

RESEARCH

Open Access



# Difference-based ridge-type estimator of parameters in restricted partial linear model with correlated errors

Jibo Wu<sup>1,2\*</sup>

\*Correspondence:  
linfen52@126.com  
<sup>2</sup> Key Laboratory of Group  
& Graph Theories  
and Applications,  
Chongqing University  
of Arts and Sciences,  
Chongqing 402160, China  
Full list of author information  
is available at the end of the  
article

## Abstract

In this article, a generalized difference-based ridge estimator is proposed for the vector parameter in a partial linear model when the errors are dependent. It is supposed that some additional linear constraints may hold to the whole parameter space. Its mean-squared error matrix is compared with the generalized restricted difference-based estimator. Finally, the performance of the new estimator is explained by a simulation study and a numerical example.

**Keywords:** Differencing estimator, Multicollinearity, Linear equality restrictions, Partial linear model, Ridge estimator

**Mathematics Subject Classification:** Primary 62J05, 62J07

## Background

Consider the following partially linear model

$$y_i = x_i' \beta + f(t_i) + \varepsilon_i, \quad i = 1, \dots, n \quad (1)$$

where  $y_i$ 's are observations,  $x_i' = (x_{i1}, \dots, x_{ip})$  and  $x_1, \dots, x_n$  are known  $p$ -dimensional with  $p \leq n$ .  $t_i$ 's are values of an extra univariate variable such as the time at which the observation is made,  $\beta = (\beta_1, \dots, \beta_p)'$  is an unknown parameter vector.  $f(\cdot)$  is an unknown smooth function, and  $\varepsilon_i$ 's are random errors supposed to be *i.i.d.*  $N(0, \sigma^2)$  distributed.

Use matrix vector notation, model (1) can be written as follows:

$$y = X\beta + f + \varepsilon \quad (2)$$

where  $y = (y_1, \dots, y_n)'$ ,  $X' = (x_1, \dots, x_n)$ ,  $f = [f(t_1), \dots, f(t_n)]'$  and  $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)'$ .

Since partially linear model has parametric and nonparametric components, and it is more flexible than linear model, many authors have been studied it, such as Ahn and Powell (1993), Wang et al. (2007).

In model (2), Yatchew (1997) mainly studied the estimation of the linear component and used differencing to eliminate bias induced from the presence of the nonparametric

component. Wang et al (2007) presented higher-order differences for optimal efficiency in estimating the linear part by using a special class of difference sequences.

In this article we will use the ridge regression concept that was presented by Hoerl and Kennard (1970) to overcome the multicollinearity in regression problem. Multicollinearity is denoted as the existence of nearly linear dependency among column vectors of the design matrix  $X$  in the linear model  $y = X\beta + \epsilon$ , where  $y$  is a  $n \times 1$  vector of observed responses,  $X$  is the observed matrix of independent variables of dimension  $n \times p$ , assumed to have full rank  $p$ ,  $\beta$  is an unknown parameter,  $\epsilon$  is an error vector with  $E(\epsilon) = 0, E(\epsilon\epsilon') = \sigma^2 I_p$ . Multicollinearity may lead to wide confidence intervals for individual parameters may produce estimates with wrong signs, etc.

The condition number is a measure of the presence of multicollinearity. The condition number of the matrix  $X$  present some information about the existence of multicollinearity, however it does not illustrate the structure of the linear dependency among the column vectors  $X_1, X_2, \dots, X_n$ . The best way of illustrating the existence and structure of multicollinearity is to see the eigenvalues of  $X'X$ . If  $X'X$  is ill-conditioned with a large condition number a ridge regression estimator can be used to estimate  $\beta$  [see e.g., Swamy et al. (1978); Sarkar (1992); Shi (2001); Zhong and Yang (2007); Zhang and Yang (2007); Tabakan and Akdeniz (2010); Akdeniz and Tabakan (2009); Roozbeh et al. (2010); Duran and Akdeniz (2012); Duran et al. (2012); Hu (2005) and Hu et al. (2015)]. In this paper, we will examine a biased estimation techniques to be followed when the matrix  $X'X$  appears to be ill-conditioned in the partial linear model. We suppose that the condition number of the parameteric component is large explain that a biased estimation procedure is desirable.

The rest of the paper is organized as follows. In section “[The model and differencing-based estimator](#)”, the model and differencing methodology are given. Section “[Generalized difference-based ridge estimator](#)” contains the definition of the generalized difference-based ridge estimator and some comparison results are given in section “[MSEM-superiority of the generalized difference-based ridge estimator  \$\hat{\beta}\_{GRD}\(k\)\$  over the the generalized restricted difference-based estimator  \$\hat{\beta}\_{GRD}\$](#) ”. The results from section “[MSEM-superiority of the generalized difference-based ridge estimator  \$\hat{\beta}\_{GRD}\(k\)\$  over the the generalized restricted difference-based estimator  \$\hat{\beta}\_{GRD}\$](#) ” are applied to a simulation study in section “[Exemplary simulation](#)” and a numerical example is given to illustrate the theoretical result in section “[A numerical example](#)”. Some conclusion remarks are given in section “[Conclusions](#)”.

### **The model and differencing-based estimator**

In this section we use a difference-based method to estimate the linear regression coefficient vector  $\beta$ . This method has been presented to remove the nonparametric component in the partially linear model by many authors (Yatchew 1997, 2000, 2003). Consider the following partially linear model

$$y = X\beta + f + \epsilon \quad (3)$$

where  $f$  is an unknown smooth function and has a bounded first derivative.

Now we present the differencing method. Let  $d = (d_0, \dots, d_m)$  be a  $m + 1$  vector, where  $m$  is the order of differencing and  $d_0, \dots, d_m$  are differencing weights satisfying the conditions

$$\sum_{j=0}^m d_j = 0, \quad \sum_{j=0}^m d_j^2 = 1 \tag{4}$$

Now, we denote the  $(n - m) \times n$  differencing matrix  $D$  whose elements satisfy Eq. (4) as follows:

$$D = \begin{pmatrix} d_0 & d_1 & \dots & d_m & 0 & 0 & \dots & 0 \\ 0 & d_0 & d_1 & \dots & d_m & 0 & \dots & 0 \\ \dots & \dots & & & & & & \\ \dots & \dots & & & & & & \\ \dots & \dots & & & & & & \\ 0 & 0 & \dots & d_1 & \dots & d_m & 0 & 0 \\ 0 & 0 & \dots & d_0 & d_1 & \dots & d_m & 0 \\ 0 & 0 & \dots & 0 & d_0 & d_1 & \dots & d_m \end{pmatrix} \tag{5}$$

This and related matrices are given, for example, in Yatchew (2003). Then we can use the differencing matrix to model (3), and this leads to direct estimation of the parametric effect. In particular, take

$$Dy = DX\beta + Df(t) + D\epsilon \tag{6}$$

Since the data have been reordered so that the  $X$ 's are close, the application of the differencing matrix  $D$  in model (6) can remove the nonparametric effect in large samples (Yatchew 2003). This ignores the presence of  $Df(t)$ . Thus, we may write Eq. (6) as

$$Dy \doteq DX\beta + D\epsilon \tag{7}$$

or

$$\tilde{y} \doteq \tilde{X}\beta + \tilde{\epsilon} \tag{8}$$

where  $\tilde{y} = Dy, \tilde{X} = DX$  and  $\tilde{\epsilon} = D\epsilon$ .

So, we can see that  $\tilde{\epsilon}$  is a  $n - m$  vector of disturbances distributed with  $E(\tilde{\epsilon}) = 0$  and  $E(\tilde{\epsilon}\tilde{\epsilon}') = \sigma^2 DD'$ .

For arbitrary differencing coefficients satisfying Eq. (8), Yatchew (1997) defines a simple differencing estimator of the parameter  $\beta$  in a partial linear model

$$\hat{\beta} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{y} \tag{9}$$

Hence, differencing allows one to perform inferences on  $\beta$  as if there were no nonparametric component  $f()$  in the model (3) (Yatchew 2003). Once  $\beta$  is estimated, a variety of nonparametric techniques could be applied to estimate  $f()$  as if  $\beta$  were known.

In order to account for the parameter  $\beta$  in Eq. (3), we propose the modified estimator of  $\sigma^2$ , defined as

$$\hat{\sigma}^2 = \frac{\tilde{y}'(I - P)\tilde{y}}{tr(D'(I - P)D)} \tag{10}$$

where  $P$  is the projection matrix and defined as

$$P = \tilde{X}(\tilde{X}'\tilde{X})^{-1}\tilde{X}' \tag{11}$$

**Generalized difference-based ridge estimator**

In this section we discuss the following partially linear model:

$$y = X\beta + f + \varepsilon \tag{12}$$

with  $E(\varepsilon) = 0$  and  $E(\varepsilon'\varepsilon) = \sigma^2V$ . So using the method we proposed in section “The model and differencing-based estimator”, we have  $\tilde{\varepsilon} = D\varepsilon$  is a  $(n - m)$ -vector of disturbances distributed with

$$E(\tilde{\varepsilon}) = 0 \quad \text{and} \quad E(\tilde{\varepsilon}'\tilde{\varepsilon}) = \sigma^2DVD' = \sigma^2V_D \tag{13}$$

where  $V_D = DVD' \neq I_{n-m}$  is a known  $(n - m) \times (n - m)$  symmetric positive definite matrix.

It is well known that adopting the linear model (12), the unbiased estimator of  $\beta$  is the following generalized difference-based estimator given by

$$\hat{\beta}_{GD} = C_D^{-1}\tilde{X}'V_D^{-1}\tilde{y}, \quad C_D = \tilde{X}'V_D^{-1}\tilde{X} \tag{14}$$

and the modified estimator  $\sigma^2$ ,

$$\hat{\sigma}^2 = \frac{\tilde{y}'V_D^{-1/2}(I - P)V_D^{-1/2}\tilde{y}}{\text{tr}(D'(I - P)D)} \tag{15}$$

where  $P$  is the projection matrix and defined as

$$P = V_D^{-1/2}\tilde{X}(\tilde{X}'V_D^{-1}\tilde{X})^{-1}\tilde{X}'V_D^{-1/2} \tag{16}$$

It is observed from Eq. (14) that the properties of the generalized difference-based estimator of  $\beta$  depends heavily on the characteristics of the information matrix  $C_D$ . If the  $C_D$  matrix is ill-conditioned, then the  $\hat{\beta}_{GD}$  leads to large sampling variances. Moreover, some of the regression coefficients may be statistically insignificant with wrong sign and meaningful statistical inference becomes difficult for the researcher. As a remedy, we consider the linear constraint

$$R\beta = 0 \tag{17}$$

for a given  $q \times p$  matrix  $R$  with rank  $q < p$ . Subject to the linear restriction (17), the generalized restricted difference-based estimator is given by

$$\hat{\beta}_{GRD} = \hat{\beta}_{GD} - C_D^{-1}R'(RC_D^{-1}R')^{-1}R\hat{\beta}_{GD} \tag{18}$$

Define  $W = C_D^{-1} - C_D^{-1}R'(RC_D^{-1}R')^{-1}RC_D^{-1}$ , we obtain

$$\hat{\beta}_{GRD} = W\tilde{X}'V_D^{-1}\tilde{y} \tag{19}$$

Now we propose a generalized difference-based ridge estimator, which is defined as

$$\hat{\beta}_{GRD}(k) = (kW + I)^{-1}\hat{\beta}_{GRD} \tag{20}$$

where  $W = C_D^{-1} - C_D^{-1}R'(RC_D^{-1}R')^{-1}RC_D^{-1}$  and  $k \geq 0$ .

Then, it is easy to see that  $\hat{\beta}_{GRD}$  and  $\hat{\beta}_{GRD}(k)$  are restricted with respect to  $R\beta = 0$ . It is also clear that for  $k = 0$ , we obtain  $\hat{\beta}_{GRD}(0) = \hat{\beta}_{GRD}$ .

**MSEM-superiority of the generalized difference-based ridge estimator  $\hat{\beta}_{GRD}(k)$  over the the generalized restricted difference-based estimator  $\hat{\beta}_{GRD}$**

In this section, our aim is to examine the difference of the mean squared error matrices (MSEM) of two estimators  $\hat{\beta}_{GRD}(k)$  and  $\hat{\beta}_{GRD}$ . Let  $b^*$  be an estimator of  $\beta$  in model  $Y = X\beta + \epsilon$ . The MSEM of  $b^*$  is defined as

$$MSEM(b^*, \beta) = E[(b^* - \beta)(b^* - \beta)'] \tag{21}$$

If we denote the covariance matrix of an estimator  $b^*$  by  $V(b^*)$ , then (21) is equivalent to

$$MSEM(b^*, \beta) = Var(b^*) + (bias(b^*))(bias(b^*))' \tag{22}$$

where  $bias(b^*) = E(b^*) - \beta$ . The scalar valued mean square error MSE is given by  $MSE(b^*, \beta) = E[(b^* - \beta)'(b^* - \beta)] = tr [MSEM(b^*, \beta)]$ .

Using Eq. (20), we obtain

$$E(\hat{\beta}_{GRD}(k)) = -k(kW + I)^{-1}W\beta \tag{23}$$

and

$$Var(\hat{\beta}_{GRD}(k)) = \sigma^2(kW + I)^{-1}W(kW + I)^{-1} \tag{24}$$

Thus,

$$Var(\hat{\beta}_{GRD}) = \sigma^2W \tag{25}$$

Then, the difference  $Var(\hat{\beta}_{GRD}) - Var(\hat{\beta}_{GRD}(k))$  can be expressed as

$$Var(\hat{\beta}_{GRD}) - Var(\hat{\beta}_{GRD}(k)) = \sigma^2(kW + I)^{-1}(k^2W^3 + 2kW^2)(kW + I)^{-1} \tag{26}$$

Since  $W$  is a nonnegative definite matrix [see Shi (2001)], we can conclude that  $Var(\hat{\beta}_{GRD}) - Var(\hat{\beta}_{GRD}(k))$  is a nonnegative definite matrix.

It is of interest to know under which conditions  $\hat{\beta}_{GRD}(k)$  is better than  $\hat{\beta}_{GRD}$ . For this, we investigate the difference  $\Delta = MSEM(\hat{\beta}_{GRD}, \beta) - MSEM(\hat{\beta}_{GRD}(k), \beta)$ , when  $\Delta$  is a nonnegative definite matrix,  $\hat{\beta}_{GRD}(k)$  is preferred to  $\hat{\beta}_{GRD}$ . Thus, for the MSE, of the generalized difference-based ridge estimator  $\hat{\beta}_{GRD}(k)$ , from (23) and (24), we obtain

$$\begin{aligned} MSEM(\hat{\beta}_{GRD}(k), \beta) &= Var(\hat{\beta}_{GRD}(k)) + (bias(\hat{\beta}_{GRD}(k)))(bias(\hat{\beta}_{GRD}(k)))' \\ &= \sigma^2(kW + I)^{-1}W(kW + I)^{-1} \\ &\quad + k^2(kW + I)^{-1}W\beta\beta'W(kW + I)^{-1} \end{aligned} \tag{27}$$

Since  $\hat{\beta}_{GRD}$  is an unbiased estimator for  $\beta$ , we have

$$MSEM(\hat{\beta}_{GRD}, \beta) = Var(\hat{\beta}_{GRD}) = \sigma^2W \tag{28}$$

Now from (27) and (28), we may write the difference  $\Delta = MSEM(\hat{\beta}_{GRD}, \beta) - MSEM(\hat{\beta}_{GRD}(k), \beta)$

$$\begin{aligned}
 \Delta &= \text{MSEM}(\hat{\beta}_{GRD}, \beta) - \text{MSEM}(\hat{\beta}_{GRD}(k), \beta) \\
 &= \sigma^2 W - \sigma^2(kW + I)^{-1} W(kW + I)^{-1} \\
 &\quad - k^2(kW + I)^{-1} W \beta \beta' W(kW + I)^{-1} \\
 &= (kW + I)^{-1}(\sigma^2 k^2 W^3 + 2\sigma^2 kW^2 - k^2 W \beta \beta' W)(kW + I)^{-1}
 \end{aligned}
 \tag{29}$$

Then, by (29),  $\Delta = \text{MSEM}(\hat{\beta}_{GRD}, \beta) - \text{MSEM}(\hat{\beta}_{GRD}(k), \beta) \geq 0$  if and only if  $\sigma^2 k^2 W^3 + 2\sigma^2 kW^2 - k^2 W \beta \beta' W \geq 0$ .

Then using Theorem (Farebrother 1976), we can conclude that if  $k > 0$ ,  $\beta' \left( W + \frac{2}{k} I \right)^{-1} \beta \leq \sigma^2$ , then  $\hat{\beta}_{GRD}(k)$  is preferred to  $\hat{\beta}_{GRD}$ .

**Theorem 4.1** Consider the two estimator  $\hat{\beta}_{GRD}$  and  $\hat{\beta}_{GRD}(k)$  of  $\beta$ . Then the biased estimator  $\hat{\beta}_{GRD}(k)$  is MSEM-superior over the  $\hat{\beta}_{GRD}$  if

$$\beta' \left( W + \frac{2}{k} I \right)^{-1} \beta \leq \sigma^2
 \tag{30}$$

is satisfied.

**Exemplary simulation**

In this section, we study the MSE of the proposed estimator. Our sampling experiment consists of different combinations of  $k$  and  $n$ . In this paper, we simulate the response from the following model:

$$y = x_{1i}\beta_1 + x_{2i}\beta_2 + x_{3i}\beta_3 + x_{4i}\beta_4 + f(t_i) + \epsilon_i
 \tag{31}$$

where  $i = 1, \dots, n$ ,  $\epsilon \sim (0, \sigma^2 V)$  which the elements of  $V$  is  $v_{ij} = (0.1)^{|i-j|}$  and  $\sigma = 0.1$ ,  $f(t_i) = \sqrt{t_i(1-t_i)} \sin \frac{2.1\pi}{t_i+0.05}$  that is called Doppler function for  $t_i = (i - 0.5)/n$  and for  $i = 1, \dots, n$ , the explanatory variables are generated by the following equation (Liu 2003):

$$x_{ij} = (1 - \gamma^2)z_{ij} + \gamma z_{i(p+1)}, \quad i = 1, \dots, n, j = 1, \dots, p$$

where  $z_{ij}$  and  $z_{i(p+1)}$  are independent standard normal pseudo-random numbers and  $\gamma$  is specified so that the correlation between any two explanatory variables is given by  $\gamma^2$ . In this paper, we consider  $n = 200$  and  $p = 4$ .

In this article we use a third-order differencing coefficients  $d_0 = 0.8502, d_1 = -0.3832, d_2 = -0.2809, d_3 = -0.1942$  in which  $m = 3$ . Now, we define the  $(200 - 3) \times 200$  differencing matrix as follows:

$$D = \begin{pmatrix} d_0 & d_1 & \dots & d_m & 0 & 0 & \dots & 0 \\ 0 & d_0 & d_1 & \dots & d_m & 0 & \dots & 0 \\ \dots & \dots & & & & & & \\ \dots & \dots & & & & & & \\ \dots & \dots & & & & & & \\ 0 & 0 & \dots & d_1 & \dots & d_m & 0 & 0 \\ 0 & 0 & \dots & d_0 & d_1 & \dots & d_m & 0 \\ 0 & 0 & \dots & 0 & d_0 & d_1 & \dots & d_m \end{pmatrix}
 \tag{32}$$

For the linear restriction (17), the  $R$  is given as follows:

$$R = (1, -2, -2, -2) \tag{33}$$

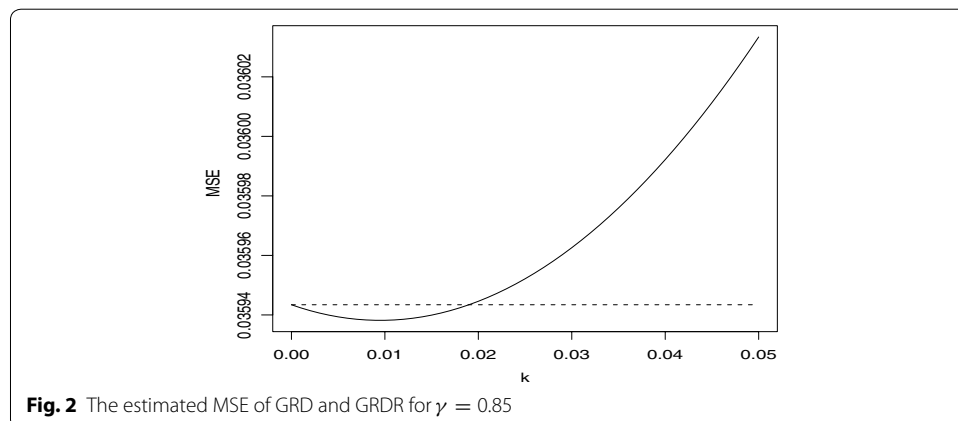
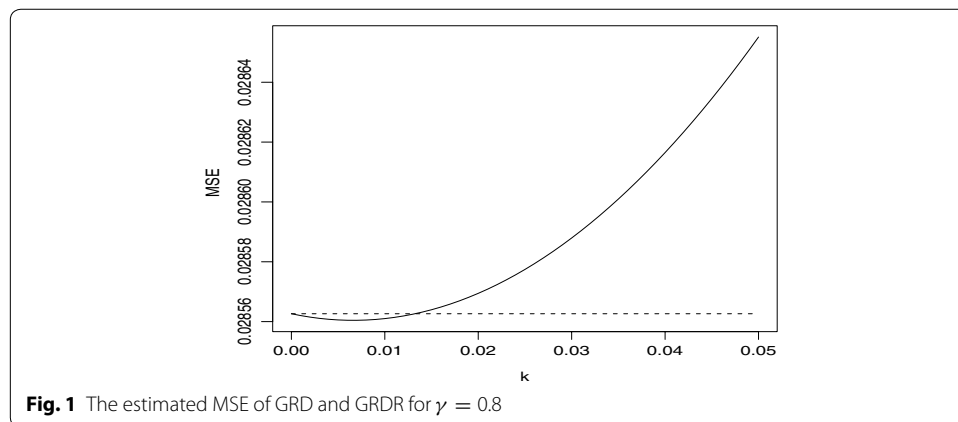
Let GRD define the generalized restricted difference-based estimator and GRDR define the generalized restricted difference-based ridge estimator and the estimated MSE of GRD and GRDR are given in Figs. 1, 2 and 3.

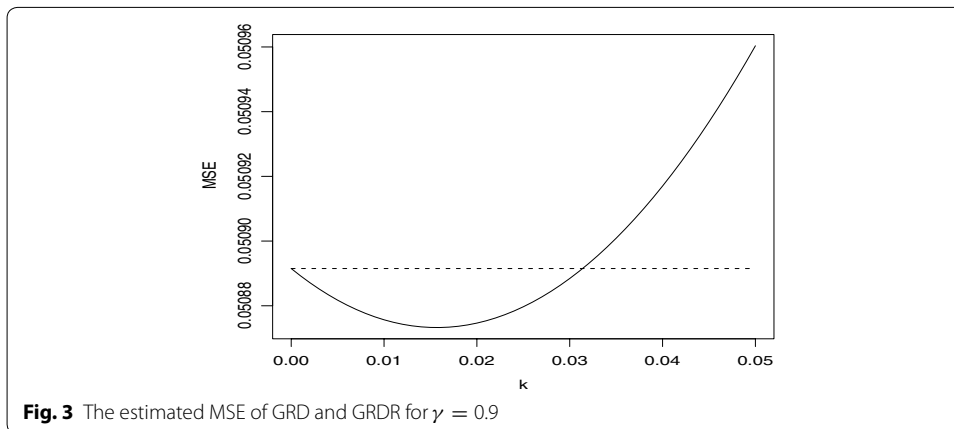
From Figs. 1 and 3, we see that as  $k$  is smaller, the new estimator is better than the generalized difference-based estimator in the mean squared error sense. And with the increase of the multicollinearity, the new estimator performs well.

### A numerical example

In this section, we consider a numerical example to explain the performance of theoretical results presented in “MSEM-superiority of the generalized difference-based ridge estimator  $\hat{\beta}_{GRD}(k)$  over the generalized restricted difference-based estimator  $\hat{\beta}_{GRD}$ ” section. The data was generated by Yatchew (2003), later discussed by Tabakan and Akdeniz (2010) and came from the survey of 81 municipal electricity distribution in Ontario, Canada, in 1993.

As we all know, the partial linear model is a simple semiparametric generalization of the Cobb–Douglas model. We consider a simple variant of the Cobb–Douglas model for the cost of distributing electricity





$$tc = f(cust) + \beta_1 wage + \beta_2 pcap + \beta_3 puc + \beta_4 kWh + \beta_5 life + \beta_6 lf + \beta_7 kmwire + \epsilon \tag{34}$$

for  $tc$  stands for the log of total cost per customer,  $cust$  denotes the log of the number of customers,  $wage$  defines the log of wage rate,  $pcap$  stands for the log price of capital,  $puc$  denotes a dummy variable for the public utility commissions that deliver additional services and may benefit from economy of scope,  $kWh$  defines the log of kilowatt hours per customer,  $life$  denotes the log of the remaining life of distribution assets,  $lf$  shows the log of the load factor and  $kmwire$  presents the log of kilometers of distribution wire per customer (Tabakan and Akdeniz 2010). It is easy to see that (34) contains both nonparametric effect and parametric effects.

Since  $V$  is seldom known, the estimation of  $V$  can be used. Trenkler (1984) gave some estimates of  $V$  as

$$V = \frac{1}{\rho^2 + 1} \begin{pmatrix} 1 + \rho^2 & \rho & 0 & \dots & 0 & 0 \\ \rho & 1 + \rho^2 & \rho & \dots & 0 & 0 \\ \cdot & \cdot & \dots & \cdot & \cdot & \cdot \\ \cdot & \cdot & \dots & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & 1 + \rho^2 & \rho \\ 0 & 0 & 0 & \dots & \rho & 1 + \rho^2 \end{pmatrix} \tag{35}$$

where the terms of the error vector are from the MA(1) process:

$$\epsilon_i = \mu_i + \rho\mu_{i-1}, \quad |\rho| < 1, \quad i = 1, 2, \dots, n$$

where  $\mu_i \sim N(0, \sigma_\mu^2)$ ,  $E(\mu_i\mu_j) = 0, i \neq j, \sigma^2 = \sigma_\mu^2(1 + \rho^2)$ .

For the linear restriction (17), the  $R$  is given as follows:

$$R = (1, -2, -2, -2, -2, -2, -2) \tag{36}$$

In this section, we study  $\rho = 0.3, \sigma_\mu^2 = 0.1$  and consider matrix  $V$  is estimated by (35). It is easy to compute the condition number is 2365.158, suggesting the presence of severe collinearity.



In this section we use the method which Hoerl and Kennard proposed to estimate  $k$ . Then we get  $MSE(\hat{\beta}_{GRD}(k), \beta) = 0.323$  and  $MSE(\hat{\beta}_{GRD}, \beta) = 0.597$ , that is to say the new estimator is better than restricted difference-based estimator.

Now we see theorem 21

$$\hat{\beta}' \left( W + \frac{2}{k} I \right)^{-1} \hat{\beta} = 0.0578 < \sigma^2 \quad (37)$$

That is to say our numerical example satisfied with theorem 4.1. This also means our method is meaningful in practice.

## Conclusions

In this article, we present a new generalized difference-based ridge estimator that can be applied in the presence of multicollinearity in a partial linear model. Its MSE is compared analytically with the generalized restricted difference-based estimator. It is shown that for small values of the ridge parameter  $k$ , the new estimator is MSEM-superior to the generalized restricted difference-based estimator over an interval depending on the design points and the unknown parameter.

## Author details

<sup>1</sup> School of Mathematics and Finances, Chongqing University of Arts and Sciences, Chongqing 402160, China. <sup>2</sup> Key Laboratory of Group & Graph Theories and Applications, Chongqing University of Arts and Sciences, Chongqing 402160, China.

## Acknowledgements

The author would like to thank the anonymous referees and the Associate Editor for their constructed suggestions which significantly improved the presentation of the article. This work was supported by the Scientific and Technological Research Program of Chongqing Municipal Education Commission (no. KJ1501114), the Natural Science Foundation Project of CQ CSTC (cstc2015cyjA00001), and the Scientific Research Foundation of Chongqing University of Arts and Sciences (no: R2013SC12, Y2015SC47).

## Competing interests

The author declares that they have no competing interests.

Received: 7 December 2015 Accepted: 12 February 2016

Published online: 25 February 2016

## References

- Ahn H, Powell J (1993) Semiparametric estimation of censored selection models with a nonparametric selection mechanism. *J Econom* 58:3–29
- Akdeniz F, Tabakan G (2009) Restricted ridge estimators of the parameters in semiparametric regression model. *Commun Stat Theory Methods* 38(11):1852–1869
- Duran EA, Akdeniz F (2012) Efficiency of the modified jackknifed Liu-type estimator. *Stat Pap* 53(2):265–280
- Duran EA, Härdle WK, Osipenko M (2012) Difference based ridge and Liu type estimators in semiparametric regression models. *J Multivar Anal* 105(1):164–175
- Farebrother RW (1976) Further results on the mean square error of ridge regression. *J R Stat Soc B* 38:248–250
- Hu HC, Yang Y, Pan X (2015) Asymptotic normality of DHD estimators in a partially linear model. *Stat Papers*. doi:10.1007/s00362-015-0666-2
- Hu HC (2005) Ridge estimation of a semiparametric regression model. *J Comput Appl Math* 176:215–222
- Hoerl AE, Kennard RW (1970) Ridge regression: biased estimation for nonorthogonal problems. *Technometrics* 12:55–67
- Liu KJ (2003) Using Liu type estimator to combat multicollinearity. *Commun Stat Theory Methods* 32(5):1009–1020
- Ruppert D, Wand MP, Holst U, Hossjer O (1997) Local polynomial variance-function estimation. *Technometrics* 39:262–272
- Roohbeh M, Arashi M, Niroumand HA (2010) Semiparametric ridge regression approach in partially linear models. *Commun Stat Simul Comput* 39:449–460
- Swamy PAVB, Mehta JS, Rapport PN (1978) Two methods of evaluating Hoerl and Kennard's ridge regression. *Commun Stat* 12:1133–1155
- Shi JH (2001) The conditional ridge-type estimation of regression coefficient in restricted linear regression model. *J Shanxi Teach Univ Natural Sci Ed* 15:10C16

- Sarkar N (1992) A new estimator combining the ridge regression and the restricted least squares methods of estimation. *Commun Stat Theory Methods* 21:1987–2000
- Tabakan G, Akdeniz F (2010) Difference-based ridge estimator of parameters in partial linear model. *Stat Pap* 51:357–368
- Trenkler G (1984) On the performance of biased estimators in the linear regression model with correlated or heteroscedastic errors. *J Econom* 25:179–190
- Wang L, Brown LD, Cai TT (2007) A difference based approach to semiparametric partial linear model. Department of Statistics, The Wharton School University of Pennsylvania, Pennsylvania, Technical report
- Yatchew A (1997) An elementary estimator of the partial linear model. *Econ Lett* 57:135–143. Additional examples contained in *Econ Lett* (1998) 59, 403–405
- Yatchew A (2000) Scale economies in electricity distribution: a semiparametric analysis. *J Appl Econom* 15(2):187–210
- Yatchew A (2003) *Semiparametric regression for the applied econometrician*. Cambridge University Press, Cambridge 123
- Zhong Z, Yang H (2007) Ridge estimation to the restricted linear model. *Commun Stat Theory Methods* 36:2099–2115
- Zhang CM, Yang H (2007) The conditional ridge-type estimation in singular linear model with linear equality restriction. *Statistics* 41(6):485–494

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

---

Submit your next manuscript at ▶ [springeropen.com](http://springeropen.com)

---