4

1 Paper Details

Title Low-rank matrix completion using alternating minimization

Authors Sujay Sanghavi, Praneeth Netrapalli, Prateek Jain

Venue STOC '13 forty-fifth annual ACM symposium on Theory of Computing

Year of Publication 2013

Problem Formulation This paper focusses on Matrix Sensing and Matrix Completion problems which are given below respectively

Matrix Sensing

$$\text{Find } X \in \mathbb{R}^{m \times n}, \text{ such that } \mathcal{A}(X) = b, \quad \text{rank}(X) \leq k \tag{1}$$

Here for the original matrix M ($\mathbb{R}^{m \times n}$), $\mathcal{A}$ is an operator acting on it defined as, $\mathcal{A}(M) = b$ where each entry of this d dimensional vector is given by $b_i = \text{trace}(A_i^* M)$ where each A_i ($\mathbb{R}^{m \times n}$) is a measurement matrix.

Matrix Completion

$$\text{Find } X \in \mathbb{R}^{m \times n}, \text{ such that } \|P_\Omega(X) = P_\Omega(M)\|, \quad \text{rank}(X) \leq k \tag{2}$$

Here for the original matrix M ($\mathbb{R}^{m \times n}$), P_Ω is an operator on it defined as below

$$P_\Omega(M) = \begin{cases} M_{ij} & \text{if } (i,j) \in \Omega \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

and $\Omega \subset \{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\}$.

Cost Function Now with the variables and operators defined, here is the *approximate* cost function with the algorithm that optimises it for each of the above problem.

Matrix Sensing solved by the algorithm `AltMinSense`.

$$\min_{U \in \mathbb{R}^{m \times k}, V \in \mathbb{R}^{n \times k}} \|\mathcal{A}(UV^*) - b\|_2^2 \tag{4}$$

Matrix Completion solved by the algorithm `AltMinComplete`.

$$\min_{U \in \mathbb{R}^{m \times k}, V \in \mathbb{R}^{n \times k}} \|P_\Omega(UV^*) - P_\Omega(M)\|_F^2 \tag{5}$$

Comparison with SVT

Advantages over SVT

- In low-rank matrix problems, k is much smaller than m, n, due to this U, V ($k(m+n)$ entries) are an order of magnitude smaller than $X(mn$ entries) which leads to more efficient solutions.
- It is easier to impose more constraints on X by imposing them on U, V such as sparse PCA where only U needs to be sparse.

Disadvantages over SVT

- Random initialisation might not be successful and instead smart initialisation is required to satisfy some property with the target subspace.
- SVT is conceptually simpler, as we directly solve a convex optimisation problem whereas the problems given above are non-convex in nature.