Approximation Algorithms for $\ell_0$-Low Rank Approximation

Karl Bringmann
Pavel Kolev
David P. Woodruff

The $\ell_0$-rank-k Problem

Input: a field $F$, a matrix $A \in F^{m \times n}$, an integer $k \in [n]$
Output: a matrix $A^* \in F^{m \times n}$ such that $A^* = \arg\min_{\text{rank}(A') = k} ||A - A'||_0$, where the number of non-zero entries.
Def: $\text{OPT}^k := ||A - A^*||_0$ and $0 \leq \text{OPT}^k \leq ||A||_0$

The Robust PCA Problem

Promise: $A = A^* + S$, rank($A^*$) = $k \ll n$, $S$ is sparse.
Goal: recover the low rank matrix $A^*$.

[3] relaxed the $\ell_0$-error measure to $\ell_1$-norm.

It is of fundamental importance for TCS to understand the theoretical guarantees for the original $\ell_0$-problem.

Example: Reals $\ell_0$-rank-1

Hard instance for algorithms that select a column:

$$A = \begin{bmatrix}
2 & 1 & \ldots & 1 \\
1 & 2 & \ldots & 1 \\
1 & 1 & \ldots & 2
\end{bmatrix}_{m \times n}$$

$$u^* = v^* = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_n$$

$OPT^k = n$

For any column $A_{\cdot i}$ the best response vector is 1, so $||A_{\cdot i}^T - A_{\cdot i}'||_0 = 2(n - 1) = 2(1 - 1/n) \text{OPT}^k$

Reals $\ell_0$-rank-1

2-approximation scheme [4] – $O(||A||_0(m + n))$ time
For every column $A_{\cdot i}$ compute best response vector $v := \arg\min ||A_{\cdot i}v^T - A||_0$. Return the best pair $(A_{\cdot i}, v)$.

Example: The Robust PCA Problem

\textbf{Theorem 1. (Sublinear)} Given $A \in R^{m \times n}$ with column adjacency arrays, $\epsilon \in (0, 1/10)$ and $\text{OPT}_R^1 \geq 1$, we can compute w.h.p. in time

$$O\left(\frac{n \log m}{\epsilon^2} + \min\{||A||_0, n + \psi_\epsilon^{-1}(\log n)\log^2 n\} \frac{\log^2 n}{\epsilon^2}\right)$$

a column $A_{\cdot j}$ and a vector $z$ such that w.h.p. $||A - A_{\cdot j}z^T||_0 \leq (2 + \epsilon) \text{OPT}^k$.

Def: $\psi_F := \text{OPT}^k/||A||_0$, $0 \leq \psi_F \leq 1$, Sublinear $o(||A||_0)$.

Boolean $\ell_0$-rank-1

\textbf{Theorem 3. (Sublinear)} Given $A \in \{0,1\}^{m \times n}$ with column adjacency arrays and with row and column sums, we can compute w.h.p. in time

$$O(\min\{||A||_0 + m + n, \psi_\epsilon^{-1}(m + n)\log^3 (mn)\})$$

vectors $u, v$ such that $||A - u_v^T||_0 \leq (1 + O(\psi_F)) \text{OPT}^d$.

\textbf{Theorem 4. (Exact)} Given $A \in \{0,1\}^{m \times n}$ with $\text{OPT}_R^1/||A||_0 \leq 1/300$, we can solve exactly the Boolean $\ell_0$-rank-1 problem in time $2^{O(\text{OPT}^d/\sqrt{||A||_0})}\text{poly}(mn)$.

\textbf{Theorem 5. (Lower Bound on Sample Complexity)} Given $A \in \{0,1\}^{m \times n}$ as in Theorem 3, and $\sqrt{\log n} \ll \phi \ll 1/100$ such that $\psi_F \leq \phi$, computing a $(1 + \phi)$-approximation to $\text{OPT}^d$ requires to read at least $\Omega(n/\phi)$ entries of $A$.

Reals $\ell_0$-rank-k

\textbf{Theorem 2. (Bicriteria Algorithm)} Given $A \in R^{m \times n}$ and $k \in [n]$ we can compute in expected time $\text{poly}(m, n)$ a subset of columns $A_{\cdot j}$ of size $\lceil |j| = O(k\log(n/k)) \rceil$ and a matrix $Z \in R^{\lceil |j| \rceil \times n}$ such that $||A - A_{\cdot j}Z||_0 \leq O(k^2\log(n/k)) \text{OPT}_R^k$.

\textbf{Lemma 1. (Structural)} For any $A \in R^{m \times n}$ and $k \in [n]$, there is a subset of columns $A_{\cdot j}$ of size $k$ and a matrix $Z \in R^{k \times n}$ such that $||A - A_{\cdot j}Z||_0 \leq (k + 1) \text{OPT}_R^k$.

\textbf{Lemma 2. (Lower Bound)} Any algorithm that selects $k$ columns of $A$ incurs at least an $\Omega(k)$-approximation.

\textbf{Proof Techniques (Theorem 2)}

A subroutine from [1] that computes in time $O(m^2 k^{\omega + 1})$ a vector $z$ such that $||Az - b||_0 \leq k \cdot \min ||Ax - b||_0$.

We extend a result from [2] using Lemma 1, to find a constant fraction of columns $C$ that satisfy $\min_x ||Ax - A_{\cdot i}||_0 \leq O(k + 1) \text{OPT}_R^k/\text{poly}(mn)$.

\textbf{References}


