A NOTE ON ALMOST UNBIASED
GENERALIZED RIDGE REGRESSION
ESTIMATOR UNDER ASYMMETRIC LOSS

ALAN T. K. WAN

Department of Management Sciences, City University of Hong Kong,
Kowloon, Hong Kong

(Received 28 September 1997)

Using the asymmetric LINEX loss function, we derive and numerically evaluate the
exact risk function of the almost unbiased feasible generalized ridge regression estimator.
Contrary to the properties of the (biased) feasible generalized ridge estimator, it is found
that regardless of the loss asymmetry, the almost unbiased feasible generalized ridge esti-
mator does not strictly dominate the traditional least squares estimator. Our numerical
results show that over a wide range of parameter values, the almost unbiased feasible gen-
eralized ridge estimator is inferior to either the least squares or the feasible generalized
ridge estimators.

Keywords and Phrases: LINEX; relative efficiency; ridge regression; risk

1. INTRODUCTION

Consider the model,

\[ Y = X\beta + \epsilon; \quad \epsilon \sim N(0, \sigma^2 I), \]

(1.1)

where \( Y \) and \( \epsilon \) are \( n \times 1 \) vectors; \( X \) is a \( n \times p \) non-stochastic matrix of
rank \( p \); and \( \beta \) is a \( p \times 1 \) vector of unknown parameters.

For purposes of analysis, we reparameterize (1.1) as,

\[ Y = H\gamma + \epsilon, \]

where \( H = XT, \ \gamma = T'\beta, \ T \) is an orthogonal matrix such that
\( T'X'XT = \Lambda \), and \( \Lambda \) is a diagonal matrix with the eigenvalues of \( X'X \)

411
as its diagonal elements. Now, the ordinary least squares (OLS) and the generalized ridge regression estimators of the $i$th element of $\gamma$ are,

$$\hat{\gamma}_i = h'_i Y / \lambda_i$$

and

$$g_i = h'_i Y / (\lambda_i + k_i)$$
$$= \lambda_i \hat{\gamma}_i / (\lambda_i + k_i)$$

respectively, where $h_i$ is the $i$th column vector of $H$, $\lambda_i$ is the $i$th diagonal element of $\Lambda$, and $k_i$ is an (unknown) positive scalar.

The estimator $g_i$ is non-operational. A feasible generalized ridge estimator (FGRE) of $\gamma_i$ is just the first step of the iterative procedure suggested by Hoerl and Kennard (1970), namely,

$$\tilde{\gamma}_i = \lambda_i \hat{\gamma}_i / (\lambda_i + \hat{k}_i),$$

where $\hat{k}_i = \hat{\sigma}^2 \hat{\gamma}_i^2$, $\hat{\sigma}^2 = (Y - H\hat{\gamma})'(Y - H\hat{\gamma})/\nu$, $\hat{\gamma} = (\hat{\gamma}_1, \hat{\gamma}_2, \ldots, \hat{\gamma}_p)'$ is the OLS estimator of $\gamma$ and $\nu = n - p$.

Hoerl et al. (1975) made an early study of the properties of $\tilde{\gamma}_i$ by Monte-Carlo simulations (see also Lawless and Wang (1976)); and Dwivedi et al. (1980) derived the exact expressions of the first two moments of $\tilde{\gamma}_i$. The latter authors showed that $\tilde{\gamma}_i$ has smaller MSE than $\hat{\gamma}_i$ if $\theta_i^2/2 = \lambda_i \gamma_i^2 / (2\sigma^2) \leq 1$, where $\theta_i^2/2$ is the non-centrality parameter associated with the non-central Beta distributed statistic $\Psi = 1/[1 + \nu \hat{\sigma}^2 / (\gamma_i^2 \lambda_i)]$. Also of relevance here are the results of Ohtani (1993) who considered the distribution and density functions of $\tilde{\gamma}_i$.

Kadiyala (1984) introduced a class of almost unbiased estimators of which the following bias corrected ridge estimator is a special case,

$$g^*_i = g_i + k_i \gamma_i / (\lambda_i + k_i).$$

The second term in (1.2) represents the negative of the bias of the generalized ridge regression estimator $g_i$. Ohtani (1986) suggested to replace $g_i$ and $\gamma_i$ by $\tilde{\gamma}_i$, and $k_i$ by $\hat{k}_i$ in (1.2) to obtain the following almost unbiased feasible generalized ridge estimator (AUFGRE),

$$\gamma^*_i = \tilde{\gamma}_i + \hat{k}_i \tilde{\gamma}_i / (\lambda_i + \hat{k}_i).$$

Contrary to common conjectures, Ohtani (1986) showed by numerical evaluations that over a wide range of the $\theta_i^2$ space, $\gamma^*_i$ is dominated by $\tilde{\gamma}_i$ in the sense of MSE.
It is interesting to note that almost all studies on ridge regression to date use the MSE, or equivalently, the (symmetric) quadratic loss as the basis of measuring estimators' performance. It is well recognised that the use of symmetric loss functions may be inappropriate in many circumstances, particularly when positive and negative errors have different consequences. Varian (1975), in a study of real estate assessment, introduced the linear exponential (LINEX) loss function which conveniently captures the asymmetry with respect to the penalty imposed on over- and under-estimation. The LINEX loss function is defined as,

$$L(\nabla) = \zeta [\exp(\alpha \nabla) - \alpha \nabla - 1]$$  \hspace{1cm} (1.3)

where $\nabla = (\bar{y}_i - y_i)/y_i$ denotes the estimation error in using $\bar{y}_i$ to estimate $y_i$, $\alpha (\neq 0)$ is a shape parameter and $\zeta > 0$ is a factor of proportionality. For small values of $|\alpha|$, the LINEX loss reduces to quadratic loss. Over (Under)-estimation is deemed to lead to more serious consequences if $\alpha$ is positive (negative). Since Zellner (1986) established the properties of the LINEX loss function, numerous authors have used this loss function in various studies. Recent examples include Takagi (1994), Cain and Janssen (1995), Ohtani (1995), Giles and Giles (1996), Zou (1997), among others. In particular, Ohtani (1995) considered the risk of the FGRE under LINEX loss. Ohtani showed that $\bar{y}_i$ can strictly dominate the OLS estimator when $\alpha$ is positive and large. His results, of course, contrast with those obtained under quadratic loss, where it is well known that neither the FGRE nor the OLS estimator strictly dominates each other (Dwivedi et al. (1980)).

In the spirit of results obtained by Ohtani (1986, 1995), this article examines the properties of the AUFGRE under the asymmetric LINEX loss function, of which quadratic loss is a special case. In the next section, we give the analytical expression of the exact risk of this estimator, the derivation of which is given in the Appendix. To interpret the derived expression, we consider some specific numerical evaluations. Our numerical results suggest that contrary to the behaviour of the FGRE, the AUFGRE does not strictly dominate the OLS estimator regardless of the loss asymmetry. Rather, over the range of the chosen parameter values, the AUFGRE is dominated by either the OLS estimator or the FGRE.
2. RESULTS

As $\zeta$ is only a factor of proportionality, without loss of generality, we let $\zeta = 1$ in our subsequent analysis.

**Lemma** Given the assumption of section 1, the risk of the AUFGRE under the LINEX loss function with $\zeta = 1$ is,

$$R(\gamma_i^*) = \sum_{j=2}^{\infty} \sum_{r=0}^{j} \sum_{s=0}^{r} \alpha^j (-1)^{j-r} / [s!(r-s)!(j-r)!!]$$

$$E\left\{ \frac{z_i^{3r} (V/\nu)^s}{\theta_i^r (z_i^2 + V/\nu)^{r+s}} \right\},$$

where $z_i = \lambda_i^{1/2} \hat{\gamma}_i / \sigma \sim N(\theta_i, 1)$, $V = \nu \hat{\sigma}^2 / \sigma^2 \sim \chi^2_v$,

$$E\left\{ \frac{z_i^{3r} (V/\nu)^s}{\theta_i^r (z_i^2 + V/\nu)^{r+s}} \right\}$$

$$= \sum_{L=0}^{\infty} \exp \left( -\frac{\theta_i^2}{2} \right) \theta_i^{-2(m-L)2^m-L\nu^2m} \Gamma \left( \frac{\nu + 1}{2} + L + m \right)$$

$$\times \left[ \Gamma \left( \frac{\nu}{2} \right) \right]^{-1} \left[ \Gamma \left( \frac{1}{2} + L \right) \right]^{-1} (L!)^{-1}$$

$$\int_0^1 f^{3m+L-1/2} (1 - f)^{\nu/2-1+s} [1 + f(\nu - 1)]^{-2m-s} df$$

(2.2)

if $r = 2m$, or,

$$E\left\{ \frac{z_i^{3r} (V/\nu)^s}{\theta_i^r (z_i^2 + V/\nu)^{r+s}} \right\}$$

$$= \sum_{L=0}^{\infty} \exp \left( -\frac{\theta_i^2}{2} \right) \theta_i^{-2(m-L)2^m-L\nu(2m+1)} \Gamma \left( \frac{\nu + 3}{2} + m + L \right)$$

$$\times \left[ \Gamma \left( \frac{\nu}{2} \right) \right]^{-1} \left[ \Gamma \left( \frac{3}{2} + L \right) \right]^{-1} (L!)^{-1}$$

$$\int_0^1 f^{3m+L+3/2} (1 - f)^{\nu/2-1+s} [1 + f(\nu - 1)]^{-2m-s-1} df$$

(2.3)

if $r = 2m + 1$. 
Proof see the Appendix.

Among other things, the risks of $\gamma_i^*$ depend on the data through $\theta_i$. Given the complexity of the risk expression, it is difficult to evaluate (2.1)–(2.3) analytically. Accordingly, we numerically evaluate these expressions using selected values of various parameters. For purposes of analysis, the risk of the FGRE, as given in Ohtani (1995), and that of the OLS estimator, are also evaluated. We consider $\nu = 20, 40, 100$, $\alpha = [-5, 5]$ and various values of $\theta_i^2$. The subroutines GAMMLN and FACTLN of Press et al. (1993) are used to calculate the Gamma functions and factorials within the risk expression, and the NAG (1993) subroutine D01AJF is used to evaluate the integrals in (2.2)–(2.3). The convergence tolerance is set to $10^{-20}$ for each of the infinite series in the risk expression. In comparing the risks, we consider relative efficiencies, defined as $\omega_1 = R(\bar{\gamma}_i)/R(\gamma_i)$ and $\omega_2 = R(\bar{\gamma}_i)/R(\gamma_i^*)$. Thus, both the FGRE and AUFGRE have smaller risks than the OLS estimator if $\omega_1 > 1$ and $\omega_2 > 1$. If $\omega_2 > \omega_1$, then the FGRE is relatively less efficient than the AUFGRE, and vice versa.

Tables I and II illustrate some results for $\nu = 20$ and 100. The loss function with $\alpha = 0.01$ exhibits very little asymmetry and so the values of $\omega_2$ match the corresponding results of Ohtani (1986), where the AUFGRE is considered under quadratic loss. Regardless of the loss asymmetry, both $\bar{\gamma}_i$ and $\gamma_i^*$ dominate the OLS estimator for small values of $\theta_i^2$. On the other hand, for large values of $\theta_i^2$, $\gamma_i^*$ is dominated by the OLS estimator. At least for the cases that we have considered, $\gamma_i^*$ has smaller risk than $\bar{\gamma}_i$ only over certain range of $\theta_i^2$ where both the FGRE and AUFGRE are inferior to the OLS estimator. When both the FGRE and AUFGRE dominate the OLS estimator, the FGRE also dominates the AUFGRE. In other words, the AUFGRE is inadmissible under the LINEX loss function. Other things being equal, for high degrees of (positive) loss asymmetry, the risk of $\gamma_i^*$ approaches that of the OLS estimator. For relatively large values of $\alpha$ (say, $\alpha > 2$), the FGRE uniformly dominates the OLS estimator. Qualitatively, the results are unaltered for other values of $\nu$ that we have considered. Quantitatively, other things being equal, $\omega_1$ increases as $\nu$ increases when $\alpha$ is relatively large. On the other hand, the efficiency of the AUFGRE with respect to the OLS estimator decreases over a wide range of parameter values as $\nu$ increases.
<table>
<thead>
<tr>
<th>$\alpha = \theta_{0}^{2}$</th>
<th>$-1.0$</th>
<th>$-0.5$</th>
<th>$0.01$</th>
<th>$0.5$</th>
<th>$1.0$</th>
<th>$2.0$</th>
<th>$3.0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$w_{1}$</td>
<td>2.375</td>
<td>1.289</td>
<td>1.970</td>
<td>1.308</td>
<td>1.810</td>
<td>1.253</td>
<td>1.879</td>
</tr>
<tr>
<td>$w_{2}$</td>
<td>1.944</td>
<td>1.296</td>
<td>1.687</td>
<td>1.240</td>
<td>1.613</td>
<td>1.188</td>
<td>1.675</td>
</tr>
<tr>
<td>$w_{1}$</td>
<td>1.655</td>
<td>1.237</td>
<td>1.497</td>
<td>1.175</td>
<td>1.468</td>
<td>1.135</td>
<td>1.531</td>
</tr>
<tr>
<td>$w_{2}$</td>
<td>1.506</td>
<td>1.175</td>
<td>1.359</td>
<td>1.121</td>
<td>1.356</td>
<td>1.093</td>
<td>1.421</td>
</tr>
<tr>
<td>$w_{1}$</td>
<td>1.313</td>
<td>1.120</td>
<td>1.254</td>
<td>1.077</td>
<td>1.267</td>
<td>1.057</td>
<td>1.334</td>
</tr>
<tr>
<td>$w_{2}$</td>
<td>0.956</td>
<td>0.951</td>
<td>0.972</td>
<td>0.945</td>
<td>1.015</td>
<td>0.950</td>
<td>1.079</td>
</tr>
<tr>
<td>$w_{1}$</td>
<td>0.822</td>
<td>0.881</td>
<td>0.855</td>
<td>0.890</td>
<td>0.904</td>
<td>0.905</td>
<td>0.963</td>
</tr>
<tr>
<td>$w_{2}$</td>
<td>0.760</td>
<td>0.855</td>
<td>0.799</td>
<td>0.870</td>
<td>0.848</td>
<td>0.888</td>
<td>0.903</td>
</tr>
<tr>
<td>$w_{1}$</td>
<td>0.732</td>
<td>0.849</td>
<td>0.772</td>
<td>0.867</td>
<td>0.819</td>
<td>0.885</td>
<td>0.871</td>
</tr>
<tr>
<td>$w_{2}$</td>
<td>0.737</td>
<td>0.902</td>
<td>0.772</td>
<td>0.915</td>
<td>0.810</td>
<td>0.926</td>
<td>0.847</td>
</tr>
<tr>
<td>$w_{1}$</td>
<td>0.826</td>
<td>0.972</td>
<td>0.848</td>
<td>0.975</td>
<td>0.871</td>
<td>0.978</td>
<td>0.894</td>
</tr>
<tr>
<td>$w_{2}$</td>
<td>0.876</td>
<td>0.989</td>
<td>0.892</td>
<td>0.990</td>
<td>0.908</td>
<td>0.991</td>
<td>0.923</td>
</tr>
<tr>
<td>$w_{1}$</td>
<td>0.960</td>
<td>0.999</td>
<td>0.965</td>
<td>0.999</td>
<td>0.970</td>
<td>0.999</td>
<td>0.975</td>
</tr>
<tr>
<td>$\tau^2_i$</td>
<td>( w_1 )</td>
<td>( w_2 )</td>
<td>( w_1 )</td>
<td>( w_2 )</td>
<td>( w_1 )</td>
<td>( w_2 )</td>
<td>( w_1 )</td>
</tr>
<tr>
<td>----------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>0.2</td>
<td>2.440</td>
<td>1.293</td>
<td>2.005</td>
<td>1.314</td>
<td>1.836</td>
<td>1.257</td>
<td>1.910</td>
</tr>
<tr>
<td>0.4</td>
<td>1.978</td>
<td>1.303</td>
<td>1.707</td>
<td>1.245</td>
<td>1.631</td>
<td>1.191</td>
<td>1.696</td>
</tr>
<tr>
<td>0.6</td>
<td>1.675</td>
<td>1.242</td>
<td>1.510</td>
<td>1.179</td>
<td>1.480</td>
<td>1.137</td>
<td>1.546</td>
</tr>
<tr>
<td>0.8</td>
<td>1.469</td>
<td>1.178</td>
<td>1.368</td>
<td>1.123</td>
<td>1.365</td>
<td>1.093</td>
<td>1.432</td>
</tr>
<tr>
<td>1.0</td>
<td>1.320</td>
<td>1.122</td>
<td>1.261</td>
<td>1.078</td>
<td>1.274</td>
<td>1.057</td>
<td>1.342</td>
</tr>
<tr>
<td>2.0</td>
<td>0.956</td>
<td>0.950</td>
<td>0.972</td>
<td>0.944</td>
<td>1.016</td>
<td>0.949</td>
<td>1.082</td>
</tr>
<tr>
<td>3.0</td>
<td>0.821</td>
<td>0.880</td>
<td>0.854</td>
<td>0.889</td>
<td>0.903</td>
<td>0.903</td>
<td>0.964</td>
</tr>
<tr>
<td>4.0</td>
<td>0.759</td>
<td>0.854</td>
<td>0.798</td>
<td>0.869</td>
<td>0.847</td>
<td>0.887</td>
<td>0.903</td>
</tr>
<tr>
<td>5.0</td>
<td>0.730</td>
<td>0.848</td>
<td>0.771</td>
<td>0.865</td>
<td>0.819</td>
<td>0.884</td>
<td>0.870</td>
</tr>
<tr>
<td>10.0</td>
<td>0.737</td>
<td>0.903</td>
<td>0.772</td>
<td>0.915</td>
<td>0.810</td>
<td>0.927</td>
<td>0.847</td>
</tr>
<tr>
<td>20.0</td>
<td>0.827</td>
<td>0.973</td>
<td>0.849</td>
<td>0.976</td>
<td>0.872</td>
<td>0.979</td>
<td>0.895</td>
</tr>
<tr>
<td>30.0</td>
<td>0.878</td>
<td>0.989</td>
<td>0.893</td>
<td>0.990</td>
<td>0.909</td>
<td>0.991</td>
<td>0.925</td>
</tr>
<tr>
<td>100.0</td>
<td>0.961</td>
<td>0.999</td>
<td>0.966</td>
<td>0.999</td>
<td>0.971</td>
<td>0.999</td>
<td>0.975</td>
</tr>
</tbody>
</table>
3. DISCUSSION

Contrary to common intuition, the AUFGRE does not strictly dominate the OLS estimator even when the loss asymmetry is positive and relatively large. This contrasts with the behaviour of the FGRE which dominates the OLS estimator when $\alpha$ is sufficiently large. Intuitively, such difference may be explained by the fact $\gamma_i$ is biased in the direction opposite to the sign of $\gamma_i$ (See Dwivedi et al. (1980)), hence when over-estimation is considered to be more serious, there are gains associated with using $\gamma_i$. Accordingly, the FGRE dominates the OLS estimator for sufficiently large values of $\alpha$. On the other hand, the AUFGRE is approximately unbiased, hence there exists no discernible gains from using this estimator whether $\alpha$ is positive or negative. It remains a task for future research to prove the inadmissibility of the AUFGRE by analytical methods. Work in process by the author extends the current analysis to other departures from quadratic loss, including the balanced loss function (Zellner (1994)) which takes into account both goodness of fit and estimation precision.

Acknowledgements

This work was supported partially by a strategic research grant from the City University of Hong Kong. The author thanks Kazuhiro Ohtani for his comments on an earlier draft of the paper.

References


**APPENDIX**

Let \( \tilde{\gamma}_i \) be an estimator of \( \gamma_i \). Under the LINEX loss function of (1.3) with \( \zeta = 1 \), the risk of \( \tilde{\gamma}_i \) is given by,

\[
R(\tilde{\gamma}_i) = E\{ \exp \{ \alpha(\tilde{\gamma}_i/\gamma_i - 1) - \alpha(\tilde{\gamma}_i/\gamma_i - 1) - 1 \} \}.
\]

As \( \exp[\alpha(\tilde{\gamma}_i/\gamma_i - 1)] = \sum_{j=0}^{\infty} \alpha^j(\tilde{\gamma}_i/\gamma_i - 1)^j/j! \) hence,

\[
R(\gamma_i) = \sum_{j=2}^{\infty} \alpha^j E((\tilde{\gamma}_i/\gamma_i - 1)^j/j! \]

\[
= \sum_{j=2}^{\infty} \sum_{r=0}^{j} \alpha^j (-1)^{j-r}/(r!(j-r)!) \} E((\tilde{\gamma}_i/\gamma_i)^r)
\]

Now, defining \( z_i = \lambda^{1/2}\tilde{\gamma}_i/\sigma \) and \( V = \nu\sigma^2/\sigma^2 \), it is readily shown that the AUFGRE can be written as,

\[
\gamma_i^* = z_i^2 \gamma_i(z_i^2 + 2V/\nu)/(\theta_i(z_i^2 + V/\nu)^2).
\]
Accordingly, the risk of \( \gamma^*_i \) is,

\[
R(\gamma^*_i) = \sum_{j=2}^{\infty} \sum_{r=0}^{j} \sum_{s=0}^{r} \alpha^{j} \left( (-1)^{j-r} / [s! (r-s)! (j-r)!] \right) E \left\{ \frac{z_i^{3r} (V/\nu)^s}{\theta_i^r (z_i^2 + V/\nu)^{r+s}} \right\}.
\]

Let's consider \( E[z_i^{3r} (V/\nu)^s / (\theta_i^r (z_i^2 + V/\nu)^{r+s})] \). Now, it is straightforward to show that,

\[
E[z_i^{3r} (V/\nu)^s / (\theta_i^r (z_i^2 + V/\nu)^{r+s})] = \int_0^\infty \int_{-\infty}^\infty \left[ z_i^{3r} (V/\nu)^s / (\theta_i^r (z_i^2 + V/\nu)^{r+s}) \right] (2\pi)^{-1/2} \times \exp \left( -(z_i^2 - \theta_i^2/2) (2

\[\exp(-2^{1/2} \Gamma(\nu/2))^{-1} V^{\nu/2-1} \exp(-V/2)dz_i dV\]

\[= \int_0^\infty \int_{-\infty}^\infty \sum_{l=0}^\infty K \exp \left( (-z_i^2 - V)/2 \right) \nu^{\nu/2-1+s} z_i^{3r+l} (z_i^2 + V/\nu)^{-r-s} dz_i dV,\]

where \( K = \exp(-\theta_i^2/2) \theta_i^{-r} (2\pi)^{-1/2} \nu^{-s}/[2^{\nu/2} \Gamma(\nu/2)!] \).

First, we consider \( r = 2m \). In this case, if \( l \) is odd, then the integral with respect to \( z_i \) is zero. So we let \( l = 2L \). Furthermore, if we make change of the variables, \( W = z_i^2 \), then \( t_1 = \nu W/V \) and \( t_2 = V \) consecutively, we obtain the following expression,

\[
E[z_i^{3r} (V/\nu)^s / (\theta_i^r (z_i^2 + V/\nu)^{r+s})] = \int_0^\infty \int_0^\infty \sum_{L=0}^\infty K \exp(-t_2(\nu + t_1)/(2\nu)) t_1^{3m+L-1/2} (1 + t_1)^{-2m-s} \times t_2^{m+L-1/2+\nu/2} \nu^{-m-L+s-1/2} dt_1 dt_2
\]

Now, if we let \( Q = t_2(\nu + t_1)/(2\nu) \) and recognizing that

\[
\int_0^\infty \exp(-Q) Q^{\nu/2+L+m-1/2} dQ = \Gamma(\nu/2 + L + m + 1/2)
\]
then we have,

\[
E[z_i^r(V/\nu)^r/(\theta_i^r(z_i^2 + V/\nu)^{r+3})]
= \int_0^\infty \sum_{L=0}^{\infty} K \Gamma(\nu/2 + L + m + 1/2)t_1^{3m+L-1/2}(1 + t_1)^{-2m-s}
\times (\nu + t_1)^{-\nu/2 - L - m - 1/2} \nu^{\nu/2 + s} 2^{\nu/2 + L + 1/2 + m} dt_1
\]

Finally, making use of the change of the variable \(f = t_1/(t_1 + \nu)\) and noting that \((2L)! \pi^{1/2} = 2^{2L} \Gamma(L + 1/2) L!\), we obtain the required expression of \(E[z_i^r(V/\nu)^r/(\theta_i^r(z_i^2 + V/\nu)^{r+3})]\) with \(r = 2m\) in the text. The corresponding expression for \(r = 2m + 1\) can be obtained in a parallel way and is left to the readers to verify.