WHY THE STOCHASTIC MV-PURE ESTIMATOR EXCELS IN HIGHLY NOISY SITUATIONS?

Tomasz Piotrowski and Isao Yamada

Department of Communications and Integrated Systems (S3-60),
Tokyo Institute of Technology, Tokyo 152-8552, Japan
Phone: +81-3-5734-2503, Fax: +81-3-5734-2905
E-mail addresses: {tpiotrowski, isao}@comm.ss.titech.ac.jp

ABSTRACT

The stochastic MV-PURE estimator has recently emerged as the robust solution for frequently occuring in practice problem of linear estimation in ill-conditioned and imperfectly known linear stochastic model. In this paper we provide theoretical results showing that the stochastic MV-PURE estimator can be used to the greatest effect in highly noisy settings. In such settings, we discuss the relation between the stochastic MV-PURE estimator and the well-known reduced rank Wiener filter. We verify the theoretical results presented by a means of numerical simulations.

Index Terms— Stochastic MV-PURE estimator, parameter estimation, reduced-rank estimation

1. INTRODUCTION

The problem of robust linear estimation of a random vector $\mathbf{x}$ in the linear stochastic model $\mathbf{y} = H\mathbf{x} + \mathbf{n}$ is of central importance in many fields of signal processing (such as modern wireless communications [1–3]), where the estimation is performed under imperfect statistical knowledge on the observable random vector $\mathbf{y}$ and/or the random vector $\mathbf{x}$ to be estimated.

In this paper we focus on the recent solution to the problem of robust estimation under imperfect model knowledge: the stochastic MV-PURE estimator [4]. This estimator has been designed as an optimal fusion of the well-known reduced rank Wiener filter [5–7], and the widely popular in wireless communication community distortionless-constrained estimator [1, 2, 8–10], which is recognized in the settings considered in [4] as the full-rank stochastic MV-PURE estimator.

The robustness of the stochastic MV-PURE estimator was demonstrated in [4], where it was employed as a linear receiver in a multiple-input multiple-output (MIMO) wireless communication system. In this simulation, we verified that under the real-world settings of limited model knowledge available at the receiver, the stochastic MV-PURE estimator achieved lower mean square error (MSE) and symbol error rate (SER) than the (MSE-optimal in the theoretical settings of perfect model knowledge) Wiener filter [8, 11], as well as than the aforementioned reduced rank Wiener filter and the distortionless-constrained estimator. Moreover, we obtained in [4] a transparent expression of the MSE of the stochastic MV-PURE estimator, which shows clearly when the MSE of the stochastic MV-PURE estimator will be significantly lower than that of its full-rank version, the distortionless-constrained estimator.

The stochastic MV-PURE estimator builds on the previously introduced in [12, 13] *minimum-variance pseudo-unbiased reduced-rank estimator (MV-PURE)*, developed for the deterministic estimation in the linear regression model. The results obtained in [12, 13] not only define conditions, under which the (deterministic) MV-PURE estimator should be optimally used, but also show that this estimator encompasses previously known reduced-rank estimators developed in the deterministic framework as its special cases. Therefore, the results of [12, 13] achieve clear positioning of the MV-PURE estimator among the reduced-rank estimators.

It is the goal of this paper to provide a similar characterization of the stochastic MV-PURE estimator. More precisely, we will validate by theoretical results, that the stochastic MV-PURE estimator can be used to the greatest effect in the ill-conditioned, highly noisy situations. Interestingly enough, we will show that precisely in such conditions, the reduced rank Wiener filter is very close to the (full-rank) Wiener filter, and hence cannot provide greater robustness than its full-rank version, if only imperfect model knowledge is available. In particular, the theoretical results presented in this paper naturally agree with the numerical simulations provided in [4], and give valuable insight into properties of the stochastic MV-PURE estimator. Moreover, we present further numerical simulations, which also confirm in practice the theoretical results presented.

Unless otherwise stated, all singular value decompositions (SVDs) [14] have singular values organized in nonincreasing order. The singular values are always denoted by the same letter (in lowercase) as the matrix containing them. Moreover, by $SVD(X) = U\Sigma V^t$ we will mean that $U\Sigma V^t$ is a SVD of X, and by $rk(X) = r$ we will mean that X has rank r.

Due to lack of space, all proofs are omitted.

2. PRELIMINARIES

2.1. Estimation in linear stochastic model

In this paper we consider the linear stochastic model of the form:

$$\mathbf{y} = H\mathbf{x} + \mathbf{n}, \tag{1}$$

where $H \in \mathbb{R}^{n\times m}$ is a known matrix of rank m, and $\mathbf{y} \in \mathbb{R}^n, \mathbf{x} \in \mathbb{R}^m, \mathbf{n} \in \mathbb{R}^n$ are random vectors. It is assumed that $\mathbf{x}$ and $\mathbf{n}$ are uncorrelated, $R_{\mathbf{xn}} = 0$, and that $R_{\mathbf{x}} \succ 0, R_{\mathbf{n}} \succ 0$ are known positive definite covariance matrices. Note that from our assumptions $R_{\mathbf{y}} = (HR_{\mathbf{x}}H^t + R_{\mathbf{n}}) \succ 0$ and $R_{\mathbf{xy}} = R_{\mathbf{x}}H^t$ are available, with $rk(R_{\mathbf{xy}}) = m$.

It can be easily verified that any two random vectors $\mathbf{x} \in \mathbb{R}^m, \mathbf{y} \in \mathbb{R}^n$, for which $R_{\mathbf{x}}, R_{\mathbf{y}}, R_{\mathbf{yx}}$ are known, where $rk(R_{\mathbf{yx}}) = m$, and for which the joint covariance matrix is positive definite, can be cast into model (1) by setting $H = R_{\mathbf{yx}}R_{\mathbf{x}}^{-1}$ and $\mathbf{n} = \mathbf{y} - R_{\mathbf{yx}}R_{\mathbf{x}}^{-1}\mathbf{x}$, see [4, 8]. Therefore, model (1) encompasses a huge variety of problems in communications[1], control, signal processing, statistics, and many other fields (see e.g. [8, 11], which are excellent references on this subject).

The problem considered is that of linear estimation of $\mathbf{x}$ given $\mathbf{y}$, under the mean square error criterion. Thus, we seek to find a fixed matrix $W \in \mathbb{R}^{m\times n}$, called here an estimator, for which the estimate of $\mathbf{x}$ given by:

$$\widehat{\mathbf{x}} = W\mathbf{y}, \tag{2}$$

is optimal with respect to a certain measure related to the mean square error of $\widehat{\mathbf{x}}$:

$$\begin{aligned}\mathrm{J}(W) = tr\left[\mathrm{E}[(\widehat{\mathbf{x}} - \mathbf{x})(\widehat{\mathbf{x}} - \mathbf{x})^t]\right] = \\ tr[WR_{\mathbf{y}}W^t] - 2tr[WR_{\mathbf{yx}}] + tr[R_{\mathbf{x}}].\end{aligned} \tag{3}$$

2.2. Known linear estimators

The well-known unique solution to the problem of minimizing (3) is the linear minimum mean square error estimator, denoted MMSE (often called the Wiener filter), given by [8, 11]:

$$W_{MMSE} = R_{\mathbf{xy}}R_{\mathbf{y}}^{-1} = R_{\mathbf{x}}H^t(HR_{\mathbf{x}}H^t + R_{\mathbf{n}})^{-1}. \tag{4}$$

Similarly, the reduced rank (Wiener) MMSE filter [5–7], denoted RR-MMSE, minimizes the MSE among reduced-rank estimators:

$$\begin{cases} \text{minimize} & \mathrm{J}(W_r) \\ \text{subject to} & W_r \in \mathcal{X}_r^{m\times n}, \end{cases} \tag{5}$$

where $\mathcal{X}_r^{m\times n} = \{W_r \in \mathbb{R}^{m\times n} : rk(W_r) \leq r\}$ is the set of rank-constrained matrices in $\mathbb{R}^{m\times n}$. Problem (5) produces as a solution:

$$W_{RR-MMSE}^r = V_rV_r^tW_{MMSE}, \tag{6}$$

where $SVD(R_{\mathbf{y}}^{-1/2}HR_{\mathbf{x}}) = U\Sigma V^t$, with $V = (v_1, \ldots, v_m)$ and $V_r = (v_1, \ldots, v_r)$ for $r \leq m$.

[1] In particular, in wireless communications employing modern techniques such as CDMA, MIMO, OFDM [1–3].

Another estimator, which is also the solution of a constrained MSE minimization, achieves perfect reconstruction of the target random vector in the noiseless case. More precisely, the distortionless-constrained estimator, denoted C-MMSE, is a solution of the following linearly constrained MSE minimization problem (see e.g. [8]):

$$\begin{cases} \text{minimize} & \mathrm{J}(W) \\ \text{subject to} & WH = I_m, \end{cases} \tag{7}$$

with the unique solution:

$$W_{C-MMSE} = (H^tR_{\mathbf{y}}^{-1}H)^{-1}H^tR_{\mathbf{y}}^{-1} = (R_{\mathbf{y}}^{-1/2}H)^{\dagger}R_{\mathbf{y}}^{-1/2}, \tag{8}$$

where $(R_{\mathbf{y}}^{-1/2}H)^{\dagger}$ denotes the Moore-Penrose pseudoinverse of $(R_{\mathbf{y}}^{-1/2}H)$ [14].

2.3. Stochastic MV-PURE estimator

The stochastic MV-PURE estimator, introduced in [4], optimally combines the reduced-rank (5) and distortionless-constrained approaches (7), by achieving smallest distortion among reduced-rank estimators.[2] More precisely, the stochastic MV-PURE estimator is defined as a solution of the following problem, for given rank constraint $r \leq m$:

$$\begin{cases} \text{minimize} & \mathrm{J}(W_r) \\ \text{subject to} & W_r \in \displaystyle\bigcap_{\iota\in\mathfrak{I}} \mathcal{P}_r^{\iota}, \end{cases} \tag{9}$$

where

$$\mathcal{P}_r^{\iota} = \arg\min_{W_r\in\mathcal{X}_r^{m\times n}} \| W_rH - I_m \|_{\iota}^2,\ \iota \in \mathfrak{I}, \tag{10}$$

where $\mathfrak{I}$ is the index set of all unitarily invariant norms (i.e., norms satisfying $\| UXV \| = \| X \|$ for all orthogonal $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{n\times n}$ and all $X \in \mathbb{R}^{m\times n}$, see [14]).

The following Theorem, cited from [4], provides a closed algebraic form of the stochastic MV-PURE estimator.

Theorem 1 ([4]) *1. Let us set rank constraint $r < m$ and let us set:*

$$K = \left(H^tR_{\boldsymbol{y}}^{-1}H\right)^{-1} - 2R_{\boldsymbol{x}}. \tag{11}$$

Moreover, let the symmetric matrix K be given with an eigenvalue decomposition $EVD(K) = E\Delta E^t$, with eigenvalues organized in nondecreasing order:

$$\delta_1 \leq \delta_2 \leq \cdots \leq \delta_m.$$

Then $W_{MV-PURE}^r \in \mathbb{R}^{m\times n}$ is a solution to problem (9) if and only if $W_{MV-PURE}^r$ is of the following form:

$$W_{MV-PURE}^r = E_rE_r^tW_{C-MMSE}, \tag{12}$$

[2] See [4] for in-depth discussion of optimality of the proposed approach. For further insight into the MV-PURE estimation, see the original papers [12, 13], where the deterministic version of the MV-PURE estimator was developed.

where W_{C-MMSE} is of the form (8), and where $E = (e_1, \dots, e_m)$ with $E_r = (e_1, \dots, e_r)$ for $r < m$. If $\delta_r \neq \delta_{r+1}$, the solution is unique. Moreover, we have:

$$\mathrm{J}(W^r_{MV-PURE}) = \sum_{i=1}^{r} \delta_i + tr[R_{\boldsymbol{x}}]. \tag{13}$$

2. For no rank constraint imposed, i.e. when $r = m$, the solution to problem (9) is uniquely given by $W^m_{MV-PURE} = W_{C-MMSE}$. In particular, we have:

$$\mathrm{J}(W^m_{MV-PURE}) = \mathrm{J}(W_{C-MMSE}) = \sum_{i=1}^{m} \delta_i + tr[R_{\boldsymbol{x}}]. \tag{14}$$

Note that immediately from (13) and (14) we obtain that $\mathrm{J}(W^r_{MV-PURE}) < \mathrm{J}(W_{C-MMSE})$ if and only if:

$$\sum_{i=r+1}^{m} \delta_i > 0. \tag{15}$$

Moreover, we showed in [4], that in a common in signal processing case of $R_{\mathbf{x}} = I_m$ (see e.g. [15]), upon setting $SVD(R_{\mathbf{y}}^{-1/2}H) = U\Sigma V^t$, with $\sigma_1 \geq \cdots \geq \sigma_m > 0$ the singular values of $R_{\mathbf{y}}^{-1/2}H$, we have for all $r < m$

$$W^r_{MV-PURE} = \widetilde{V}_r \widetilde{V}_r^t W_{MMSE}, \tag{16}$$

where $\widetilde{V} = (v_1/\sigma_1, \dots, v_m/\sigma_m) \in \mathbb{R}^{m\times m}$ with $\widetilde{V}_r = (v_1/\sigma_1, \dots, v_r/\sigma_r) \in \mathbb{R}^{m\times r}$ for $r < m$, and where W_{MMSE} is of the form (4) for $R_{\mathbf{x}} = I_m$. If $\sigma_r \neq \sigma_{r+1}$, the solution is unique. Moreover, we showed in [4] that:

$$\mathrm{J}(W^r_{MV-PURE}) = \sum_{i=1}^{r} \sigma_i^{-2} - 2r + m, \tag{17}$$

and for $r = m$ we have $W^m_{MV-PURE} = W_{C-MMSE}$ with:

$$\mathrm{J}(W^m_{MV-PURE}) = \mathrm{J}(W_{C-MMSE}) = \sum_{i=1}^{m} \sigma_i^{-2} - m. \tag{18}$$

Note that from (17)-(18), we obtain immediately that:

$$\mathrm{J}(W_{C-MMSE}) - \mathrm{J}(W^r_{MV-PURE}) = \sum_{i=r+1}^{m} \sigma_i^{-2} + 2(r-m). \tag{19}$$

3. WHY THE STOCHASTIC MV-PURE ESTIMATOR EXCELS IN HIGHLY NOISY SITUATIONS?

In this section, we provide first the main theoretical results of this paper.

Let $\delta \in \mathbb{R}_+$, where $\mathbb{R}_+$ is the set of nonnegative real numbers, and let us define:

$$R_{\mathbf{y}}(\delta) := HR_{\mathbf{x}}H^t + \delta R_{\mathbf{n}}. \tag{20}$$

Then, we have:

$$\lim_{\delta\to\infty} H^t(R_{\mathbf{y}}(\delta))^{-1}H = 0 \in \mathbb{R}^{m\times m}. \tag{21}$$

The following Proposition holds.

Proposition 1 *Let reduced rank Wiener filter $W^r_{RR-MMSE} = V_r V_r^t W_{MMSE}$ be given as in (6), where $SVD(R_{\boldsymbol{y}}^{-1/2}HR_{\boldsymbol{x}}) = U\Sigma V^t$, with $V = (v_1, \dots, v_m)$ and $V_r = (v_1, \dots, v_r)$ for $r < m$. Note that we obtain thus*

$$EVD(R_{\boldsymbol{x}}H^tR_{\boldsymbol{y}}^{-1}HR_{\boldsymbol{x}}) = V(\Sigma^t\Sigma)V^t.$$

Then, for all $r = 1, \dots, m-1$, we have:

$$\forall\, \iota \in \mathfrak{I} \quad \| W_{MMSE} - W^r_{RR-MMSE} \|_\iota \leq \left(1/\sqrt{\lambda_n(R_{\boldsymbol{y}})}\right) \| \begin{pmatrix} \sigma_{r+1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_m \end{pmatrix} \|_\iota, \tag{22}$$

where $\lambda_n(R_{\boldsymbol{y}})$ is the smallest singular value of $R_{\boldsymbol{y}}$, and where $\mathfrak{I}$ is the index set of all unitarily invariant norms.

We consider now the common in signal processing case of $R_{\mathbf{x}} = I_m$, and we focus on the behaviour of the estimators under consideration in highly noisy situations.

We note first that as δ increases, all eigenvalues of $R_{\mathbf{y}}(\delta)$ (20) grow unboundedly, thus in particular $1/\sqrt{\lambda_n(R_{\mathbf{y}})}$ converges to 0. Moreover, as δ increases, from (21) we obtain that the matrix $H^tR_{\mathbf{y}}^{-1}H$ becomes severely ill-conditioned, with its eigenvalues $\sigma_1^2 \geq \cdots \geq \sigma_m^2 > 0$ pulling toward 0.[3]

Thus, from (19) we obtain immediately that in highly noisy settings of large δ, the stochastic MV-PURE estimator achieves much better performance than the C-MMSE estimator, since the term $\sum_{i=r+1}^{m} \sigma_i^{-2}$ in (19) may be extremely large.

It is thus very interesting to note that precisely in such highly noisy settings, from Proposition 1 we conclude that as δ increases, the reduced rank Wiener filter $W^r_{RR-MMSE}$ converges to the Wiener filter W_{MMSE}, for all rank constraints $r < m$. Therefore, the reduced rank Wiener filter does not provide gain in performance over the full-rank Wiener filter in highly noisy situations.

We conclude therefore that in highly noisy situations, the stochastic MV-PURE estimator achieves huge gain in performance over the C-MMSE estimator, and that the RR-MMSE estimator provides very similar performance to the MMSE estimator.

4. NUMERICAL EXAMPLE

In a numerical example provided in [4], the linear stochastic model (1) represented a MIMO wireless communication system, where we considered $m = 8$ and $n = 8$ transmit and receive antennas, respectively. Thus, to check the performance

[3]Note that, with our usual notation, we have $SVD(R_{\mathbf{y}}^{-1/2}H) = U\Sigma V^t \Rightarrow EVD(H^tR_{\mathbf{y}}^{-1}H) = V(\Sigma^t\Sigma)V^t$.

of the stochastic MV-PURE estimator in a different settings, we chose a much larger $m = 100, n = 200$ system in the current simulations, where we illustrate the theoretical results provided in section 3. The entries of H are realizations of iid. zero-mean Gaussians with unity variance, and the resulting H is a well-conditioned matrix with singular values approximately equal to $23.39, 22.76, \ldots, 4.86, 4.47$. Similarly, the entries of the input random vector **x** were iid. Gaussians with unity variance, and we considered $Q = 400$ realizations of **x**. We assumed correlated Gaussian noise, with covariance matrix $\delta R_{\mathbf{n}} \succ 0$, where $\delta > 0$. The SNR was defined as:

$$SNR := \frac{tr[HH^t]}{\delta tr[R_{\mathbf{n}}]}.$$

In the simulations depicted in Fig.1, we assumed that the matrix H is available, but that the noise covariance matrix $\delta R_{\mathbf{n}}$ is not. Thus, in order to construct the estimators under consideration, the covariance matrix $R_{\mathbf{y}}$ was approximated in their definitions using its finite sample estimate $R_{\mathbf{y}} \approx \widehat{R_{\mathbf{y}}} = \frac{1}{Q}\sum_{q=1}^{Q} \mathbf{y}_q \mathbf{y}_q^t$, where $Q = 400$ is the length of the data transmitted. The rank of the stochastic MV-PURE estimator was chosen in order to minimize (17)-(18), and we used the form of the stochastic MV-PURE estimator given in (16). Moreover, since we are not aware of a similarly simple rank-selection criterion for the RR-MMSE estimator, we always used the same rank for both the stochastic MV-PURE and RR-MMSE estimators.

By evaluating the performance of the estimators under consideration, Fig. 1 demonstrates that the theoretical results of section 3 are applicable also for the case, where the covariance matrix $R_{\mathbf{y}}$ is replaced by its finite sample estimate. Indeed, the discussion presented in section 3 provides full explanation of the results shown in Fig. 1.

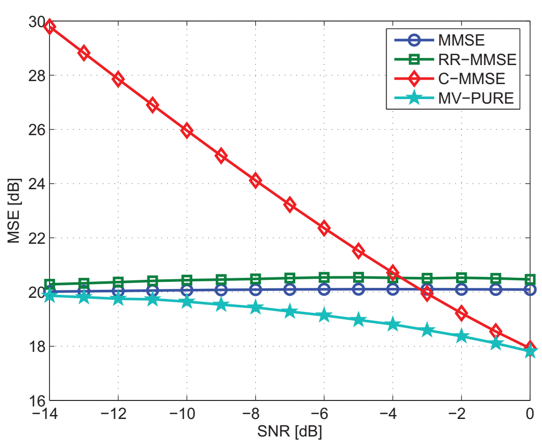

Fig. 1. Performance of MMSE, RR-MMSE, C-MMSE and stochastic MV-PURE estimators for the case of imperfectly known covariance matrix $R_{\mathbf{y}}$.

Acknowledgement: The authors would like to express their deep gratitude to Prof. Kohichi Sakaniwa of Tokyo Institute of Technology for helpful discussions.

This work was supported in part by the JSPS Grants-in-Aid (C-19500186).

5. REFERENCES

[1] X. Wang and H. V. Poor, *Wireless Communication Systems*, Prentice Hall, Upper Saddle River, 2004.

[2] A. B. Gershman and N. D. Sidiropoulos, Eds., *Space-Time Processing for MIMO Communications*, John Wiley & Sons, Chichester, 2005.

[3] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*, Cambridge University Press, New York, 2003.

[4] T. Piotrowski, R. L. G. Cavalcante, and I. Yamada, "Stochastic MV-PURE estimator," *IEEE Trans. Signal Process.*, accepted.

[5] D. R. Brillinger, *Time Series: Data Analysis and Theory*, Holt, Rinehart and Winston, New York, 1975.

[6] L. L. Scharf, "The SVD and reduced rank signal processing," *Signal Process.*, vol. 25, pp. 113–133, 1991.

[7] Y. Hua, M. Nikpour, and P. Stoica, "Optimal reduced-rank estimation and filtering," *IEEE Trans. Signal Process.*, vol. 49, no. 3, pp. 457–469, Mar. 2001.

[8] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, Prentice Hall, New Jersey, 2000.

[9] X. Wang and A. Høst-Madsen, "Group-blind multiuser detection for uplink CDMA," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 11, pp. 1971–1984, Nov. 1999.

[10] S. Shahbazpanahi, M. Beheshti, A. B. Gershman, M. Gharavi-Alkhansari, and K. M. Wong, "Minimum-variance linear receivers for multiaccess MIMO wireless systems with space-time block coding," *IEEE Trans. Signal Process.*, vol. 52, no. 12, pp. 3306–3313, Dec. 2004.

[11] D. G. Luenberger, *Optimization by Vector Space Methods*, John Wiley & Sons, New York, 1969.

[12] I. Yamada and J. Elbadraoui, "Minimum-variance pseudo-unbiased low-rank estimator for ill-conditioned inverse problems," in *Proc. ICASSP*, Toulouse, France, May 2006, pp. 325–328.

[13] T. Piotrowski and I. Yamada, "MV-PURE estimator: Minimum-variance pseudo-unbiased reduced-rank estimator for linearly constrained ill-conditioned inverse problems," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3408–3423, Aug. 2008.

[14] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985.

[15] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, 1995.