2023 Ninth Indian Control Conference (ICC)
18 - 20 December, 2023
GITAM University, Visakhapatnam, India

Adaptive estimation of random vectors with bandit feedback: A mean-squared error viewpoint

Dipayan Sen[1], Prashanth L.A.[1] and Aditya Gopalan[2]

***Abstract*— We consider the problem of sequentially learning to estimate, in the mean-squared error (MSE) sense, a Gaussian K-vector of unknown covariance by observing only $m < K$ of its entries in each round. We first establish a concentration bound for MSE estimation. We then frame the estimation problem with bandit feedback, and we propose a variant of the successive elimination algorithm. We also derive a minimax lower bound to understand the fundamental limit on the sample complexity of this problem.**

I. Introduction

Many real-world applications involve collecting local measurements of a physical phenomenon and leveraging correlation structures to estimate the phenomenon over a larger area. Examples include monitoring temperature over a region [1], estimating the cellular network load using base stations [2], and detecting water contamination using sensors [3].

In this paper, we use the mean-squared error (MSE) to capture the correlation structure. Formally, we consider a jointly Gaussian K-vector X with mean zero and covariance matrix Σ. A Gaussian modeling assumption is shown to be practically valid in a cellular network application [2], where the goal is to estimate traffic load across all base stations by collecting measurements from a few. For a m-subset A, the MSE $\psi(A)$ is given by

$$\psi(A) = \mathrm{Tr}\left(\Sigma_{A'A'} - \Sigma_{A'A}\big(\Sigma_{AA}\big)^{-1}\Sigma_{AA'}\right), \qquad (1)$$

where A' is $[K] \setminus A$, $\Sigma_{AA}, \Sigma_{A'A'}, \Sigma_{A'A}, \Sigma_{AA'}$ are sub-matrices of Σ in obvious notation, and $\mathrm{Tr}(A)$ denotes the trace of the matrix A.

Non-adaptive estimation We address the problem of estimating the MSE for a subset A when provided with i.i.d. samples for each of the sub-matrices defined in (1). This approach is non-adaptive as each sub-matrix entry is sampled equally.

The 'sample-average' estimator $\widehat{\Sigma}_{AA}$ may not always be invertible, even though it is positive definite with high probability. To handle invertibility, we form the matrix $\widehat{\Sigma_{A'A'}}^{+}$ by performing an eigen-decomposition of $\widehat{\Sigma}_{AA}$, followed by a projection of eigenvalues to the positive side. Formally, for $i = 1, \dots, m$, let $\hat{\lambda}_i$ denote the eigenvalue of $\widehat{\Sigma}_{AA}$, with corresponding eigenvector v_i. The estimator $\widehat{\Sigma}^{+}_{AA}$ is defined by

$$\widehat{\Sigma}^{+}_{AA} \triangleq \sum_{i=1}^{m} \hat{\lambda}^{+}_i v_i v_i^{\intercal}, \qquad (2)$$

where $\hat{\lambda}^{+}_i = \begin{cases} \hat{\lambda}_i & \text{if } |\hat{\lambda}_i| \geq \zeta, \\ \zeta & \text{otherwise,} \end{cases}$ for $i = 1, \dots, m$ with $\zeta > 0$. The MSE $\psi(A)$ associated with set A is then estimated as follows:

$$\widehat{\psi}(A) \triangleq \mathrm{Tr}\left(\widehat{\Sigma_{A'A'}} - \widehat{\Sigma_{A'A}}\big(\widehat{\Sigma_{AA}}^{+}\big)^{-1}\widehat{\Sigma_{AA'}}\right). \qquad (3)$$

In Section III of [4], we present a concentration bound for the MSE estimator (3), which shows that the tail decay is sub-Gaussian.

Adaptive estimation. In this setting, we focus on non-uniform sampling of the covariance matrix to optimize the reuse of samples for estimating MSE across various subsets.

When concerned with estimating the MSE for a specific subset, one can use the estimate in (3). However, if one need to estimate the MSE for multiple subsets using the same samples efficiently, one can maintain estimates for each entry of the covariance matrix and then extract the necessary information from the sample covariance matrix to form MSE estimates for any given subset. Based on this approach, for a subset $A = \{i_1, \dots, i_m\}$, the MSE $\psi(A)$, can be re-written as follows:

$$\psi(A) = \sum_{j=1}^{K} \left[\sigma_j^2 - C_j(\Sigma_{AA}^{-1})C_j^{\intercal}\right], \qquad (4)$$

where $C_j = [\rho_{ji_1}\sigma_{i_1}\sigma_j \ \dots \ \rho_{ji_m}\sigma_{i_m}\sigma_j]$. To estimate the MSE as described in (4), we need to estimate the variances and correlation coefficients. We're provided with n_i samples for variance σ_i^2 and n_{ij} samples for the correlation coefficient ρ_{ij}, with $i, j \in [K], i \neq j$.

For each $j = 1, \dots, K$ and $k = 1, \dots, m$, we compute $\hat{\sigma_j}^2$ and $\hat{\rho}_{ji_k}$. These are formed using n_j and n_{ji_k} samples, respectively, as follows:

$\hat{\sigma_j}^2 = \overline{X_j}^2$, where $\overline{X_j}^2 = \frac{1}{n_j}\sum_{t=1}^{n_j} X_{jt}^2$.

$\hat{\rho}_{ji_k} = \frac{\overline{X_j X_{i_k}}}{\hat{\sigma} j \hat{\sigma} i_k}$, where $\overline{X_j X_{i_k}} = \frac{1}{n_{ji_k}}\sum_{t=1}^{n_{ji_k}} X_{jt}X_{i_k t}$.

Using these estimates of variance and correlation coefficients, we estimate the MSE $\psi(A)$ as follows:

[1]Department of Computer Science and Engineering, Indian Institute of Technology Madras, Chennai 600036, India, {cs18s012,prashla}@cse.iitm.ac.in
[2]Department of Electrical Communication Engineering, Indian Institute of Science, Bengaluru 560012, India, aditya@iisc.ac.in

2023 Ninth Indian Control Conference (ICC) | 979-8-3503-9525-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICC61519.2023.10442119

$$\widehat{\psi}(A) = \sum_{j=1}^{K} \left[\hat{\sigma_j}^2 - \widehat{C_j} \big(\widehat{\Sigma_{AA}}^{+} \big)^{-1} \widehat{C_j}^{\intercal} \right], \qquad (5)$$

where $\widehat{C_j} = \left[\hat{\rho}_{ji_1} \hat{\sigma}_{i_1} \hat{\sigma}_j \ldots \hat{\rho}_{ji_m} \hat{\sigma}_{i_m} \hat{\sigma}_j \right]$, $\widehat{\Sigma}_{AA}$ formed by using the relevant sample correlation coefficients $\hat{\rho}_{i_k i_l}, i_k, i_l \in A$, and sample variances $\hat{\sigma}^2_{i_k}, i_k \in A$, and $\widehat{\Sigma}^{+}_{AA}$ is defined in (2).

In Section IV of [4], we provide a tail bound for the MSE estimator (5), which establishes exponential concentration of this estimator around the true MSE.

Adaptive estimation with bandit feedback. The goal here is to find the optimal subset with the lowest MSE, denoted as A^*:

$$A^* \in \arg\min_{A \in \mathcal{A}} \psi(A).$$

For finding A^*, we aim to develop a δ-PAC algorithm that efficiently finds the best m-subset with high probability. For a given confidence parameter $\delta \in (0, 1)$, a δ-PAC algorithm stops after τ rounds and returns a set A_τ such that $\mathbb{P}\left(A_\tau \neq A^*\right) \leq \delta$. We prefer the algorithm with the lowest sample complexity $\mathbb{E}[\tau]$.

For any set A, define

$$\Delta(A) \triangleq \psi(A) - \psi(A^*), \text{ and } \Delta = \min_{A \in \mathcal{A}} \Delta(A). \qquad (6)$$

where $\Delta(A)$ represents the difference in MSE for a subset A, and Δ is the smallest such difference.

Successive elimination for correlated bandits. We consider the fixed confidence variant of the best-arm identification framework [5]. In the fixed confidence setting that we consider, a naive approach, based on Algorithm 1 in [6], would be to sample each subset an equal number of times. This uniform sampling is effective when all subsets provide similar information about others, i.e., when correlations and variances are similar. However, in the presence of varying correlations, uniform sampling becomes less suitable. To minimize the error in identifying the best m-subset, it's essential to sample the set of candidates for the most informative subset more frequently, and the successive elimination technique [6] embodies this idea.

We propose a modified successive elimination algorithm tailored for finding the best m-subset based on the MSE objective. The algorithm begins with an active set, initially comprising all m-subsets, denoted as $\mathcal{A}$. In each round t, the algorithm samples each active m-subset once and estimates its MSE using (5). Subsequently, the algorithm eliminates subsets whose confidence intervals are clearly separated from the confidence interval of the subset with the lowest estimated MSE seen thus far. The algorithm continues until only one m-subset remains in the active set, which inevitably occurs with probability one.

For the algorithm mentioned above, we derive a sample complexity bound of the form $O\left(\frac{\binom{K}{m}}{\Delta} \log\left(\frac{\binom{K}{m}}{\delta}\right)\right)$. The reader is referred to Section V of [4] for the details.

Lower bound. We consider a special case of the adaptive estimation problem, where the goal is to identify the best pair of arms, i.e.,

$$(i_1^*, i_2^*) \in \arg\min_{(i,j) \in [K] \times [K], i \neq j} \psi(\{i, j\}).$$

Let $Alg(\delta, K)$ denote the class of algorithms that are δ-PAC for the best pair identification problem.

In Section VI of [4], we prove that, for any δ-PAC algorithm, there exists a bandit problem instance governed by a covariance matrix Σ such that the sample complexity $\mathbb{E}_\Sigma[\tau_\delta]$ of this algorithm satisfies

$$\mathbb{E}_\Sigma[\tau_\delta] \geq \frac{\log(\frac{1}{2.4\delta})}{\Delta}. \qquad (7)$$

where Δ denotes the smallest gap on the problem instance governed by Σ.

The reader is referred to the longer version of this paper, available in [4], for the following: (i) the concentration bounds for MSE estimators presented in (3) and (5); (ii) a detailed description and a sample complexity analysis for a bandit algorithm designed to find the most correlated subset (in the MSE sense) under the fixed confidence best-arm identification framework; and (iii) the proof of the lower bound presented in (7).

References

[1] C. Guestrin, A. Krause, and A. P. Singh, "Near-optimal sensor placements in gaussian processes," in *International conference on Machine learning*, 2005, pp. 265–272.

[2] U. Paul, L. Ortiz, S. R. Das, G. Fusco, , and M. M. Buddhikot, "Learning probabilistic models of cellular network traffic with applications to resource management," in *IEEE International Symposium on Dynamic Spectrum Access Networks*, 2014, pp. 82–91.

[3] A. Krause, J. Leskovec, C. Guestrin, J. VanBriesen, and C. Faloutsos, "Efficient Sensor Placement Optimization for Securing Large Water Distribution Networks," *Journal of Water Resources Planning and Management*, vol. 134, no. 6, pp. 516–526, 2008.

[4] D. Sen, L. Prashanth, and A. Gopalan, "Adaptive estimation of random vectors with bandit feedback," *arXiv preprint arXiv:2203.16810*, 2023.

[5] T. Lattimore and C. Szepesvári, *Bandit algorithms*. Cambridge University Press, 2020.

[6] E. Even-Dar, S. Mannor, and Y. Mansour, "PAC bounds for multi-armed bandit and Markov decision processes," in *International Conference on Computational Learning Theory*. Springer, 2002, pp. 255–270.