APPLICATIONES MATHEMATICAE
50,2 (2023), pp. 97–106

Alexander Zaigraev (Toruń)

# A NOTE ON OPTIMAL JOINT PREDICTION OF ORDER STATISTICS

*Abstract.* The problem of prediction of several future order statistics, based on previous ones, is considered. An optimal predictor is defined as one minimizing the determinant of the covariance matrix of the predictor or of the predictive error vector. It is shown that the Lagrange multipliers method works well in all cases, despite some statements in the papers by Balakrishnan et al. [Metrika 85 (2022), 253–267; J. Multivariate Anal. 188 (2022), art. 104854; Statistics 57 (2023), 1239–1250].

**1. Introduction.** The problem of optimal joint prediction of several future order statistics, based on previous order statistics, is classical, and much literature is devoted to this topic. In this context it is worth mentioning the paper by Goldberger [G], whose ideas were developed by Kaminsky and Nelson [KN1]. They established the so-called marginal best linear unbiased predictor, that is, the optimal predictor for a single future order statistic. A review can be found in [KN2].

In the last two years Balakrishnan and Bhattacharya [BB1]–[BB3] and Balakrishnan and Mukerjee [BM1]–[BM3] published a series of papers on that problem in different settings.

Here, we are going to discuss the setting described in [BB1, BB2, BM1], concentrating on the results of [BB1].

Let $X_1, \dots, X_n$ be a sample from a location-scale family of distributions with unknown location parameter $\mu$ and unknown scale parameter $\sigma > 0$, and let $X_{1:n}, \dots, X_{n:n}$ be order statistics from this sample. Let $Z_{i:n} := (X_{i:n} - \mu)/\sigma$ be the corresponding order statistics corresponding to the stan-

2020 *Mathematics Subject Classification*: Primary 62G30; Secondary 62F10.
*Key words and phrases*: optimal prediction, order statistics, Lagrange multipliers.
Received 24 November 2023; revised 6 February 2024.
Published online 27 February 2024.

DOI: 10.4064/am2506-2-2024  

dard distribution with zero location and unit scale, and let $\alpha_i := \mathbb{E}Z_{i:n}$, $i = 1, \ldots, n$.

Given $2 \leqslant r < n$, define

$$\boldsymbol{X} := (X_{1:n}, \ldots, X_{r:n})', \quad \boldsymbol{Z} := (Z_{1:n}, \ldots, Z_{r:n})', \quad \boldsymbol{\alpha} := (\alpha_1, \ldots, \alpha_r)'.$$

Evidently, $\mathbb{E}\boldsymbol{X} = \mu\mathbf{1}_r + \sigma\boldsymbol{\alpha}$, where $\mathbf{1}_r$ denotes the column vector consisting of $r$ units. In addition, we denote by $\Sigma$ the covariance matrix of $\boldsymbol{Z}$. Then the covariance matrix of the vector $\boldsymbol{X}$ equals $\sigma^2\Sigma$.

We are interested in prediction of the vector

$$\boldsymbol{X}_* := (X_{u_1:n}, \ldots, X_{u_p:n})',$$

where $1 \leqslant p \leqslant n-r$, $r < u_1 < \cdots < u_p \leqslant n$. As in [BB1], the main attention will be paid to the cases $p = 2$ or $p = 3$; other cases can be considered similarly.

We consider linear predictors $A\boldsymbol{X}$ of $\boldsymbol{X}_*$ with $p \times r$ matrix $A$ satisfying the *unbiasedness* conditions, i.e. $\mathbb{E}A\boldsymbol{X} = \mathbb{E}\boldsymbol{X}_*$ for any $\mu$ and $\sigma$, which means that $A\mathbf{1}_r = \mathbf{1}_p$ and $A\boldsymbol{\alpha} = \boldsymbol{\alpha_*}$, where $\boldsymbol{\alpha}_* := (\alpha_{u_1}, \ldots, \alpha_{u_p})'$.

In order to determine the optimality of the predictor $A\boldsymbol{X}$, we take into account two matrices:

- the covariance matrix of the predictor, $\mathrm{Cov}(A\boldsymbol{X}) = \sigma^2 A\Sigma A'$,
- the covariance matrix of the predictive error vector,

$$\mathrm{Cov}(\boldsymbol{X}_* - A\boldsymbol{X}) = \sigma^2[A\Sigma A' - AW' - WA' + \mathrm{Cov}\boldsymbol{Z}_*],$$

  where $\boldsymbol{Z}_* := (Z_{u_1:n}, \ldots, Z_{u_p:n})'$, whereas the $p \times r$ matrix $W$ is the cross-covariance matrix of the vectors $\boldsymbol{Z}_*$ and $\boldsymbol{Z}$, i.e. $W := \mathrm{Cov}(\boldsymbol{Z}_*, \boldsymbol{Z})$.

As in [BB1], we use the $D$-optimality criterion to choose an optimal predictor. That is, we consider two problems:

(a) find a matrix $A$ such that $\det \mathrm{Cov}(A\boldsymbol{X})$ is minimal,
(b) find a matrix $B$ such that $\det \mathrm{Cov}(\boldsymbol{X}_* - B\boldsymbol{X})$ is minimal.

It should be noted that a more general optimality criterion is considered in [BM1], which allows the results obtained to be generalized.

We apply the well-known method of *Lagrange multipliers*, as is usual for an optimization problem with constraints. This method was also used in [BB1], although some of the statements there are invalid. In the later papers [BB2, BM1], the authors claimed that the method of Lagrange multipliers did not work in those special cases. We show that, on the contrary, the method, correctly applied, works in all those cases, although the proof is not straightforward.

**2. Optimal prediction of order statistics.** We start by recalling the well-known solution of both problems (a) and (b) for a single future order statistic to be predicted, i.e. for the case $p = 1$ (see, e.g., [KN1]). Denote

$$\Delta := \begin{bmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{12} & \Delta_{22} \end{bmatrix} := \begin{bmatrix} \mathbf{1}_r'\Sigma^{-1}\mathbf{1}_r & \mathbf{1}_r'\Sigma^{-1}\boldsymbol{\alpha} \\ \mathbf{1}_r'\Sigma^{-1}\boldsymbol{\alpha} & \boldsymbol{\alpha}'\Sigma^{-1}\boldsymbol{\alpha} \end{bmatrix}, \quad \det \Delta > 0.$$

In (a) the optimal predictor of $X_{u_1:n}$ has the form

$$\boldsymbol{a}'\boldsymbol{X} = \widehat{\mu} + \widehat{\sigma}\alpha_{u_1},$$

where

$$\widehat{\mu} := \frac{\Delta_{22}}{\det \Delta}\mathbf{1}_r'\Sigma^{-1}\boldsymbol{X} - \frac{\Delta_{12}}{\det \Delta}\boldsymbol{\alpha}'\Sigma^{-1}\boldsymbol{X},$$

$$\widehat{\sigma} := \frac{\Delta_{11}}{\det \Delta}\boldsymbol{\alpha}'\Sigma^{-1}\boldsymbol{X} - \frac{\Delta_{12}}{\det \Delta}\mathbf{1}_r'\Sigma^{-1}\boldsymbol{X}$$

(note that $\widehat{\mu}$ and $\widehat{\sigma}$ are the best linear unbiased estimators of $\mu$ and $\sigma$, respectively). In (b) the optimal predictor of $X_{u_1:n}$ has the form

$$\boldsymbol{b}'\boldsymbol{X} = \widehat{\mu} + \widehat{\sigma}\alpha_{u_1} + \boldsymbol{w}_{u_1}'\Sigma^{-1}(\boldsymbol{X} - \widehat{\mu}\mathbf{1}_r - \widehat{\sigma}\boldsymbol{\alpha}),$$

where $\boldsymbol{w}_{u_1} := \mathrm{Cov}(Z_{u_1:n}, \boldsymbol{Z})$.

As in [BB1, BM1], if $p > 1$, then we define a predictor of the vector $\boldsymbol{X}_*$ to be the *marginal predictor* if each of its components looks like $\boldsymbol{a}'\boldsymbol{X}$ in the case of (a) or $\boldsymbol{b}'\boldsymbol{X}$ in case of (b). That is, e.g., for $p = 2$ the marginal predictor has the form

$$\begin{bmatrix} \widehat{\mu} + \widehat{\sigma}\alpha_{u_1} \\ \widehat{\mu} + \widehat{\sigma}\alpha_{u_2} \end{bmatrix} \text{ in (a)}, \qquad \begin{bmatrix} \widehat{\mu} + \widehat{\sigma}\alpha_{u_1} + \boldsymbol{w}_{u_1}'\Sigma^{-1}(\boldsymbol{X} - \widehat{\mu}\mathbf{1}_r - \widehat{\sigma}\boldsymbol{\alpha}) \\ \widehat{\mu} + \widehat{\sigma}\alpha_{u_2} + \boldsymbol{w}_{u_2}'\Sigma^{-1}(\boldsymbol{X} - \widehat{\mu}\mathbf{1}_r - \widehat{\sigma}\boldsymbol{\alpha}) \end{bmatrix} \text{ in (b)}.$$

THEOREM 2.1. *In* (a), *the optimal predictor is not uniquely determined for* $p \geqslant 3$; *the marginal predictor is one of the solutions.*

REMARK 1. In [BB1] the authors, proving their Theorem 1, failed to notice that the marginal predictor is optimal, giving a complicated formula for an optimal predictor.

REMARK 2. In [BB1, Section 3.1] the authors claimed that there is no solution in (a) for $p = 3$. In the later paper [BM1] Balakrishnan and Mukerjee stated more cautiously that it was not possible to solve (a) with the Lagrange multipliers method. However, both statements are erroneous: the problem can be solved with this method, as will be seen from the proof below. It is worth emphasizing that for $p = 3$ the proof is in fact similar to that for $p = 2$.

*Proof of Theorem 2.1.* We prove the theorem for the cases $p = 2$ and $p = 3$ only; the proof for $p > 3$ is similar to that for $p = 3$. Considering a linear predictor $A\boldsymbol{X}$ of $\boldsymbol{X}_*$ in case $p = 2$, let us write

$$A := \begin{bmatrix} \boldsymbol{a}_1' \\ \boldsymbol{a}_2' \end{bmatrix}$$

with $r \times 1$ vectors $\boldsymbol{a}_1, \boldsymbol{a}_2$. Denote the matrix $A\Sigma A'$ by $V$, that is,

$$\begin{bmatrix} \boldsymbol{a}_1'\Sigma\boldsymbol{a}_1 & \boldsymbol{a}_1'\Sigma\boldsymbol{a}_2 \\ \boldsymbol{a}_1'\Sigma\boldsymbol{a}_2 & \boldsymbol{a}_2'\Sigma\boldsymbol{a}_2 \end{bmatrix} =: \begin{bmatrix} V_{11} & V_{12} \\ V_{12} & V_{22} \end{bmatrix},$$

and consider the following optimization problem:

$$\begin{aligned} Q(\boldsymbol{a}_1, \boldsymbol{a}_2) := \det V &- 2\lambda_1(\boldsymbol{a}_1'\mathbf{1}_r - 1) - 2\lambda_1^*(\boldsymbol{a}_1'\boldsymbol{\alpha} - \alpha_{u_1}) \\ &- 2\lambda_2(\boldsymbol{a}_2'\mathbf{1}_r - 1) - 2\lambda_2^*(\boldsymbol{a}_2'\boldsymbol{\alpha} - \alpha_{u_2}) \to \min_{\boldsymbol{a}_1,\boldsymbol{a}_2} \end{aligned}$$

under the unbiasedness conditions

$$\boldsymbol{a}_1'\mathbf{1}_r = 1, \quad \boldsymbol{a}_1'\boldsymbol{\alpha} = \alpha_{u_1}, \quad \boldsymbol{a}_2'\mathbf{1}_r = 1, \quad \boldsymbol{a}_2'\boldsymbol{\alpha} = \alpha_{u_2}.$$

Taking partial derivatives of $Q$ with respect to $\boldsymbol{a}_1$ and with respect to $\boldsymbol{a}_2$, and equating them to 0, we obtain

$$\begin{cases} V_{22}\Sigma\boldsymbol{a}_1 - V_{12}\Sigma\boldsymbol{a}_2 = \lambda_1\mathbf{1}_r + \lambda_1^*\boldsymbol{\alpha}, \\ V_{11}\Sigma\boldsymbol{a}_2 - V_{12}\Sigma\boldsymbol{a}_1 = \lambda_2\mathbf{1}_r + \lambda_2^*\boldsymbol{\alpha}. \end{cases} \tag{2.1}$$

Multiplying each equation in (2.1) by $\boldsymbol{a}_1'$ and $\boldsymbol{a}_2'$, we obtain four linear equations with four unknowns, $\lambda_1, \lambda_1^*, \lambda_2, \lambda_2^*$. The solution is unique, but there is no need to write it down.

Instead, let us obtain expressions for $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$. First, we multiply both sides of both equations of (2.1) by $\Sigma^{-1}$:

$$\begin{cases} V_{22}\boldsymbol{a}_1 - V_{12}\boldsymbol{a}_2 = \lambda_1\Sigma^{-1}\mathbf{1}_r + \lambda_1^*\Sigma^{-1}\boldsymbol{\alpha}, \\ V_{11}\boldsymbol{a}_2 - V_{12}\boldsymbol{a}_1 = \lambda_2\Sigma^{-1}\mathbf{1}_r + \lambda_2^*\Sigma^{-1}\boldsymbol{\alpha}. \end{cases} \tag{2.2}$$

Second, we solve (2.2) for $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$:

$$\begin{cases} \boldsymbol{a}_1 = \dfrac{V_{11}\lambda_1 + V_{12}\lambda_2}{\det V}\Sigma^{-1}\mathbf{1}_r + \dfrac{V_{11}\lambda_1^* + V_{12}\lambda_2^*}{\det V}\Sigma^{-1}\boldsymbol{\alpha}, \\ \\ \boldsymbol{a}_2 = \dfrac{V_{12}\lambda_1 + V_{22}\lambda_2}{\det V}\Sigma^{-1}\mathbf{1}_r + \dfrac{V_{12}\lambda_1^* + V_{22}\lambda_2^*}{\det V}\Sigma^{-1}\boldsymbol{\alpha}, \end{cases} \tag{2.3}$$

($\det V \neq 0$, since the rank of $V$ is 2). This is not the final result yet, since $V_{11}, V_{12}, V_{22}$ depend on $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$.

Applying the unbiasedness conditions to (2.3), we obtain four equations, which help us to get the final result:

$$\begin{cases} \dfrac{V_{11}\lambda_1 + V_{12}\lambda_2}{\det V}\Delta_{11} + \dfrac{V_{11}\lambda_1^* + V_{12}\lambda_2^*}{\det V}\Delta_{12} = 1, \\ \\ \dfrac{V_{11}\lambda_1 + V_{12}\lambda_2}{\det V}\Delta_{12} + \dfrac{V_{11}\lambda_1^* + V_{12}\lambda_2^*}{\det V}\Delta_{22} = \alpha_{u_1}, \end{cases} \tag{2.4}$$

$$
\left\{
\begin{aligned}
&\frac{V_{12}\lambda_1 + V_{22}\lambda_2}{\det V}\Delta_{11} + \frac{V_{12}\lambda_1^* + V_{22}\lambda_2^*}{\det V}\Delta_{12} = 1,\\
&\frac{V_{12}\lambda_1 + V_{22}\lambda_2}{\det V}\Delta_{12} + \frac{V_{12}\lambda_1^* + V_{22}\lambda_2^*}{\det V}\Delta_{22} = \alpha_{u_2}.
\end{aligned}
\right. \tag{2.5}
$$

Solving both systems of linear equations for

$$
\frac{V_{11}\lambda_1 + V_{12}\lambda_2}{\det V}, \quad \frac{V_{11}\lambda_1^* + V_{12}\lambda_2^*}{\det V},
$$

and

$$
\frac{V_{12}\lambda_1 + V_{22}\lambda_2}{\det V}, \quad \frac{V_{12}\lambda_1^* + V_{22}\lambda_2^*}{\det V},
$$

we obtain

$$
\frac{V_{11}\lambda_1 + V_{12}\lambda_2}{\det V} = \frac{\Delta_{22} - \alpha_{u_1}\Delta_{12}}{\det \Delta}, \quad \frac{V_{11}\lambda_1^* + V_{12}\lambda_2^*}{\det V} = \frac{\alpha_{u_1}\Delta_{11} - \Delta_{22}}{\det \Delta}, \tag{2.6}
$$

$$
\frac{V_{12}\lambda_1 + V_{22}\lambda_2}{\det V} = \frac{\Delta_{22} - \alpha_{u_2}\Delta_{12}}{\det \Delta}, \quad \frac{V_{12}\lambda_1^* + V_{22}\lambda_2^*}{\det V} = \frac{\alpha_{u_2}\Delta_{11} - \Delta_{22}}{\det \Delta}. \tag{2.7}
$$

After substitution of (2.6) and (2.7) into (2.3), we get the final result:

$$
\boldsymbol{a}_1 = \frac{\Delta_{22} - \alpha_{u_1}\Delta_{12}}{\det \Delta}\Sigma^{-1}\mathbf{1}_r + \frac{\alpha_{u_1}\Delta_{11} - \Delta_{12}}{\det \Delta}\Sigma^{-1}\boldsymbol{\alpha}, \tag{2.8}
$$

$$
\boldsymbol{a}_2 = \frac{\Delta_{22} - \alpha_{u_2}\Delta_{12}}{\det \Delta}\Sigma^{-1}\mathbf{1}_r + \frac{\alpha_{u_2}\Delta_{11} - \Delta_{12}}{\det \Delta}\Sigma^{-1}\boldsymbol{\alpha}. \tag{2.9}
$$

Thus, the optimal predictor for $\boldsymbol{X}_*$ in (a) has the form $A\boldsymbol{X} = \begin{bmatrix} \widehat{\mu} + \widehat{\sigma}\alpha_{u_1} \\ \widehat{\mu} + \widehat{\sigma}\alpha_{u_2} \end{bmatrix}$.

If $p = 3$, then

$$
A := \begin{bmatrix} \boldsymbol{a}_1' \\ \boldsymbol{a}_2' \\ \boldsymbol{a}_3' \end{bmatrix},
$$

and the optimization problem can be written as

$$
\begin{aligned}
&Q(\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)\\
&\quad := \det V - 2\lambda_1(\boldsymbol{a}_1'\mathbf{1}_r - 1) - 2\lambda_1^*(\boldsymbol{a}_1'\boldsymbol{\alpha} - \alpha_{u_1}) - 2\lambda_2(\boldsymbol{a}_2'\mathbf{1}_r - 1)\\
&\qquad - 2\lambda_2^*(\boldsymbol{a}_2'\boldsymbol{\alpha} - \alpha_{u_2}) - 2\lambda_3(\boldsymbol{a}_3'\mathbf{1}_r - 1) - 2\lambda_3^*(\boldsymbol{a}_3'\boldsymbol{\alpha} - \alpha_{u_3}) \to \min_{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3}
\end{aligned}
$$

under the unbiasedness conditions

$$
\boldsymbol{a}_1'\mathbf{1}_r = 1,\ \boldsymbol{a}_1'\boldsymbol{\alpha} = \alpha_{u_1},\ \boldsymbol{a}_2'\mathbf{1}_r = 1,\ \boldsymbol{a}_2'\boldsymbol{\alpha} = \alpha_{u_2},\ \boldsymbol{a}_3'\mathbf{1}_r = 1,\ \boldsymbol{a}_3'\boldsymbol{\alpha} = \alpha_{u_3}.
$$

Here

$$V := \begin{bmatrix} V_{11} & V_{12} & V_{13} \\ V_{12} & V_{22} & V_{23} \\ V_{13} & V_{23} & V_{33} \end{bmatrix} := \begin{bmatrix} \boldsymbol{a}_1'\Sigma\boldsymbol{a}_1 & \boldsymbol{a}_1'\Sigma\boldsymbol{a}_2 & \boldsymbol{a}_1'\Sigma\boldsymbol{a}_3 \\ \boldsymbol{a}_1'\Sigma\boldsymbol{a}_2 & \boldsymbol{a}_2'\Sigma\boldsymbol{a}_2 & \boldsymbol{a}_2'\Sigma\boldsymbol{a}_3 \\ \boldsymbol{a}_1'\Sigma\boldsymbol{a}_3 & \boldsymbol{a}_2'\Sigma\boldsymbol{a}_3 & \boldsymbol{a}_3'\Sigma\boldsymbol{a}_3 \end{bmatrix}.$$

Taking partial derivatives of $Q$ with respect to $\boldsymbol{a}_1, \boldsymbol{a}_2$ and $\boldsymbol{a}_3$, and equating them to 0, we obtain a system of equations, which is the analogue of (2.1):

$$(2.10)\quad \begin{cases} [V_{22}V_{33}-V_{23}^2]\Sigma\boldsymbol{a}_1+[V_{13}V_{23}-V_{12}V_{33}]\Sigma\boldsymbol{a}_2+[V_{12}V_{23}-V_{13}V_{22}]\Sigma\boldsymbol{a}_3 \\ \qquad = \lambda_1\mathbf{1}_r + \lambda_1^*\boldsymbol{\alpha}, \\ [V_{13}V_{23}-V_{12}V_{33}]\,\Sigma\boldsymbol{a}_1+\left[V_{11}V_{33}-V_{13}^2\right]\,\Sigma\boldsymbol{a}_2+[V_{12}V_{13}-V_{23}V_{11}]\,\Sigma\boldsymbol{a}_3 \\ \qquad = \lambda_2\mathbf{1}_r + \lambda_2^*\boldsymbol{\alpha}, \\ [V_{12}V_{23}-V_{13}V_{22}]\,\Sigma\boldsymbol{a}_1+[V_{12}V_{13}-V_{23}V_{11}]\,\Sigma\boldsymbol{a}_2+\left[V_{11}V_{22}-V_{12}^2\right]\,\Sigma\boldsymbol{a}_3 \\ \qquad = \lambda_3\mathbf{1}_r + \lambda_3^*\boldsymbol{\alpha}. \end{cases}$$

A simple analysis of (2.10) leads to the following conclusion: all $\lambda_i, \lambda_i^*$, $i = 1, 2, 3$, are zeroes. Indeed, if we multiply each equation in (2.10) on the left by $\boldsymbol{a}_1'$, by $\boldsymbol{a}_2'$ and by $\boldsymbol{a}_3'$, we obtain nine equalities for the six parameters $\lambda_1, \lambda_1^*, \lambda_2, \lambda_2^*, \lambda_3, \lambda_3^*$ :

$$\begin{array}{lll} \det V = \lambda_1 + \lambda_1^*\alpha_{u_1}, & 0 = \lambda_1 + \lambda_1^*\alpha_{u_2}, & 0 = \lambda_1 + \lambda_1^*\alpha_{u_3}, \\ 0 = \lambda_2 + \lambda_2^*\alpha_{u_1}, & \det V = \lambda_2 + \lambda_2^*\alpha_{u_2}, & 0 = \lambda_2 + \lambda_2^*\alpha_{u_3}, \\ 0 = \lambda_3 + \lambda_3^*\alpha_{u_1}, & 0 = \lambda_3 + \lambda_3^*\alpha_{u_2}, & \det V = \lambda_3 + \lambda_3^*\alpha_{u_3}. \end{array}$$

As a solution, we get $\lambda_i = \lambda_i^* = 0$, $i = 1, 2, 3$, and, in addition, $\det V = 0$. This means that the unconditional minimum of $\det V$, which is 0, lies on the surface defined by the constraints. Therefore, the vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ are linearly dependent and belong either to a two-dimensional or to a one-dimensional subspace of $\mathbb{R}^r$. Assume that they belong to a two-dimensional subspace, that is, there exist two linearly independent vectors $\boldsymbol{\zeta}, \boldsymbol{\nu} \in \mathbb{R}^r$ such that

$$\boldsymbol{a}_1 = s_1\boldsymbol{\zeta} + s_2\boldsymbol{\nu}, \qquad \boldsymbol{a}_2 = t_1\boldsymbol{\zeta} + t_2\boldsymbol{\nu}, \qquad \boldsymbol{a}_3 = y_1\boldsymbol{\zeta} + y_2\boldsymbol{\nu}.$$

The vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ should satisfy the unbiasedness conditions $\boldsymbol{a}_1'\mathbf{1}_r = 1$, $\boldsymbol{a}_2'\mathbf{1}_r = 1$, $\boldsymbol{a}_3'\mathbf{1}_r = 1$, $\boldsymbol{a}_1'\boldsymbol{\alpha} = \alpha_{u_1}$, $\boldsymbol{a}_2'\boldsymbol{\alpha} = \alpha_{u_2}$, $\boldsymbol{a}_3'\boldsymbol{\alpha} = \alpha_{u_3}$. From those equalities we calculate $s_1, s_2, t_1, t_2, y_1, y_2$, and obtain the final result under the condition that $\boldsymbol{\zeta}'\mathbf{1}_r \cdot \boldsymbol{\nu}'\boldsymbol{\alpha} \neq \boldsymbol{\zeta}'\boldsymbol{\alpha} \cdot \boldsymbol{\nu}'\mathbf{1}_r$:

$$(2.11)\qquad \boldsymbol{a}_1 = \frac{\boldsymbol{\nu}'\boldsymbol{\alpha} - \alpha_{u_1}\boldsymbol{\nu}'\mathbf{1}_r}{\boldsymbol{\zeta}'\mathbf{1}_r \cdot \boldsymbol{\nu}'\boldsymbol{\alpha} - \boldsymbol{\zeta}'\boldsymbol{\alpha} \cdot \boldsymbol{\nu}'\mathbf{1}_r}\boldsymbol{\zeta} + \frac{\alpha_{u_1}\boldsymbol{\zeta}'\mathbf{1}_r - \boldsymbol{\zeta}'\boldsymbol{\alpha}}{\boldsymbol{\zeta}'\mathbf{1}_r \cdot \boldsymbol{\nu}'\boldsymbol{\alpha} - \boldsymbol{\zeta}'\boldsymbol{\alpha} \cdot \boldsymbol{\nu}'\mathbf{1}_r}\boldsymbol{\nu},$$

$$(2.12)\qquad \boldsymbol{a}_2 = \frac{\boldsymbol{\nu}'\boldsymbol{\alpha} - \alpha_{u_2}\boldsymbol{\nu}'\mathbf{1}_r}{\boldsymbol{\zeta}'\mathbf{1}_r \cdot \boldsymbol{\nu}'\boldsymbol{\alpha} - \boldsymbol{\zeta}'\boldsymbol{\alpha} \cdot \boldsymbol{\nu}'\mathbf{1}_r}\boldsymbol{\zeta} + \frac{\alpha_{u_2}\boldsymbol{\zeta}'\mathbf{1}_r - \boldsymbol{\zeta}'\boldsymbol{\alpha}}{\boldsymbol{\zeta}'\mathbf{1}_r \cdot \boldsymbol{\nu}'\boldsymbol{\alpha} - \boldsymbol{\zeta}'\boldsymbol{\alpha} \cdot \boldsymbol{\nu}'\mathbf{1}_r}\boldsymbol{\nu},$$

$$(2.13)\qquad \boldsymbol{a}_3 = \frac{\boldsymbol{\nu}'\boldsymbol{\alpha} - \alpha_{u_3}\boldsymbol{\nu}'\mathbf{1}_r}{\boldsymbol{\zeta}'\mathbf{1}_r \cdot \boldsymbol{\nu}'\boldsymbol{\alpha} - \boldsymbol{\zeta}'\boldsymbol{\alpha} \cdot \boldsymbol{\nu}'\mathbf{1}_r}\boldsymbol{\zeta} + \frac{\alpha_{u_3}\boldsymbol{\zeta}'\mathbf{1}_r - \boldsymbol{\zeta}'\boldsymbol{\alpha}}{\boldsymbol{\zeta}'\mathbf{1}_r \cdot \boldsymbol{\nu}'\boldsymbol{\alpha} - \boldsymbol{\zeta}'\boldsymbol{\alpha} \cdot \boldsymbol{\nu}'\mathbf{1}_r}\boldsymbol{\nu}.$$

Similarly, when the vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ are assumed to belong to a one-dimensional subspace of $\mathbb{R}^r$, we come to the conclusion that this is not possible since the unbiasedness conditions cannot be satisfied. So, the general solution of problem (a) for $p = 3$ is given by (2.11)–(2.13).

If we take $\zeta = \Sigma^{-1}\mathbf{1}_r$ and $\nu = \Sigma^{-1}\boldsymbol{\alpha}$, then (2.11)–(2.13) can be rewritten in the form

$$\boldsymbol{a}_1 = \frac{\Delta_{22} - \alpha_{u_1}\Delta_{12}}{\det \Delta}\Sigma^{-1}\mathbf{1}_r + \frac{\alpha_{u_1}\Delta_{11} - \Delta_{12}}{\det \Delta}\Sigma^{-1}\boldsymbol{\alpha}, \tag{2.14}$$

$$\boldsymbol{a}_2 = \frac{\Delta_{22} - \alpha_{u_2}\Delta_{12}}{\det \Delta}\Sigma^{-1}\mathbf{1}_r + \frac{\alpha_{u_2}\Delta_{11} - \Delta_{12}}{\det \Delta}\Sigma^{-1}\boldsymbol{\alpha}, \tag{2.15}$$

$$\boldsymbol{a}_3 = \frac{\Delta_{22} - \alpha_{u_3}\Delta_{12}}{\det \Delta}\Sigma^{-1}\mathbf{1}_r + \frac{\alpha_{u_3}\Delta_{11} - \Delta_{12}}{\det \Delta}\Sigma^{-1}\boldsymbol{\alpha}, \tag{2.16}$$

and we obtain a particular solution, where an optimal predictor for $\boldsymbol{X}_*$ has the form

$$A\boldsymbol{X} = \begin{bmatrix} \widehat{\mu} + \widehat{\sigma}\alpha_{u_1} \\ \widehat{\mu} + \widehat{\sigma}\alpha_{u_2} \\ \widehat{\mu} + \widehat{\sigma}\alpha_{u_3} \end{bmatrix}. \ \blacksquare$$

REMARK 3. Balakrishnan and Bhattacharia [BB1, Section 3.2] claimed that there is no solution of (a) for $p = 2$ in the case of a scale family of distributions. Again, it is an erroneous statement. We can solve the problem of optimal prediction with the Lagrange multipliers method in the same way as above, and obtain the following result (non-unique): $\min_{\boldsymbol{a}_1,\boldsymbol{a}_2} \det V = 0$ and it is attained when $\boldsymbol{a}_1, \boldsymbol{a}_2$ are such that $\boldsymbol{a}_1 = (\alpha_{u_1}/\alpha_{u_2})\boldsymbol{a}_2$, while $\boldsymbol{a}_2$ is any vector satisfying $\boldsymbol{a}_2'\boldsymbol{\alpha} = \alpha_{u_2}$. In particular, one can take $\boldsymbol{a}_2 = (\alpha_{u_2}/\Delta_{22})\Sigma^{-1}\boldsymbol{\alpha}$ and, as a consequence,

$$A\boldsymbol{X} = \begin{bmatrix} \widehat{\sigma}\alpha_{u_1} \\ \widehat{\sigma}\alpha_{u_2} \end{bmatrix},$$

where $\widehat{\sigma} := \boldsymbol{\alpha}'\Sigma^{-1}\boldsymbol{X}/\Delta_{22}$ is the best linear unbiased estimator of $\sigma$ in this case.

THEOREM 2.2. *In* (b), *the optimal predictor is uniquely determined and it is the marginal predictor.*

*Proof.* We prove the statement for $p = 2$ only; for $p > 2$ the proof is similar. Consider linear predictors $B\boldsymbol{X}$ of $\boldsymbol{X}_*$, and write

$$B := \begin{bmatrix} \boldsymbol{b}_1' \\ \boldsymbol{b}_2' \end{bmatrix}$$

with $r \times 1$ vectors $\boldsymbol{b}_1, \boldsymbol{b}_2$. Denote the matrix $B\Sigma B' - BW' - WB' + \mathrm{Cov}\boldsymbol{Z}_*$

by $H$, that is,

$$\begin{bmatrix} \boldsymbol{b}_1'\Sigma\boldsymbol{b}_1 - 2\boldsymbol{b}_1'\boldsymbol{w}_{u_1} + z_{11} & \boldsymbol{b}_1'\Sigma\boldsymbol{b}_2 - \boldsymbol{b}_1'\boldsymbol{w}_{u_2} - \boldsymbol{b}_2'\boldsymbol{w}_{u_1} + z_{12} \\ \boldsymbol{b}_1'\Sigma\boldsymbol{b}_2 - \boldsymbol{b}_1'\boldsymbol{w}_{u_2} - \boldsymbol{b}_2'\boldsymbol{w}_{u_1} + z_{12} & \boldsymbol{b}_2'\Sigma\boldsymbol{b}_2 - 2\boldsymbol{b}_2'\boldsymbol{w}_{u_2} + z_{22} \end{bmatrix} =: \begin{bmatrix} H_{11} & H_{12} \\ H_{12} & H_{22} \end{bmatrix},$$

where $\boldsymbol{w}_{u_1} := \mathrm{Cov}(\boldsymbol{Z}_{u_1:n},\ \boldsymbol{Z}), \boldsymbol{w}_{u_2} := \mathrm{Cov}(\boldsymbol{Z}_{u_2:n}, \boldsymbol{Z})$, whereas $\{z_{ij}\}_{i,j=1}^2$ are the elements of the matrix $\mathrm{Cov}\,\boldsymbol{Z}_*$.

In what follows, we repeat the solution given in Theorem 2.1. Aiming to minimize $\det H$ under the unbiasedness conditions, we create the function $Q = Q(\boldsymbol{b}_1, \boldsymbol{b}_2)$. Equating to 0 its partial derivatives with respect to $\boldsymbol{b}_1$ and with respect to $\boldsymbol{b}_2$, we obtain

$$(2.1') \qquad \begin{cases} H_{22}(\Sigma\boldsymbol{b}_1 - \boldsymbol{w}_{u_1}) - H_{12}(\Sigma\boldsymbol{b}_2 - \boldsymbol{w}_{u_2}) = \lambda_1\mathbf{1}_r + \lambda_1^*\boldsymbol{\alpha}, \\ H_{11}(\Sigma\boldsymbol{b}_2 - \boldsymbol{w}_{u_2}) - H_{12}(\Sigma\boldsymbol{b}_1 - \boldsymbol{w}_{u_1}) = \lambda_2\mathbf{1}_r + \lambda_2^*\boldsymbol{\alpha}, \end{cases}$$

$$(2.2') \qquad \begin{cases} H_{22}(\boldsymbol{b}_1 - \Sigma^{-1}\boldsymbol{w}_{u_1}) - H_{12}(\boldsymbol{b}_2 - \Sigma^{-1}\boldsymbol{w}_{u_2}) = \lambda_1\Sigma^{-1}\mathbf{1}_r + \lambda_1^*\Sigma^{-1}\boldsymbol{\alpha}, \\ H_{11}(\boldsymbol{b}_2 - \Sigma^{-1}\boldsymbol{w}_{u_2}) - H_{12}(\boldsymbol{b}_1 - \Sigma^{-1}\boldsymbol{w}_{u_1}) = \lambda_2\Sigma^{-1}\mathbf{1}_r + \lambda_2^*\Sigma^{-1}\boldsymbol{\alpha}, \end{cases}$$

$$(2.3') \qquad \begin{cases} \boldsymbol{b}_1 = \Sigma^{-1}\boldsymbol{w}_{u_1} + \dfrac{H_{11}\lambda_1 + H_{12}\lambda_2}{\det H}\Sigma^{-1}\mathbf{1}_r + \dfrac{H_{11}\lambda_1^* + H_{12}\lambda_2^*}{\det H}\Sigma^{-1}\boldsymbol{\alpha}, \\ \boldsymbol{b}_2 = \Sigma^{-1}\boldsymbol{w}_{u_2} + \dfrac{H_{12}\lambda_1 + H_{22}\lambda_2}{\det H}\Sigma^{-1}\mathbf{1}_r + \dfrac{H_{12}\lambda_1^* + H_{22}\lambda_2^*}{\det H}\Sigma^{-1}\boldsymbol{\alpha}, \end{cases}$$

($\det H \neq 0$, since the rank of the matrix $H$ is 2). From the unbiasedness conditions we have

$$(2.4') \qquad \begin{cases} \dfrac{H_{11}\lambda_1 + H_{12}\lambda_2}{\det H}\Delta_{11} + \dfrac{H_{11}\lambda_1^* + H_{12}\lambda_2^*}{\det H}\Delta_{12} = 1 - \mathbf{1}_r'\Sigma^{-1}\boldsymbol{w}_{u_1}, \\ \dfrac{H_{11}\lambda_1 + H_{12}\lambda_2}{\det H}\Delta_{12} + \dfrac{H_{11}\lambda_1^* + H_{12}\lambda_2^*}{\det H}\Delta_{22} = \alpha_{u_1} - \boldsymbol{\alpha}'\Sigma^{-1}\boldsymbol{w}_{u_1}, \end{cases}$$

$$(2.5') \qquad \begin{cases} \dfrac{H_{12}\lambda_1 + H_{22}\lambda_2}{\det H}\Delta_{11} + \dfrac{H_{12}\lambda_1^* + H_{22}\lambda_2^*}{\det H}\Delta_{12} = 1 - \mathbf{1}_r'\Sigma^{-1}\boldsymbol{w}_{u_2}, \\ \dfrac{H_{12}\lambda_1 + H_{22}\lambda_2}{\det H}\Delta_{12} + \dfrac{H_{12}\lambda_1^* + H_{22}\lambda_2^*}{\det H}\Delta_{22} = \alpha_{u_2} - \boldsymbol{\alpha}'\Sigma^{-1}\boldsymbol{w}_{u_2}. \end{cases}$$

Solving both systems of linear equations we obtain

$$(2.6') \qquad \begin{cases} \dfrac{H_{11}\lambda_1 + H_{12}\lambda_2}{\det H} = \dfrac{(1 - \mathbf{1}_r'\Sigma^{-1}\boldsymbol{w}_{u_1})\Delta_{22} - (\alpha_{u_1} - \boldsymbol{\alpha}'\Sigma^{-1}\boldsymbol{w}_{u_1})\Delta_{12}}{\Delta}, \\ \dfrac{H_{11}\lambda_1^* + H_{12}\lambda_2^*}{\det H} = \dfrac{(\alpha_{u_1} - \boldsymbol{\alpha}'\Sigma^{-1}\boldsymbol{w}_{u_1})\Delta_{11} - (1 - \mathbf{1}_r'\Sigma^{-1}\boldsymbol{w}_{u_1})\Delta_{12}}{\Delta}, \end{cases}$$

$$(2.7') \quad \begin{cases} \dfrac{H_{12}\lambda_1+H_{22}\lambda_2}{\det H} = \dfrac{(1-\mathbf{1}_r'\Sigma^{-1}\boldsymbol{w}_{u_2})\Delta_{22}-(\alpha_{u_2}-\boldsymbol{\alpha}'\Sigma^{-1}\boldsymbol{w}_{u_2})\Delta_{12}}{\Delta}, \\ \dfrac{H_{12}\lambda_1^*+H_{22}\lambda_2^*}{\det H} = \dfrac{(\alpha_{u_2}-\boldsymbol{\alpha}'\Sigma^{-1}\boldsymbol{w}_{u_2})\Delta_{11}-(1-\mathbf{1}_r'\Sigma^{-1}\boldsymbol{w}_{u_2})\Delta_{12}}{\Delta}. \end{cases}$$

After substitution of (2.6′) and (2.7′) to (2.3′), we get the final result:

$$(2.8') \quad \begin{aligned} \boldsymbol{b}_1 = \Sigma^{-1}\boldsymbol{w}_{u_1} &+ \frac{(1-\mathbf{1}_r'\Sigma^{-1}\boldsymbol{w}_{u_1})\Delta_{22}-(\alpha_{u_1}-\boldsymbol{\alpha}'\Sigma^{-1}\boldsymbol{w}_{u_1})\Delta_{12}}{\Delta}\Sigma^{-1}\mathbf{1}_r \\ &+ \frac{(\alpha_{u_1}-\boldsymbol{\alpha}'\Sigma^{-1}\boldsymbol{w}_{u_1})\Delta_{11}-(1-\mathbf{1}_r'\Sigma^{-1}\boldsymbol{w}_{u_1})\Delta_{12}}{\Delta}\Sigma^{-1}\boldsymbol{\alpha}, \end{aligned}$$

$$(2.9') \quad \begin{aligned} \boldsymbol{b}_2 = \Sigma^{-1}\boldsymbol{w}_{u_2} &+ \frac{(1-\mathbf{1}_r'\Sigma^{-1}\boldsymbol{w}_{u_2})\Delta_{22}-(\alpha_{u_2}-\boldsymbol{\alpha}'\Sigma^{-1}\boldsymbol{w}_{u_2})\Delta_{12}}{\Delta}\Sigma^{-1}\mathbf{1}_r \\ &+ \frac{(\alpha_{u_2}-\boldsymbol{\alpha}'\Sigma^{-1}\boldsymbol{w}_{u_2})\Delta_{11}-(1-\mathbf{1}_r'\Sigma^{-1}\boldsymbol{w}_{u_2})\Delta_{12}}{\Delta}\Sigma^{-1}\boldsymbol{\alpha}. \end{aligned}$$

So, the optimal predictor $B_*\boldsymbol{X}$ of $\boldsymbol{X}_*$ in the case of (b) has the form

$$B_*\boldsymbol{X} = \begin{bmatrix} \widehat{\mu}+\widehat{\sigma}\alpha_{u_1}+\boldsymbol{w}_{u_1}'\Sigma^{-1}(\boldsymbol{X}-\widehat{\mu}\mathbf{1}_r-\widehat{\sigma}\boldsymbol{\alpha}) \\ \widehat{\mu}+\widehat{\sigma}\alpha_{u_2}+\boldsymbol{w}_{u_2}'\Sigma^{-1}(\boldsymbol{X}-\widehat{\mu}\mathbf{1}_r-\widehat{\sigma}\boldsymbol{\alpha}) \end{bmatrix}. \ \blacksquare$$

REMARK 4. In general, $\det \mathrm{Cov}(\boldsymbol{X}_*-B_*\boldsymbol{X}) \neq 0$, since for all $\boldsymbol{c} \neq \mathbf{0} \in \mathbb{R}^p$,

$$\begin{aligned} \boldsymbol{c}'\mathrm{Cov}(\boldsymbol{X}_*-B_*\boldsymbol{X})\boldsymbol{c} &= \sigma^2\boldsymbol{c}'\mathbb{E}(\boldsymbol{Z}_*-B_*\boldsymbol{Z})(\boldsymbol{Z}_*'-\boldsymbol{Z}'B_*')\boldsymbol{c} \\ &= \sigma^2\mathbb{E}(\boldsymbol{c}'\boldsymbol{Z}_*-\boldsymbol{c}'B_*\boldsymbol{Z})^2 > 0. \end{aligned}$$

The exceptional case is $W = \mathbf{0}$, $\mathrm{Cov}\boldsymbol{Z}_* = \mathbf{0}$ for $p \geqslant 3$.

REMARK 5. From the solution in (b) one can obtain the solution of (a) by formally setting $W = \mathbf{0}$, $\mathrm{Cov}\,\boldsymbol{Z}_* = \mathbf{0}$.

It is also worth pointing out that in Section 4 of [BB1], devoted to numerical computations, the authors compare the optimal predictor in (a) with the non-optimal predictor, namely with the marginal predictor obtained for (b), aiming to show the advantage of the first predictor if the criterion is the determinant of the covariance matrix of the predictor. But since the optimal predictor in (a), under $D$-optimality, coincides with that under $A$-optimality, which can also be proved by the Lagrange multipliers method (and it also follows immediately from the more general result proved by Balakrishnan and coathors in [BB2, BM1]), it does not make sense to define 'efficiency loss' as 1 minus trace-efficiency, and then to define 'overall efficiency gain' as 'efficiency gain' minus 'efficiency loss' [BB1, Section 4].

**Acknowledgements.** The author is grateful to the reviewer for helpful corrections that improved the paper.

**References**

[BB1] N. Balakrishnan and R. Bhattacharya, *D-optimal joint best linear unbiased prediction of order statistics*, Metrika 85 (2022), 253–267.

[BB2] N. Balakrishnan and R. Bhattacharya, *On simultaneous best linear unbiased prediction of future order statistics and associated properties*, J. Multivariate Anal. 188 (2022), art. 104854, 11 pp.

[BB3] N. Balakrishnan and R. Bhattacharya, *Simultaneous best linear invariant prediction of future order statistics for location-scale and associated optimality properties*, Statistics 56 (2022), 1409–1426.

[BM1] N. Balakrishnan and R. Mukerjee, *On optimal joint prediction of order statistics*, Statistics 57 (2023), 1239–1250.

[BM2] N. Balakrishnan and R. Mukerjee, *Joint linear estimation and prediction based on order statistics in a linear model*, preprint, 2023.

[BM3] N. Balakrishnan and R. Mukerjee, *An explicit multiple case deletion formula for a linear model with correlation and a resulting property of the BLUP of a multivariate predictand*, preprint, 2023.

[G] A. S. Goldberger, *Best linear unbiased prediction in the generalized linear regression model*, J. Amer. Statist. Assoc. 57 (1962), 369–375.

[KN1] K. S. Kaminsky and P. I. Nelson, *Best linear unbiased prediction of order statistics in location and scale families*, J. Amer. Statist. Assoc. 70 (1975), 145–150.

[KN2] K. S. Kaminsky and P. I. Nelson, *Prediction of order statistics*, in: Handbook of Statistics – 17: Order Statistics – Applications, N. Balakrishnan and C. R. Rao (eds.), Chapter 17, North-Holland, Amsterdam, 1998, 431–450.

Alexander Zaigraev
Faculty of Mathematics and Computer Science
Nicolaus Copernicus University
87-100 Toruń, Poland
E-mail: alzaig@mat.umk.pl
ORCID: 0000-0002-6261-031X