A RANK SELECTION OF MV-PURE WITH AN UNBIASED PREDICTED-MSE CRITERION AND ITS EFFICIENT IMPLEMENTATION IN IMAGE RESTORATION

Masao Yamagishi and Isao Yamada

†Department of Communications and Computer Engineering, Tokyo Institute of Technology, Japan
Email:{myamagi, isao}@sp.ce.titech.ac.jp

ABSTRACT

The Minimum-Variance Pseudo-Unbiased Reduced-rank Estimator (MV-PURE) is designed, as a natural reduced-rank extension of the Gauss-Markov estimator, for the unknown deterministic vector in ill-conditioned linear regression model. In this paper, we propose a novel rank-selection for the MV-PURE to achieve a small Mean Square Error (MSE). The proposed rank-selection is realized by minimizing an unbiased estimate of the predicted-MSE, not of the MSE. Our unbiased estimate can be applicable to any noise distribution with zero mean and a finite covariance matrix, while Stein-type unbiased criteria cannot in general. We apply the proposed selection to an image restoration problem and introduce its efficient $\mathcal{O}(m \log m)$ implementation by using a special structure found in typical blur matrices, where the blur matrix is of size $m \times m$. A numerical example demonstrates that the MV-PURE with the proposed rank-selection achieves a MSE comparable with the minimal MSE for the unknown vector among all possible ranks.

Index Terms— linear model, ill-conditioned, reduced-rank estimator, unbiased predicted-MSE criterion, image restoration

1. INTRODUCTION

Ill-conditioned linear parameter estimation problems arise in wide range signal processing applications. A main goal of the problems is to estimate an unknown deterministic vector $\beta \in \mathbb{R}^n$ by using an observation $y \in \mathbb{R}^m$ modeled by the multiplication of β and a known model matrix $L \in \mathbb{R}^{m \times n}$ having a small singular value, with additive noise. Unfortunately, the Mean Square Error (MSE) of β, a standard criterion of estimate, cannot be minimized without complete knowledge on the unknown vector. For this reason, many estimators have been proposed to suppress somehow the MSE. One of the most well-known classical techniques is the Best Linear Unbiased Estimator (BLUE, or Gauss-Markov Estimator) [1, 2, 3], which works poorly because small perturbation in the observation y may result in unacceptably large error in the estimate.

For the ill-conditioned problems, the Minimum Variance Pseudo-Unbiased Reduced-rank Estimator (MV-PURE) has been proposed [4, 5]. The MV-PURE is designed, as a natural reduced-rank extension of the BLUE, to achieve a small MSE. The MV-PURE includes a classical reduced rank estimator [6, 7] as its special example. In many situations, by imposing an appropriate rank constraint, the MV-PURE outperforms commonly used regularization techniques [8]. Practical applications of the MV-PURE are found in e.g. [9, 10].

In this paper, we propose a rank selection criterion of the MV-PURE based on an unbiased estimate of the predicted-MSE, i.e., the MSE for $L\beta$. Suppressing our proposed criterion is necessary because the criterion is a lower bound of the MSE for β up to a constant factor in the mean value and is much more robust in terms of the variance than an unbiased criterion of the MSE for β against the ill-conditioned case. Thanks to focusing on the MV-PURE (more generally, affine estimators), our proposed unbiased predicted-MSE criterion can be applicable to any noise distribution with zero mean and a finite covariance matrix, while Stein-type unbiased criteria (e.g., [11, 12]) cannot in general. To examine the performance, the MV-PURE with the proposed rank selection is applied to an image restoration problem [13, 14]. Fortunately, in typical scenarios, we obtain an efficient $\mathcal{O}(m \log m)$ implementation of the MV-PURE with the rank minimizing the proposed criterion, where we assume as in [15] the model matrix of size $m \times m$ is diagonalizable by a computationally efficient orthogonal matrix (e.g. DCT, FFT, etc). The numerical example demonstrates that (i) the proposed unbiased predicted-MSE criterion is much more robust than an unbiased criterion of the MSE against the ill-conditioned blur matrix, (ii) the rank minimizing the proposed criterion is almost identical to the one minimizing the MSE, and (iii) the MV-PURE with the proposed rank selection can be implemented efficiently and achieves a MSE comparable with the minimal MSE for the unknown vector among all possible ranks.

2. PRELIMINARIES

2.1. Linear regression model

In the linear regression model, it is assumed that we can observe a data vector $y \in \mathbb{R}^n$ modeled by

$$y = L\beta + \epsilon, \tag{1}$$

where $L \in \mathbb{R}^{n \times m}$ is a known model matrix (of full column rank m) with an SVD[1]

$$L = U\Sigma V^t = \sum_{i=1}^{m} \sigma_i u_i v_i^t, \tag{2}$$

This work was supported in part by JSPS Grants-in-Aid (B-21300091) and JSPS Grant-in-Aid for Research Activity start-up (24800022).

[1]The *singular value decomposition* (SVD) of a matrix $X \in \mathbb{R}^{n \times m}$ is given by

$$X = P\Sigma Q^t = \sum_{i=1}^{\min(m,n)} \sigma_i p_i q_i^t,$$

where $P = (p_1, p_2, \ldots, p_n) \in \mathbb{R}^{n \times n}, Q = (q_1, q_2, \ldots, q_m) \in \mathbb{R}^{m \times m}$ are orthogonal matrices and $\Sigma \in \mathbb{R}^{n \times m}$ contains on its main diagonal the singular values $\sigma_1 \ldots, \sigma_{\min(m,n)}$ of X and 0's elsewhere. Without losing generality, we assume that all SVDs considered have singular values organized in nonincreasing order, that is, $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_{\mathrm{rk}(X)} > 0$ and $\sigma_s = 0$ for $s > \mathrm{rk}(X)$, where $\mathrm{rk}(X)$ stands for the rank of X.

$\beta \in \mathbb{R}^m$ is an unknown deterministic vector to be estimated and $\epsilon \in \mathbb{R}^n$ is a random noise vector with zero mean and positive definite covariance matrix $E[\epsilon\epsilon^t] = \sigma^2 Q \in \mathbb{R}^{n\times n}$, where E denotes expectation and $\sigma > 0$. Assume that we have auxiliary knowledge that the unknown vector $\beta \in \mathbb{R}^m$ is an element of the null set of a matrix $A \in \mathbb{R}^{s\times m}$ ($\mathrm{rk}(A) \le m$)

$$\mathcal{N}(A) := \{\tilde{\beta} \in \mathbb{R}^m \mid A\tilde{\beta} = 0\} \neq \varnothing$$

onto which the metric projection is given as

$$P_{\mathcal{N}(A)} = N \begin{pmatrix} I_{m-\mathrm{rk}(A)} & O \\ O & O \end{pmatrix} N^t,$$

where $N \in \mathbb{R}^{m\times m}$ is an orthonormal matrix defined by eigen-decomposition of A^tA.

The goal of the linear estimation problem is to estimate the unknown vector $\beta = P_{\mathcal{N}(A)} z$ ($\forall z \in \beta + \mathcal{N}(A)^\perp$) by

$$\hat{\beta} := \Phi y,$$

where $\Phi \in \mathbb{R}^{m\times n}$ is a constant matrix called here an estimator. More precisely, the major goal is to find Φ suppressing the Mean Square Error (MSE) of Φy

$$\begin{aligned} J(\Phi) :=& E(\|\Phi y - \beta\|^2) \\ =& \underbrace{\sigma^2 \operatorname{tr}(\Phi Q \Phi^t)}_{\text{variance}} + \underbrace{\|(\Phi L P_{\mathcal{N}(A)} - P_{\mathcal{N}(A)})z\|^2}_{\text{bias}^2} \\ & (\forall z \in \beta + \mathcal{N}(A)^\perp), \end{aligned} \tag{3}$$

where E denotes expectation and $\|\cdot\|$ denotes the Euclidean norm. The critical problem is that β is unknown, hence it is impossible to minimize J directly.

2.2. Minimum-Variance Pseudo-Unbiased Reduced-Rank Estimator (MV-PURE)

The Minimum-Variance Pseudo-Unbiased Reduced-rank Estimator (MV-PURE) has been derived as a natural extension of the Gauss-Markov estimator to the case of reduced-rank estimator, with the unknown vector possibly subjected to linear (more generally, affine) constraints. The MV-PURE is the solution of the following problem: For a given rank $r \in \{0, 1, \ldots, m\}$,

$$\begin{aligned} &\min \operatorname{tr}[\Phi_r^* Q (\Phi_r^*)^t] \\ &\text{s.t. } \Phi_r^* \in \bigcap_{\iota\in\mathcal{J}} \mathcal{P}_r^\iota, \end{aligned} \tag{4}$$

where $\mathcal{P}_r^\iota := \operatorname*{arg\,min}_{\Phi_r \in \mathcal{X}_r^{m\times n}} \|\Phi_r L P_{\mathcal{N}(A)} - P_{\mathcal{N}(A)}\|_\iota^2$, $\mathcal{X}_r^{m\times n} := \{X \in \mathbb{R}^{m\times n} \mid \mathrm{rk}(X) \le r\}$, and $\mathcal{J}$ is the index set of all unitarily invariant norms[2].

A closed-form solution of (4) [5] is given as

$$\Phi_r^{\mathrm{MVP}} := N \begin{pmatrix} S_r S_r^t & O \\ O & O \end{pmatrix} N^t (L')^\dagger Q^{-1/2}, \tag{5}$$

where $L' := Q^{-1/2} L P_{\mathcal{N}(A)}$, $S_r \in \mathbb{R}^{(m-\mathrm{rk}(A))\times r}$ is an semi-orthonormal matrix whose range is r-minor eigensubspace of[3] $[N^t(L')^\dagger(L'^t)^\dagger N]_{\mathrm{sub}(m-\mathrm{rk}(A))\times(m-\mathrm{rk}(A))}$. Here, $X_{\mathrm{sub}(\varrho\times\varrho)}$ is the $\varrho\times\varrho$ principal submatrix of $X \in \mathbb{R}^{m\times m}$.

Selection of the rank r of the MV-PURE is important to suppress the MSE of the estimate $\Phi_r^{\mathrm{MVP}}(y)$. To see this clearly, let's consider the typical situation where we do not have any auxiliary knowledge on β, i.e., $A = O$, and the noise ϵ is white, i.e., $Q = I_n$. If we simply set $r = m$, the MV-PURE reproduces the BLUE, of which resulting MSE is

$$J(\Phi_{\mathrm{BLUE}}) = \sigma^2 \sum_{i=1}^n \frac{1}{\sigma_i^2}. \tag{6}$$

From (6), we observe that the BLUE yields inherently a drastic inadequacy when the linear regression model (1) is ill-conditioned, i.e., when L possesses very small singular values.

3. PROPOSED RANK SELECTION AND APPLICATION IN IMAGE RESTORATION

In this paper, considering the ill-conditioned nature of L, we propose a rank selection criterion of the MV-PURE based on an unbiased estimate of the MSE for $L\beta$.

3.1. An unbiased weighted-MSE criterion

Our general criterion is the following[4]:

Proposition 1 *For any user-defined matrix $G \in \mathbb{R}^{n\times n}$: $G \neq O$, a criterion*

$$\begin{aligned} \tilde{J}_y^G :& \mathbb{R}^{m\times n} \to \mathbb{R} \\ \tilde{J}_y^G(\Phi) :=& \|GL\Phi y - Gy\|^2 + 2\sigma^2 \operatorname{tr}(GL\Phi Q G^t) \\ & - \sigma^2 \operatorname{tr}(GQG^t) \end{aligned} \tag{7}$$

is an unbiased estimate of the so-called weighted MSE as well as an lower bound of the MSE up to a scale factor in the mean value, i.e.,

$$\begin{aligned} E\left[\tilde{J}_y^G(\Phi)\right] &= E\left[\|GL\Phi y - GL\beta\|^2\right] \\ &\le (\|GL\|_2)^2 J(\Phi),\ \forall \Phi \in \mathbb{R}^{m\times n}. \end{aligned} \tag{8}$$

Remark that the criterion (7) produces an unbiased MSE criterion by $G = (L^tL)^{-1}L^t$, or of an unbiased predicted-MSE criterion by $G = I_n$. Proposition 1 can be applicable to any zero mean distribution with positive definite covariance matrix $\sigma^2 Q$, while each of these criteria is identical to the SURE [11] or the predicted-SURE [17, 18] under the Gaussian noise case.

3.2. Rank selection with an unbiased predicted-MSE criterion

We propose to select a rank of the MV-PURE by minimizing the criterion (7) with $(G, \Phi) = (I_n, \Phi_r^{\mathrm{MVP}})$, i.e.,

$$r^* \in \operatorname*{arg\,min}_{0\le r\le m} \tilde{J}_y^{I_n}(\Phi_r^{\mathrm{MVP}}),$$

[2] We denote by $\mathcal{J}$ the index set of all unitarily invariant norms on $\mathbb{R}^{n\times m}$ (i.e., norm satisfying $\|UXV\| = \|X\|$ for all orthogonal $U \in \mathbb{R}^{n\times n}, V \in \mathbb{R}^{m\times m}$, and all $X \in \mathbb{R}^{n\times m}$, see, e.g., [16]). In particular, note that the following widely used norms are unitarily invariant: the Frobenius norm $\|X\|_F = \sqrt{\operatorname{tr}[X^tX]} = \sqrt{\sum_{i=1}^{\min(m,n)} \gamma_i}$, the spectral norm $\|X\|_2 = \max\{\sqrt{\lambda} \mid \lambda \text{ an eigenvalue of } X^tX\} = \gamma_1$, and the trace (nuclear) norm $\|X\|_{\mathrm{tr}} = \sum_{i=1}^{\min(m,n)} \gamma_i$.

[3] For a given matrix $M \in \mathbb{R}^{m\times n}$, $M^\dagger$ represents the Moore-Penrose pseudo inverse.

[4] In this paper, for simplicity, Proposition 1 is given only for linear estimators. However, this result can be extended for affine estimators straightforwardly, which implies that our rank selection can be extended for the MV-PURE with affine constraint case (see [5, Sec.III A-2]).

where

$$\tilde{J}_y^{I_n}(\Phi_r^{\mathrm{MVP}}) = \|L\Phi_r^{\mathrm{MVP}}(y) - y\|^2 + 2\sigma^2 \operatorname{tr}(L\Phi_r^{\mathrm{MVP}}Q) - \sigma^2 \operatorname{tr}(Q). \tag{9}$$

The inequality (8) with $(G, \Phi) = (I_n, \Phi_r^{\mathrm{MVP}})$ shows that suppressing $E\left[\tilde{J}_y^{I_n}(\Phi_r^{\mathrm{MVP}})\right]$ is necessary for suppressing $J(\Phi_r^{\mathrm{MVP}})$. Under the assumption that no auxiliary knowledge on β is available, i.e., $A = O$, the noise ϵ is white Gaussian with $Q = I_n$, and L is a symmetric matrix, the result in [19, Prop. 5] (or [20, Property 1]) implies that the variance of the criterion (9) is

$$\begin{aligned} &E[(\tilde{J}_y^{I_n}(\Phi_r^{\mathrm{MVP}}) - \|L\Phi_r^{\mathrm{MVP}}(y) - L\beta\|^2)^2] \\ &= 2\sigma^4 n + 4\sigma^2 \|(L\Phi_r^{\mathrm{MVP}} - I_n)L\beta\|^2, \end{aligned} \tag{10}$$

which is much lower than the variance of $\tilde{J}_y^{(L^tL)^{-1}L^t}$ (see Remark below).

Remark (Variance of an unbiased MSE criterion):
Under the same assumption, the criterion (7) with $(G, \Phi) = ((L^tL)^{-1}L^t, \Phi_r^{\mathrm{MVP}})$, i.e.,

$$\begin{aligned} \tilde{J}_y^{(L^tL)^{-1}L^t}(\Phi_r^{\mathrm{MVP}}) = \|(\Phi_r^{\mathrm{MVP}} - (L^tL)^{-1}L^t)y\|^2 \\ + 2\sigma^2 \operatorname{tr}(\Phi_r^{\mathrm{MVP}} L(L^tL)^{-1}) \\ - \sigma^2 \operatorname{tr}((L^tL)^{-1}) \end{aligned} \tag{11}$$

has a variance

$$\begin{aligned} &E[(\tilde{J}_y^{(L^tL)^{-1}L^t}(\Phi_r^{\mathrm{MVP}}) - \|\Phi_r^{\mathrm{MVP}}(y) - \beta\|^2)^2] \\ &= 2\sigma^4 \operatorname{tr}[\Sigma^{-4}] + 4\sigma^2 \|L^{-1}(\Phi_r^{\mathrm{MVP}} - L^{-1})L\beta\|^2, \end{aligned}$$

which becomes large if Σ is ill-conditioned. On the other hand, the variance (10) suggests that the proposed criterion (9) is much more robust against the small singular values of L than (11). Hence the proposed criterion (9) is more reliable than (11) under the ill-conditioned case.

3.3. An efficient implementation of MV-PURE applied in image restoration

We apply the MV-PURE Φ_r^{MVP}, minimizing $\tilde{J}_y^{I_n}(\Phi_r^{\mathrm{MVP}})$, to an image restoration problem formulated as in (1), where y is a known target image (in a vector expression), L is a known symmetric blur matrix, and β is an unknown desired image. For simplicity, we assume that no auxiliary knowledge on β is available, i.e., $A = O$, and the noise ϵ is white, i.e., $Q = I_n$. In addition, we assume as in [15] that L is diagonalized by a computationally efficient orthogonal matrix U, i.e., $L = U\Gamma U^t$ (Here, assume that the eigenvalues on the main diagonal of Γ organized in nonincreasing order of the absolute values). A typical example of U is the Discrete Cosine Transform (DCT) matrix.

For this case, the MV-PURE is reduced to

$$\Phi_r^{\mathrm{MVP}} = U[\operatorname{trun}_r(\Gamma^{-1})]U^t,$$

where $\operatorname{trun}_\varrho \colon \mathbb{R}^{m\times m} \to \mathbb{R}^{m\times m}$ is defined by

$$\operatorname{trun}_\varrho(X) = \begin{pmatrix} X_{\mathrm{sub}(\varrho\times\varrho)} & O \\ O & O \end{pmatrix}$$

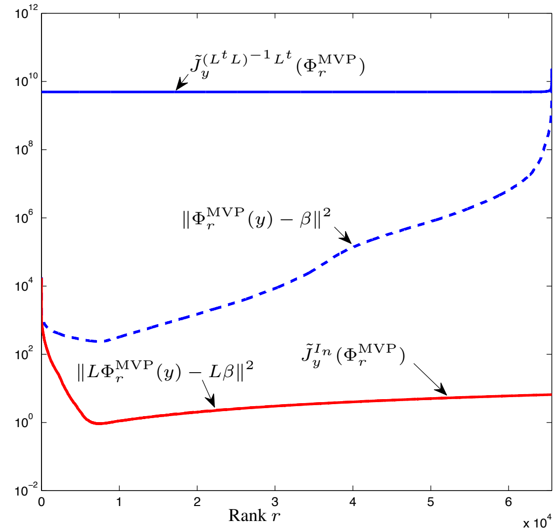

Fig. 1. A comparison of $\tilde{J}_y^{I_n}(\Phi_r^{\mathrm{MVP}})$ in (9) and $\tilde{J}_y^{(L^tL)^{-1}L^t}(\Phi_r^{\mathrm{MVP}})$ in (11), where the vertical axis indicates the value of each criterion for a given rank r. The criterion $\tilde{J}_y^{I_n}(\Phi_r^{\mathrm{MVP}})$ in (9) provides an accurate estimate of $\|L\Phi_r^{\mathrm{MVP}}(y) - L\beta\|^2$, while $\tilde{J}_y^{(L^tL)^{-1}L^t}(\Phi_r^{\mathrm{MVP}})$ in (11) fails in estimating $\|\Phi_r^{\mathrm{MVP}}(y) - \beta\|^2$.

with $\varrho \in \{0, 1, \ldots, m\}$. Thanks to the special assumption on U, we can compute the MV-PURE with $\mathcal{O}(m \log m)$. In addition, we can minimize with $\mathcal{O}(m \log m)$ multiplications the criterion (9)

$$\left\|(\operatorname{trun}_r(I_m) - I_m)U^t y\right\|^2 + 2\sigma^2 r - \sigma^2 m,$$

by the following steps:

1. Compute $\xi := (\xi_1, \xi_2, \ldots, \xi_m)^t := U^t y \in \mathbb{R}^m$.
2. Compute $\alpha_r := \|[\operatorname{trun}_r(I_m) - I_m]\xi\|^2$ for each $r \in \{0, 1, \ldots, m\}$ recursively:
$$\alpha_m = 0,$$
$$\alpha_{j-1} = \alpha_j + \xi_j^2,$$
for $j = m, m-1, \ldots, 1$.
3. Find $r^* \in \arg\min\{\alpha_r + 2\sigma^2 r\}_{r=0}^m$.

The complexity for these three steps is also $\mathcal{O}(m \log m)$.

4. NUMERICAL EXAMPLE

We consider a problem for restoration of an image from the observation degraded by Gaussian blur and contaminated with white Gaussian noise. We use 'cameraman' image (256 × 256 [pixels], see Fig. 2) as the unknown desired image $\beta \in \mathbb{R}^m$. $y \in \mathbb{R}^m$ is a known target degraded image, $L \in \mathbb{R}^{m\times m}$ is a known symmetric blur matrix whose kernel is Gaussian (standard deviation is 2) and L is diagonalized by the DCT matrix (see [15]); Then L is ill-conditioned with condition number of 6.5538×10^6. The additive noise is $\mathcal{N}(0, 10^{-4})$.

Fig. 1 shows a comparison between $\tilde{J}_y^{I_n}(\Phi_r^{\mathrm{MVP}})$ in (9) and $\tilde{J}_y^{(L^tL)^{-1}L^t}(\Phi_r^{\mathrm{MVP}})$ in (11). We observe that $\tilde{J}_y^{I_n}(\Phi_r^{\mathrm{MVP}})$ succeeds

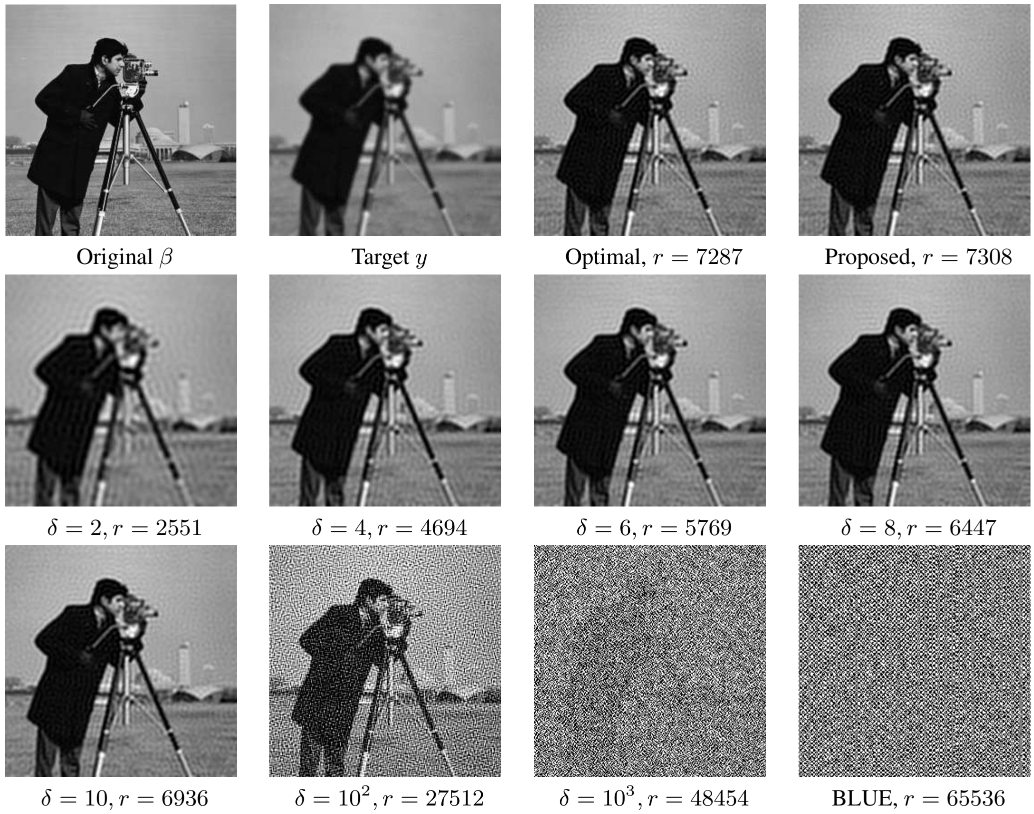

Fig. 2. Resulting images of the MV-PURE. "Proposed" is the resulting image of the MV-PURE minimizing (9). As the optimal rank, we choose a rank achieving the smallest distance to the desired image among all ranks. The parameter δ is utilized for selecting a rank by (12).

Table 1. CPU time of our implementation of the MV-PURE.

	CPU time [sec]
$\Phi_r^{\mathrm{MVP}}(y)$ for a given rank r	0.044
a minimizer of (9) and $\Phi_r^{\mathrm{MVP}}(y)$	0.077

in estimating $\|L\Phi_r^{\mathrm{MVP}}(y) - L\beta\|^2$, while $\tilde{J}_y^{(L^tL)^{-1}L^t}(\Phi_r^{\mathrm{MVP}})$ fails in estimating $\|\Phi_r^{\mathrm{MVP}}(y) - \beta\|^2$. Moreover, the rank minimizing $\tilde{J}_y^{I_n}(\Phi_r^{\mathrm{MVP}})$ in (9) can approximate well the one minimizing $\|\Phi_r^{\mathrm{MVP}}(y) - \beta\|^2$, which shows that the proposed criterion is effective to select a proper rank.

Fig. 2 depicts the resulting images by the MV-PURE. We choose the rank achieving the smallest distance to the desired image among all ranks, as the optimal rank. The MV-PURE minimizing $\tilde{J}_y^{I_n}(\Phi_r^{\mathrm{MVP}})$ is comparable with the MV-PURE with the optimal rank. Although a heuristic rank selection suggested in [5]

$$\arg\max_r \left\{ \frac{\sigma_1}{\sigma_r} < \delta \right\} \tag{12}$$

also selects an appropriate rank for $\delta \in \{6, 8, 10\}$, our rank selection is advanced because it does not require tuning any parameters.

Finally, CPU time of our implementation is shown in Table 1. We measured on a laptop computer equipped with Intel Core 2 duo 1.6Ghz processor and 4GB of RAM. For given y and user-defined rank r, our implementation requires only 0.044 sec to compute $\Phi_r^{\mathrm{MVP}}(y)$. Moreover, for given y, our implementation requires 0.077 sec to obtain the result $\Phi_r^{\mathrm{MVP}}(y)$ with r minimizing $\tilde{J}_y^{I_n}(\Phi_r^{\mathrm{MVP}})$. That is, we can find a minimizer of (9) in time shorter than the computation of $\Phi_r^{\mathrm{MVP}}(y)$ for two different ranks.

5. CONCLUDING REMARKS

In this paper, we have proposed a rank selection criterion, based on an unbiased estimate of the predicted-MSE, of the MV-PURE to achieve a small MSE. The proposed criterion is necessary to suppress because it is a lower bound of the MSE in the mean value up to a scale factor. Moreover, the proposed criterion is robust to ill-conditioned cases, compared to the unbiased MSE criterion. Our unbiased criterion can be applicable to any noise distribution with zero mean and a finite covariance matrix, while Stein-type unbiased criteria (e.g., [11, 12]) cannot. Hence the MV-PURE with our proposed rank selection is applicable to many fields, e.g., image restoration, function approximation [21], electroencephalography [9], and magnetoencephalography [10]. In image restoration setting, an efficient $\mathcal{O}(m \log m)$ implementation have been introduced. A numerical example demonstrated that the rank minimizing the proposed criterion can approximate well the one minimizing the MSE, and the MV-PURE with our rank selection achieves a MSE comparable with the minimal MSE among all possible ranks.

It should be noted that the authors and Shimamura presented preliminary results of this work in [21, 22] in Japanese. Recently, an efficient iterative computation for the stochastic MV-PURE proposed in [23], which also can be extended to the MV-PURE. Such extension will be discussed elsewhere.

6. REFERENCES

[1] D. G. Luenberger, *Optimization by Vector Space Methods*, Wiley, 1969.

[2] H. V. Poor, *An Introduction to Signal Detection and Estimation*, Springer-Verlag, New York, 1994.

[3] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, Englewood Cliffs, NJ:Prentice-Hall, 2000.

[4] I. Yamada and J. Elbadraoui, "Minimum-variance pseudo-unbiased low-rank estimator for ill-conditioned inverse problems," *in Proc. IEEE ICASSP 2006*, May. 2006.

[5] T. Piotrowski and I. Yamada, "MV-PURE estimator: Minimum-Variance Pseudo-Unbiased Reduced-Rank Estimator for linearly constrained ill-conditioned inverse problems," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3408–3423, Aug. 2008.

[6] D. W. Marquardt, "Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation," *Technometrics*, vol. 12, pp. 591–612, 1970.

[7] J. S. Chipman, "Linear restrictions, rank reduction, and biased estimation in linear regression," *Linear Algebra Appl.*, vol. 289, pp. 55–74, 1999.

[8] A. N. Tikhonov, "Solution of incorrectly formulated problems and the regularization method," *Soviet Math. Doktl*, vol. 5, pp. 1035–1038, 1963.

[9] T. Piotrowski, C. C. Zaragoza-Martinez, D. Gutierrez, and I. Yamada, "MV-PURE estimator of dipole source signals in EEG," *in Proc. IEEE ICASSP 2013*, 2013, (accepted).

[10] Y. Uno, L. Qi, K. Amano, and T. Takeda, "Stabilization of multipole expansion in the signal space separation," *in Proc. Annual Meeting of the Society for Neuroscience*, 2012.

[11] C. Stein, "Estimation of the mean of a multivariate normal distribution," *Ann. Statist.*, vol. 9, pp. 1135–1151, 1981.

[12] Y. C. Eldar, "Generalized SURE for exponential families: Applications to regularization," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 471–481, Feb. 2009.

[13] H. C. Andrews and B. R. Hunt, *Digital Image Restoration*, Prentice-Hall, Englewood Cliffs, NJ, 1977.

[14] H. Stark, Ed., *Image Recovery: Theory and Application*, Academic, San Diego, CA, 1987.

[15] P. C. Hansen, J. G. Nagy, and D. P. O'Leary, *Deblurring Images: Matrices, Spectra, and Filtering*, Society for Industrial and Applied Mathematic, 2006.

[16] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses — Theory and Applications*, Springer, 2003.

[17] J. A. Rice, "Choice of smoothing parameter in deconvolution problems," *Contemporary Mathematics*, vol. 59, pp. 137–151, 1986.

[18] A. M. Thompson, J. C. Brown, J. W. Kay, and D. M. Titterington, "A study of methods of choosing the smoothing parameter in image restoration by regularization," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 13, no. 4, pp. 326–339, 1991.

[19] J.-C. Pesquet, A. Benazza-Benyahia, and C. Chaux, "A SURE approach for digital signal/image deconvolution problems," *IEEE Transactions on Signal Processing*, vol. 57, no. 12, pp. 4616–4632, 2009.

[20] F. Luisier, *The SURE-LET Approach to Image Denoising*, Ph.D. thesis, EPFL, Lausanne, 2009.

[21] S. Shimamura, M. Yamagishi, and I. Yamada, "A rank-selection criterion for MV-PURE and its applications to ill-conditioned inverse problems," *IEICE Technical Report.*, Mar. 2011, (in Japanese).

[22] I. Yamada, "Signal processing, optimization, inverse problems — Fun in the interdisciplinary perspective around signal processing —," *IEICE Fundamentals Review*, vol. 5, no. 1, pp. 68–79, 2011, (in Japanese).

[23] T. Piotrowski and I. Yamada, "Efficient parallel computation of the stochastic MV-PURE estimator by the hybrid steepest descent method," *Lecture Notes in Computer Science*, vol. 7268, pp. 404–412, 2012.