Nystrom-based Localization in Precision Agriculture Sensors

Samuel Lichtenberg[1], Abiy Tasissa[1]

[1] Department of Mathematics, Tufts University, Medford, MA 02155, USA.

A paper from the Proceedings of the
16th International Conference on Precision Agriculture
21-24 July 2024
Manhattan, Kansas, United States

Abstract.

Wireless sensor networks play a pivotal role in a myriad of applications, including agriculture, health monitoring, tracking and structural health monitoring. One crucial aspect of these applications involves accurately determining the positions of the sensors. In this paper, we study a novel Nystrom based sampling protocol in which a selected group of anchor nodes, with known locations, establish communication with only a subset of the remaining sensor nodes. Leveraging partial distance information, we present an efficient algorithm for estimating sensor locations. To demonstrate the effectiveness of our approach, we provide empirical results using synthetic data and underscore the practical advantages of our sampling technique for precision agriculture.

Keywords.
Precision agriculture, Nystrom method, Sensor localization, Distance geometry, Optimization.

Introduction

Precision agriculture refers to the discipline that uses technology and scientific principles to improve farming practices, aiming to boost crop yields and protect environmental health (Bayih et al., 2022, Pierce et al., 1999). Key technologies in this field include GPS, wireless sensor networks (WSNs), remote sensing, and fertilizers (Oliver et al., 2013, Bhakta et al., 2019). This paper focuses on WSNs, which are employed for tasks such as irrigation scheduling and environmental condition monitoring (Ojha et al., 2015). WSNs typically consist of battery-powered sensors interconnected wirelessly to perform specific tasks (Akyildize et al., 2010). In many applications, it is crucial to determine the sensors' locations using cost-effective and energy-effective protocols, which leads to the sensor localization problem. One method relies on the distance information between sensors (Rudafshani et al., 2007). Specifically, assume that there are $\boldsymbol{n}+\boldsymbol{m}$ sensors, and the locations of $\boldsymbol{m}$ of these sensors are known. The sensors with known locations are called anchors. For the remaining sensors, whose positions are unknown, we have partial pairwise distance information between some of them. The localization problem is then to estimate the positions of these $\boldsymbol{n}$ sensors using the anchors and the available distance information.

One approach for localization is the Nystrom method (Platt, 2005, Williams et al., 2000). Here on, sensor nodes without position information are referred to as mobile nodes. The Nystrom method assumes that the distances from each mobile node to all $\boldsymbol{m}$ anchor nodes are known, but there is no distance information between any pairs of mobile nodes. Let $\boldsymbol{p_1}, \boldsymbol{p_2}, \dots, \boldsymbol{p_m}$ represent the positions of the $\boldsymbol{m}$ anchor nodes and let $\boldsymbol{p_{m+1}}, \dots, \boldsymbol{p_{m+n}}$ represent the positions of the $\boldsymbol{n}$ mobile nodes. In the Nystrom method, the distance matrix is structured as follows:

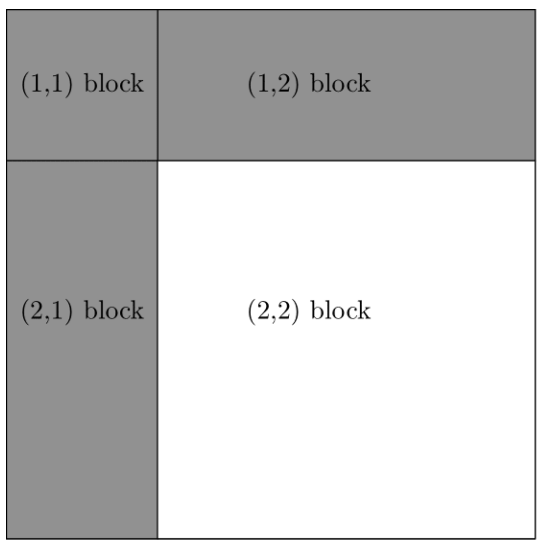

Fig.1. Illustration of the distance matrix in the Nystrom method. The (1,1) block is the anchor-anchor distance matrix. The (1,2) and (2,1) blocks are the anchor-mobile and mobile-anchor distance matrices respectively. The (2,2) block is the mobile-mobile distance matrix. Grey indicates the distances are known in that block while white denotes the distance information is missing in that block.

One advantage of Nystrom localization is that, under mild assumptions about the anchor nodes and assuming all available distances are precise, it provides the exact positions for the n mobile nodes (Kumar et al., 2009). Typically, the anchors are assumed to make up a small percentage of the total number of sensors. However, the requirement to know all distances between a given mobile node and all anchors can be restrictive due to factors like weather and geographical limitations. In this paper, we relax this assumption and instead assume that only some of the distances between mobile nodes and anchors are known. Additionally, to formulate a well-posed mathematical problem, we make one additional assumption: we know the distances of the mobile nodes from one specific anchor. Unlike the Nystrom method, this approach only requires complete distance information to one anchor node. For simplicity, we assume that this complete distance information pertains to the m-th anchor. This method was introduced in our previous work (Lichtenberg et al., 2023), where the theoretical analysis was established. This paper focuses on its application. In our proposed method, the distance matrix has the following structure:

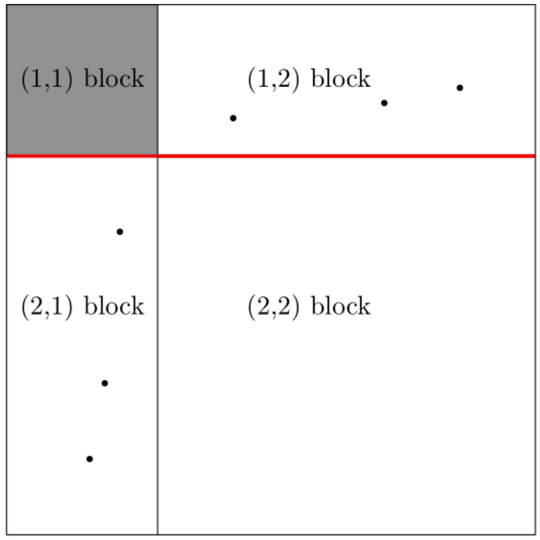

Fig.2. Illustration of the distance matrix in our proposed method. The (1,1) block is the anchor-anchor distance matrix. The (1,2) and (2,1) blocks are the anchor-mobile and mobile-anchor distance matrices respectively. The black dots indicate observed distance entries in these blocks. The (2,2) block is the mobile nodes - mobile nodes distance matrix. Grey indicates the distances are known in that block while white denotes the distance information is missing in that block. The red line indicates the last anchor node from which we know the distances to all mobile nodes.

In the next section, we outline the main concept behind our method which estimates the positions of the n mobile nodes.

Proposed Approach

An essential object in our approach is the Gram matrix. The Gram matrix is an $(n+m)\times(n+m)$ matrix where the (i,j)-th entry is $G_{i,j} = p_i^T p_j$. If the points lie in an r-dimensional space, the rank of the Gram matrix is r. Note that the rank is independent of the number of points, and this fact is useful for optimization. In the context of our proposed method (see Fig. 2), the Gram matrix has the following block structure:

$$G = \begin{matrix} A & B \\ B^T & C \end{matrix} \tag{1}$$

In the above block-structure of G, A is $m \times m$, B is $m \times n$ and C is $n \times n$. Since the anchors are known, A is the Gram matrix corresponding to the anchors and is thus known. In the standard Nystrom problem, it is assumed that the distances from every mobile sensor to all anchor sensors are known. In that setting, the Nystrom method first computes B and then uses the following closed form formula to estimate C (Platt, 2005, Williams et al., 2000):

$$C_* = B^T A^+ B, \tag{2}$$

where A^+ denotes the pseudinverse of A. As noted in the introduction, under mild assumptions, $C_* = C$ (Kumar et al., 2009). However, (2) cannot be directly applied when we do not have access to all distances between a mobile node and all the anchors. In that case, B cannot be explicitly computed, which necessitates an alternative approach. Let $X = [p_1, p_2, \dots, p_m]$ and $Y = [p_{m+1}, \dots, p_{m+n}]$ where X is $r \times m$ and Y is $r \times n$. Motivated by the theoretical analysis developed in (Lichtenberg et al., 2023), the proposed problem is:

$$Find\ Y \in R^{r\times n}\ \ subject\ to\ (X^TY)_{i,j} - (X^TY)_{m,j} = H_{i,j} \quad (i,j) \in \Omega, \tag{3}$$

where Ω consists of the indices of the available distances between the anchors and the mobile nodes and the entries $H_{i,j}$ are known constants related to the known distances. For simplicity of presentation, we have omitted the explicit form of $H_{i,j}$ and we refer the interested reader to (Lichtenberg et al., 2023) for details. We highlight a few advantages of our proposed approach. First, our method does not require knowing all the distances between a mobile node and all the anchors. Second, (3) is a convex optimization program and one can employ existing solvers to estimate the positions of the mobile nodes efficiently. We note that our approach is connected to low rank matrix completion problem (Candes et al., 2010, Keshavan et al., 2010, Recht, 2011) and Euclidean distance geometry (EDG) problem (Dokmanic et al., 2015, Liberti et al., 2017). Our approach differs by using a novel sampling model that extends the Nystrom method. To our knowledge, such a model is new to the matrix completion and EDG literature.

Results

To demonstrate the feasibility of our method for agricultural applications, we conducted the following numerical experiments. We sample $m+n$ random points from a uniform $[0.0, 1.0)$ distribution in 2 dimensions, where $n = 1000,\ m = 21$. Let P denote the $(m+n) \times 2$ point matrix consisting of these points. We assume that the points are corrupted by a Gaussian noise with a mean 0 and standard deviation 0.001. Next, we compute the full distance matrix of these points. Let γ denote the sampling rate. For each mobile node, we assume that we know the distances to $(m-1) \times \gamma$ anchors, which are selected uniformly at random. We use $m-1$ since we assume the distance from each mobile node to all anchor nodes is known (see Fig. 2).

To solve the convex optimization problem in (3), we use CVXPY (Agrawal et al., 2018, Diamond et al., 2016). Let Y_* denote the optimal solution, which corresponds to the estimates of the positions of the mobile nodes. Since any solution is unique only up to rotations, we use an orthogonal Procrustes alignment algorithm to compare Y_* with the original mobile nodes and compute the root mean squared error (RMSE). For each choice of sampling rate, we perform 25 trials, and report the mean and standard deviation of the RMSE values across those trials. All experiments are performed on an M1 2020 MacBook Air with 8GB RAM. The results of our numerical experiments are summarized in Table 1. We observe that, even when each mobile node has distance information to only 2 randomly selected anchors, the positions of the mobile nodes are estimated accurately. This suggests that the proposed method is effective in low-resource settings where it is essential to minimize power consumption and cost.

Table 1. RMSE statistics of reconstructed mobile nodes as a function of sampling rate.

γ (Sampling rate)	Distance samples per mobile node	Mean of RMSE	Standard deviation of RMSE
0.05	1	4.806	4.000
0.1	2	5.977×10^{-1}	3.355×10^{-1}
0.2	4	8.031×10^{-14}	2.457×10^{-13}
0.3	6	6.927×10^{-16}	2.503×10^{-16}

Conclusion or Summary

In this paper, we apply a method proposed in (Lichtenberg et al., 2023) for a precision agriculture application. Our framework requires knowing the location of a few sensors and estimates the positions of the remaining sensors relying on limited distance information from the anchors. Our empirical experiments on synthetic data demonstrate that the proposed method can be beneficial for localization in resource-constrained settings.

Acknowledgments

Abiy Tasissa acknowledges support from NSF through grant DMS-2208392.

References

Agrawal, A., Verschueren, R., Diamond, S., & Boyd, S. (2018). A rewriting system for convex optimization problems. *Journal of Control and Decision*, *5*(1), 42-60.

Akyildiz, I. F., & Vuran, M. C. (2010). *Wireless sensor networks*. John Wiley & Sons.

Bayih, A. Z., Morales, J., Assabie, Y., & de By, R. A. (2022). Utilization of internet of things and wireless sensor networks for sustainable smallholder agriculture. *Sensors*, *22*(9), 3273.

Bhakta, I., Phadikar, S., & Majumder, K. (2019). State-of-the-art technologies in precision agriculture: a systematic review. *Journal of the Science of Food and Agriculture*, *99*(11), 4878-4888.

Candes, E. J., & Plan, Y. (2010). Matrix completion with noise. *Proceedings of the IEEE*, *98*(6), 925-936.

Diamond, S., & Boyd, S. (2016). CVXPY: A Python-embedded modeling language for convex optimization. *Journal of Machine Learning Research*, *17*(83), 1-5.

Dokmanic, I., Parhizkar, R., Ranieri, J., & Vetterli, M. (2015). Euclidean distance matrices: essential theory, algorithms, and applications. *IEEE Signal Processing Magazine*, *32*(6), 12-30.

Keshavan, R. H., Montanari, A., & Oh, S. (2010). Matrix completion from a few entries. *IEEE transactions on information theory*, *56*(6), 2980-2998.

Kumar, S., Mohri, M., & Talwalkar, A. (2009, June). On sampling-based approximate spectral decomposition. In *Proceedings of the 26th annual international conference on machine learning* (pp. 553-560).

Liberti, L., & Lavor, C. (2017). *Euclidean distance geometry* (Vol. 3). Berlin: Springer.

Lichtenberg, S., & Tasissa, A. (2023). A Nystr om method with missing distances. *arXiv preprint arXiv:2311.18076*.

Ojha, T., Misra, S., & Raghuwanshi, N. S. (2015). Wireless sensor networks for agriculture: The state-of-the-art in practice and future challenges. *Computers and electronics in agriculture*, *118*, 66-84.

Oliver, M. A., Bishop, T. F., & Marchant, B. P. (Eds.). (2013). *Precision agriculture for sustainability and environmental protection* (pp. 1-283). Abingdon: Routledge.

Pierce, F. J., & Nowak, P. (1999). Aspects of precision agriculture. *Advances in agronomy*, *67*, 1-85.

Platt, J. (2005, January). Fastmap, metricmap, and landmark mds are all nyström algorithms. In *International Workshop on Artificial Intelligence and Statistics* (pp. 261-268). PMLR.

Recht, B. (2011). A simpler approach to matrix completion. *Journal of Machine Learning Research*, *12*(12).

Rudafshani, M., & Datta, S. (2007, April). Localization in wireless sensor networks. In *Proceedings of the 6th international conference on Information processing in sensor networks* (pp. 51-60).

Williams, C., & Seeger, M. (2000). Using the Nyström method to speed up kernel machines. *Advances in neural information processing systems*, *13*.