Performance Assessment of the Modified Kernel Ridge Predictors in the Partially Linear Mixed Measurement Error Models via Covid-19 Data Analysis

Ölçüm Hatalı Kısmi Lineer Karma Modellerde Modified Kernel Ridge Öntahmin Edicilerin Covid-19 Veri Analizi Yoluyla Performans Değerlendirmesi

Özge Kuran *, Seçil Yalaz *

Dicle Üniversitesi Fen Fakültesi İstatistik Bölümü, Diyarbakır, TÜRKİYE
Corresponding Author / Sorumlu Yazar *: ozge.kuran@dicle.edu.tr; ozge-kuran@hotmail.com

Abstract

In this article we describe new predictors under multicollinearity situation in the partially linear mixed measurement error models. In order to achieve this aim, we refer to some preliminary information and use it in order to suggest the modified Kernel ridge predictors in the partially linear mixed measurement error models. In addition, we also attain some mean square error comparisons between our new described modified Kernel ridge predictors and predictors previously described in literature for the partially linear mixed measurement error model. In conclusion, the article showcases real data analysis and a simulation study to illustrate our theoretical findings.

Keywords: Linear Mixed Model, Measurement Error, Multicollinearity, Partially Linear Model, Ridge Predictor

Öz

Bu çalışmada, ölçüm hatalı kısmi lineer karma modellerde çoklu iç ilişki durumunda yeni öntahmin ediciler tanımlanmaktadır. Bu amaçla ulaşmak için, bazı ön bilgiler ele alınmıştır ve bu bilgi hesaba katılarak, ölçüm hatalı kısmi lineer karma modellerdeki modified Kernel ridge öntahmin edicileri oluşturulmuştur. İlk olarak, ölçüm hatalı kısmi lineer karma model literatüründe daha önce tanımlanan öntahmin ediciler ile yeni tanımlanan modified Kernel ridge öntahmin ediciler arasında bazı hatalar ortalamalari karşılaşılmaktan da yapılmasını. Daha sonra, teorik bulgularımızı kanıtlamak için gerçek bir veri analizi ve simulasyon çalışması ile makale sonlandırılmıştır.

Anahtar Kelimeler: Lineer Karma Model, Ölçüm Hatası, Çoklu İç İlişki, Kısmi Lineer Model, Ridge Öntahmin Ediciler

1. Introduction

Linear mixed model (LMM) [1] is an expanded version of linear model (LM). LMM has both fixed and random effects, and are especially employed to study clustered data such as longitudinal data, repeated measures data, multilevel data and etc. Another commonly studied statistical model in literature is nonparametric model (NPM) under the measurement error problem [2]. This model introduces the functional form of LM where heterogeneity is not handled. For the purpose of taking advantage of the favorable ideas of these two favored models together, partially linear mixed measurement error model (PLMMeM), a combination of LMM and the NPM under the measurement error problem, is defined by [3].

The PLMMeM based on a sample of size n with measured error in parametric part component $X_i$ is considered as

$$
Y_i = X_i^T \beta + g(T_i) + Z_i^T b_i + \epsilon_i,
$$
$$
W_i = X_i + U_i,
$$

(1)

where fixed effects design matrices are $X_i = (x_{i1}, ..., x_{ip})^T$ and for $t_{1i}, ..., t_{id}$ defined on $[0, 1], T_i = (t_{1i}, ..., t_{id})^T$, random effects design matrix is $Z_i = (z_{i1}, ..., z_{iq})^T$, a parameter vector of fixed effects design matrix is $\beta = (\beta_1, ..., \beta_p)^T$, an unknown function defined from $\mathbb{R}^d$ to $\mathbb{R}^2$ is $g(.)$, independent and identically distributed (i.i.d.), unobservable vector of the random effects design matrix is $b_i$ and i.i.d. random vector of errors is $\epsilon_i$. Independent $b_i$ and $\epsilon_i$ are chosen from a Gaussian process with mean zero and covariance matrix $\Sigma_b$ and $\Sigma_\epsilon$, respectively.

When $X_i$’s are observable, the conditional distribution of $Y_i$ for a given $b_i$ is $Y_i|b_i \sim N(X_i^T \beta + g(T_i) + Z_i^T b_i, \Sigma_i)$. However, we observe $W_i$ instead of observing $X_i$ in model (1), assuming that the measurement error $U_i$ has a known i.i.d. with mean zero and covariance matrix $\Sigma_{uu}$ and independent of $(Y_i, X_i, T_i, Z_i)$.

If we introduce the conditional expectations also known as the kernel regressions of $Y, X$ and $Z$ with bandwidth $h$, respectively, as

$$
\omega_y(T_i) = E(Y_i|T_i),
$$
$$
\omega_x(T_i) = E(X_i|T_i),
$$
$$
\omega_z(T_i) = E(Z_i|T_i),
$$

then the matrix form of model (1) is obtained as
The Kernel ridge and Kernel Liu predictors given by Eqs. (5-8) are biased prediction approaches that are suggested by using some prior information [9] in order to eliminate the negative effects of the multicollinearity problem in PLMMeMs. In addition to these two approaches, our goal in this article is to propose a new biased prediction approach in PLMMeMs by taking a convex combination of Kernel ridge and Kernel Liu predictors as prior information. This new approach called the modified Kernel ridge is such a convex combined approach that it results in unifying the advantages of the Kernel ridge prediction and Kernel Liu prediction. Since it is a combination of both Kernel ridge and Kernel Liu approaches, it is thought to be more successful than Kernel ridge and Kernel Liu approaches in minimizing the negative effects of multicollinearity. Then, the rest of this paper is structured as follows: Section 2, the new predictors in PLMMeMs are characterized. In Section 3, we make some mean square error comparisons and Covid-19 data analysis under known measurement errors and covariance matrix is done in Section 4. In Section 5, a simulation study is also done. Finally, concluding remarks are given in Section 6.

2. The Modified Kernel Ridge Predictors

Our aim in this section is to suggest the modified Kernel ridge prediction approach for PLMMeMs using the idea of the modified ridge estimation in linear models [9] and in LMMs [12]. We know that under model (2),

$$\mathbf{y}^\top \mathbf{y} \sim N\left(\mathbf{0}, \mathbf{I}D_{\mathcal{D}}^2\right)$$

where \(\mathbf{y}\) and \(\mathbf{\beta}\) are jointly Gaussian distributed. For a given \(\mathbf{b}\) the conditional distribution of \(\mathbf{\beta}\) is given as \(\mathbf{\beta}^{\mathcal{D}} \sim N(\mathbf{0}, \mathbf{\Sigma}^\mathcal{D})\). Then, the joint density of \(\mathbf{y}\) and \(\mathbf{b}\) is

$$f(\mathbf{y}, \mathbf{b}) = f(\mathbf{y}|\mathbf{b})f(\mathbf{b}) = (2\pi)^{-(n+q)/2}|\Sigma^\mathcal{D}|^{-1/2} \exp\left(-\frac{1}{2}\left[(\mathbf{y} - \mathbf{\beta} - \mathbf{Zb})^\top \mathbf{\Sigma}^{-1}(\mathbf{y} - \mathbf{\beta} - \mathbf{Zb}) + b^\top D^{-1}b\right]\right)$$

where \(|\cdot|\) denotes the determinant of a matrix. logf(\(\mathbf{y}, \mathbf{b}\)) is derived by dropping the constant term as

$$\log f(\mathbf{y}, \mathbf{b}) + \log f(\mathbf{b})$$

and so, a penalization term with regularization parameter \(\delta = -\frac{1}{2} \geq 0\) is added to logf(\(\mathbf{y}, \mathbf{b}\)).

$$\log f(\mathbf{y}, \mathbf{b}) - \frac{1}{2} k(1 + d) \mathbf{\beta}^\top \mathbf{\beta}. \quad (9)$$

Here, we use the prior information from [9] and [13] to chose the stochastic linear restriction \(0 = \sqrt{k(1 + d)} \mathbf{\beta} + \varepsilon\). The partial derivatives of Eq. (9) with respect to the elements of \(\mathbf{\beta}\) and \(\mathbf{b}\) are taken equal to zero, then, by switching \(\mathbf{\beta}\) and \(\mathbf{b}\) by \(\mathbf{\beta}_{k,\mathcal{D}}\) and \(\mathbf{\beta}_{k,\mathcal{D}}\) respectively.

$$\mathbf{X}_\mathcal{D}^\top \mathbf{\Sigma}^{-1}(\mathbf{y} - \mathbf{\beta}_{k,\mathcal{D}}) - k(1 + d) \mathbf{\beta}_{k,\mathcal{D}} = 0. \quad (10)$$

$$2\mathbf{Z}^\top \mathbf{\Sigma}^{-1}(\mathbf{Zb}_{k,\mathcal{D}}) - (Z^\top \mathbf{\Sigma}^{-1}Z + D^{-1}) \mathbf{b}_{k,\mathcal{D}} = 0. \quad (11)$$

are obtained. Eqs. (10) and (11) are also equal the matrix form given by

$$\mathbf{X}_\mathcal{D}^\top \mathbf{\Sigma}^{-1} + k(1 + d) \mathbf{I}_p = \mathbf{X}_\mathcal{D}^\top \mathbf{\Sigma}^{-1} Z + D^{-1} \mathbf{b}_{k,\mathcal{D}} = 0. \quad (12)$$
\[
\begin{align*}
\hat{\beta}_{k,d}(t) &= \sum_{j=1}^{p} \omega_{nj}(t)(Y_j - W_{k,d}^T Z_{k,d}). \\
\end{align*}
\]

3. Some Mean Square Error Comparisons

Under specific matrices \( L \in \mathbb{R}^{s \times s} \) and \( M \in \mathbb{R}^{s \times s} \), we demonstrate the prediction of PLMMs as \( \mu = L\beta + M\beta \) [16, 17] for \( s = 1 \). By using Eqs. (3)-(8) and Eqs. (15) and (16), the predictors of \( \mu \) under the kernel, the kernel ridge, the Kernel Liu and the modified kernel ridge predictors are definable, respectively, as

\[
\hat{\mu} = L^T \hat{\beta} + M^T \hat{\beta} = Q\hat{\beta} + M^T D\hat{V}^T \hat{\beta},
\]

\[
\hat{\mu}_k = L^T \hat{\beta}_k + M^T \hat{\beta}_k = Q\hat{\beta}_k + M^T D\hat{V}^T \hat{\beta}_k,
\]

\[
\hat{\mu}_d = L^T \hat{\beta}_d + M^T \hat{\beta}_d = Q\hat{\beta}_d + M^T D\hat{V}^T \hat{\beta}_d,
\]

\[
\hat{\mu}_{k,d} = L^T \hat{\beta}_{k,d} + M^T \hat{\beta}_{k,d} = Q\hat{\beta}_{k,d} + M^T D\hat{V}^T \hat{\beta}_{k,d},
\]

where \( Q = L^T - M^T D\hat{V}^T \hat{\beta} \).

The matrix mean square error (MMSE) criterion is used to compare the betterness of \( \hat{\mu}, \hat{\mu}_k, \hat{\mu}_d \) and \( \hat{\mu}_{k,d} \). By following [18], the MMSES for \( \hat{\mu}, \hat{\mu}_k, \hat{\mu}_d \) and \( \hat{\mu}_{k,d} \) are calculated, respectively, as

\[
\begin{align*}
\text{MMSE}(\hat{\mu}) &= Q\text{MMSE}(\hat{\beta})Q^T + M^T(D - D\hat{V}^T \hat{\beta} D)M, \\
\text{MMSE}(\hat{\mu}_k) &= Q\text{MMSE}(\hat{\beta}_k)Q^T + M^T(D - D\hat{V}^T \hat{\beta}_k D)M, \\
\text{MMSE}(\hat{\mu}_d) &= Q\text{MMSE}(\hat{\beta}_d)Q^T + M^T(D - D\hat{V}^T \hat{\beta}_d D)M, \\
\text{MMSE}(\hat{\mu}_{k,d}) &= Q\text{MMSE}(\hat{\beta}_{k,d})Q^T + M^T(D - D\hat{V}^T \hat{\beta}_{k,d} D)M,
\end{align*}
\]

where \( \text{MMSE}(\hat{\beta}) = N^{-1}, \)

\[
\begin{align*}
\text{MMSE}(\hat{\beta}_k) &= N_k^{-1}NN_k^{-1} + k^2N_k^{-1}BD^2N_k^{-1}, \\
\text{MMSE}(\hat{\beta}_d) &= N_k^{-1}NN_k^{-1}N_d N_d^{-1} + (1 - d)N_k^{-1}BD^2N_k^{-1}, \\
\text{MMSE}(\hat{\beta}_{k,d}) &= NN_{k,d}^{-1}N_k^{-1}N_d N_d^{-1}N_{k,d}^{-1} - (N_k^{-1} - N_d^{-1})(N_d^{-1} - N_{k,d}^{-1}) \beta < 1.
\end{align*}
\]
UNKNOWN. Secondly, we specify the vaccine efficacy rate of each vaccine (COM=0.95, AZ=0.7469, MOD=0.8941, JANSS=0.7735, SPU=0.9760, BECNGB=0.7934 and UNK=0.8583 are computed as by taking the geometric mean of the COM, AZ, MOD, BECNGB, JANSS, SPU vaccines). And lastly, we create the nonparametric part by taking the geometric mean of the efficacy rates of the vaccines. Since the regions are randomly selected from 17 countries, random effect is explained as the regions. We log-transform the variables to make the data conform more closely to the normal distribution and to improve the model fit since the distribution of Covid-19 data is right skewed. Then, the PLMMeM is written as, for \( i = 1, \ldots, 17, j = 1, \ldots, 11 \), 
\[
y_{ij} = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + b_1 + b_2 \text{time}_{ij} + g(t_{ij}) + \epsilon_{ij},
\]
where the \( i \)th observation of the \( j \)th region of the explanatory variable \( x_{ij} \) is indicated as \( x_{ij} \) the \( i \)th observation of the \( j \)th region of the response is indicated as \( y_{ij} \) and time corresponding to \( y_{ij} \) is demonstrated as \( \text{time}_{ij} \). It is chosen using optimal bandwidth selection rule given by [21]. We use the quartic kernel function \( K(u) = (15/16)(1 - u^2)^2 I(|u| \leq 1) \) for kernel smoothing regression with the measurement error which has normal distribution \( U \sim N(0,0.25) \).

In our real data analysis, we select the restricted maximum likelihood (REML) method which has the smallest AIC/BIC values for all models from Table 1. The UN(1) variance-covariance model under AIC and BIC is seen as the best model in modeling the variance-covariance matrix structure with respect to the response.

Table 1. Variance-covariance matrix results

<table>
<thead>
<tr>
<th>Cov. struct.</th>
<th>Est. Met. for Cov. Par.</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>UN</td>
<td>ML</td>
<td>568.68</td>
<td>594.35</td>
</tr>
<tr>
<td></td>
<td>REML</td>
<td>552.50</td>
<td>578.35</td>
</tr>
<tr>
<td>UN(1)</td>
<td>REML</td>
<td>567.91</td>
<td>590.38</td>
</tr>
<tr>
<td>VC</td>
<td>REML</td>
<td>551.94</td>
<td>574.56</td>
</tr>
<tr>
<td>CS</td>
<td>REML</td>
<td>570.43</td>
<td>589.68</td>
</tr>
</tbody>
</table>


\( \hat{D}_{REML} \) and \( \hat{\Sigma}_{REML} \) values given in Table 2. Then, via\( V = \hat{Z}\hat{D}\hat{Z}^T + \Sigma \) formula, \( \hat{V}_{REML} \) values are found.

Table 2. Covariance structures estimates

| \( \hat{D}_{REML} \) | \( \begin{bmatrix} 0.0742 & 0 \\ 0 & 0.4591 \end{bmatrix} \) |
|----------------------|
| \( \hat{\Sigma}_{REML} \) | 0.8855 \( \lambda_{\text{est}} \) |

\( \lambda_1 = 0.0162e^{+03} \), \( \lambda_2 = 2.1292e^{+03} \), \( \lambda_3 = 0.6513e^{+03} \) and \( \lambda_4 = 0.1312e^{+03} \) are obtained from the matrix \( \hat{X}^T\hat{D}_{REML}^{-1}\hat{X} \). The condition number calculated as \( \lambda_{\text{max}}/\lambda_{\text{min}} \) = 132.076 is used to measure the extent of multicolinearity and \( \lambda_{\text{max}}/\lambda_{\text{min}} > 100 \) shows moderate multicolinearity.

We determine the estimators of the biasing parameters \( k \) and \( d \) with a computational algorithm as follows:

1. By using for each \( \lambda_i \) value, the \( k \) value is estimated from [7] for PLMMeMs as \( \hat{k} = k_{\text{LM}} = \frac{p}{\hat{\lambda}_{\text{max}}/\hat{\lambda}_{\text{min}}} = 12.4055 \) when the covariance parameters are estimated by REML.
2. After the \( \hat{k} \) value is found from the point 1, the Liu biasing parameter \( d \) is selected as \( \hat{d}_L \) which is given by Theorem 4.2 [11] where \( h \) is determined as multiplying the upper bound defined in Theorem 4.2 by 0.99 if \( \hat{\Sigma}_{i=1} = \lambda_i(1+\lambda_i) > \frac{2}{\hat{\lambda}_{\text{max}}(1+\lambda_i)} \).
3. If \( \hat{\Sigma}_{i=1} = \lambda_i(1+\lambda_i) < \frac{2}{\hat{\lambda}_{\text{max}}(1+\lambda_i)} \), we determined arbitrarily \( \hat{d}_L \) as 0.9705.

The estimates of the fixed parameters and nonparametric function, the predictions of the random parameters and the scalar mean square error (SME) values for Kernel, Kernel ridge, Kernel Liu and modified Kernel ridge cases under PLMMeM are presented in Table 3.

We see that in Table 3 the modified Kernel ridge estimator has better results in the sense of SME for \( \hat{k}_{\text{LM}} = 12.4055 \) and \( \hat{d}_L = 0.9705 \) than the Kernel, Kernel ridge and Kernel Liu estimators. Moreover, we calculate the conditions given by Theorems 3.1, 3.2 and 3.3, respectively, as -0.3274, -0.3273 and -0.3296, which are smaller than 1. Thus, we also say that the modified Kernel ridge estimator dominates the Kernel, Kernel ridge and Kernel Liu estimators on the MMSE criterion.

Figure 1. Comparison of the finite sample and asymptotic distributions of the estimators

Additionally, comparison between the asymptotic distributions of Kernel (green), Kernel ridge (red), Kernel Liu (blue), modified Kernel ridge (magenta) estimators and the finite sample properties are also examined. In Figure 1 where the abscissa is \( Z = (\text{Var}(g(t,h_a))^{1/2}(g(t,h_a) - E(g(t,h_a))) \) and the ordinate is probability. The empirical cumulative distribution functions (CDFs) of the estimators agree very well with the normal CDFs.
### Table 3. Data analysis results

<table>
<thead>
<tr>
<th>Kernel</th>
<th>Kernel ridge</th>
<th>Kernel Liu</th>
<th>Modified Kernel ridge</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\beta}_0$</td>
<td>$\text{4.058E}^{-10}$</td>
<td>$\text{2.303E}^{-10}$</td>
<td>$\text{4.051E}^{-10}$</td>
</tr>
<tr>
<td>$\hat{\beta}_1$</td>
<td>0.0186</td>
<td>0.0168</td>
<td>0.0186</td>
</tr>
<tr>
<td>$\hat{\beta}_2$</td>
<td>-0.0172</td>
<td>-0.0161</td>
<td>-0.0172</td>
</tr>
<tr>
<td>$\hat{\beta}_3$</td>
<td>-0.0095</td>
<td>-0.0094</td>
<td>-0.0095</td>
</tr>
<tr>
<td>$\hat{b}_1$</td>
<td>$\text{3.563E}^{-16}$</td>
<td>$\text{1.6294E}^{-10}$</td>
<td>$\text{6.323E}^{-13}$</td>
</tr>
<tr>
<td>$\hat{b}_2$</td>
<td>-0.6394</td>
<td>-0.6401</td>
<td>-0.6394</td>
</tr>
<tr>
<td>$\hat{\gamma}$</td>
<td>6.3563</td>
<td>6.3727</td>
<td>6.3563</td>
</tr>
</tbody>
</table>

### Table 4. Estimated and predicted MSE values with $g_1(t)$ function

<table>
<thead>
<tr>
<th>m</th>
<th>$\gamma^2$</th>
<th>$\hat{\beta}$</th>
<th>$\hat{\beta}_{ktw}$</th>
<th>$\hat{\beta}_{\hat{d}}$</th>
<th>$\hat{\beta}_{ktw,d}$</th>
<th>$\hat{b}$</th>
<th>$\hat{b}_{ktw}$</th>
<th>$\hat{b}_{\hat{d}}$</th>
<th>$\hat{b}_{ktw,d}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.90</td>
<td>1.309997</td>
<td>1.309915</td>
<td>1.309914</td>
<td>1.283515</td>
<td>0.008436</td>
<td>0.008435</td>
<td>0.008431</td>
<td>0.008316</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.95</td>
<td>0.779188</td>
<td>0.779710</td>
<td>0.761362</td>
<td>0.009774</td>
<td>0.009773</td>
<td>0.009769</td>
<td>0.009614</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>0.830462</td>
<td>0.830393</td>
<td>0.811062</td>
<td>0.010974</td>
<td>0.010973</td>
<td>0.010969</td>
<td>0.009924</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>0.947033</td>
<td>0.947018</td>
<td>0.936671</td>
<td>0.002307</td>
<td>0.002306</td>
<td>0.002304</td>
<td>0.002219</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.95</td>
<td>0.696424</td>
<td>0.696413</td>
<td>0.689989</td>
<td>0.001687</td>
<td>0.001686</td>
<td>0.001682</td>
<td>0.001537</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>0.633088</td>
<td>0.633077</td>
<td>0.626945</td>
<td>0.003710</td>
<td>0.003709</td>
<td>0.003705</td>
<td>0.003582</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>0.976395</td>
<td>0.976391</td>
<td>0.971870</td>
<td>0.456936\times 10^{-3}</td>
<td>0.456935\times 10^{-3}</td>
<td>0.456931\times 10^{-3}</td>
<td>0.456721\times 10^{-3}</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>0.95</td>
<td>0.912487</td>
<td>0.912484</td>
<td>0.908417</td>
<td>0.420959\times 10^{-3}</td>
<td>0.420958\times 10^{-3}</td>
<td>0.420957\times 10^{-3}</td>
<td>0.420860\times 10^{-3}</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>0.804819</td>
<td>0.804815</td>
<td>0.800992</td>
<td>0.426339\times 10^{-3}</td>
<td>0.426339\times 10^{-3}</td>
<td>0.426335\times 10^{-3}</td>
<td>0.426144\times 10^{-3}</td>
<td></td>
</tr>
</tbody>
</table>

### Table 5. Estimated and predicted MSE values with $g_2(t)$ function

<table>
<thead>
<tr>
<th>m</th>
<th>$\gamma^2$</th>
<th>$\hat{\beta}$</th>
<th>$\hat{\beta}_{ktw}$</th>
<th>$\hat{\beta}_{\hat{d}}$</th>
<th>$\hat{\beta}_{ktw,d}$</th>
<th>$\hat{b}$</th>
<th>$\hat{b}_{ktw}$</th>
<th>$\hat{b}_{\hat{d}}$</th>
<th>$\hat{b}_{ktw,d}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.90</td>
<td>1.063705</td>
<td>1.063638</td>
<td>1.063631</td>
<td>1.043362</td>
<td>0.008777</td>
<td>0.008776</td>
<td>0.008772</td>
<td>0.008604</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.95</td>
<td>0.987899</td>
<td>0.987841</td>
<td>0.987840</td>
<td>0.026092</td>
<td>0.026091</td>
<td>0.026090</td>
<td>0.025986</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>1.017534</td>
<td>1.017466</td>
<td>1.017465</td>
<td>0.013327</td>
<td>0.013326</td>
<td>0.013324</td>
<td>0.013227</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>0.840638</td>
<td>0.840625</td>
<td>0.840623</td>
<td>0.001965</td>
<td>0.001964</td>
<td>0.001962</td>
<td>0.001945</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.95</td>
<td>0.800622</td>
<td>0.800609</td>
<td>0.800605</td>
<td>0.002950</td>
<td>0.002949</td>
<td>0.002946</td>
<td>0.002845</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>0.626170</td>
<td>0.626158</td>
<td>0.626157</td>
<td>0.003990</td>
<td>0.003989</td>
<td>0.003985</td>
<td>0.003872</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>0.985321</td>
<td>0.985317</td>
<td>0.985316</td>
<td>0.980775</td>
<td>0.419944\times 10^{-3}</td>
<td>0.419944\times 10^{-3}</td>
<td>0.419941\times 10^{-3}</td>
<td>0.419842\times 10^{-3}</td>
</tr>
<tr>
<td>60</td>
<td>0.95</td>
<td>0.896444</td>
<td>0.896441</td>
<td>0.896440</td>
<td>0.892368</td>
<td>0.421393\times 10^{-3}</td>
<td>0.421393\times 10^{-3}</td>
<td>0.421391\times 10^{-3}</td>
<td>0.421290\times 10^{-3}</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
<td>0.843971</td>
<td>0.843968</td>
<td>0.843967</td>
<td>0.839810</td>
<td>0.475318\times 10^{-3}</td>
<td>0.475317\times 10^{-3}</td>
<td>0.475311\times 10^{-3}</td>
<td>0.475303\times 10^{-3}</td>
</tr>
</tbody>
</table>
5. A Simulation Study

In this section, we will investigate the performances of Kernel, Kernel ridge, Kernel Liu and modified Kernel ridge estimators in the sense of the estimated mean square error (EMSE) and the performances of Kernel, Kernel ridge, Kernel Liu and modified Kernel ridge predictors in the sense of the predicted mean square error (PMSE) under known covariance matrix.

By following [22], the fixed effects are calculated as

\[ x_{ijk} = (1 - \gamma^2) \frac{1}{2} w_{ijk} + \gamma w_{ijk} + 1, \quad i = 1, \ldots, m, \]

\[ j = 1, \ldots, n_i, k = 1, \ldots, p, \]

where \( w_{ijk} \) are independent standard normal pseudo-random numbers and \( \gamma \) is specified so that the correlation between any two fixed effects is given by \( \gamma^2 = 0.90, 0.95, 0.99 \). And, the fixed effects number size is selected as \( p = 3 \).

We think \( m \) = 15,30,60 subjects and \( n_i = 10 \) observation per subject and then, we report the simulation results with the sample sizes of \( n = \sum_{i=1}^{m} n_i = 150,300,600 \). The parameter vector \( \beta = (\beta_1, \ldots, \beta_p)^T \) is chosen as the normalized eigenvector corresponding to the largest eigenvalue of \( \mathbf{X}^T \mathbf{X}^{-1} \mathbf{X} \) so that \( \beta^T \beta = 1 \) (see [23]). Then, the underlying model takes the following form with \( q = 2 \) random effects

\[ y_{ij} = \sum_{l=1}^{q} (1 - \gamma^2) \frac{1}{2} \sum_{k=1}^{p} \beta_l x_{ijl} + \beta_l \sum_{k=1}^{p} b_{ijl} + b_{ij}, \]

where \( D = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \) is the AR(1) process with \( \rho = 0.99 \) and time \( t_{ij} \) shows time which was taken as the same set of occasions, \( \{ t_{ij} = j \) for \( i = 1, \ldots, m, j = 1, \ldots, n_i \}, k \) and \( d \) are selected as used in the Covid-19 data analysis.

We think two functions that the first is the piecewise linear continuous function \( g_l(t) = S(t) \) as an example of the ordinarily smooth nonparametric function and the second is the error function \( g_2(t) = \text{erf}(t) \) as an example of the supersmooth nonparametric function. Supposing that \( T \sim \text{Uniform}[0,1] \), \( h_n^{-1} = 1.2(\ln n)^{0.25} \) and using the Gaussian Kernel function

\[ K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}, \]

for the models have nonparametric part, we examine these models with the measurement error which has normal distribution \( U \sim N(0, 0.5) \).

For each choice of \( m, \gamma \) and \( g(t) \), the experiment is replicated 500 times by generating response variable and the EMSE for any estimator \( \hat{\beta} \) of \( \beta \) and the PMSE for any predictor \( \hat{b} \) of \( b \) are calculated, respectively, as

\[ \text{EMSE}(\hat{\beta}) = \frac{1}{500} \sum_{r=1}^{500} (\hat{\beta}_r - \beta)^T (\hat{\beta}_r - \beta), \]

\[ \text{PMSE}(\hat{b}) = \frac{1}{500} \sum_{r=1}^{500} (\hat{b}_r - b)^T (\hat{b}_r - b), \]

where the subscript \( r \) refers to the \( r \)th replication.

The simulation results are summarized in Tables 4 and 5. When we examine the results of Tables 4 and 5, we see that EMSE values of the modified Kernel ridge estimator and PMSE values of the modified Kernel ridge predictor are smaller than the others in all conditions. However, this superiority situation is more clearly be seen in large sample (for 600) and high correlation value (for 0.99). Additionally, we can also say that the superiority of the estimators/predictors over each other may vary depending on the selection of the biasing parameters.

6. Concluding Remarks

In this article, the modified Kernel ridge predictors have been studied with their MMSE comparisons under multicollinearity in PLMMeMs. To show the theoretical results, a Covid-19 analysis and a simulation study are given and these analyses demonstrate that although the modified Kernel ridge estimator is better than the Kernel ridge and Kernel Liu estimators, the superiority of the modified Kernel ridge estimator depends on the chosen values of the biasing parameters.

Ethics committee approval and conflict of interest statement

This article does not require ethics committee approval.

This article has no conflicts of interest with any individual or institution.

Acknowledgment

The authors would like to thank the editor, associate editor, and anonymous referees for their constructive comments.

Author Contribution Statement

All authors contributed equally to this manuscript and they have accepted responsibility for the entire content of this manuscript and approved its submission.

References


