RESEARCH

Open Access

Greedy selection of sensors with measurements under correlated noise



Yoon Hak Kim^{1*}

*Correspondence: yhk@chosun.ac.kr

¹ Department of Electronic Engineering, College of IT Convergence Engineering, Chosun University, 309 Pilmun-daero, Dong-gu, Gwangju 61452, Republic of Korea

Abstract

We address the sensor selection problem where linear measurements under correlated noise are gathered at the selected nodes to estimate the unknown parameter. Since finding the best subset of sensor nodes that minimizes the estimation error requires a prohibitive computational cost especially for a large number of nodes, we propose a greedy selection algorithm that uses the log-determinant of the inverse estimation error covariance matrix as the metric to be maximized. We further manipulate the metric by employing the QR and LU factorizations to derive a simple analytic rule which enables an efficient selection of one node at each iteration in a greedy manner. We also make a complexity analysis of the proposed algorithm and compare with different selection methods, leading to a competitive complexity of the proposed algorithm. For performance evaluation, we conduct numerical experiments using randomly generated measurements under correlated noise and demonstrate that the proposed algorithm achieves a good estimation accuracy with a reasonable selection complexity as compared with the previous novel selection methods.

Keywords: Sensor selection, Greedy algorithm, QR factorization, LU factorization, Correlated noise

1 Introduction

In wireless sensor networks, a large number of sensor nodes are spatially distributed and transmit their noise-corrupted measurements which are typically formulated by a linear combination of the known observation matrix **H** and the parameter to be estimated. The sensor selection is conducted so as to optimize the estimation accuracy [1–7]: Convex relaxation [1]- and cross-entropy optimization [2]-based methods were presented with a prohibitive computational complexity for large sensor networks. Noting that greedy approach yields a feasible complexity, the sensor selection problem has been mostly tackled in a greedy manner. To guarantee near-optimality with regard to the mean squared estimation error (MSE), a submodular cost function called the frame potential was devised and a greedy removal method was proposed to find optimal sensor locations [3]. A QR factorization-based greedy selection method was proposed to minimize the estimation error [4]. To further reduce the complexity of the selection process, the log-determinant of the inverse estimation error covariance matrix was employed as a metric to present simple greedy methods instead of directly minimizing the estimation



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http:// creativeCommons.org/licenses/by/4.0/.

error [5, 6]. Specifically, analytic selection rules were established by proving the metric to be a monotone submodular function [5] and by employing the QR factorization [6], respectively. An efficient greedy algorithm was also developed for sensor placement by maximally projecting onto the minimum eigenspace of the dual observation matrix [7]. Clearly, the selection of the best subset of p sensor nodes corresponds to the construction of the matrix with the most informative p rows selected from the matrix **H**.

It should be noticed that most of the selection methods have been derived under the assumption of uncorrelated noise: That is, off-diagonal elements of the noise covariance matrix are assumed to be approximately zero or ignorable. The sensor selection problem in the presence of correlated noise is challenging since the Fisher information (equivalently, the inverse estimation error covariance matrix) is no longer linear in the selected sensors [8-10]. To solve the problem, various selection algorithms have been previously presented in [8-12]. A multi-step sensor selection strategy was proposed for linear dynamic systems under correlated noise [8]. Sparsity-aware sensor selection approaches in centralized and distributed versions were developed in the presence of correlated measurements using convex relaxation for the maximum likelihood estimation (MLE) [9]. Two selection algorithms were derived for estimation of random parameters to minimize the MSE based on convex relaxation and greedy approach [10], respectively. Recently, a greedy selection algorithm maximizing the log-determinant of the inverse estimation error covariance matrix has been presented to provide an additional increase in estimation accuracy [11, 12]. Whereas the sensor selection problem has been extensively studied for the additive Gaussian linear measurement model which is considered in this work, it has been also addressed for nonlinear measurement model [13, 14]. In [13], the sensor selection algorithm for target tracking was proposed based on the extended Kalman filtering for an additive Gaussian nonlinear model. In [14], sensor selection methods for a general nonlinear model were developed with an aid of convex relaxation technique for estimation of unknown parameters.

In this paper, we consider the scenario where a given number of sensor nodes with measurements corrupted by correlated noise are selected to estimate the parameter. We aim to find the subset of sensor nodes that minimizes the estimation error. We first formulate the estimation error covariance matrix as a function of the matrix H_S with rows selected from H and the covariance matrix K_S of noise samples at the selected nodes. To expedite the selection procedure, we adopt the log-determinant of the inverse estimation error covariance matrix as a metric to be maximized, which allows us to avoid a large matrix inversion. We choose one node at each iteration in a greedy manner until the selection set of the desired cardinality is constructed. To simplify the metric, we factor the matrices H_S and K_S based on the QR and LU factorizations, respectively, to obtain the upper triangular matrix **R** and the lower triangular matrix **M**. We show that the factored matrices **R** and **M** can be updated by simply appending the newly computed column or row vector to the last positions of those matrices at the previous iteration, respectively. This simple computation of the key matrices **R** and **M** leads to an analytic rule which enables an efficient selection of such a maximizing node at each iteration. We also make a complexity analysis of the proposed algorithm, resulting in a competitive complexity of the same order as that of [4]. We evaluate the performance of the proposed method through numerical experiments using randomly generated measurements under correlated noise. We demonstrate the advantage

of the proposed technique over previous selection methods in terms of the estimation performance and the selection complexity.

Compared with [4, 11, 12], we have the new contributions in this paper:

- In [4], the MSE is minimized by applying the QR factorization to \mathbf{H}_S in the assumption of uncorrelated noise. However, in the presence of correlated noise, the MSE is not manageable using only the QR factorization. The main contributions are given in two aspects: First, in order to yield a simple selection process, the final selection rule should be expressed as the first term related with the previous iteration plus the second term determined by only the current selection. The LU factorization of the covariance matrix \mathbf{K}_S is proved to generate the desired expression of the final rule. Second, the overall derivation process is completely different from [4] except for the QR factorization of \mathbf{H}_S .
- The only similarity with [11, 12] is to optimize the same metric which is the log-determinant of the inverse estimation error covariance matrix. Except for that, the optimization process to derive the proposed selection rule in this paper is totally different from [11, 12]. Furthermore, the proposed method is shown to operate faster than the method in [11, 12] in numerical experiments while preserving the same estimation performance.

This paper is organized as follows. The problem is formulated in Sect. 2 in which the metric is shown to be expressed in terms of the sampled matrix H_S , the covariance matrix K_S and the statistics of the parameter. In Sect. 3, the QR and LU factorizations are employed to simplify the metric and a simple selection criterion is derived. The complexity analysis of the proposed algorithm is provided in Sect. 4.1. Extensive experiments are presented in Sect. 4.2 and conclusions in Sect. 5.

2 Problem Formulation

Suppose that *N* wireless sensors deployed in a sensor field generate the measurements $\mathbf{y} \in \mathbb{R}^N$. We seek to estimate the parameter vector $\theta \in \mathbb{R}^p$ with a Gaussian distribution $\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\theta})$ by using n(< N) measurements gathered by the *n* selected sensors in the set *S*. We assume that the measurements are corrupted by the additive correlated Gaussian noise $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \mathbf{K}) \in \mathbb{R}^N$ independent of θ where the covariance matrix **K** is positive definite and symmetric. We also assume that the measurements are produced by a linear model with the known observation $N \times p$ full column rank matrix **H** with *N* row vectors $\mathbf{h}_i^\top, i \in \mathcal{V} = \{1 \cdots N\}$: Specifically,

$$\mathbf{y} = \mathbf{H}\boldsymbol{\theta} + \mathbf{w} \tag{1}$$

We denote the estimator by $\hat{\theta}(\mathbf{y}_S)$ where \mathbf{y}_S is the $|S| \times 1$ column vector with entries of \mathbf{y} indexed by *S*. In this work, we use an optimal Bayesian linear estimator (e.g., maximum a posteriori (MAP) estimator or minimum mean squared error (MMSE) estimator). Note that the MAP estimator is equivalent to the MMSE estimator in our formulation where the parameter and the noise are assumed to be Gaussian. Then, we can derive the estimator $\hat{\theta}(\mathbf{y}_S)$ and the estimation error covariance matrix $\boldsymbol{\Sigma}(S)$ as follows [5, 15]:

$$\hat{\theta}(\mathbf{y}_S) = \left(\mathbf{\Sigma}_{\theta}^{-1} + \mathbf{H}_S^{\mathsf{T}} \mathbf{K}_S^{-1} \mathbf{H}_S\right)^{-1} \mathbf{H}_S^{\mathsf{T}} \mathbf{K}_S^{-1} \mathbf{y}_S$$
(2)

$$\boldsymbol{\Sigma}(S) = \left(\boldsymbol{\Sigma}_{\theta}^{-1} + \mathbf{H}_{S}^{\top}\mathbf{K}_{S}^{-1}\mathbf{H}_{S}\right)^{-1}$$
(3)

where \mathbf{H}_{S} is the matrix with rows of the matrix \mathbf{H} indexed by S and \mathbf{K}_{S} the $|S| \times |S|$ covariance matrix of the noise vector \mathbf{w}_{S} with entries of \mathbf{w} indexed by S. Notice that the estimation of deterministic and unknown parameters can be also conducted by letting $\mathbf{\Sigma}_{\theta}^{-1} = \mathbf{0}$ in (2) and (3).

In this work, we aim to find the best subset S with |S| = p that maximizes the log-determinant of the inverse estimation error covariance matric given by log det $(\Sigma(S)^{-1})$ which has been employed as a metric for the sensor selection problem in previous work [1, 5, 11, 12] and shown to produce a good estimation performance with a reduced complexity. We also seek to select one node at each iteration in a greedy manner. More specifically, given the set S_i of i nodes already selected until the ith iteration, we select one node from the set of the remaining nodes, $S_i^C \equiv (\mathcal{V} - S_i)$ that maximizes the intermediate metric log det $(\Sigma(S_{i+1})^{-1})$ at the (i + 1)th iteration where $S_{i+1} = S_i + \{j\}, j \in S_i^C$. The selection process is expressed by

$$j^* = \arg \max_{S_{i+1} = S_i + \{j\}, j \in S_i^C} \log \det \left(\mathbf{\Sigma}_{\theta}^{-1} + \mathbf{H}_{S_{i+1}}^{\top} \mathbf{K}_{S_{i+1}}^{-1} \mathbf{H}_{S_{i+1}} \right)$$
(4)

where $\mathbf{H}_{S_{i+1}}$ is simply created from \mathbf{H}_{S_i} and the *j*th row vector \mathbf{h}_i^{\top} selected from **H**:

$$\mathbf{H}_{S_{i+1}}^{\top} = \begin{bmatrix} \mathbf{H}_{S_i}^{\top} \ \mathbf{h}_j \end{bmatrix}$$
(5)

and $\mathbf{K}_{S_{i+1}}$ is also constructed from \mathbf{K}_{S_i} and the *j*th node selected from S_i^C :

$$\mathbf{K}_{S_{i+1}} = \begin{bmatrix} \mathbf{K}_{S_i} & \mathbf{k} \\ \mathbf{k}^\top & \mathbf{k} \end{bmatrix}$$

$$\mathbf{k}_{i+1}^\top \equiv \begin{bmatrix} \mathbf{k}^\top & \mathbf{k} \end{bmatrix} = \begin{bmatrix} k_{(1)j} & k_{(2)j} & \cdots & k_{(i)j} & k_{jj} \end{bmatrix}$$
 (6)

where the subscript (*i*) indicates the number of the node selected at the *i*th iteration and \mathbf{k}_{i+1}^{\top} denotes the 1 × (*i* + 1) row vector with $\mathbf{k} = [k_{(1)j} \cdots k_{(i)j}]$ and $k = k_{jj}$. Notice that $k_{(i)j}$ represents the covariance of the noise samples at the (*i*)th and *j*th nodes. The process in (4) is performed repeatedly until the selection set S with |S| = p is constructed.

3 Method: efficient sampling algorithm

In this section, we present an analytic result for a simple selection process by manipulating the matrix $\mathbf{H}_{S_{i+1}}^{\top}$ and the covariance matrix $\mathbf{K}_{S_{i+1}}$ based on the QR and LU factorizations, respectively. Note that the Householder transformation is employed to perform the QR factorization since it has strength in terms of complexity and sensitivity to rounding error in comparison with the Gram–Schmidt orthogonalization [16]. First, noting that the covariance matrix $\mathbf{K}_{S_{i+1}}$ is symmetric, we can factor it as follows:

$$\mathbf{K}_{S_{i+1}} = \mathbf{M}_{S_{i+1}} \mathbf{M}_{S_{i+1}}^{\top} \tag{7}$$

where $\mathbf{M}_{S_{i+1}}$ is the $(i + 1) \times (i + 1)$ lower triangular matrix. We also do the QR factorization of $\mathbf{H}_{S_{i+1}}^{\top} = \mathbf{Q}\bar{\mathbf{R}}^{i+1}$ where \mathbf{Q} is the $p \times p$ orthogonal matrix with p column vectors $\mathbf{q}_{j,j} = 1, \ldots, p$ and $\bar{\mathbf{R}}^{i+1}$ the $p \times (i + 1)$ matrix. Then, we manipulate the metric in (4) to

derive a simpler form with assumption of $\Sigma_{\theta} = \sigma_{\theta}^2 \mathbf{I}_p$ where \mathbf{I}_p indicates the $p \times p$ identity matrix. Specifically,

$$\log \det \left(\boldsymbol{\Sigma}(S_{i+1})^{-1} \right)$$

$$= \log \det \left(\frac{1}{\sigma_{\theta}^{2}} \mathbf{I}_{p} + \mathbf{Q} \bar{\mathbf{R}}^{i+1} \left(\mathbf{M}_{S_{i+1}} \mathbf{M}_{S_{i+1}}^{\top} \right)^{-1} (\bar{\mathbf{R}}^{i+1})^{\top} \mathbf{Q}^{\top} \right)$$

$$= \log \det \left(\mathbf{Q} \left(\frac{1}{\sigma_{\theta}^{2}} \mathbf{I}_{p} + \bar{\mathbf{R}}^{i+1} \left(\mathbf{M}_{S_{i+1}} \mathbf{M}_{S_{i+1}}^{\top} \right)^{-1} (\bar{\mathbf{R}}^{i+1})^{\top} \right) \mathbf{Q}^{\top} \right)$$

$$= \log \det \left(\frac{1}{\sigma_{\theta}^{2}} \mathbf{I}_{p} + \bar{\mathbf{R}}^{i+1} \left(\mathbf{M}_{S_{i+1}} \mathbf{M}_{S_{i+1}}^{\top} \right)^{-1} (\bar{\mathbf{R}}^{i+1})^{\top} \right)$$
(8)

where (8) follows from the notion that det \mathbf{Q} det $\mathbf{Q}^{\top} = 1$. Note that $\mathbf{\bar{R}}^{i+1}$ can be written by

$$\bar{\mathbf{R}}^{i+1} = \begin{bmatrix} \mathbf{R}^{i+1} \\ \mathbf{0}_{(p-i-1)\times(i+1)} \end{bmatrix}$$
(9)

where \mathbf{R}^{i+1} is the $(i + 1) \times (i + 1)$ upper triangular matrix and $\mathbf{0}_{a \times b}$ indicates the $a \times b$ zero matrix. Then, plugging (9) into (8) yields

$$\begin{pmatrix}
\frac{1}{\sigma_{\theta}^{2}}\mathbf{I}_{p} + \bar{\mathbf{R}}^{i+1} \left(\mathbf{M}_{S_{i+1}}\mathbf{M}_{S_{i+1}}^{\top}\right)^{-1} (\bar{\mathbf{R}}^{i+1})^{\top} \\
= \begin{bmatrix}
\mathbf{R}^{i+1} \left(\mathbf{M}_{S_{i+1}}\mathbf{M}_{S_{i+1}}^{\top}\right)^{-1} (\mathbf{R}^{i+1})^{\top} + \frac{1}{\sigma_{\theta}^{2}}\mathbf{I}_{i+1} \ \mathbf{0}_{(p-i-1)\times(i+1)}^{\top} \\
\mathbf{0}_{(p-i-1)\times(i+1)} & \frac{1}{\sigma_{\theta}^{2}}\mathbf{I}_{p-i-1}\end{bmatrix}$$
(10)

Hence, we can further simplify the log-determinant in (8) as follows:

$$\log \det \left(\boldsymbol{\Sigma}(S_{i+1})^{-1} \right)$$

= $\log \det \left(\mathbf{R}^{i+1} \left(\mathbf{M}_{S_{i+1}} \mathbf{M}_{S_{i+1}}^{\top} \right)^{-1} (\mathbf{R}^{i+1})^{\top} + \frac{1}{\sigma_{\theta}^2} \mathbf{I}_{i+1} \right)$
+ $\log \det \left(\frac{1}{\sigma_{\theta}^2} \mathbf{I}_{p-i-1} \right)$ (11)

$$\propto \log \det \left(\mathbf{R}^{i+1} \left(\mathbf{M}_{S_{i+1}} \mathbf{M}_{S_{i+1}}^{\top} \right)^{-1} (\mathbf{R}^{i+1})^{\top} + \frac{1}{\sigma_{\theta}^2} \mathbf{I}_{i+1} \right)$$
(12)

where (12) follows since the second term in (11) is irrelevant in finding the maximizing node at (i + 1)th iteration. In this paper, noting that det(**A**) + det(**B**) \leq det(**A** + **B**) for positive definite matrices **A** and **B** [17], we propose to maximize the lower bound of (12) to yield a low-complexity selection process: Specifically, the selection process in (4) is approximated by

$$j^* \approx \arg\max_{j \in S_i^C} \log \left(\det \left(\mathbf{R}^{i+1} \left(\mathbf{M}_{S_{i+1}} \mathbf{M}_{S_{i+1}}^\top \right)^{-1} (\mathbf{R}^{i+1})^\top \right) + \det \left(\frac{1}{\sigma_{\theta}^2} \mathbf{I}_{i+1} \right) \right)$$
(13)

$$= \arg \max_{j \in S_i^C} \log \det \left(\mathbf{R}^{i+1} \left(\mathbf{M}_{S_{i+1}} \mathbf{M}_{S_{i+1}}^\top \right)^{-1} (\mathbf{R}^{i+1})^\top \right)$$
(14)

where (13) follows from the notion that \mathbf{R}^{i+1} and $\mathbf{M}_{S_{i+1}}$ are positive definite since the covariance matrix **K** is positive definite and **H** a full column rank matrix and (14) from the irrelevancy of det $\left(\frac{1}{\sigma_{\theta}^2}\mathbf{I}_{i+1}\right)$ in (13) to the selection at (i + 1)th iteration. In the optimization perspective, the selection process in (14) corresponds to the methodology which makes the worst case as good as possible for the estimation of the parameters with equal variances (i.e., $\Sigma_{\theta} = \sigma_{\theta}^2 \mathbf{I}_p$). Notably, the proposed process in (14) is equivalent to the selection process for the estimation of unknown and deterministic parameters (i.e., $\Sigma_{\theta}^{-1} = \mathbf{0}$ in (3)). Later in Sect. 4.2, the assumption of $\Sigma_{\theta} = \sigma_{\theta}^2 \mathbf{I}_p$ is relieved to evaluate the proposed method in numerical experiments.

We prove a theorem which presents a simple criterion to enable an efficient selection of the node that maximizes the log-determinant in (14) at each iteration.

Theorem Let the (i + 1)th column vectors \mathbf{r}_{i+1} of \mathbf{R}^{i+1} and \mathbf{k}_{i+1} of $\mathbf{K}_{S_{i+1}}$ be given by $\mathbf{r}_{i+1}^{\top} = [\mathbf{r}^{\top} \quad r]$ and $\mathbf{k}_{i+1}^{\top} = [\mathbf{k}^{\top} \quad k]$, respectively. Let the (i + 1)th row vector \mathbf{m}_{i+1}^{\top} of $\mathbf{M}_{S_{i+1}}$ be given by $\mathbf{m}_{i+1}^{\top} = [\mathbf{m}^{\top} \quad m]$. Then, the node at the (i + 1)th iteration that maximizes the log-determinant in (14) is simply selected from the set S_i^C consisting of (N - i) remaining nodes:

$$j^* = \arg\max_{j \in S_i^C} \frac{r^2}{m^2}$$
(15)

where $m^2 = k - \| \mathbf{m} \|^2$.

Proof We continue to simplify the metric in (14) to yield

$$\log \det \left(\mathbf{\Sigma}(S_{i+1})^{-1} \right) \approx \log \left(\det \left(\mathbf{R}^{i+1} \right)^2 / \det \left(\mathbf{M}_{S_{i+1}} \right)^2 \right)$$
(16)

where (16) follows from det(\mathbf{A}) = det(\mathbf{A}^{\top}) and det(\mathbf{A}^{-1}) = 1/det(\mathbf{A}). In order to compute det (\mathbf{R}^{i+1}) and det ($\mathbf{M}_{S_{i+1}}$), it is noted that the matrices \mathbf{R}^{i+1} and $\mathbf{M}_{S_{i+1}}$ can be simply obtained from those at the previous iteration. Specifically, for each $j \in S_i^C$ we compute $\mathbf{K}_{S_{i+1}}$ from (6) and also construct \mathbf{R}^{i+1} and $\mathbf{M}_{S_{i+1}}$ as follows:

$$\mathbf{R}^{i+1} = \begin{bmatrix} \mathbf{R}^i & \mathbf{r} \\ \mathbf{0}^\top & \mathbf{r} \end{bmatrix}$$
(17)

$$\mathbf{M}_{S_{i+1}} = \begin{bmatrix} \mathbf{M}_{S_i} & \mathbf{0} \\ \mathbf{m}^\top & m \end{bmatrix}$$
(18)

Here, \mathbf{r}_{i+1} is given by $\mathbf{r}_{i+1}^{\top} = [\mathbf{r}^{\top} \quad r] = [\mathbf{q}_1^{\top} \mathbf{h}_j \quad \mathbf{q}_2^{\top} \mathbf{h}_j \quad \cdots \quad \mathbf{q}_{i+1}^{\top} \mathbf{h}_j]$ where $r = \mathbf{q}_{i+1}^{\top} \mathbf{h}_j$. Also, the row vector $\mathbf{m}_{i+1}^{\top} = [\mathbf{m}^{\top} \quad m]$ can be easily computed by the LU factorization of the symmetric matrix $\mathbf{K}_{S_{i+1}}$:

$$\mathbf{K}_{S_{i+1}} = \begin{bmatrix} \mathbf{K}_{S_i} & \mathbf{k} \\ \mathbf{k}^\top & k \end{bmatrix} = \begin{bmatrix} \mathbf{M}_{S_i} & \mathbf{0} \\ \mathbf{m}^\top & m \end{bmatrix} \begin{bmatrix} \mathbf{M}_{S_i}^\top & \mathbf{m} \\ \mathbf{0}^\top & m \end{bmatrix}$$
(19)

Then, we have from (19)

$$\mathbf{M}_{S_i}\mathbf{m} = \mathbf{k}, \quad m^2 = k - \|\mathbf{m}\|^2 \tag{20}$$

Therefore, the log-determinant in (16) can be further simplified as follows:

$$\log\left(\det\left(\mathbf{R}^{i+1}\right)^{2}/\det\left(\mathbf{M}_{S_{i+1}}\right)^{2}\right)$$

$$=\log\left(\det\left(\mathbf{R}^{i}\right)^{2}/\det\left(\mathbf{M}_{S_{i}}\right)^{2}\right)+\log\frac{r^{2}}{m^{2}}$$
(21)

$$\propto \log \frac{r^2}{m^2}$$
 (22)

where (22) follows since the first term in (21) is irrelevant in finding the maximizing node at the (i + 1)th iteration.

It should be noted that since \mathbf{M}_{S_i} is lower triangular, **m** in (20) can be directly obtained by the forward substitution without computing the inverse of \mathbf{M}_{S_i} . However, (20) requires *i* floating-point divisions repeated over (N - i) remaining nodes at the (i + 1)th iteration. Further, the division operation typically spends much more time than the multiplication. Hence, we avoid those divisions by computing **m** as follows:

$$\mathbf{m} = \mathbf{M}_{S_i}^{-1} \mathbf{k}, \quad m^2 = k - \| \mathbf{m} \|^2$$
 (23)

where $\mathbf{M}_{S_{i+1}}^{-1}$ can be easily computed from $\mathbf{M}_{S_i}^{-1}$ without inversion of large matrices by using

$$\mathbf{M}_{S_{i+1}}^{-1} = \begin{bmatrix} \mathbf{M}_{S_i}^{-1} & \mathbf{0} \\ -\frac{\mathbf{m}^{\top} \mathbf{M}_{S_i}^{-1}}{m} & \frac{1}{m} \end{bmatrix}$$
(24)

Note that (24) can be verified by multiplying $\mathbf{M}_{S_{i+1}}^{-1}$ by $\mathbf{M}_{S_{i+1}}$ in (18). It is obvious that (23) and (24) runs faster than (20) since (23) without divisions is repeated over (N - i) remaining nodes and (24) with one division is conducted once at each iteration.

Initially, $\mathbf{q}_1 = \mathbf{h}_j / \| \mathbf{h}_j \|$, $\mathbf{K}_{S_1} = k = k_{jj}$ with $\mathbf{k} = 0$ for $j \in \mathcal{V}$. Then we have $r = \mathbf{q}_1^{\top} \mathbf{h}_j = \| \mathbf{h}_j \|$ and $m^2 = k$ with $\mathbf{m} = 0$ and the first node is selected as follows:

$$j^* = \arg\max_{j\in\mathcal{V}} \frac{r^2}{m^2} = \arg\max_{j\in\mathcal{V}} \frac{\|\mathbf{h}_j\|^2}{k_{jj}}$$
(25)

Once the first node is selected, we update $\mathbf{M}_{S_{i+1}}^{-1}$ by (24) (e.g., $\mathbf{M}_{S_1}^{-1} = 1/m$) and \mathbf{Q} by the QR factorization. The updated matrices $\mathbf{M}_{S_{i+1}}^{-1}$ and \mathbf{Q} are used at the next iteration. In what follows, the proposed selection algorithm is briefly explained.

Algorithm 1 Efficient greedy sensor selection algorithm

Input: $\mathbf{H} \in \mathbb{R}^{N \times p}$ and $\mathbf{K} \in \mathbb{R}^{N \times N}$ **Output:** set of sensor nodes, S of cardinality |S| = p1. Initialize $S_0 = \emptyset$, $\mathbf{k} = 0$ and $\mathbf{m} = 0$. 2. Select the first node by using (25) and set $S_1 = \{j^*\}$. 3. For i = 1, ..., p - 1 do 1) For each $j \in S_i^C$ where $S_i^C = \mathcal{V} - S_i$, - compute \mathbf{r}_{i+1} where $\mathbf{r}_{i+1}^{\top} = [\mathbf{r}^{\top} \quad r]$ with $(i \times 1)$ column vector \mathbf{r} and the (i+1)-th entry r. - construct $\mathbf{k}_{i+1}^{\top} = \begin{bmatrix} \mathbf{k} & k \end{bmatrix}$ by using (6). - compute $\mathbf{m}_{i+1}^{\top} = [\mathbf{m} \ m]$ by using (23). 2) select the node that maximizes the log-determinant by using (15) and set $S_{i+1} = S_i + \{j^*\}.$ 3) for the j^* -th node selected, update **Q** by using the QR factorization and compute $\mathbf{M}_{S_{i+1}}^{-1}$ by using (24). 4. end for 5. Return $S = S_p$

4 Results and discussion

4.1 Complexity analysis of proposed algorithm

Given the covariance matrix **K** of the correlated noise vector **w** in (1), the proposed algorithm aims to construct the best set S^* by conducting two main tasks. First, computation of the last column vector \mathbf{r}_{i+1} of \mathbf{R}^{i+1} is performed by using the QR factorization of $\mathbf{H}_{S_{i+1}}^{\top} = \mathbf{Q}\bar{\mathbf{R}}^{i+1}$. Specifically, given **Q** and \mathbf{R}^i at the previous iteration, the last entry of \mathbf{r}_{i+1} is computed for each $h_{j,j} \in S_i^C$. Once the maximizing node is selected, **Q** is updated for the next iteration by the Householder transformation. The second is to compute \mathbf{m}_{i+1} by using (23) given $\mathbf{M}_{S_i}^{-1}$ and \mathbf{k}_{i+1} for each node $j \in S_i^C$. After the selection of the maximizing node, $\mathbf{M}_{S_i+1}^{-1}$ is updated for the next iteration by (24). Since these two tasks are repeated |S| - 1 times, the proposed algorithm yields the complexity of $O(Np|S|^2)$ which is the same complexity order as that of [4].

We evaluate three previous selection methods for the performance comparison. First, we consider the sensor selection method denoted by greedy sensor selection (GSS) [5] which seeks to maximize the log-determinant of the inverse error covariance matrix $\Sigma(S)^{-1}$. GSS assumed uncorrelated noise (equivalently, $\mathbf{K} = \mathbf{I}_N$) and the metric is given by log det $\Sigma(S)^{-1} \approx \log \det \mathbf{H}_S^\top \mathbf{H}_S$ under the assumption of $\Sigma_{\theta}^{-1} = \mathbf{0}$. Secondly, We compare with the method denoted by greedy sensor selection based on QR factorization (GSS-QR) in [4] which has been developed by applying the QR factorization of the matrix \mathbf{H}_S^\top in the MSE in (3). Unlike GSS, it aims to directly minimize the MSE(S) with $\mathbf{K} = \mathbf{I}_N$. For this case, the MSE(S) is given by $tr\left[\left(\mathbf{H}_S^\top \mathbf{H}_S\right)^+\right] = ||\mathbf{H}_S^+||_F^2$ when $\Sigma_{\theta}^{-1} = \mathbf{0}$ is assumed. Thirdly, we evaluate the method denoted by greedy sensor selection and algorithm. In Table 1, the complexity and the metrics of the selection methods are

summarized for comparison. As seen, the proposed algorithm, GSS-QR and GSS-CN offer the same complexity order for the case of |S| = p. In the following section, the estimation performance of the methods is experimentally investigated for the case of |S| = p

4.2 Experimental results

In this section, we conduct the extensive experiments to validate the proposed sensor selection algorithm by comparing with various selection methods in three different cases as follows:

- Case 1: Random matrix **H** with Gaussian iid entries, $h_{ij} \sim \mathcal{N}(0, 0.1^2)$.
- Case 2: Random matrix **H** with Bernoulli iid entries, *h_{ij}* which take binary values (0 or 1) with the probability 0.5.
- Case 3: Random matrix **H** and covariance matrix **K** generated by using the linear reduced-order modeling [11, 12].

In Case 1 and 2, we generate 50 different realizations for each type of $\mathbf{H} \in \mathbb{R}^{N \times p}$ with N = 1000 and $p = 10, 15, \ldots, 40$. For each realization of \mathbf{H} , we generate 100 test samples of the parameter vector θ and the correlated noise \mathbf{w} drawn from $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_{\theta})$ and $\mathcal{N}(\mathbf{0}, \mathbf{K})$, respectively. The covariance matrix \mathbf{K} is constructed from random data generated from uniform distribution over $[0 \ 1]$. In Case 3, 100 measurement vectors $\mathbf{y} \in \mathbb{R}^{1000}$ are generated from the normal distribution $\mathcal{N}(0, 1)$ and \mathbf{H} and \mathbf{K} are constructed by using the linear reduced-order modeling with the order equal to p = |S| (see [11, 12] for the details). From the matrices \mathbf{H} and \mathbf{K} , we construct the set S with the cardinality |S| = p by using the various sensor selection methods such as GSS, GSS-QR, GSS-CN and the proposed algorithm. We then select the measurements $y_i, i \in S$ from the measurement vector \mathbf{y} and estimate the parameter vector θ by employing the optimal linear estimator in (2). For evaluation of the estimation performance, we compare the MSEs of the selection methods given by $\mathbb{E} \frac{\|\theta - \hat{\theta}\|^2}{\|\theta\|^2}$.

4.2.1 Performance evaluation with respect to parameter dimension

In this experiment, we generate the parameters by assuming $\Sigma_{\theta} = \sigma_{\theta}^2 \mathbf{I}_p$, $\sigma_{\theta} = 0.1$ and run the selection methods in order to construct the sets *S* with |S| = p by varying the parameter dimension $p = 10, 15, \ldots, 40$ and compare the MSEs which are plotted in Figs. 1, 2 and 3. It is noted that the proposed method shows better performance than GSS and GSS-QR since unlike the methods, the proposed algorithm takes into account the correlation of

Method	Optimality criteria	Decomposition	Operation count
GSS [5]	log det $(\mathbf{H}_{S}^{T}\mathbf{H}_{S})$	-	$O(Np^2 S)$
GSS-QR [4]	$\ \mathbf{H}_{S}^{+}\ _{F}^{2}$	QR	$O(Np S ^2)$
GSS-CN [11, 12]	$\log \det \left(\mathbf{H}_{S}^{\top} \mathbf{K}_{S}^{-1} \mathbf{H}_{S} \right)$	-	$O(N S ^3)$
Proposed method	$\log \det \left(\mathbf{H}_{S}^{T} \mathbf{K}_{S}^{-1} \mathbf{H}_{S} \right)$	QR, LU	$O(Np S ^2)$

Table 1 Comparison with various greedy selection methods



Fig. 1 Evaluation of estimation performance in Case 1 (random Gaussian matrix): the proposed algorithm is compared with different selection methods with $\sigma_{\theta} = 0.1$ by varying the dimension of the parameter, p = |S|



Fig. 2 Evaluation of estimation performance in Case 2 (random Bernoulli matrix): the proposed algorithm is compared with different selection methods with $\sigma_{\theta} = 0.1$ by varying the dimension of the parameter, p = |S|

the noise to construct the set *S*. As expected, the proposed algorithm and GSS-CN yield identical MSEs because both of them optimize the same metric. Notably, as the parameter dimension increases, the estimation accuracy becomes improved because the optimal linear estimator using more measurements can suppress the effect of correlated noise more efficiently. Furthermore, we investigate the estimation performance of the methods with p = |S| = 50 for different signal-to-noise ratios (SNRs) by varying the standard deviation of the parameter, σ_{θ} from 0.1 to 0.5. In Fig. 4, the MSEs are demonstrated in Case 1 (random Gaussian matrix **H**). Obviously, the selection methods yield more accurate estimation



Fig. 3 Evaluation of estimation performance in Case 3 (linear reduced-order model): the proposed algorithm is compared with different selection methods by varying the dimension of the parameter, p = |S|



Fig. 4 Evaluation of estimation performance in Case 1 (random Gaussian matrix): the proposed algorithm is compared with different selection methods with p = |S| = 50 by varying the standard deviation of the parameter, σ_{θ}

performance with higher SNR. We also test the methods by relaxing the assumption of $\Sigma_{\theta} = \sigma_{\theta}^2 \mathbf{I}_p$ and generate the parameters from the non-diagonal covariance matrix Σ_{θ} constructed by using random data uniformly distributed over [0 1]. Notice that the *p* diagonal entries of the covariance matrix Σ_{θ} take much higher values than $\sigma_{\theta}^2 = 0.1^2$ used in the experiment for the case of $\Sigma_{\theta} = \sigma_{\theta}^2 \mathbf{I}_p$. The MSEs are provided in figure 5, showing a robust estimation performance of the proposed method for the case of the non-diagonal covariance matrix. The complexity of the selection methods is experimentally examined by



Fig. 5 Evaluation of estimation performance for non-diagonal covariance matrix of the parameter in Case 1 (random Gaussian matrix): the proposed algorithm is compared with different selection methods by varying the dimension of the parameter, p = |S|



Fig. 6 Evaluation of complexity in Case 1 (random Gaussian matrix): the various selection algorithms are evaluated in terms of the execution time in second by varying the dimension of the parameter, p = |S|

measuring the execution time in second. Figure 6 plots the execution times of the methods in Case 1 (random Gaussian matrix **H**) with respect to the parameter dimension p = |S| = 10, ..., 40. The proposed method offers a competitive complexity as compared with GSS-QR and GSS-CN. Specifically, the proposed method operates about twice as fast as GSS-CN in all three cases.

5 Conclusions

In a situation where the measurements at sensor nodes are contaminated by correlated noise, we considered a problem of finding the best set of sensor nodes which minimizes the estimation error computed by the signal values at the selected nodes. Instead of directly minimizing the estimation error, we focused on maximizing the log-determinant of the inverse estimation error covariance matrix and applied the QR and LU factorizations so as to present a simple analytic selection rule. We also provided a complexity analysis of the proposed algorithm to reveal a competitive complexity when compared with the previous novel selection methods. We finally investigated the performance of the proposed algorithm in various cases in terms of estimation performance and execution time and demonstrated that the proposed algorithm yields more accurate estimation with a reasonable complexity than the previous methods.

Abbreviations

MSE	Mean squared estimation error
MLE	Maximum likelihood estimation
MAP	Maximum a posteriori
MMSE	Minimum mean squared error
GSS	Greedy sensor selection
GSS-QR	Greedy sensor selection based on QR factorization
GSS-CN	Greedy sensor selection under correlated noise
SNR	Signal-to-noise ratio

Acknowledgements

We would like to thank the Editor and the anonymous reviewers for their careful reading of the manuscript and valuable suggestions which led to significant improvements in the manuscript.

Author Contributions

The author read and approved the final manuscript.

Funding

This study was supported by research fund from Chosun University, 2022.

Availability of data and materials

All experiments are described in detail within a reproducible signal processing framework. Please contact the author for data requests.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication Not applicable.

Competing interests

The authors declare that they have no competing interests

Received: 2 October 2023 Accepted: 23 February 2024 Published online: 11 March 2024

References

- 1. S. Joshi, S. Boyd, Sensor selection via convex optimization. IEEE Trans. Signal Process. 57(2), 451–462 (2009)
- M. Nareem, S. Xue, D. Lee, Cross-entropy optimization for sensor selection problems. In: International Symposium on Communication and Information Technology (ISCIT), pp. 396–401 (2009)
- J. Ranieri, A. Chebira, M. Vetterli, Near-optimal sensor placement for linear inverse problems. IEEE Trans. Signal Process. 62(5), 1135–1146 (2014)
- 4. Y.H. Kim, Greedy sensor selection based on QR factorization. EURASIP J. Adv. Signal Process. 2021(117), 1–13 (2021)
- 5. M. Shamaiah, S. Banerjee, H. Vikalo, Greedy sensor selection: leveraging submodularity. In: 49th IEEE International Conference on Decision and Contron (CDC), pp. 2572–2577 (2010)
- Y.H. Kim, Low-complexity sensor selection based on QR factorization. J. Korea Inst. Inform. Commun. Eng. 27(1), 103–108 (2023)

- C. Jiang, Y.C. Soh, H. Li, Sensor placement by maximal projection on minimum eigenspace for linear inverse problems. IEEE Trans. Signal Process. 64(21), 5595–5610 (2016)
- Y. Mo, R. Ambrosino, B. Sinopoli, Sensor selection strategies for state estimation in energy constrained wireless sensor networks. Automatica 47(7), 1330–1338 (2011)
- 9. H. Jamali-Rad, A. Simonetto, G. Leus, X. Ma, Sparsity-aware sensor selection for correlated noise. In: 17th International Conference on Information Fusion (FUSION), pp. 1–7 (2014)
- S. Liu, S.P. Chepuri, M. Fardad, E. Masazade, G. Leus, P.K. Varshney, Sensor selection for estimation with correlated measurement noise. IEEE Trans. Signal Process. 64(13), 3509–3522 (2016)
- 11. K. Yamada, Y. Saito, K. Nankai, T. Nonomura, K. Asai, D. Tsubakino, Fast greedy optimization of sensor selection in measurement with correlated noise. Mech. Syst. Signal Process. **158**(107619) (2021)
- K. Yamada, Y. Saito, T. Nonomura, K. Asai, Greedy sensor selection for weighted linear least squares estimation under correlated noise. IEEE Access 10, 79356–79364 (2022)
- E. Masazade, M. Fardad, P.K. Varshney, Sparsity-promoting extended Kalman filtering for target tracking in wireless sensor networks. IEEE Signal Process. Lett. 19(12), 845–848 (2012)
- 14. S.P. Chepuri, G. Leus, Sparsity-promoting sensor selection for non-linear measurement models. IEEE Trans. Signal Process. 63(3), 684–698 (2015)
- 15. S.M. Kay, Fundamentals of statistical signal processing: estimation theory (Prentice Hall, New Jersey, 1993)
- 16. G. Strang, Linear algebra and its applications, 3rd edn. (Harcourt Brace Jovanovich College Publishers, Florida, 1988)
- 17. R.A. Horn, C.R. Johnson, *Matrix analysis* (Cambridge University Press, New York, 1985)

Publisher's Note

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