Estimating the matching map between two sets of high-dimensional noisy features

Online Seminar Research Unit 5381

A. Dalalyan (with A. Minasyan, T. Galstyan, S. Hunanyan)

CREST, ENSAE, Institut Polytechnique de Paris

June 2, 2022

Introduction

- Finding the best match between two clouds of points is a problem encountered in many real problems.
- In computer vision, one can look for correspondences between two sets of local descriptors extracted from two images.
- In text analysis, one can be interested in matching vector representations of the words of two similar texts, potentially in two different languages.

- The goal of the present work is to gain theoretical understanding of the statistical limits of the matching problem.

Mathematical formulation of the problem (no outliers)

- Notation: $[n] = \{1, \dots, n\}$ for any integer n and $\|\cdot\|$ the Euclidean norm in $\mathbb{R}^d$
- Assume 2 independent sequences $\boldsymbol{X} = (X_i; i \in [n])$ and $\boldsymbol{Y} = (Y_i; i \in [n])$ of independent vectors, such that $X_i, Y_i \sim P_i$ on $\mathbb{R}^d$.
- We observe $\boldsymbol{X}$ and a shuffled version $\boldsymbol{X}^{\#}$ of $\boldsymbol{Y}$.
  That is $X_i^{\#} = Y_{\pi^*(i)}$ for some unobserved permutation π^*.
- The goal is to recover the π^* from data $(\boldsymbol{X}, \boldsymbol{X}^{\#})$.

We consider the simplest case when $P_i = \mathcal{N}_d(\theta_i, \sigma^2 \mathbf{I}_d)$, leading to

$$\begin{cases} X_i = \theta_i + \sigma\xi_i \ , \\ X_i^{\#} = \theta_i^{\#} + \sigma\xi_i^{\#}, \end{cases} \quad i = 1, \dots, n \quad \text{and} \quad \boxed{\theta^{\#}_{\pi^*(i)} = \theta_i} \quad (1)$$

Mathematical formulation of the problem with outliers

- We observe $\{X_i\}_{i\in[n]}$ and $\{X_i^{\#}\}_{i\in[n]}$ such that

$$\begin{cases} X_i = \theta_i + \sigma\xi_i \ , \\ X_i^{\#} = \theta_i^{\#} + \sigma\xi_i^{\#}, \end{cases} \tag{2}$$

  such that there exists $S \subset [n]$ and an injective mapping $\pi^* : S \to [n]$ such that $\theta^{\#}_{\pi^*(i)} = \theta_i$ for every $i \in S$.

- Relevant quantity for matching: the signal-to-noise ratio

$$\kappa \triangleq \min_{j\neq\pi^*(i)} \|\theta_i - \theta_j^{\#}\|_2/\sigma. \tag{3}$$

- If $\kappa = 0$, then it is impossible to recover π^*.
- **Question**: what is the smallest value of κ for which consistent recovery of π^* is possible?

Illustration of the considered framework described in (2)

Known S [Optimal detection of the feature matching map in presence of noise and outliers (arXiv:2106.07044)]

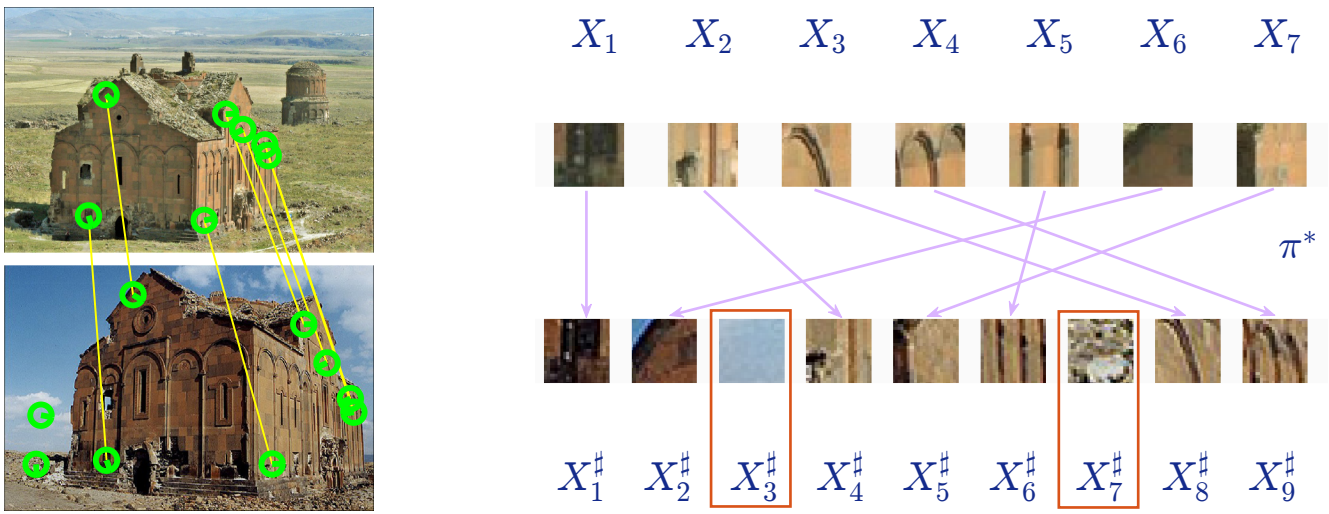

Figure: We wish to match a set of 7 patches extracted from the first image to the 9 patches from the second image. The picture on the left shows the locations of patches as well as the true matching map π^* (the yellow lines).

Summary of main results

- We prove that if $\kappa = c\{(d \log n)^{1/4} \vee \log^{1/2} n\}$ then it is impossible to recover π^*. That is, for each estimator $\widehat{\pi}$, there is a collection of vectors θ_i with signal to noise ratio κ and a permutation π^* such that $\mathbf{P}(\widehat{\pi} \neq \pi^*) > 0.1$.
- Let $\alpha \in (0, 1)$ and

$$\lambda(n, d, \alpha) = 4\Big(\big(d \log(4n^2/\alpha)\big)^{1/4} \vee \big(8 \log(4n^2/\alpha)^{1/2}\big) \Big). \qquad (4)$$

  We also prove that if $\kappa > \lambda$, then consistent recovery is possible, that is there is a procedure $\widehat{\pi}_n$ satisfying $\mathbf{P}(\widehat{\pi}_n \neq \pi^*) \leq \alpha$.
- The procedure attaining the bound above is computationally tractable and combines ideas from model selection and least sum of squares.

Warm up 1: The case $S = [n]$ (no outlier)

Collier and D., Minimax rates in permutation estimation for feature matching, JMLR 2016

- The optimal rate is $\left(d \log(4n^2/\alpha)\right)^{1/4} \vee \left(8 \log(4n^2/\alpha)^{1/2}\right)$.
- The LSS estimator is optimal

$$\widehat{\pi}^{\mathsf{LSS}} \in \operatorname*{argmin}_{\pi} \sum_{i \in [n]} \|X_i - X^{\#}_{\pi(i)}\|_2^2. \tag{5}$$

- This can be rewritten as

$$\widehat{\Pi}^{\mathsf{LSS}} \in \operatorname*{argmin}_{\Pi} \sum_{i,j \in [n]} \|X_i - X^{\#}_{j}\|_2^2 \Pi_{i,j} \tag{6}$$

  where Π is a bistochastic matrix (positive entries, sums of rows and columns equal to 1).
- The problem is also called "assignment problem" and can be solved using the Hungarian algorithm. Complexity is $O(n^3)$.
- Interestingly, the LSS procedure is sub-optimal when the noise is heteroscedastic.

Warm up 2: The case of known $S \subset [n]$ (no outlier at left)

Optimal detection of the feature matching map in presence of noise and outliers, arXiv:2106.07044

- The LSS estimator is still optimal

$$\widehat{\pi}^{\mathsf{LSS}} \in \operatorname*{argmin}_{\pi} \sum_{i \in [n]} \|X_i - X^{\#}_{\pi(i)}\|_2^2. \tag{7}$$

- The problem is also called "imperfect assignment problem" and can be solved using the extended Hungarian algorithm.
- Interestingly, when the noise is heteroscedastic with unknown variances, the rate degrades to $\left(d \log(4n^2/\alpha)\right)^{1/2}$.

Warm up 3: partly known $k^* = |S|$ (number of inliers)

- We can use the LSS procedure

$$\widehat{\pi}_k^{\mathsf{LSS}} \in \operatorname*{argmin}_{\pi \in \mathcal{P}_k} \sum_{i \in S_\pi} \|X_i - X^{\#}_{\pi(i)}\|_2^2, \tag{8}$$

where S_π denotes the support of function π and $\mathcal{P}_k$ is defined by

$$\mathcal{P}_k := \Big\{ \pi : S \to [n] \text{ s.t. } S \subset [n], |S| = k \text{ and } \pi \text{ is injective} \Big\}. \tag{9}$$

- We show that this problem can indeed be solved efficiently with complexity $\tilde{\mathcal{O}}(\sqrt{k}\, n^2)$ using the min-cost flow algorithm.

Theorem 1 Let $\alpha \in (0,1)$ and

$$\lambda(n,d,\alpha) = 4\Big(\big(d \log(4n^2/\alpha)\big)^{1/4} \vee \big(8 \log(4n^2/\alpha)^{1/2}\big) \Big). \tag{10}$$

Denote $\widehat{S} \triangleq \mathbf{supp}(\widehat{\pi})$ for $\widehat{\pi} = \widehat{\pi}_k^{\mathsf{LSS}}$ defined by (8). If $k \le k^*$ and $\kappa \ge \lambda(n,d,\alpha)$ then, with probability at least $1-\alpha$,

$$\mathbf{P}\big(\widehat{S} \subset S^* \text{ and } \widehat{\pi}(i) = \pi^*(i), \forall i \in \widehat{S}\big) \ge 1 - \alpha. \tag{11}$$

Warm up 4: known σ (level of noise)

- We still use the LSS procedure

$$\widehat{\pi}_k^{\mathsf{LSS}} \in \operatorname*{argmin}_{\pi \in \mathcal{P}_k} \sum_{i \in S_\pi} \|X_i - X^{\#}_{\pi(i)}\|_2^2, \tag{12}$$

- We also set

$$\widehat{\Phi}(k) = \min_{\pi \in \mathcal{P}_k} \sum\nolimits_{i \in S_\pi} \|X_i - X^{\#}_{\pi(i)}\|_2^2. \tag{13}$$

- Define

$$\widehat{k}(\alpha) = 1 + \max\left\{k < n : \widehat{\Phi}(k+1) - \widehat{\Phi}(k) \leq (d + \lambda_n(\alpha)^2/4)\sigma^2\right\}$$

Theorem 2 If for some $\alpha \in (0,1)$ we have $\kappa > \lambda(n,d,\alpha)$ then it holds that $\mathbf{P}\big(\widehat{k}(\alpha) = k^* \text{ and } \widehat{\pi}_{\widehat{k}(\alpha)} = \pi^*\big) \geq 1 - \alpha$. Therefore, $\lambda(n,d,\alpha)$ is an upper bound on the separation distance in the case of unknown k^* as well.

Final result: unknown k^* and σ

For every k we can define the estimator

$$\bar{\sigma}_k^2 = \frac{1}{(1-\gamma)kd} \min_{\pi \in \mathcal{P}_k} \sum\nolimits_{i \in S_\pi} \|X_i - X^{\#}_{\pi(i)}\|_2^2, \tag{14}$$

The algorithm reads as follows: initialize $k \leftarrow k_{\min}$

1. Compute $\bar{\sigma}_k^2$ using (14).
2. If $k = n$ or $\widehat{\Phi}(k+1) - \widehat{\Phi}(k) > (d + \lambda(\alpha))\bar{\sigma}_k^2$, then output $(k, \bar{\sigma}_k^2, \widehat{\pi}_k^{\mathsf{LSS}})$.
3. Otherwise, put $k \leftarrow k+1$ and go to Step 1

Theorem 3 If for some $\alpha \in (0,1)$ we have $\kappa > 2\lambda(n,d,\alpha)$ then it holds that $\mathbf{P}\big(\widehat{k}(\alpha) = k^* \text{ and } \widehat{\pi} = \pi^*\big) \geq 1 - \alpha$.

Conclusion

- We have analyzed the problem of matching map recovery between two sets of feature vectors when the number k^* of true matches is unknown.
- First, assuming a lower bound on k^* is available, we proved that that the k-LSS procedure with high probability makes no mistake under the weakest possible condition on the signal-to-noise ratio.
- Secondly, we proposed a procedure for estimating unknown matching size k^*, even when noise levels σ and $\sigma^{\#}$ are unknown. We proved that this procedure finds the correct value of k^* and the true matching map π^* with high probability.

Thank you for listening!

https://www.crest.science