge Regression: Applications to Non-orthogonal Problems, Technometrics, Vol. 12, No. 1, p. 69-82.

James, W. and C. Stein (1961). Estimation with Quadratic Loss, Proc. 4th Berkeley Symposium 1, p. 361-379.

Moulaert, F. (1976). Ridge Regression: a Geometrical Revisitation, Regional Science Research Paper no. 15, Centrum voor Economische Studiën van de Katholieke Universiteit te Leuven.

Newhouse, J.P. and S.D. Oman (1971). The Evaluation of Ridge Estimators. Report R-716-PR prepared for project Rand.

Smith, A.F.M. and M. Goldstein (19). Ridge Regression: Some Comments on a Paper of Stone and Conniffe.

Stein, C.M. (1960). Multiple Regression. Chapter 37 in Essays in Honour of Harold Hotelling. Stanford Univ. Press.

Vinod, H.D. (1976-a). Canonical Ridge and Econometrics of Joint Production, Journal of Econometrics 4, p. 147-166.

Vinod, H.D. (1977-b). Application of New Ridge Regression Methods to a Study of Bell System Scale Economies, JASA, Vol. 71, No. 356, p. 835-841.

Appendix A

Consider the random variable X, which is Normally distributed with mean θ_1 and variance θ_2^2, and the random variable Y which is defined as

$$\text{(A.1)} \quad \begin{cases} Y = X & \text{for } |X| < 2\theta_2 \\ Y = \tfrac{1}{2}X\left(1 + \sqrt{1 - 4\theta_2^2 / X^2}\right) & \text{for } |X| > 2\theta_2 \end{cases}$$

The Mean Square Error of Y with respect to θ_1 is defined as

$$\text{(A.2)} \quad \pi(Y;\theta_1) = E\,[Y - \theta_1]^2$$

and the Mean Square Error of X with respect to θ_1 is

$$\text{(A.3)} \quad \pi(X;\theta_1) = E\,[X - \theta_1]^2 = \theta_2^2$$

We shall now derive an integral expression for the ratio $\pi(Y;\theta_1) / \pi(X;\theta_1)$ and show that this ratio is a function of θ_1/θ_2 only.

From (A.2) and (A.3) it follows that

$$\text{(A.4)} \quad \frac{\pi(Y;\theta_1)}{\pi(X;\theta_1)} = \frac{1}{\theta_2^2}\, E\,[Y - \theta_1]^2$$

and since X is normally distributed we may write for (A.4)

$$\text{(A.5)} \quad \frac{\pi(Y;\theta_1)}{\pi(X;\theta_1)} = \frac{1}{\theta_2^2} \cdot \frac{1}{\theta_2\sqrt{2\pi}} \left\{ \int_{|x| < 2\theta_2} (x - \theta_1)^2 \exp\left\{-\tfrac{1}{2}\left(\frac{x-\theta_1}{\theta_2}\right)^2\right\} dx \right.$$

$$\left. + \int_{|x| > 2\theta_2} \left[\tfrac{1}{2}x\left(1 + \sqrt{1 - 4\theta_2^2 / x^2}\right) - \theta_1\right]^2 \exp\left\{-\tfrac{1}{2}\left(\frac{x-\theta_1}{\theta_2}\right)^2\right\} dx \right\}$$

This expression can be rewritten if we apply the transformation $z = x/\theta_2$; we then have

$$\text{(A.6)} \qquad \frac{\pi(Y;\theta_1)}{\pi(X;\theta_1)} = \frac{1}{\sqrt{2\pi}} \left\{ \int_{|z|<2} (z - \frac{\theta_1}{\theta_2})^2 \exp\{-\tfrac{1}{2}(z - \frac{\theta_1}{\theta_2})^2\} \, dz \, + \right.$$

$$+ \int_{|z|>2} [\tfrac{1}{2}z(1 + \sqrt{1 - 4/z^2}) - \frac{\theta_1}{\theta_2}]^2 \exp\{-\tfrac{1}{2}(z - \frac{\theta_1}{\theta_2})^2\} \, dz$$

$$\equiv \varphi_1(\theta), \qquad \text{with } \theta = \theta_1/\theta_2$$

It can easily be seen that the function $\varphi_1(\theta)$ is symmetric about the origin, therefore we confirm ourselves to $\theta > 0$ for its numerical evaluation. The function $\varphi_1(\theta)$ has been tabulated in appendix C.

Appendix B

Consider the independently distributed random variables X and Z which have as marginal distributions a normal distribution with mean θ_1 and variance θ_2^2 and a chi-square distribution with m degrees of freedom respectively. Then the joint density of X and Z is

$$
\text{(B.1)} \quad \begin{cases} f(x,z) = c.z^{\frac{m}{2}-1} \exp\{-\frac{z}{2} - \frac{1}{2}(\frac{x-\theta_1}{\theta_2})^2\} & \text{for } -\infty < x < \infty,\ z > 0 \\ \quad = 0 & \text{elsewhere} \\ \text{with } C = [\sigma\sqrt{2\pi}\,\Gamma(\frac{m}{2})\ z^{m/2}]^{-1} & \end{cases}
$$

In this appendix we shall investigate the mean square error with respect to θ_1 of the following function of X and Z:

$$
\text{(B.2)} \quad \begin{cases} Y = X & \text{for } |X| < 2\theta_2\sqrt{Z/m} \\ Y = \frac{1}{2}X\ (1 + \sqrt{1 - 4\theta_2^2\ Z/mX^2} & \text{for } |X| > 2\theta_2\sqrt{Z/m} \end{cases}
$$

and we shall consider again the mean square error of Y as a proportion of the mean square error of X; we are thus looking for

$$
\text{(B.3)} \quad \frac{\pi(Y;\theta_1)}{\pi(X;\theta_1)} = \frac{1}{\theta_2^2}\ E\,[Y - \theta_1]^2 =
$$

$$
\frac{C}{\theta_2^2}\Bigg[\int_{-\infty}^{\infty}\int_{\frac{mx^2}{4\theta_2^2}}^{\infty} (x-\theta_1)^2\ z^{\frac{m}{2}-1} \exp\{-\frac{z}{2} - \frac{1}{2}(\frac{x-\theta_1}{\theta_2})^2\}dzdx\ +
$$

$$
+ \int_{-\infty}^{\infty}\int_{0}^{\frac{mx^2}{4\theta_2^2}} [\tfrac{1}{2}x\ (1 + \sqrt{1 - 4\theta_2^2 z/mx^2}) - \theta_1]^2\ z^{\frac{m}{2}-1}\ \bullet
$$

$$
\bullet \exp\{-\frac{z}{2} - \frac{1}{2}(\frac{x-\theta_1}{\theta_2})^2\}\,dzdx\Bigg]
$$

After the transformation $w = x/\theta_2$, the above expression can be rewritten as

$$C'\left[\int_{-\infty}^{\infty}\int_{0}^{\frac{mw^2}{4}}[\tfrac{1}{2}w\,(1+\sqrt{1-4z/mw^2})-\theta]^2\, z^{\frac{m}{2}-1}\exp\{-\tfrac{z}{2}-\tfrac{1}{2}(w-\theta)^2\}\,dzdw\right.$$

$$-\int_{-\infty}^{\infty}\int_{0}^{\frac{mw^2}{4}}(w-\theta)^2\, z^{\frac{m}{2}-1}\exp\{-\tfrac{z}{2}-\tfrac{1}{2}(w-\theta)^2\}\,dzdw$$

$$\left.+\int_{-\infty}^{\infty}\int_{0}^{\infty}(w+\theta)^2\, z^{\frac{m}{2}-1}\exp\{-\tfrac{z}{2}-\tfrac{1}{2}(w-\theta)^2\}\,dzdw\right]$$

with $\theta = \theta_1/\theta_2$

and $C' = [\sqrt{2\pi}\,\Gamma(\frac{m}{2})\, z^{m/2}]^{-1}$

It may readily be seen that the third integral in this expression equals $(C')^{-1}$ and that the first two integrals may be taken together. If, moreover, we apply the transformation

$$y = \frac{4}{mw^2}\,z$$

we obtain

$$\text{(B.4)}\qquad \frac{\pi(Y;\theta_1)}{\pi(X;\theta_1)} = 1 + C''\int_0^1 y^{\frac{m}{2}-1}\exp\{-\tfrac{1}{2}\,\frac{\theta^2 my}{my+4}\}\int_{-\infty}^{\infty} h(w,y;\theta)|\,w|^m\,\cdot$$

$$\cdot\exp\{-\tfrac{1}{2}(\,\frac{w-\mu}{\sigma}\,)^2\}\,dwdy$$

$$\equiv \varphi_2(\theta,m)$$

with $\theta = \theta_1/\theta_2$

$\mu = 4\theta/(my+4)$

$\sigma = 4/(my+4)$

$h = (w,y;\theta) = [\{ \frac{1}{2}w\ (1 + \sqrt{1-y})-\theta\}^2 - \{w-\theta\}^2]$

$C'' = C'(\frac{m}{4})^{m/2}$

Hence $\pi(Y;\theta_1)\ /\ \pi(\acute{X};\theta_1)$ can be written as a function of $\theta = \theta_1/\theta_2$ and m alone. This function labelled as $\varphi_2(\theta,m)$ has been tabulated in appendix C for positive values of θ and even values of m[1].

----

1) It should be noted that the integral expression (B.4) is relatively easy to calculate for even m and since the function $\varphi_2(\theta,m)$ is not sensitive to m, it did not seem worthwhile to devote more time and effort to the calculation of $\varphi_2(\theta,m)$ for odd values of m.

Appendix C

θ	$\varphi_1(\theta)$	$\varphi_2(\theta,m)$									
		m = 2	m = 6	m = 10	m = 14	m = 18	m = 22	m = 26	m = 30	m = 40	m = 50
0.00	0.891	0.891	0.883	0.885	0.886	0.886	0.887	0.888	0.888	0.889	0.889
0.20	0.886	0.890	0.881	0.881	0.882	0.882	0.883	0.883	0.884	0.884	0.884
0.40	0.872	0.886	0.872	0.871	0.871	0.871	0.871	0.871	0.871	0.871	0.871
0.60	0.850	0.881	0.860	0.856	0.854	0.853	0.853	0.852	0.852	0.852	0.851
0.80	0.825	0.875	0.846	0.838	0.835	0.833	0.831	0.830	0.830	0.829	0.828
1.00	0.800	0.871	0.833	0.821	0.816	0.813	0.811	0.809	0.808	0.806	0.805
1.20	0.780	0.868	0.823	0.809	0.801	0.797	0.794	0.792	0.791	0.788	0.786
1.40	0.770	0.869	0.819	0.803	0.795	0.790	0.786	0.784	0.782	0.779	0.777
1.60	0.773	0.874	0.824	0.807	0.799	0.794	0.790	0.788	0.786	0.783	0.781
1.80	0.792	0.884	0.838	0.823	0.815	0.811	0.808	0.805	0.804	0.801	0.799
2.00	0.828	0.900	0.862	0.850	0.845	0.841	0.839	0.837	0.836	0.834	0.833
2.20	0.877	0.920	0.895	0.888	0.885	0.884	0.882	0.882	0.881	0.880	0.880
2.40	0.937	0.944	0.935	0.935	0.935	0.935	0.935	0.935	0.935	0.936	0.936
2.60	1.001	0.972	0.980	0.986	0.989	0.991	0.992	0.994	0.994	0.996	0.997
2.80	1.063	1.001	1.026	1.038	1.043	1.047	1.050	1.052	1.053	1.055	1.057
3.00	1.119	1.032	1.071	1.086	1.094	1.099	1.103	1.105	1.107	1.110	1.111
3.20	1.165	1.061	1.111	1.129	1.138	1.144	1.147	1.150	1.152	1.155	1.157
3.40	1.198	1.088	1.145	1.164	1.173	1.178	1.182	1.184	1.186	1.189	1.191
3.60	1.219	1.113	1.172	1.189	1.197	1.202	1.205	1.207	1.208	1.211	1.212
3.80	1.227	1.134	1.190	1.205	1.211	1.215	1.217	1.219	1.220	1.222	1.223
4.00	1.226	1.150	1.200	1.211	1.216	1.219	1.220	1.221	1.222	1.223	1.224

-continued-

θ	$\varphi_1(\theta)$	$\varphi_2(\theta,m)$									
		m = 2	m = 6	m = 10	m = 14	m = 18	m = 22	m = 26	m = 30	m = 40	m = 50
4.20	1.218	1.163	1.204	1.211	1.214	1.215	1.216	1.216	1.217	1.217	1.218
4.40	1.206	1.171	1.201	1.205	1.206	1.206	1.206	1.206	1.206	1.206	1.206
4.60	1.191	1.176	1.195	1.195	1.195	1.194	1.194	1.193	1.193	1.192	1.192
4.80	1.175	1.177	1.185	1.183	1.181	1.180	1.179	1.179	1.178	1.177	1.177
5.00	1.159	1.175	1.174	1.170	1.167	1.166	1.164	1.164	1.163	1.162	1.162
5.50	1.126	1.161	1.144	1.137	1.134	1.132	1.131	1.130	1.130	1.129	1.128
6.00	1.101	1.141	1.117	1.111	1.108	1.106	1.105	1.105	1.104	1.103	1.103
6.50	1.083	1.120	1.096	1.091	1.088	1.087	1.086	1.086	1.085	1.085	1.084
7.00	1.070	1.102	1.080	1.076	1.074	1.073	1.072	1.072	1.072	1.071	1.071
7.50	1.059	1.086	1.068	1.064	1.063	1.062	1.062	1.061	1.061	1.061	1.060
8.00	1.051	1.074	1.058	1.055	1.054	1.054	1.053	1.053	1.053	1.052	1.052
8.50	1.045	1.064	1.051	1.048	1.047	1.047	1.046	1.046	1.046	1.046	1.045
9.00	1.039	1.056	1.045	1.043	1.042	1.041	1.041	1.041	1.040	1.040	1.040
9.50	1.035	1.049	1.040	1.038	1.037	1.037	1.036	1.036	1.036	1.036	1.036
10.00	1.031	1.044	1.035	1.034	1.033	1.033	1.032	1.032	1.032	1.032	1.032
15.00	1.013	1.018	1.015	1.014	1.014	1.014	1.014	1.014	1.014	1.013	1.013
20.00	1.007	1.010	1.008	1.008	1.007	1.007	1.007	1.007	1.007	1.007	1.007
25.00	1.004	1.006	1.005	1.005	1.005	1.005	1.004	1.004	1.004	1.004	1.004
30.00	1.003	1.004	1.003	1.003	1.003	1.003	1.003	1.003	1.003	1.003	1.003
35.00	1.002	1.003	1.002	1.002	1.002	1.002	1.002	1.002	1.002	1.002	1.002
40.00	1.001	1.002	1.002	1.002	1.001	1.001	1.001	1.001	1.001	1.001	1.001

Remark: $\varphi_1(\theta) = 0$ for $\theta = 0.25966799$