

RESEARCH

Open Access



Admissibility of simultaneous prediction for actual and average values in finite population

Chao Bai^{1,2*}  and Haiqi Li¹

*Correspondence:
bbc144@163.com

¹School of Finance and Statistics,
Hunan University, Changsha, China

²College of Science, Central South
University of Forestry and
Technology, Changsha, China

Abstract

This paper studies the admissibility of simultaneous prediction of actual and average values of the regressand in the generalized linear regression model under the quadratic loss function. Necessary and sufficient conditions are derived for the simultaneous prediction to be admissible in classes of homogeneous and nonhomogeneous linear predictors, respectively.

MSC: 62C15; 62M20; 62J12; 15A39

Keywords: Admissibility; Simultaneous prediction; Quadratic loss function; Generalized linear regression model

1 Introduction

We begin this paper with some notation first. For an $m \times n$ matrix \mathbf{A} , the symbols $\mathcal{M}(\mathbf{A})$, \mathbf{A}' , \mathbf{A}^- and \mathbf{A}^+ denote the column space, the transpose, the generalized inverse and Moore–Penrose inverse of \mathbf{A} , respectively. For squared matrices \mathbf{B} and \mathbf{C} , $\mathbf{B} \geq \mathbf{C}$ means that $\mathbf{B} - \mathbf{C}$ is a symmetric nonnegative definite matrix and $\mathbf{B} - \mathbf{C} \geq \mathbf{0}$. Let $\text{rk}(\mathbf{A})$ be the rank of \mathbf{A} and $\text{tr}(\mathbf{A})$ be the trace of \mathbf{A} when \mathbf{A} is a squared matrix. \mathbf{I} is an identity matrix with an appropriate order and \mathbb{R}^n denotes the n dimensional vector set.

Consider the following generalized linear regression model:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad E(\boldsymbol{\varepsilon}) = \mathbf{0}, \quad \text{Cov}(\boldsymbol{\varepsilon}) = \boldsymbol{\Sigma}, \quad (1)$$

where \mathbf{y} is the $n \times 1$ observable vector of regressand, \mathbf{X} is the $n \times p$ observable matrix of regressor, $\boldsymbol{\beta}$ is a $p \times 1$ unknown vector of regression coefficient, $\boldsymbol{\varepsilon}$ is the $n \times 1$ vector of disturbances and $\boldsymbol{\Sigma} \geq \mathbf{0}$. Suppose $\text{rk}(\mathbf{X}) \leq p$.

Given the matrix of regressor \mathbf{X}_0 (which is correlated with new observations), the relationship between the unobservable random vector \mathbf{y}_0 and \mathbf{X}_0 is

$$\mathbf{y}_0 = \mathbf{X}_0\boldsymbol{\beta} + \boldsymbol{\varepsilon}_0, \quad E(\boldsymbol{\varepsilon}_0) = \mathbf{0}, \quad \text{Cov}(\boldsymbol{\varepsilon}_0) = \boldsymbol{\Sigma}_0, \quad (2)$$

where \mathbf{y} is the $m \times 1$ vector of the regressand to be predicted, \mathbf{X}_0 is the $m \times p$ matrix of prediction regressor, $\boldsymbol{\beta}$ is the same as that in model (1), $\boldsymbol{\varepsilon}_0$ is the $m \times 1$ vector of prediction

disturbances and $\Sigma_0 \geq \mathbf{0}$. Assume \mathbf{y} and \mathbf{y}_0 are correlated and $\text{Cov}(\boldsymbol{\varepsilon}_0, \boldsymbol{\varepsilon}') = \mathbf{V}$. Suppose the finite population is composed of \mathbf{y} and \mathbf{y}_0 . Thus, combining models of (1) and (2), we have

$$\mathbf{y}_T = \mathbf{X}_T \boldsymbol{\beta} + \boldsymbol{\varepsilon}_T, \tag{3}$$

where

$$\mathbf{y}_T = \begin{pmatrix} \mathbf{y} \\ \mathbf{y}_0 \end{pmatrix}, \quad \mathbf{X}_T = \begin{pmatrix} \mathbf{X} \\ \mathbf{X}_0 \end{pmatrix}, \quad \boldsymbol{\varepsilon}_T = \begin{pmatrix} \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon}_0 \end{pmatrix},$$

$$E\boldsymbol{\varepsilon}_T = \mathbf{0}, \quad \text{Cov } \boldsymbol{\varepsilon}_T = \text{Cov} \begin{pmatrix} \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon}_0 \end{pmatrix} = \begin{pmatrix} \Sigma & \mathbf{V}' \\ \mathbf{V} & \Sigma_0 \end{pmatrix} = \Sigma_T.$$

For the prediction in model (3), [1] obtained the best linear unbiased predictor (BLUP) of \mathbf{y}_0 . [2] considered the optimal Stein-rule prediction. Reference [3] investigated the admissibility of linear predictors with inequality constraints under the quadratic loss function. [4] reviewed the existing theory of minimum mean squared error (MSE) predictors and made an extension based on the principle of equivariance. [5] derived the BLUP and the admissible predictor under the matrix loss function. Under the MSE loss function, the optimal predictor of \mathbf{y}_0 is the conditional expectation $E(\mathbf{y}_0|\mathbf{X}_0) = \mathbf{X}_0\boldsymbol{\beta}$, which relates naturally to the plug-in estimators of $\boldsymbol{\beta}$. [6] proposed the simple projection predictor (SPP) of $\mathbf{X}_0\boldsymbol{\beta}$ by plugging in the best linear unbiased estimator (BLUE) of $\boldsymbol{\beta}$. The plug-in approach spawned a large literature for the derivation of combined prediction; see [7, 8], etc.

Generally, predictions are investigated either for \mathbf{y}_0 or for $E\mathbf{y}_0$ at a time. However, sometimes in the fields of medicine and economics, people would like to know the actual value of \mathbf{y}_0 and its average value $E\mathbf{y}_0$ simultaneously. For example, in the financial markets, some investors may want to know the actual profit while others would be more interested in the average profit. Therefore, in order to meet different requirements, the market manager should acquire both the prediction of the actual profit and the prediction of the average profit at the same time, and can assign different weights to each prediction to provide a more comprehensive combined prediction of the profit. Under these circumstances, we consider the following target function:

$$\boldsymbol{\delta} = \lambda\mathbf{y}_0 + (1 - \lambda)E\mathbf{y}_0,$$

where $\lambda \in [0, 1]$ is a non-stochastic scalar representing the preference to the prediction of the actual and average values of the studied variable. A prediction of $\boldsymbol{\delta}$ is the simultaneous prediction of the actual and average values of \mathbf{y}_0 . Note that $\boldsymbol{\delta} = \mathbf{y}_0$ if $\lambda = 1$ and $\boldsymbol{\delta} = E\mathbf{y}_0$ if $\lambda = 0$, which means the studied $\boldsymbol{\delta}$ is more sensitive and inclusive.

Studies on the simultaneous prediction of the actual and average values of the studied variable (namely prediction of $\boldsymbol{\delta}$) have been carried out in the literature from various perspective. The properties of the predictors by plugging in Stein-rule estimators have been concerned by [9] and [10]. [11] investigated the Stein-rule prediction for $\boldsymbol{\delta}$ in linear regression model when the error covariance matrix was positive definite yet unknown. References [12, 13] and [14] considered predictors for $\boldsymbol{\delta}$ in linear regression models with stochastic or non-stochastic linear constraints on the regression coefficients. The issues of simultaneous prediction in measurement error models have been addressed in [15] and

[16], [17] considered a matrix multiple of the classical forecast vector for the simultaneous prediction of and discussed the performance properties.

This paper aims to study the admissibility of simultaneous prediction of actual and average values of the unobserved regressand in finite population under the quadratic loss function. Admissibility is an interesting problem in statistical theory and received much attention. [18, 19] and [20] discussed the admissibility of predictions of \mathbf{y}_T . [21, 22] and [23] studied the admissibility of estimations of $\boldsymbol{\beta}$. We discuss the admissible predictors of δ in classes of homogeneous and nonhomogeneous linear predictors, respectively. Necessary and sufficient conditions for the simultaneous prediction to be admissible are provided.

The rest of this paper is organized as follows. In Sect. 2, we give some preliminaries. In Sect. 3, we obtain the homogeneous linear admissible simultaneous predictors for the actual and average values of the unobserved regressand. In Sect. 4, we derive the necessary and sufficient conditions for linear simultaneous prediction to be admissible in class of nonhomogeneous linear predictors. Concluding remarks are placed in Sect. 5.

2 Preliminaries

Suppose d is the predictor of δ and denote $R(d; \beta)$ as the risk of d under the quadratic loss function, then for model (3)

$$R(d; \beta) = E(d - \delta)'(d - \delta) = E[d - \lambda y_0 - (1 - \lambda)X_0\beta]'[d - \lambda y_0 - (1 - \lambda)X_0\beta].$$

Denote the classes of homogeneous and nonhomogeneous linear predictors, respectively, by

$$\begin{aligned} \mathcal{LH} &= \{\mathbf{C}\mathbf{y} \mid \mathbf{C} \text{ is an } m \times n \text{ matrix}\}, \quad \text{and} \\ \mathcal{LN} &= \{\mathbf{C}\mathbf{y} + \mathbf{u} \mid \mathbf{C} \text{ is an } m \times n \text{ matrix and } \mathbf{u} \text{ is an } m \times 1 \\ &\quad \text{nonstochastic vector and } \mathbf{u} \neq \mathbf{0}\}. \end{aligned}$$

The nonhomogeneous linear predictor is actually an adjustment of the homogeneous linear predictor. We study the admissibility of the prediction of δ in \mathcal{LH} and \mathcal{LN} . Before the discussion begins, we first present some important preliminaries and basic results.

Definition 2.1 A predictor \mathbf{d} is said to be admissible for δ under the quadratic loss function, denoted $\mathbf{d} \sim \delta$, iff there exists no other predictor \mathbf{d}^* such that $R[\mathbf{d}^*; \boldsymbol{\beta}] \leq R[\mathbf{d}; \boldsymbol{\beta}]$ with strict inequality holding at least at one value of $\boldsymbol{\beta}$. \mathbf{d} is called better than \mathbf{d}^* iff $R[\mathbf{d}; \boldsymbol{\beta}] \leq R[\mathbf{d}^*; \boldsymbol{\beta}]$ with strict inequality holding at least at one value of $\boldsymbol{\beta}$.

Definition 2.2 \mathbf{d} is an unbiased predictor of δ if $E\mathbf{d} = E\delta$.

Lemma 2.1 ([24]) *If $\mathbf{C}\mathbf{y} \sim \mathbf{y}_0$, then for, any $l \times m$ matrix \mathbf{L} , $\mathbf{L}\mathbf{C}\mathbf{y} \sim \mathbf{L}\mathbf{y}_0$.*

Lemma 2.2 ([25]) *If the square matrix \mathbf{B} is not symmetrical, then there exists an orthogonal matrix \mathbf{P} such that $\text{tr}(\mathbf{PB}) = \text{tr}(\mathbf{B})$.*

Lemma 2.3 *Suppose $\mathbf{C}\mathbf{y}$ is an arbitrary predictor of δ in model (3). Let*

$$\tilde{\mathbf{C}} = \mathbf{C}\mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+ + \lambda\mathbf{V}\mathbf{T}^+[I - \mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+],$$

where $\mathbf{T} = \mathbf{X}\mathbf{X}' + \Sigma$. Then under the quadratic loss function

$$R(\mathbf{C}\mathbf{y}; \beta) \geq R(\tilde{\mathbf{C}}\mathbf{y}; \beta)$$

for every $\beta \in \mathbb{R}^p$ and the equation holds if and only if either of the following two conditions holds:

- (1) $\mathbf{C}\Sigma = \tilde{\mathbf{C}}\Sigma$;
- (2) $\mathcal{M}(\Sigma\mathbf{C}' - \lambda\mathbf{V}') \subseteq \mathcal{M}(\mathbf{X})$.

Proof By direct calculation,

$$\begin{aligned} R(\mathbf{C}\mathbf{y}; \beta) &= E[\mathbf{C}\mathbf{y} - \lambda\mathbf{y}_0 - (1 - \lambda)\mathbf{X}_0\beta]'[\mathbf{C}\mathbf{y} - \lambda\mathbf{y}_0 - (1 - \lambda)\mathbf{X}_0\beta] \\ &= \text{tr}(\mathbf{C}\Sigma\mathbf{C}' + \lambda^2\Sigma_0 - 2\lambda\mathbf{C}\mathbf{V}') + \beta'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)\beta \\ &= E\{\lambda(\mathbf{C}\mathbf{y} - \mathbf{y}_0)'(\mathbf{C}\mathbf{y} - \mathbf{y}_0) + (1 - \lambda)(\mathbf{C}\mathbf{y} - \mathbf{X}_0\beta)'(\mathbf{C}\mathbf{y} - \mathbf{X}_0\beta)\} + (\lambda^2 - \lambda) \text{tr} \Sigma_0 \\ &= E\{\lambda(\mathbf{C}\mathbf{y} - \tilde{\mathbf{C}}\mathbf{y} + \tilde{\mathbf{C}}\mathbf{y} - \mathbf{y}_0)'(\mathbf{C}\mathbf{y} - \tilde{\mathbf{C}}\mathbf{y} + \tilde{\mathbf{C}}\mathbf{y} - \mathbf{y}_0) \\ &\quad + (1 - \lambda)(\mathbf{C}\mathbf{y} - \tilde{\mathbf{C}}\mathbf{y} + \tilde{\mathbf{C}}\mathbf{y} - \mathbf{X}_0\beta)'(\mathbf{C}\mathbf{y} - \tilde{\mathbf{C}}\mathbf{y} + \tilde{\mathbf{C}}\mathbf{y} - \mathbf{X}_0\beta)\} + (\lambda^2 - \lambda) \text{tr} \Sigma_0 \\ &= E\{(\mathbf{C}\mathbf{y} - \tilde{\mathbf{C}}\mathbf{y})'(\mathbf{C}\mathbf{y} - \tilde{\mathbf{C}}\mathbf{y}) + \lambda(\tilde{\mathbf{C}}\mathbf{y} - \mathbf{y}_0)'(\tilde{\mathbf{C}}\mathbf{y} - \mathbf{y}_0) + (1 - \lambda)(\tilde{\mathbf{C}}\mathbf{y} - \mathbf{X}_0\beta)'(\tilde{\mathbf{C}}\mathbf{y} - \mathbf{X}_0\beta) \\ &\quad + 2\lambda(\mathbf{C}\mathbf{y} - \tilde{\mathbf{C}}\mathbf{y})'(\tilde{\mathbf{C}}\mathbf{y} - \mathbf{y}_0) + 2(1 - \lambda)(\mathbf{C}\mathbf{y} - \tilde{\mathbf{C}}\mathbf{y})'(\tilde{\mathbf{C}}\mathbf{y} - \mathbf{X}_0\beta)\} + (\lambda^2 - \lambda) \text{tr} \Sigma_0 \\ &= E(\mathbf{C}\mathbf{y} - \tilde{\mathbf{C}}\mathbf{y})'(\mathbf{C}\mathbf{y} - \tilde{\mathbf{C}}\mathbf{y}) + R(\tilde{\mathbf{C}}\mathbf{y}) \\ &\quad + E\{2\lambda(\mathbf{C}\mathbf{y} - \tilde{\mathbf{C}}\mathbf{y})'(\tilde{\mathbf{C}}\mathbf{y} - \mathbf{y}_0) + 2(1 - \lambda)(\mathbf{C}\mathbf{y} - \tilde{\mathbf{C}}\mathbf{y})'(\tilde{\mathbf{C}}\mathbf{y} - \mathbf{X}_0\beta)\}. \end{aligned}$$

Note that

$$\begin{aligned} E\mathbf{y}\mathbf{y}' &= \Sigma + \mathbf{X}\beta\beta'\mathbf{X}' = \mathbf{T} - \mathbf{X}\mathbf{X}' + \mathbf{X}\beta\beta'\mathbf{X}', \\ E\mathbf{y}_0\mathbf{y}' &= \mathbf{V} + \mathbf{X}_0\beta\beta'\mathbf{X}', \\ (\mathbf{C} - \tilde{\mathbf{C}})\mathbf{X} &= (\mathbf{C} - \lambda\mathbf{V}\mathbf{T}^+)[I - \mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+]\mathbf{X} = \mathbf{0}, \end{aligned}$$

then

$$\begin{aligned} &E\{2\lambda(\mathbf{C}\mathbf{y} - \tilde{\mathbf{C}}\mathbf{y})'(\tilde{\mathbf{C}}\mathbf{y} - \mathbf{y}_0) + 2(1 - \lambda)(\mathbf{C}\mathbf{y} - \tilde{\mathbf{C}}\mathbf{y})'(\tilde{\mathbf{C}}\mathbf{y} - \mathbf{X}_0\beta)\} \\ &= E\{2(\mathbf{C}\mathbf{y} - \tilde{\mathbf{C}}\mathbf{y})'(\mathbf{C}\mathbf{y} - \mathbf{y}_0) + 2(1 - \lambda)(\mathbf{C}\mathbf{y} - \tilde{\mathbf{C}}\mathbf{y})'(\mathbf{y}_0 - \mathbf{X}_0\beta)\} \\ &= \text{tr}\{[\mathbf{C}\mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}' - \lambda\mathbf{V}\mathbf{T}^+(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}' - \mathbf{C}\mathbf{X}\mathbf{X}' + \mathbf{C}\mathbf{X}\beta\beta'\mathbf{X}' - \mathbf{X}_0\beta\beta'\mathbf{X}'](\mathbf{C} - \tilde{\mathbf{C}})'\} \\ &= \text{tr}\{[\mathbf{C}\mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1} - \lambda\mathbf{V}\mathbf{T}^+(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1} - \mathbf{C}\mathbf{X} + \mathbf{C}\mathbf{X}\beta\beta' - \mathbf{X}_0\beta\beta'][(\mathbf{C} - \tilde{\mathbf{C}})\mathbf{X}']\} \\ &= 0. \end{aligned}$$

Thus, we have

$$R(\mathbf{C}\mathbf{y}; \boldsymbol{\beta}) = E(\mathbf{C}\mathbf{y} - \tilde{\mathbf{C}}\mathbf{y})'(\mathbf{C}\mathbf{y} - \tilde{\mathbf{C}}\mathbf{y}) + R(\tilde{\mathbf{C}}\mathbf{y}; \boldsymbol{\beta}) \geq R(\tilde{\mathbf{C}}\mathbf{y}; \boldsymbol{\beta}),$$

and the equation holds for every $\boldsymbol{\beta} \in \mathbb{R}^p$ if and only if

$$E(\mathbf{C}\mathbf{y} - \tilde{\mathbf{C}}\mathbf{y})'(\mathbf{C}\mathbf{y} - \tilde{\mathbf{C}}\mathbf{y}) = 0.$$

Since

$$\begin{aligned} E(\mathbf{C}\mathbf{y} - \tilde{\mathbf{C}}\mathbf{y})'(\mathbf{C}\mathbf{y} - \tilde{\mathbf{C}}\mathbf{y}) &= 0 \\ \Leftrightarrow [(\mathbf{C} - \tilde{\mathbf{C}})\mathbf{X}\boldsymbol{\beta}]'[(\mathbf{C} - \tilde{\mathbf{C}})\mathbf{X}\boldsymbol{\beta}] + \text{tr}(\mathbf{C} - \tilde{\mathbf{C}})' \boldsymbol{\Sigma} (\mathbf{C} - \tilde{\mathbf{C}}) &= 0 \\ \Leftrightarrow (\mathbf{C} - \tilde{\mathbf{C}}) \boldsymbol{\Sigma}^{\frac{1}{2}} &= \mathbf{0} \\ \Leftrightarrow (\mathbf{C} - \tilde{\mathbf{C}}) \boldsymbol{\Sigma} &= \mathbf{0}, \end{aligned}$$

the equation holds for every $\boldsymbol{\beta} \in \mathbb{R}^p$ if and only if

$$\mathbf{C}\boldsymbol{\Sigma} = \tilde{\mathbf{C}}\boldsymbol{\Sigma}.$$

Now we prove conditions (1) and (2) are equivalent. First, as $\mathbf{C}\boldsymbol{\Sigma} = \tilde{\mathbf{C}}\boldsymbol{\Sigma}$,

$$\begin{aligned} \boldsymbol{\Sigma}\mathbf{C}' &= \boldsymbol{\Sigma}\tilde{\mathbf{C}}' \\ &= \boldsymbol{\Sigma}\mathbf{T}^+\mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{C}' + \lambda\boldsymbol{\Sigma}[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+]' \mathbf{T}^+\mathbf{V}' \\ &= (\mathbf{T} - \mathbf{X}\mathbf{X}')\mathbf{T}^+\mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{C}' \\ &\quad + \lambda(\mathbf{T} - \mathbf{X}\mathbf{X}')[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+]' \mathbf{T}^+\mathbf{V}' \\ &= [\mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}' - \mathbf{X}\mathbf{X}'] + \lambda\mathbf{V}' - \lambda\mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{V}' \Rightarrow \\ \boldsymbol{\Sigma}\mathbf{C}' - \lambda\mathbf{V}' &= \mathbf{X}[(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{C}' - \mathbf{X}'\mathbf{C}' - \lambda(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{V}'], \end{aligned}$$

which shows that $\mathcal{M}(\boldsymbol{\Sigma}\mathbf{C}' - \lambda\mathbf{V}') \subseteq \mathcal{M}(\mathbf{X})$.

Second, as $\mathcal{M}(\boldsymbol{\Sigma}\mathbf{C}' - \lambda\mathbf{V}') \subseteq \mathcal{M}(\mathbf{X})$, there exists a matrix \mathbf{H} such that

$$\boldsymbol{\Sigma}\mathbf{C}' - \lambda\mathbf{V}' = \mathbf{X}\mathbf{H}, \quad \text{or} \quad \lambda\mathbf{V}' = \mathbf{C}\boldsymbol{\Sigma} + \mathbf{H}'\mathbf{X}'.$$

Then

$$\begin{aligned} \tilde{\mathbf{C}}\boldsymbol{\Sigma} &= \mathbf{C}\mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+(\mathbf{T} - \mathbf{X}\mathbf{X}') + \lambda\mathbf{V}\mathbf{T}^+[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+](\mathbf{T} - \mathbf{X}\mathbf{X}') \\ &= (\mathbf{C}\boldsymbol{\Sigma} - \mathbf{H}'\mathbf{X}')\mathbf{T}^+[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+](\mathbf{T} - \mathbf{X}\mathbf{X}') \\ &\quad + \mathbf{C}\mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}' - \mathbf{C}\mathbf{X}\mathbf{X}' \\ &= \mathbf{C}\mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}' - \mathbf{C}\mathbf{X}\mathbf{X}' + \mathbf{C}\boldsymbol{\Sigma} - \mathbf{C}(\mathbf{T} - \mathbf{X}\mathbf{X}')\mathbf{T}^+\mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}' \\ &= \mathbf{C}\boldsymbol{\Sigma}. \end{aligned}$$

Therefore, we complete the proof. □

Lemma 2.4 *If $C\Sigma = \tilde{C}\tilde{\Sigma}$, then the risk of Cy under the quadratic loss function is*

$$R(Cy; \beta) = R(\tilde{C}y; \beta) = \text{tr}[CXQX'C' + \lambda^2\Sigma_0 - 2\lambda CX(X'T^+X)^-X'T^+V' - \lambda^2V(T^+ - T^+X(X'T^+X)^-X'T^+)V'] + \beta'(CX - X_0)'(CX - X_0)\beta,$$

where $Q = (X'T^+X)^- - I = (X'T^+X)^-X'T^+\Sigma T^+X(X'T^+X)^-$.

Proof From Lemma 2.3, when $C\Sigma = \tilde{C}\tilde{\Sigma}$,

$$R(Cy; \beta) = R(\tilde{C}y; \beta) = \text{tr}(\tilde{C}\tilde{\Sigma}\tilde{C}' + \lambda^2\Sigma_0 - 2\lambda\tilde{C}\tilde{V}') + \beta'(\tilde{C}X - X_0)'(\tilde{C}X - X_0)\beta.$$

Since $\tilde{C} = CX(X'T^+X)^-X'T^+ + \lambda VT^+[I - X(X'T^+X)^-X'T^+]$ and $\tilde{\Sigma} = T - XX'$,

$$\begin{aligned} \tilde{C}\tilde{\Sigma}\tilde{C}' &= CXQX'C' + \lambda^2V(T^+ - T^+X(X'T^+X)^-X'T^+)V', \\ Q &= (X'T^+X)^- - I = (X'T^+X)^-X'T^+\Sigma T^+X(X'T^+X)^-, \\ \tilde{C}\tilde{V}' &= CX(X'T^+X)^-X'T^+V' + \lambda VT^+[I - X(X'T^+X)^-X'T^+]V', \\ \tilde{C}X &= CX. \end{aligned}$$

The lemma is easily proved by substitution of these equations. □

3 Admissibility of homogeneous linear predictors

In this section, we derive the necessary and sufficient conditions for the admissibility of simultaneous prediction in class of the homogeneous linear predictors. The best linear unbiased predictor of δ is obtained. Examples are presented to give some admissible predictors.

Theorem 3.1 *Let $l' = q'C$, where C is a matrix and l, q are vectors with appropriate dimensions. If $Cy \sim \delta$ under the quadratic loss function, then*

$$l'XQX'l \leq \lambda(l'X - q'X_0)(X'T^+X)^-X'T^+V'q + l'XQX'q,$$

where $Q = (X'T^+X)^- - I = (X'T^+X)^-X'T^+\Sigma T^+X(X'T^+X)^-$.

Proof Since $Cy \sim \delta$ and $l' = q'C$, by Lemma 2.1, $l'y \sim q'\delta$ for any $1 \times m$ vector q . Suppose k be a real constant and $0 < k < 1$. Let

$$l'_k = [kl'X + (1 - k)q'X_0](X'T^+X)^-X'T^+ + \lambda q'VT^+[I - X(X'T^+X)^-X'T^+].$$

The risk of l'_ky is

$$\begin{aligned} R(l'_ky; \beta) &= [kl'X + (1 - k)q'X_0]Q[kl'X + (1 - k)q'X_0]' + \lambda^2q'\Sigma_0q \\ &\quad - 2\lambda[kl'X + (1 - k)q'X_0](X'T^+X)^-X'T^+V'q \\ &\quad - \lambda^2q'VT^+[I - X(X'T^+X)^-X'T^+]V'q + k^2[\beta'(l'X - q'X_0)]^2. \end{aligned}$$

Since $l'y \sim q'y_0$ and by Lemma 2.3, Lemma 2.4 and $l' = q'C$, the risk of $l'y$ is

$$R(l'y; \beta) = l'XQX'l + \lambda^2 q' \Sigma_0 q - 2\lambda l'X(X'T^+X)^- X'T^+V'q - \lambda^2 q'V(T^+ - T^+X(X'T^+X)^- X'T^+)V'q + [\beta'(l'X - q'X_0)]^2,$$

and for any $k \in (0, 1)$, $R(l'y; \beta) \leq R(l'_k y; \beta)$. It means that

$$\begin{aligned} R(l'y; \beta) - R(l'_k y; \beta) &= (1 - k^2)l'XQX'l - 2(1 - k)\lambda l'X(X'T^+X)^- X'T^+V'q \\ &\quad - 2k(1 - k)q'X_0QX'l - (1 - k)^2 q'X_0QX'_0q \\ &\quad + 2(1 - k)\lambda q'X_0(X'T^+X)^- X'T^+V'q \leq 0. \end{aligned}$$

Divide both sides of the above inequality by $1 - k$ and then let $k \rightarrow 1$, we have

$$l'XQX'l \leq \lambda(l'X - q'X_0)(X'T^+X)^- X'T^+V'q + l'XQX'_0q. \tag{□}$$

Theorem 3.2 For the model (3), $Cy \sim \delta$ in \mathcal{LH} under the quadratic loss function if and only if

- (1) $C\Sigma = \tilde{C}\Sigma$ (equivalently $\mathcal{M}(\Sigma C' - \lambda V') \subseteq \mathcal{M}(X)$), and
- (2) $CXQX'C' \leq \lambda(CX - X_0)(X'T^+X)^- X'T^+V' + CXQX'_0$, where $Q = (X'T^+X)^- - I = (X'T^+X)^- X'T^+ \Sigma T^+ X(X'T^+X)^-$, and
- (3) $\mathcal{M}(CX - X_0) = \mathcal{M}(G)$, where $G = (CX - X_0)Q(CX - X_0)' + (CX - X_0)[\lambda(X'T^+X)^- X'T^+V' - QX'C']$.

Proof Necessity:

- (i) The condition (1) is shown in Lemma 2.3;
- (ii) Since $Cy \sim \delta$, then $\eta'Cy \sim \eta'\delta$ for any $\eta \in \mathbb{R}^m$ by Lemma 2.2. Let $l' = \eta'C$ and $q' = \eta'$ in Theorem 3.1, we have

$$\eta'CXQX'C'\eta \leq \lambda\eta'(CX - X_0)(X'T^+X)^- X'T^+V'\eta + \eta'CXQX'_0\eta. \tag{4}$$

To prove condition (2), we can only prove that $\lambda(CX - X_0)(X'T^+X)^- X'T^+V' + CXQX'_0$ is symmetric from (4). Using reduction to absurdity and suppose $\lambda(CX - X_0)(X'T^+X)^- X'T^+V' + CXQX'_0$ is not a symmetric matrix. With this assumption and the fact that $CXQX'C'$ and $(CX - X_0)Q(CX - X_0)'$ are both symmetric positive semi-definite matrices, we have

$$\begin{aligned} &\lambda(CX - X_0)(X'T^+X)^- X'T^+V' + CXQX'_0 - CXQX'C' \\ &= -(CX - X_0)QX'_0 - (CX - X_0)Q(CX - X_0)' \\ &\quad + \lambda(CX - X_0)(X'T^+X)^- X'T^+V' \end{aligned}$$

is not symmetric and hence

$$-(CX - X_0)QX'_0 + \lambda(CX - X_0)(X'T^+X)^- X'T^+V'$$

is not symmetric. By Lemma 2.2, there exists an orthogonal matrix \mathbf{P} such that

$$\begin{aligned} & \text{tr}\{\mathbf{P}[-(\mathbf{C}\mathbf{X} - \mathbf{X}_0)\mathbf{Q}\mathbf{X}'_0 + \lambda(\mathbf{C}\mathbf{X} - \mathbf{X}_0)(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{V}']\} \\ & > \text{tr}\{-(\mathbf{C}\mathbf{X} - \mathbf{X}_0)\mathbf{Q}\mathbf{X}'_0 + \lambda(\mathbf{C}\mathbf{X} - \mathbf{X}_0)(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{V}'\}. \end{aligned} \tag{5}$$

Let

$$\begin{aligned} \mathbf{H} &= \mathbf{X}_0(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+ + \lambda\mathbf{V}\mathbf{T}^+[I - \mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+] \\ & \quad - \mathbf{P}\{\mathbf{X}_0(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+ + \lambda\mathbf{V}\mathbf{T}^+[I - \mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+] - \mathbf{C}\} \end{aligned}$$

and

$$\tilde{\mathbf{H}} = \mathbf{H}\mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+ + \lambda\mathbf{V}\mathbf{T}^+[I - \mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+].$$

It is easy to prove that $\mathbf{H}\mathbf{\Sigma} = \tilde{\mathbf{H}}\mathbf{\Sigma}$. Then with Lemma 2.4 and (5), we have

$$\begin{aligned} \mathbf{R}(\mathbf{H}\mathbf{y}; \boldsymbol{\beta}) &= \mathbf{R}(\tilde{\mathbf{H}}\mathbf{y}; \boldsymbol{\beta}) \\ &= \text{tr}\{(C\mathbf{X} - X_0)\mathbf{Q}(\mathbf{C}\mathbf{X} - \mathbf{X}_0)' + 2\mathbf{P}(\mathbf{C}\mathbf{X} - \mathbf{X}_0)\mathbf{Q}\mathbf{X}'_0 + \mathbf{X}_0\mathbf{Q}\mathbf{X}'_0 \\ & \quad - 2\lambda\mathbf{P}(\mathbf{C}\mathbf{X} - \mathbf{X}_0)(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{V} - 2\lambda\mathbf{X}_0(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{V} + \lambda^2\boldsymbol{\Sigma}_0 \\ & \quad - \lambda^2\mathbf{V}\mathbf{T}^+[I - \mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+]\} + \boldsymbol{\beta}'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)\boldsymbol{\beta} \\ & < \boldsymbol{\beta}'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)\boldsymbol{\beta} + \text{tr}\{\mathbf{C}\mathbf{X}\mathbf{Q}\mathbf{X}'\mathbf{C}' - 2\lambda\mathbf{C}\mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{V} \\ & \quad - \lambda^2\mathbf{V}\mathbf{T}^+[I - \mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+] + \lambda^2\boldsymbol{\Sigma}_0\} \\ &= \mathbf{R}(\mathbf{C}\mathbf{y}; \boldsymbol{\beta}), \end{aligned}$$

which contradicts the admissibility of $\mathbf{C}\mathbf{y}$. Thus,

$$-(\mathbf{C}\mathbf{X} - \mathbf{X}_0)\mathbf{Q}\mathbf{X}'_0 + \lambda(\mathbf{C}\mathbf{X} - \mathbf{X}_0)(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{V}'$$

is symmetric and then

$$\lambda(\mathbf{C}\mathbf{X} - \mathbf{X}_0)(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{V}' + \mathbf{C}\mathbf{X}\mathbf{Q}\mathbf{X}'_0$$

is symmetric. From (4), condition (2) is proved.

(iii) With the expression of \mathbf{G} , it is obviously seen that

$$\mathcal{M}(\mathbf{G}) \subseteq \mathcal{M}(\mathbf{C}\mathbf{X} - \mathbf{X}_0). \tag{6}$$

If $\mathbf{C}\mathbf{X} = \mathbf{X}_0$, $\mathcal{M}(\mathbf{C}\mathbf{X} - \mathbf{X}_0) \subseteq \mathcal{M}(\mathbf{G})$. If $\mathbf{C}\mathbf{X} \neq \mathbf{X}_0$, we also use reduction to absurdity and suppose that, for any nonzero vector $\boldsymbol{\eta}$, $\boldsymbol{\eta}'\mathbf{G}\boldsymbol{\eta} = 0$, namely

$$\boldsymbol{\eta}'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)\mathbf{Q}(\mathbf{C}\mathbf{X} - \mathbf{X}_0)'\boldsymbol{\eta} + \boldsymbol{\eta}'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)[\lambda(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{V}' - \mathbf{Q}\mathbf{X}'\mathbf{C}']\boldsymbol{\eta} = 0.$$

Since $(\mathbf{C}\mathbf{X} - \mathbf{X}_0)[\lambda(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{V}' - \mathbf{Q}\mathbf{X}'\mathbf{C}']$ is symmetric and nonnegative definite by the proof of condition (2), $\mathbf{G} \geq \mathbf{0}$ and

$$\begin{aligned} \eta'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)\mathbf{Q}(\mathbf{C}\mathbf{X} - \mathbf{X}_0)'\eta &= 0, \\ \eta'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)[\lambda(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{V}' - \mathbf{Q}\mathbf{X}'\mathbf{C}']\eta &= 0, \end{aligned}$$

which are equivalent to

$$\begin{aligned} \eta'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)\mathbf{Q} &= \mathbf{0}, \\ \eta'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)[\lambda(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{V}'] &= \mathbf{0}. \end{aligned} \tag{7}$$

Let $\xi'\mathbf{y}$ be a predictor of $\eta'\mathbf{y}_0$, where

$$\xi' = \eta' \{ \mathbf{X}_0(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+ + \lambda\mathbf{V}\mathbf{T}^+[I - \mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+] \}.$$

It can be verified by Lemma 2.1 that $R(\xi'\mathbf{y}; \beta) \geq R(\eta'\mathbf{C}\mathbf{y}; \beta)$ since $\eta'\mathbf{C}\mathbf{y} \sim \eta'\mathbf{y}_0$. From Lemma 2.4 and (7), we obtain

$$\begin{aligned} R(\eta'\mathbf{C}\mathbf{y}; \beta) - R(\xi'\mathbf{y}; \beta) &= \eta' \{ \mathbf{C}\mathbf{X}\mathbf{Q}\mathbf{X}'\mathbf{C}' - \mathbf{X}_0\mathbf{Q}\mathbf{X}'_0 - 2\lambda(\mathbf{C}\mathbf{X} - \mathbf{X}_0)(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{V} \\ &\quad + \beta'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)\beta \} \eta \\ &= \eta'\beta'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)\beta \eta \geq 0. \end{aligned}$$

As $\mathbf{C}\mathbf{X} \neq \mathbf{X}_0$, there at least exists one $\beta_0 \in \mathbb{R}^p$ such that $R(\eta'\mathbf{C}\mathbf{y}; \beta_0) - R(\xi'\mathbf{y}; \beta_0) = \eta'\beta'_0(\mathbf{C}\mathbf{X} - \mathbf{X}_0)'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)\beta_0\eta > 0$, which is contradicted to the admissibility of $\eta'\mathbf{C}\mathbf{y}$. Therefore, if $\mathbf{C}\mathbf{X} \neq \mathbf{X}_0$, then $\eta'\mathbf{G}\eta \neq 0$. Equivalently, if $\eta'\mathbf{G}\eta = 0$, then $\mathbf{C}\mathbf{X} = \mathbf{X}_0$. Because $\eta'\mathbf{G}\eta = 0 \Leftrightarrow \eta'\mathbf{G} = \mathbf{0}$, then if $\eta'\mathbf{G} = \mathbf{0}$, $\eta'(\mathbf{C}\mathbf{X} - \mathbf{X}_0) = \mathbf{0}$, namely $\mathcal{M}(\mathbf{C}\mathbf{X} - \mathbf{X}_0) \subseteq \mathcal{M}(\mathbf{G})$. Together with (6), $\mathcal{M}(\mathbf{C}\mathbf{X} - \mathbf{X}_0) = \mathcal{M}(\mathbf{G})$.

Sufficiency: Let $\mathbf{M}\mathbf{y}$ be an arbitrary predictor of \mathbf{y}_0 , by Lemma 2.3, we only need to prove that if $\mathbf{M}\Sigma = \tilde{\mathbf{M}}\Sigma$, where $\tilde{\mathbf{M}} = \mathbf{M}\mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+ + \lambda\mathbf{V}\mathbf{T}^+[I - \mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+]$, then $\mathbf{M}\mathbf{y}$ cannot be better than $\mathbf{C}\mathbf{y}$. Or, we can only prove that there exists at least one β^* such that $R(\mathbf{M}\mathbf{y}; \beta^*) > R(\mathbf{C}\mathbf{y}; \beta^*)$. We divide the issue into two circumstances:

- (i) $\mathbf{C}\mathbf{X} = \mathbf{X}_0$:
 - If $\mathbf{M}\mathbf{X} \neq \mathbf{X}_0$, it is obviously that $\mathbf{M}\mathbf{y}$ cannot be better than $\mathbf{C}\mathbf{y}$ by computing their risks;
 - If $\mathbf{M}\mathbf{X} = \mathbf{X}_0$, $R(\mathbf{M}\mathbf{y}; \beta) = R(\mathbf{C}\mathbf{y}; \beta)$.
- (ii) $\mathbf{C}\mathbf{X} \neq \mathbf{X}_0$. There are two cases to be discussed:
 - (a) $\mathbf{M}\mathbf{X} = \mathbf{X}_0$. By condition (2),

$$\begin{aligned} R(\mathbf{M}\mathbf{y}; \mathbf{0}) - R(\mathbf{C}\mathbf{y}; \mathbf{0}) - \text{tr}(\mathbf{G}) &= \text{tr}[\mathbf{X}_0\mathbf{Q}\mathbf{X}'_0 - \mathbf{C}\mathbf{X}\mathbf{Q}\mathbf{X}'_0 + \lambda(\mathbf{C}\mathbf{X} - \mathbf{X}_0)(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{V}'] \\ &\geq \text{tr}(\mathbf{X}_0\mathbf{Q}\mathbf{X}'_0 - \mathbf{C}\mathbf{X}\mathbf{Q}\mathbf{X}'_0 + \mathbf{C}\mathbf{X}\mathbf{Q}\mathbf{X}'\mathbf{C}' - \mathbf{C}\mathbf{X}\mathbf{Q}\mathbf{X}'_0) \\ &= \text{tr}[(\mathbf{C}\mathbf{X} - \mathbf{X}_0)\mathbf{Q}(\mathbf{C}\mathbf{X} - \mathbf{X}_0)'] \geq 0. \end{aligned}$$

Since $\mathbf{C}\mathbf{X} \neq \mathbf{X}_0$ and by condition (3), $\text{rk}(\mathbf{G}) = \text{rk}(\mathbf{C}\mathbf{X} - \mathbf{X}_0) \geq 1$, and $\text{tr}(\mathbf{G}) > 0$. Thus, $R(\mathbf{M}\mathbf{y}; \mathbf{0}) > R(\mathbf{C}\mathbf{y}; \mathbf{0})$ and $\mathbf{M}\mathbf{y}$ cannot be better than $\mathbf{C}\mathbf{y}$.

(b) $\mathbf{MX} \neq \mathbf{X}_0$. If $\mathbf{MX} = \mathbf{CX}$, then $R(\mathbf{My}; \beta) = R(\mathbf{Cy}; \beta)$. So we only need to discuss this issue on the case that $\mathbf{MX} \neq \mathbf{CX}$. By Lemma 2.4 and for any β , we only need to consider the admissibility under this circumstance that

$$\beta'(\mathbf{MX} - \mathbf{X}_0)'(\mathbf{MX} - \mathbf{X}_0)\beta \leq \beta'(\mathbf{CX} - \mathbf{X}_0)'(\mathbf{CX} - \mathbf{X}_0)\beta,$$

namely

$$(\mathbf{MX} - \mathbf{X}_0)'(\mathbf{MX} - \mathbf{X}_0) \leq (\mathbf{CX} - \mathbf{X}_0)'(\mathbf{CX} - \mathbf{X}_0). \tag{8}$$

It can be shown from (6) that $\mathcal{M}[(\mathbf{MX} - \mathbf{X}_0)'] \subset \mathcal{M}[(\mathbf{CX} - \mathbf{X}_0)']$, then

$$(\mathbf{MX} - \mathbf{X}_0)' = (\mathbf{CX} - \mathbf{X}_0)' \mathbf{Z}, \tag{9}$$

where

$$\mathbf{Z} = [(\mathbf{CX} - \mathbf{X}_0)(\mathbf{CX} - \mathbf{X}_0)']^+ (\mathbf{CX} - \mathbf{X}_0)(\mathbf{MX} - \mathbf{X}_0)'$$

By (8), $\mathbf{ZZ}' \leq [(\mathbf{CX} - \mathbf{X}_0)(\mathbf{CX} - \mathbf{X}_0)']^+ (\mathbf{CX} - \mathbf{X}_0)(\mathbf{CX} - \mathbf{X}_0)' \leq \mathbf{I}$. By (9) and condition (2), we have

$$\begin{aligned} & R(\mathbf{My}; \mathbf{0}) - R(\mathbf{Cy}; \mathbf{0}) \\ &= \text{tr}[(\mathbf{MX} - \mathbf{X}_0)\mathbf{Q}(\mathbf{MX} - \mathbf{X}_0)' - 2\lambda(\mathbf{MX} - \mathbf{X}_0)\mathbf{X}'\mathbf{T}^+\mathbf{X} - \mathbf{X}'\mathbf{T}^+\mathbf{V}' \\ &\quad + 2\mathbf{X}_0\mathbf{Q}(\mathbf{MX} - \mathbf{X}_0)' - (\mathbf{CX} - \mathbf{X}_0)\mathbf{Q}(\mathbf{CX} - \mathbf{X}_0)' \\ &\quad + 2\lambda(\mathbf{CX} - \mathbf{X}_0)\mathbf{X}'\mathbf{T}^+\mathbf{X} - \mathbf{X}'\mathbf{T}^+\mathbf{V}' - 2\mathbf{X}_0\mathbf{Q}(\mathbf{CX} - \mathbf{X}_0)'] \\ &= \text{tr}[(\mathbf{I} - \mathbf{Z})(\mathbf{I} - \mathbf{Z})'\mathbf{X}_0\mathbf{Q}\mathbf{X}'_0 + 2\mathbf{Z}(\mathbf{I} - \mathbf{Z})'\mathbf{X}_0\mathbf{Q}\mathbf{X}'\mathbf{C}' \\ &\quad + 2\lambda(\mathbf{I} - \mathbf{Z})'(\mathbf{CX} - \mathbf{X}_0)(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{V}' - (\mathbf{I} - \mathbf{ZZ}')\mathbf{C}\mathbf{X}\mathbf{Q}\mathbf{X}'\mathbf{C}'] \\ &\geq \text{tr}[(\mathbf{I} - \mathbf{Z})(\mathbf{I} - \mathbf{Z})'\mathbf{X}_0\mathbf{Q}\mathbf{X}'_0 + 2\mathbf{Z}(\mathbf{I} - \mathbf{Z})'\mathbf{X}_0\mathbf{Q}\mathbf{X}'\mathbf{C}' \\ &\quad + 2\lambda(\mathbf{I} - \mathbf{Z})'(\mathbf{CX} - \mathbf{X}_0)(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{V}' \\ &\quad - \lambda(\mathbf{I} - \mathbf{ZZ}')(\mathbf{CX} - \mathbf{X}_0)(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{V}' - (\mathbf{I} - \mathbf{ZZ}')\mathbf{X}_0\mathbf{Q}\mathbf{X}'\mathbf{C}'] \\ &= \text{tr}[(\mathbf{I} - \mathbf{Z})(\mathbf{I} - \mathbf{Z})'\mathbf{X}_0\mathbf{Q}\mathbf{X}'_0 - 2(\mathbf{I} - \mathbf{Z})(\mathbf{I} - \mathbf{Z})'\mathbf{X}_0\mathbf{Q}\mathbf{X}'\mathbf{C}' \\ &\quad + \lambda(\mathbf{I} - 2\mathbf{Z}' + \mathbf{ZZ}')(\mathbf{CX} - \mathbf{X}_0)(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{V}' + \mathbf{X}_0\mathbf{Q}\mathbf{X}'\mathbf{C}'] \\ &= \text{tr}[(\mathbf{I} - \mathbf{Z})(\mathbf{I} - \mathbf{Z})'\mathbf{X}_0\mathbf{Q}\mathbf{X}'_0 - 2(\mathbf{I} - \mathbf{Z})(\mathbf{I} - \mathbf{Z})'\mathbf{X}_0\mathbf{Q}\mathbf{X}'\mathbf{C}' \\ &\quad + \lambda(\mathbf{I} - \mathbf{Z} - \mathbf{Z}' + \mathbf{ZZ}')(\mathbf{CX} - \mathbf{X}_0)(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{V}' + \mathbf{X}_0\mathbf{Q}\mathbf{X}'\mathbf{C}'] \\ &= \text{tr}[(\mathbf{I} - \mathbf{Z})'\mathbf{G}(\mathbf{I} - \mathbf{Z})]. \end{aligned}$$

Notice that $(\mathbf{CX} - \mathbf{X}_0)'(\mathbf{I} - \mathbf{Z}) = (\mathbf{CX} - \mathbf{X}_0)' - (\mathbf{MX} - \mathbf{X}_0)' = \mathbf{CX} - \mathbf{MX} \neq \mathbf{0}$ and together with condition (3), $\mathbf{G}(\mathbf{I} - \mathbf{Z}) \neq \mathbf{0}$. Moreover, as $\mathbf{G} \geq \mathbf{0}$, then

$$R(\mathbf{My}; \mathbf{0}) - R(\mathbf{Cy}; \mathbf{0}) \geq \text{tr}[(\mathbf{I} - \mathbf{Z})'\mathbf{G}(\mathbf{I} - \mathbf{Z})] > 0,$$

which means \mathbf{My} cannot be better than \mathbf{Cy} .

In summary, the proof is completed. □

Corollary 3.1 For the model (3), if $\Sigma > \mathbf{0}$, then $\mathbf{C}\mathbf{y} \sim \delta$ under the quadratic loss function if and only if

- (1) $\mathbf{C} = \mathbf{C}\mathbf{X}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1} + \lambda\mathbf{V}\Sigma^{-1}[\mathbf{I} - \mathbf{X}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}]$, and
- (2) $\mathbf{C}\mathbf{X}\mathbf{Q}\mathbf{X}'\mathbf{C}' \leq \lambda(\mathbf{C}\mathbf{X} - \mathbf{X}_0)(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{V}' + \mathbf{C}\mathbf{X}\mathbf{Q}\mathbf{X}'_0$.

Proof If $\Sigma > \mathbf{0}$, then $\mathbf{T} > \mathbf{0}$ and

$$\begin{aligned} \mathbf{T}^+ &= \mathbf{T}^{-1} = \Sigma^{-1} - \Sigma^{-1}\mathbf{X}(\mathbf{I} + \mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}, \\ \mathbf{X}'\Sigma^{-1}\mathbf{X}(\mathbf{I} + \mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1} &= \mathbf{I} - (\mathbf{I} + \mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}. \end{aligned} \tag{10}$$

Therefore,

$$\mathbf{X}'\mathbf{T}^{-1} = (\mathbf{I} + \mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1} \tag{11}$$

and

$$\begin{aligned} \mathbf{X}(\mathbf{X}'\mathbf{T}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^{-1} &= \mathbf{X}[(\mathbf{I} + \mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{X}]^{-1}(\mathbf{I} + \mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1} \\ &= \mathbf{X}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}. \end{aligned} \tag{12}$$

By (10), (11) and (12), conditions (1), (2) in Corollary 3.1 hold and condition (3) in Theorem 3.2 holds naturally since $\text{rk}(\mathbf{C}\mathbf{X} - \mathbf{X}_0) = \text{rk}(\mathbf{G})$. □

Remark 3.1 As $\delta = \mathbf{y}_0$ when $\lambda = 1$ and $\delta = \mathbf{E}\mathbf{y}_0$ when $\lambda = 0$, it is convenient to obtain the sufficient and necessary conditions for the predictors of \mathbf{y}_0 and $\mathbf{E}\mathbf{y}_0$ to be admissible from Theorem 3.2 and Corollary 3.1.

Corollary 3.2 For model (3) and under the quadratic loss function, the best linear unbiased predictors of δ , \mathbf{y}_0 and $\mathbf{E}\mathbf{y}_0$ are

$$\begin{aligned} \hat{\delta}_{\text{BLUP}} &= \mathbf{C}\mathbf{y} = \mathbf{X}_0(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{y} + \lambda\mathbf{V}\mathbf{T}^+[\mathbf{y} - \mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{y}], \\ \hat{\mathbf{y}}_{0\text{BLUP}} &= \mathbf{X}_0(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{y} + \mathbf{V}\mathbf{T}^+[\mathbf{y} - \mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{y}], \\ \hat{\mathbf{E}}\mathbf{y}_{0\text{BLUP}} &= \mathbf{X}_0(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+\mathbf{y}. \end{aligned}$$

Proof Let $\mathbf{C}\mathbf{y} + \mathbf{u}$ be the linear unbiased predictor of δ . $\mathbf{E}(\mathbf{C}\mathbf{y} + \mathbf{u}) = \mathbf{E}\delta$ gives that $\mathbf{C}\mathbf{X} = \mathbf{X}_0$ and $\mathbf{u} = \mathbf{0}$. Let $\mathbf{C} = \mathbf{X}_0(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+ + \lambda\mathbf{V}\mathbf{T}^+[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{T}^+\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}^+]$, the corollary is easily proved by verifying the conditions in Theorem 3.2. □

We give some admissible predictors in the following examples.

Example 3.1 Suppose $\Sigma > \mathbf{0}$, the best linear unbiased predictor of \mathbf{y}_0 in [1] is $\hat{\mathbf{y}}_{0\text{BLUP}} = \mathbf{X}_0(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{y} + \mathbf{V}\Sigma^{-1}[\mathbf{I} - \mathbf{X}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}]\mathbf{y}$. Let $\hat{\beta}_{\text{BLUE}} = (\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{y}$. $\hat{\mathbf{y}}_{0\text{SPP}} = \mathbf{X}_0\hat{\beta}_{\text{BLUE}}$ is the simple projective predictor of \mathbf{y}_0 in [6]. Note that $\mathbf{E}\hat{\mathbf{y}}_{0\text{BLUP}} = \mathbf{E}\delta = \mathbf{E}\mathbf{y}_{0\text{SPP}}$, which means $\hat{\mathbf{y}}_{0\text{BLUP}}$ and $\hat{\mathbf{y}}_{0\text{SPP}}$ are also unbiased predictors of δ . If $\lambda = 1$, then by Corollary 3.1, $\hat{\mathbf{y}}_{0\text{BLUP}} \sim \mathbf{y}_0$, and therefore $\hat{\mathbf{y}}_{0\text{BLUP}} \approx \delta$ if $\lambda \neq 1$. If $\lambda = 0$, then by Corollary 3.1, $\hat{\mathbf{y}}_{0\text{SPP}} \sim \mathbf{E}\mathbf{y}_0$, which means if $\lambda \neq 0$, $\hat{\mathbf{y}}_{0\text{SPP}} \approx \delta$.

Remark 3.2 Example 3.1 indicates that the unbiased predictors of δ are also the unbiased predictors of \mathbf{y}_0 and $\mathbf{E}\mathbf{y}_0$, and the unbiased predictors of \mathbf{y}_0 and $\mathbf{E}\mathbf{y}_0$ are also the unbiased predictors of δ since $\mathbf{E}\delta = \mathbf{E}\mathbf{y}_0 = \mathbf{E}(\mathbf{E}\mathbf{y}_0) = \mathbf{X}_0\boldsymbol{\beta}$. However, admissibility of those predictors for each studied variables are different.

Example 3.2 Suppose $\mathbf{V} = \mathbf{0}$, $\boldsymbol{\Sigma} > \mathbf{0}$ and $\text{rk}(\mathbf{X}) < p$ in model (3). Suppose t_{\max} be the maximum eigenvalue of $\mathbf{X}'\mathbf{X}$, non-stochastic scalars $k > 0$ and $\theta = \frac{t_{\max}+k}{t_{\max}}$. Let

$$\mathbf{D} = \mathbf{B} - \mathbf{B}[\mathbf{I}_n - \theta\mathbf{X}(\mathbf{X}'\mathbf{X} + k\mathbf{I}_p)^{-1}\mathbf{X}']^{-1}[\mathbf{I}_n - \theta\mathbf{X}(\mathbf{X}'\mathbf{X} + k\mathbf{I}_p)^{-1}\mathbf{X}'],$$

where \mathbf{B} is an $m \times n$ arbitrary matrix. If $\mathbf{X}_0 = \theta\mathbf{D}\mathbf{X}$, then by Corollary 3.1 without tedious calculations, $\mathbf{X}_0(\mathbf{X}'\mathbf{X} + k\mathbf{I}_p)^{-1}\mathbf{X}'\mathbf{y} \sim \delta$.

Remark 3.3 Denote $\hat{\boldsymbol{\beta}}_{\text{ridge}}$ as the ridge estimator of $\boldsymbol{\beta}$ in model (1) when $\boldsymbol{\Sigma} > \mathbf{0}$ and $\text{rk}(\mathbf{X}) < p$. Example 3.2 indicates that in particular linear regression model, if \mathbf{X}_0 and \mathbf{X} have some relations, we can use the ridge predictor $\mathbf{X}_0\hat{\boldsymbol{\beta}}_{\text{ridge}}$ as the admissible predictor for δ , especially for \mathbf{y}_0 and $\mathbf{E}\mathbf{y}_0$.

4 Admissibility of nonhomogeneous linear predictors

In this section, we investigate the admissibility of simultaneous prediction in class of non-homogeneous linear predictors, and we obtain the necessary and sufficient conditions. Studies show the admissibility of simultaneous prediction in the class of nonhomogeneous linear predictors is based on the admissibility of simultaneous prediction in the class of homogeneous linear predictors.

Theorem 4.1 *For the model (3), $\mathbf{C}\mathbf{y} + \mathbf{u} \sim \delta$ in $\mathcal{L}\mathcal{N}$ under the quadratic loss function if and only if*

- (1) $\mathbf{C}\mathbf{y} \sim \delta$ in $\mathcal{L}\mathcal{H}$, and
- (2) $\mathbf{u} \in \mathcal{M}(\mathbf{C}\mathbf{X} - \mathbf{X}_0)$.

Proof Necessity: Suppose by contradiction that $\mathbf{u} \notin \mathcal{M}(\mathbf{C}\mathbf{X} - \mathbf{X}_0)$ and $\mathbf{P} = (\mathbf{C}\mathbf{X} - \mathbf{X}_0)[(\mathbf{C}\mathbf{X} - \mathbf{X}_0)'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)]^{-1}(\mathbf{C}\mathbf{X} - \mathbf{X}_0)'$, then $(\mathbf{I} - \mathbf{P})\mathbf{u} \neq \mathbf{0}$ and

$$\begin{aligned} \mathbf{R}(\mathbf{C}\mathbf{y} + \mathbf{u}; \boldsymbol{\beta}) &= \mathbf{E}[(\mathbf{C}\mathbf{y} + \mathbf{u} - \lambda\mathbf{y}_0 - (1 - \lambda)\mathbf{X}_0\boldsymbol{\beta})'(\mathbf{C}\mathbf{y} + \mathbf{u} - \lambda\mathbf{y}_0 - (1 - \lambda)\mathbf{X}_0\boldsymbol{\beta})] \\ &= \text{tr}[\mathbf{C}\boldsymbol{\Sigma}\mathbf{C}' + \lambda^2\boldsymbol{\Sigma}_0 - 2\lambda(\mathbf{C}\mathbf{V}')] + \boldsymbol{\beta}'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)\boldsymbol{\beta} \\ &\quad + \mathbf{u}'\mathbf{u} + 2\mathbf{u}'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)\boldsymbol{\beta} \\ &= \text{tr}[\mathbf{C}\boldsymbol{\Sigma}\mathbf{C}' + \lambda^2\boldsymbol{\Sigma}_0 - 2\lambda(\mathbf{C}\mathbf{V}')] + \boldsymbol{\beta}'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)\boldsymbol{\beta} \\ &\quad + \mathbf{u}'\mathbf{P}\mathbf{u} + \mathbf{u}'(\mathbf{I} - \mathbf{P})\mathbf{u} + 2\mathbf{u}'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)\boldsymbol{\beta} \\ &\geq \text{tr}[\mathbf{C}\boldsymbol{\Sigma}\mathbf{C}' + \lambda^2\boldsymbol{\Sigma}_0 - 2\lambda(\mathbf{C}\mathbf{V}')] + \boldsymbol{\beta}'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)\boldsymbol{\beta} \\ &\quad + \mathbf{u}'\mathbf{P}\mathbf{u} + 2\mathbf{u}'\mathbf{P}(\mathbf{C}\mathbf{X} - \mathbf{X}_0)\boldsymbol{\beta} \\ &= \mathbf{R}(\mathbf{C}\mathbf{y} + \mathbf{P}\mathbf{u}; \boldsymbol{\beta}). \end{aligned} \tag{13}$$

This contradicts the fact that $\mathbf{C}\mathbf{y} + \mathbf{u} \sim \delta$, and therefore $\mathbf{u} \in \mathcal{M}(\mathbf{C}\mathbf{X} - \mathbf{X}_0)$.

Now suppose $\mathbf{u} \in \mathcal{M}(\mathbf{CX} - \mathbf{X}_0)$ while \mathbf{Cy} is not admissible, then there exists a predictor $\mathbf{Hy} \in \mathcal{LH}$ of \mathbf{y}_0 such that $R(\mathbf{Hy}; \boldsymbol{\beta}) \leq R(\mathbf{Cy}; \boldsymbol{\beta})$ for any $\boldsymbol{\beta}$ and the inequality holds at least at some $\boldsymbol{\beta}_0$. As $\mathbf{u} \in \mathcal{M}(\mathbf{CX} - \mathbf{X}_0)$, there exists some vector \mathbf{p} such that $\mathbf{u} = (\mathbf{CX} - \mathbf{X}_0)\mathbf{p}$. Then, similar to the derivation of (13),

$$\begin{aligned} &R[\mathbf{Hy} + (\mathbf{HX} - \mathbf{X}_0)\mathbf{p}; \boldsymbol{\beta}] \\ &= \text{tr}[\mathbf{H}\boldsymbol{\Sigma}\mathbf{H}' + \lambda^2\boldsymbol{\Sigma}_0 - 2\lambda(\mathbf{H}\mathbf{V}')] + \boldsymbol{\beta}'(\mathbf{HX} - \mathbf{X}_0)'(\mathbf{HX} - \mathbf{X}_0)\boldsymbol{\beta} \\ &\quad + \mathbf{p}'(\mathbf{HX} - \mathbf{X}_0)'(\mathbf{HX} - \mathbf{X}_0)\mathbf{p} + 2\mathbf{p}'(\mathbf{HX} - \mathbf{X}_0)'(\mathbf{HX} - \mathbf{X}_0)\boldsymbol{\beta} \\ &= \text{tr}[\mathbf{H}\boldsymbol{\Sigma}\mathbf{H}' + \lambda^2\boldsymbol{\Sigma}_0 - 2\lambda(\mathbf{H}\mathbf{V}')] + (\boldsymbol{\beta} + \mathbf{p})'(\mathbf{HX} - \mathbf{X}_0)'(\mathbf{HX} - \mathbf{X}_0)(\boldsymbol{\beta} + \mathbf{p}) \\ &= R(\mathbf{Hy}; \boldsymbol{\beta} + \mathbf{p}) \\ &\leq R(\mathbf{Cy}; \boldsymbol{\beta} + \mathbf{p}) \\ &= \text{tr}[\mathbf{C}\boldsymbol{\Sigma}\mathbf{C}' + \lambda^2\boldsymbol{\Sigma}_0 - 2\lambda(\mathbf{C}\mathbf{V}')] + (\boldsymbol{\beta} + \mathbf{p})'(\mathbf{CX} - \mathbf{X}_0)'(\mathbf{CX} - \mathbf{X}_0)(\boldsymbol{\beta} + \mathbf{p}) \\ &= R(\mathbf{Cy} + \mathbf{p}; \boldsymbol{\beta}) \end{aligned}$$

at any $\boldsymbol{\beta}$ and the inequality holds at $\boldsymbol{\beta} = \boldsymbol{\beta}_0 - \mathbf{p}$, which contradicts $\mathbf{Cy} + \mathbf{u} \sim \delta$. Therefore, \mathbf{Cy} is admissible.

Sufficiency: Suppose $\mathbf{Ny} + \mathbf{v}$ is a predictor of \mathbf{y}_0 and for any $\boldsymbol{\beta} \in \mathbb{R}^p$, $R(\mathbf{Ny} + \mathbf{v}; \boldsymbol{\beta}) \leq R(\mathbf{Cy} + \mathbf{u}; \boldsymbol{\beta})$. Let \mathbf{P}_v and \mathbf{P}_u be the orthogonal projection matrices onto $\mathcal{M}(\mathbf{NX} - \mathbf{X}_0)$ and $\mathcal{M}(\mathbf{CX} - \mathbf{X}_0)$ respectively. Since $\mathbf{u} \in \mathcal{M}(\mathbf{CX} - \mathbf{X}_0)$, there exist $\boldsymbol{\beta}_v$ and $\boldsymbol{\beta}_u$ in \mathbb{R}^p such that

$$\begin{aligned} \mathbf{P}_v\mathbf{v} &= (\mathbf{NX} - \mathbf{X}_0)\boldsymbol{\beta}_v, \\ \mathbf{u} &= \mathbf{P}_u\mathbf{u} = (\mathbf{CX} - \mathbf{X}_0)\boldsymbol{\beta}_u. \end{aligned}$$

According to the derivation of (13), we have

$$R(\mathbf{Ny} + \mathbf{P}_v\mathbf{v}; \boldsymbol{\beta}) \leq R(\mathbf{Ny} + \mathbf{v}; \boldsymbol{\beta}) \leq R(\mathbf{Cy} + \mathbf{u}; \boldsymbol{\beta}) = R(\mathbf{Cy} + \mathbf{P}_u\mathbf{u}; \boldsymbol{\beta}). \tag{14}$$

Equation (14) holds for any $\boldsymbol{\beta}$ so that

$$(\mathbf{NX} - \mathbf{X}_0)'(\mathbf{NX} - \mathbf{X}_0) \leq (\mathbf{CX} - \mathbf{X}_0)'(\mathbf{CX} - \mathbf{X}_0), \tag{15}$$

$$R(\mathbf{Ny} + \mathbf{P}_v\mathbf{v}; \boldsymbol{\beta}) = R(\mathbf{Ny}; \boldsymbol{\beta} + \boldsymbol{\beta}_v) \leq R(\mathbf{Cy}; \boldsymbol{\beta} + \boldsymbol{\beta}_u) = R(\mathbf{Cy} + \mathbf{P}_u\mathbf{u}; \boldsymbol{\beta}). \tag{16}$$

Let $\boldsymbol{\beta} = -\boldsymbol{\beta}_u$ in (16), we have

$$R(\mathbf{Ny}; \mathbf{0}) \leq R(\mathbf{Hy}; \boldsymbol{\beta}_v - \boldsymbol{\beta}_u) \leq R(\mathbf{Cy}; \mathbf{0}), \tag{17}$$

and by (15) and (17),

$$\begin{aligned} R(\mathbf{Ny}; \boldsymbol{\beta}) &= R(\mathbf{Ny}; \mathbf{0}) + \boldsymbol{\beta}'(\mathbf{NX} - \mathbf{X}_0)'(\mathbf{NX} - \mathbf{X}_0)\boldsymbol{\beta} \\ &\leq R(\mathbf{Cy}; \mathbf{0}) + \boldsymbol{\beta}'(\mathbf{CX} - \mathbf{X}_0)'(\mathbf{CX} - \mathbf{X}_0)\boldsymbol{\beta} = R(\mathbf{Cy}; \boldsymbol{\beta}). \end{aligned} \tag{18}$$

Since $\mathbf{C}\mathbf{y} \sim \delta$ and β is arbitrary, (18), (17) and (15) successively gives

$$\begin{aligned} R(\mathbf{N}\mathbf{y}; \beta) &= R(\mathbf{C}\mathbf{y}; \beta), \\ R(\mathbf{N}\mathbf{y}; \mathbf{0}) &= R(\mathbf{H}\mathbf{y}; \beta_v - \beta_u) = R(\mathbf{C}\mathbf{y}; \mathbf{0}), \end{aligned} \tag{19}$$

$$(\mathbf{N}\mathbf{X} - \mathbf{X}_0)'(\mathbf{N}\mathbf{X} - \mathbf{X}_0) = (\mathbf{C}\mathbf{X} - \mathbf{X}_0)'(\mathbf{C}\mathbf{X} - \mathbf{X}_0). \tag{20}$$

By (19), we have

$$(\beta_v - \beta_u)'(\mathbf{N}\mathbf{X} - \mathbf{X}_0)'(\mathbf{N}\mathbf{X} - \mathbf{X}_0)(\beta_v - \beta_u) = 0,$$

namely $(\mathbf{N}\mathbf{X} - \mathbf{X}_0)(\beta_v - \beta_u) = \mathbf{0}$. In addition, by (20),

$$\begin{aligned} R(\mathbf{N}\mathbf{y}; \beta + \beta_v) &= R(\mathbf{N}\mathbf{y}; \mathbf{0}) + (\beta + \beta_v)'(\mathbf{N}\mathbf{X} - \mathbf{X}_0)'(\mathbf{N}\mathbf{X} - \mathbf{X}_0)(\beta + \beta_v) \\ &= R(\mathbf{N}\mathbf{y}; \mathbf{0}) + (\beta + \beta_u)'(\mathbf{N}\mathbf{X} - \mathbf{X}_0)'(\mathbf{N}\mathbf{X} - \mathbf{X}_0)(\beta + \beta_u) \\ &= R(\mathbf{C}\mathbf{y}; \mathbf{0}) + (\beta + \beta_u)'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)'(\mathbf{C}\mathbf{X} - \mathbf{X}_0)(\beta + \beta_u) \\ &= R(\mathbf{C}\mathbf{y}; \beta + \beta_u), \end{aligned}$$

and hence (16) $\Leftrightarrow R(\mathbf{N}\mathbf{y}; \beta + \beta_v) = R(\mathbf{C}\mathbf{y}; \beta + \beta_u) \Leftrightarrow R(\mathbf{N}\mathbf{y} + \mathbf{P}_v\mathbf{v}; \beta) = R(\mathbf{C}\mathbf{y} + \mathbf{P}_u\mathbf{u}; \beta)$. Consequently, (14) means that, for any $\beta \in \mathbb{R}^p$ and any predictor $\mathbf{N}\mathbf{y} + \mathbf{v} \in \mathcal{LN}$ of \mathbf{y}_0 , if $R(\mathbf{N}\mathbf{y} + \mathbf{v}; \beta) \leq R(\mathbf{C}\mathbf{y} + \mathbf{u}; \beta)$, then $R(\mathbf{N}\mathbf{y} + \mathbf{v}; \beta) = R(\mathbf{C}\mathbf{y} + \mathbf{u}; \beta)$, namely $\mathbf{C}\mathbf{y} + \mathbf{u} \sim \delta$. \square

Remark 4.1 Theorem 4.1 shows the relation between the admissible homogeneous and nonhomogeneous linear predictors, and indicates that the admissibility of the homogeneous linear predictor is more significant. To derive an admissible predictor $\mathbf{C}\mathbf{y} + \mathbf{u}$ in \mathcal{LN} , we can derive the admissible predictor $\mathbf{C}\mathbf{y}$ in \mathcal{LH} beforehand.

5 Concluding remarks

In this paper, we investigate the admissibility of linear prediction in the generalized linear regression model under the quadratic loss function. Predictions are based on a composite target function that allows one to predict actual and average values of the unobserved regressand simultaneously, according to some practical needs. Necessary and sufficient conditions for the simultaneous prediction to be admissible are obtained in classes of homogeneous and nonhomogeneous linear predictors, respectively. Although the unbiased predictors of the composite target function are the unbiased predictors of the actual and average values of the unobserved regressand and vice versa, yet the admissibility of these predictors for each studied variables are different. Under some circumstances, the ridge predictor is admissible although it is biased. However, whether the admissible linear prediction is minimax under quadratic loss function is unclear. Further research on the minimaxity of admissible simultaneous prediction is in progress.

Acknowledgements

The authors are grateful to the responsible editor and the anonymous reviewers for their valuable comments and suggestions, which have greatly improved this paper. This research is supported by the Scientific Research Fund of Hunan Provincial Education Department (13C1139), the Youth Scientific Research Foundation of Central South University of Forestry and Technology of China (QJ2012013A) and the Natural Science Foundation of Hunan Province (2015JJ4090).

Abbreviations

BLUP, Best linear unbiased predictor; MSE, Mean squared error; SPP, Simple projection predictor; BLUE, Best linear unbiased estimation.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors contributed equally to the writing of this paper. All authors read and approved the final manuscript.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 24 January 2018 Accepted: 25 April 2018 Published online: 16 May 2018

References

- Goldberger, A.S.: Best linear unbiased prediction in the generalized linear regression model. *J. Am. Stat. Assoc.* **57**, 369–375 (1962)
- Gotway, C.A., Cressie, N.: Improved multivariate prediction under a general linear model. *J. Multivar. Anal.* **45**, 56–72 (1993)
- Xu, L.W.: Admissible linear predictors in the superpopulation model with respect to inequality constraints. *Commun. Stat., Theory Methods* **38**, 2528–2540 (2009)
- Teunissen, P.J.G.: Best prediction in linear models with mixed integer/real unknowns: theory and application. *J. Geod.* **81**(12), 759–780 (2007)
- Xu, L.W., Lu, M., Jiang, C.F.: Optimal prediction in finite populations under matrix loss. *J. Stat. Plan. Inference* **141**, 2503–2512 (2011)
- Bolfarine, H., Rodrigues, J.: On the simple projection predictor in finite populations. *Aust. J. Stat.* **30**(3), 338–341 (1988)
- Hendry, D.F., Clements, M.P.: Pooling of forecasts. *Econom. J.* **5**, 1–26 (2002)
- Timmermann, A.: Forecast combinations. In: Elliott, G., Granger, C.W.J., Timmermann, A. (eds.) *Handbook of Economic Forecasting*, vol. 1, pp. 135–196. Elsevier, North Holland (2006)
- Chaturvedi, A., Singh, S.P.: Stein rule prediction of the composite target function in a general linear regression model. *Stat. Pap.* **41**(3), 359–367 (2000)
- Chaturvedi, A., Kesarwani, S., Chandra, R.: Simultaneous prediction based on shrinkage estimator. In: Shalabh Heumann, C. (ed.) *Recent Advances in Linear Models and Related Areas*, pp. 181–204. Springer, New York (2008)
- Chaturvedi, A., Wan, A.T.K., Singh, S.P.: Improved multivariate prediction in a general linear model with an unknown error covariance matrix. *J. Multivar. Anal.* **83**(1), 166–182 (2002)
- Toutenburg, H., Shalabh: Predictive performance of the methods of restricted and mixed regression estimators. *Biom. J.* **38**(8), 951–959 (1996)
- Shalabh, H.T.: Improved predictions in linear regression models with stochastic linear constraints. *Biom. J.* **42**(1), 71–86 (2000)
- Manocha, M.D.V.: Simultaneous prediction in restricted regression models. *J. Appl. Statist. Sc.* **11**(4), 277–288 (2002)
- Shalabh Heumann, C.: Simultaneous prediction of actual and average values of response variable in replicated measurement error models. In: Shalabh Heumann, C. (ed.) *Recent Advances in Linear Models and Related Areas*, pp. 105–133. Springer, New York (2008)
- Garg, G., Shalabh: Simultaneous predictions under exact restrictions in ultrastructural model. *Journal of Statistical Research* **45**(2), 139–154 (2011)
- Shalabh: A revisit to efficient forecasting in linear regression models. *J. Multivar. Anal.* **114**, 161–170 (2013)
- Yu, S.H.: Admissibility of linear predictor in multivariate linear model with arbitrary rank. *Sankhya, Ser. B* **66**, 621–633 (2004)
- He, D.J., Xu, X.Z.: Admissibility of linear predictors in the superpopulation model with respect to inequality constraints under matrix loss function. *Commun. Stat., Theory Methods* **21**, 3789–3799 (2011)
- Xu, L.W., Yu, S.H.: Admissible prediction in superpopulation models with random regression coefficients under matrix loss function. *J. Multivar. Anal.* **103**, 68–76 (2012)
- Hu, G.K., Peng, P.: Admissibility for linear estimators of regression coefficient in a general Gauss markoff model under balanced loss function. *J. Stat. Plan. Inference* **140**, 3365–3375 (2010)
- Hu, G.K., Peng, P.: All admissible linear estimators of a regression coefficient under a balanced loss function. *J. Multivar. Anal.* **102**, 1217–1224 (2011)
- Zhang, S.L., Gui, W.H.: Admissibility in general linear model with respect to an inequality constraint under balanced loss. *J. Inequal. Appl.* **2014**, 70 (2014)
- Rao, C.R.: Estimation of parameters in a linear model. *Ann. Stat.* **4**(6), 1023–1037 (1976)
- Chen, X.R.: *An Introduction of Mathematical Statistics*. Science Press, Beijing (1981)