Optimal Analysis of Subset-Selection Based ℓ_p Low Rank Approximation

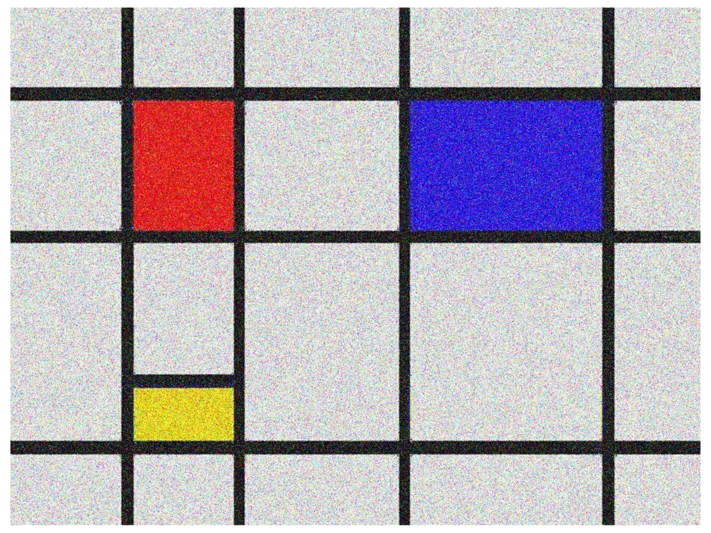
Chen Dan*, Hong Wang†, Hongyang Zhang‡, Yuchen Zhou§, Pradeep Ravikumar*

*CMU**, *Princeton*†, *TTIC*‡, *University of Wisconsin-Madison*§

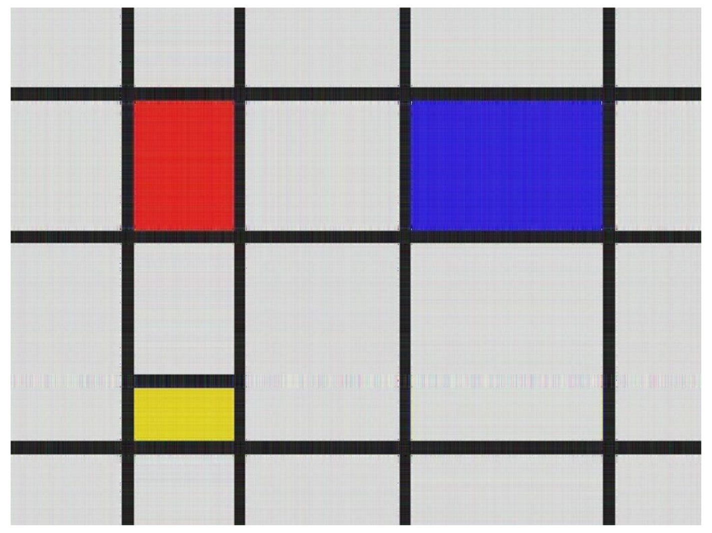
Overview

- Approximation algorithms for low rank approximation problem in entry-wise ℓ_p loss.
- **Optimal** (up to constant 1!) analysis for **Column Subset Selection** (CSS) algorithm, improved [1] and generalized [2] in both upper and lower bounds.
- Introduced **Riesz-Thorin interpolation theorem** from harmonic analysis as the main technical tool. First work in Machine Learning / Theoretical Computer Science community.

ℓ_p Low Rank Approximation Problem



Noisy Mondrian



Low Rank Mondrian

- Solve the **Low Rank Approximation** problem for matrix A:

$$\mathsf{OPT} = \min_{X:\mathrm{rank}(X)\le k} ||X - A||.$$

- Frobenious norm: poly-time algorithm by SVD.
- This paper: entry-wise ℓ_p norm:

$$||X - A||_p = \left(\sum_{i,j} |X_{ij} - A_{ij}|^p\right)^{1/p}$$

  Frobenius norm $\Leftrightarrow p = 2$.

- When $p \neq 2$: computationally hard!
- **Goal 1**: Design **α-approximation algorithms**: output Y, s.t.

$$||Y - A||_p \le \alpha \cdot \mathsf{OPT}$$

- **Goal 2**: Provide the **best possible analysis** for the approximation ratio of algorithms.

Column Subset Selection (CSS) Algorithm

- For each subset of k columns A_S ($\binom{n}{k}$ such subsets in total):
- Solve the (convex) ℓ_p linear regression problem:

$$\min_{V_S} ||A - A_S V_S||_p.$$

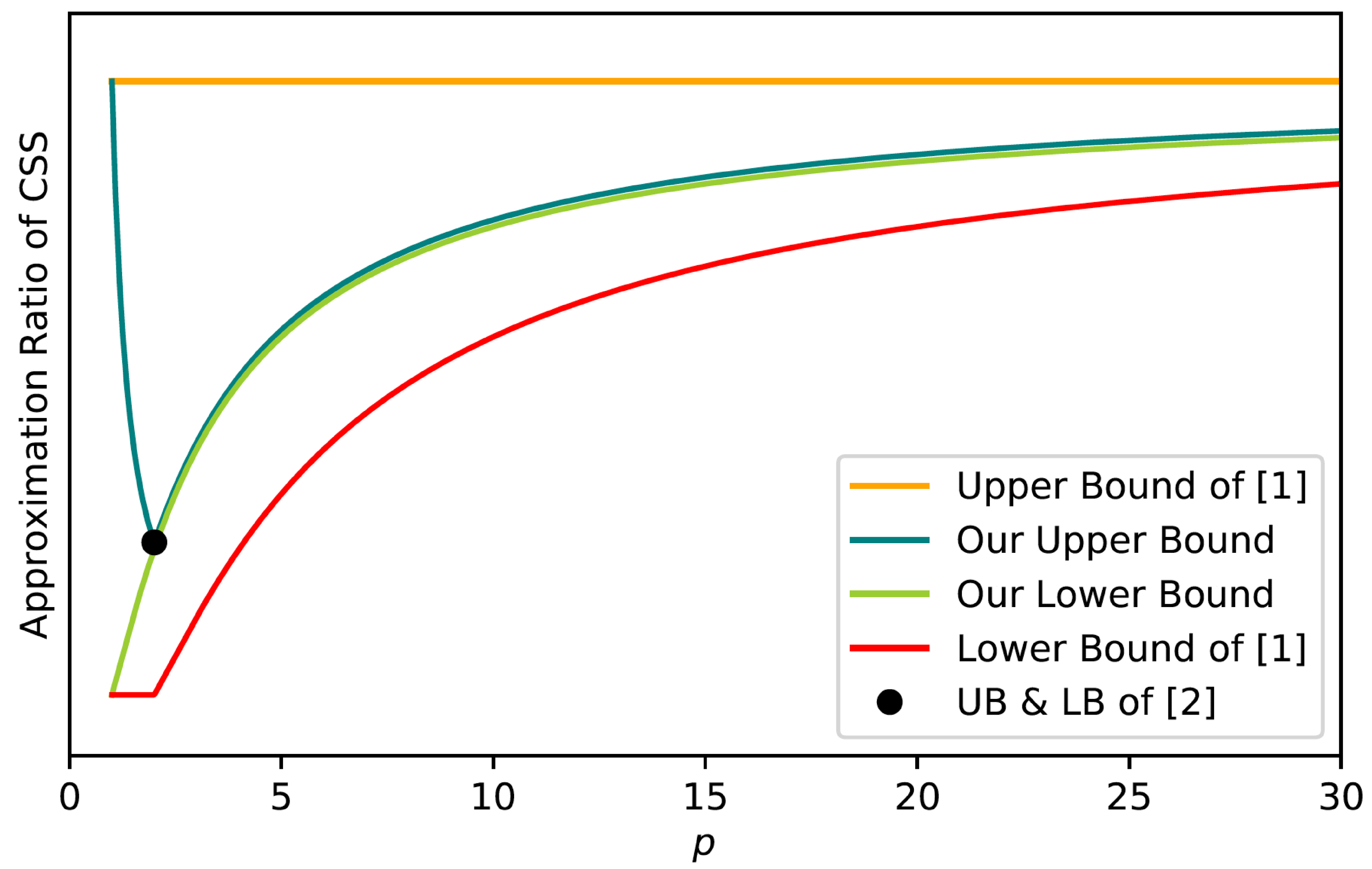
- Return $Y = A_{S^*} V_{S^*}$, where S^* is the best subset.
- Natural heuristic and algorithmic foundation of many efficient algorithms, e.g. random selection [1], volume sampling [2].

Main Results

1. Approximation ratio analysis of CSS algorithm:

	$1 \le p < 2$	$p = 2$	$p > 2$
Existing Upper Bound	$(k+1)$ [1]	$\sqrt{k+1}$ [2]	$(k+1)$ [1]
Our Upper Bound	$(k+1)^{\frac{1}{p}}$	$\sqrt{k+1}$	$(k+1)^{1-\frac{1}{p}}$
Existing Lower Bound	-	$\sqrt{k+1}$ [2]	$(k+1)^{1-\frac{2}{p}}$ [1]
Our Lower Bound	$(k+1)^{\frac{1}{p}}$	$\sqrt{k+1}$	$(k+1)^{1-\frac{1}{p}}$

Our analysis is **tight** when $p \ge 2$!



2. Approximation ratio analysis of efficient algorithms:

Algorithm	$1 \le p \le 2$	$p \ge 2$
[1], Bi-Criteria	$O(k+1)$	$O(k+1)$
This work, Bi-Criteria	$O((k+1)^{\frac{1}{p}})$	$O((k+1)^{1-\frac{1}{p}})$
[1], Exact Rank	$\tilde{O}((k+1)^4)$	$\tilde{O}((k+1)^4)$
This work, Exact Rank	$\tilde{O}((k+1)^{1+\frac{3}{p}})$	$\tilde{O}((k+1)^{4-\frac{3}{p}})$

Technique: Riesz-Thorin Interpolation

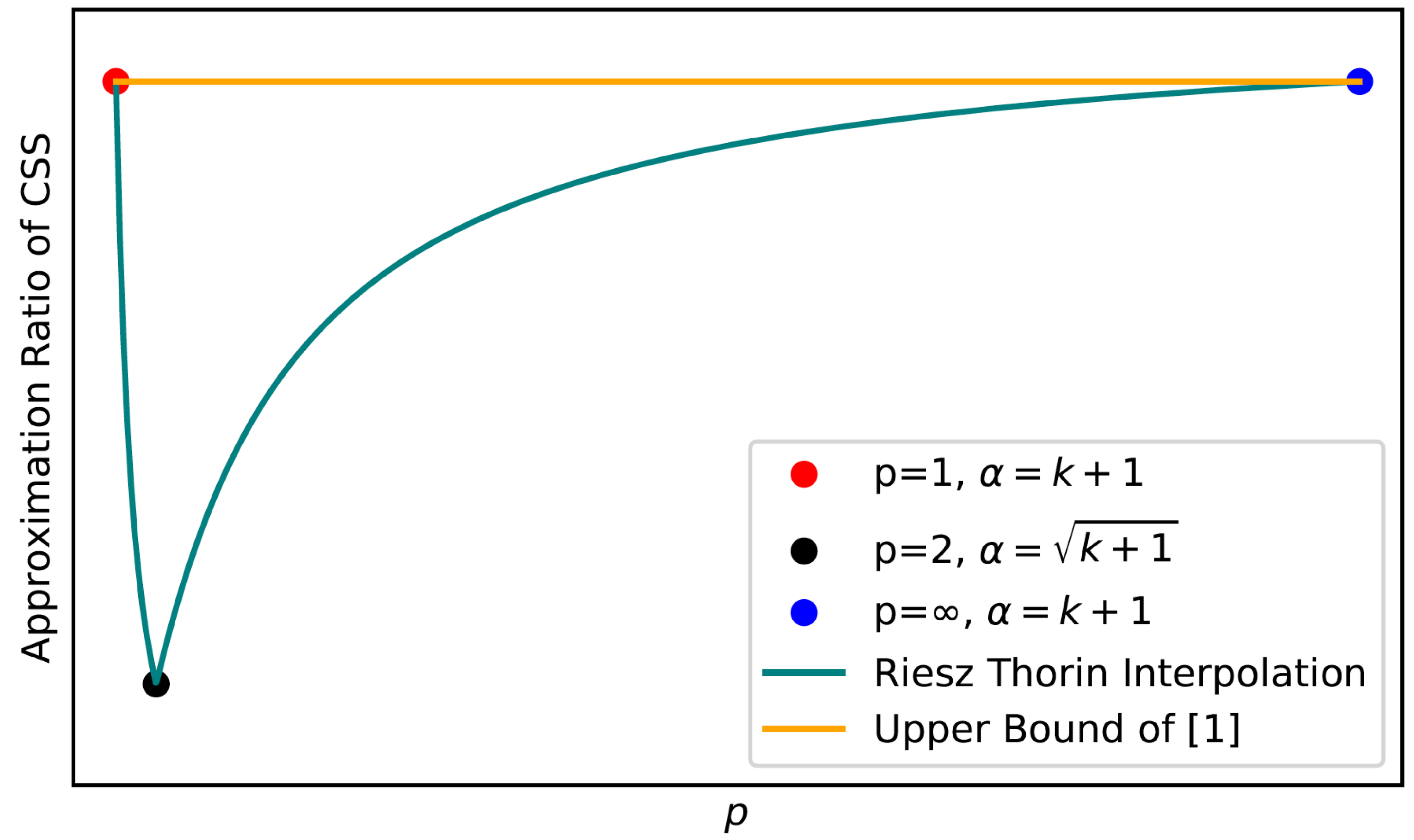
- Classical result from harmonic analysis in 1930s.
- Sharp bound on the ℓ_p operator norm of bi-linear forms.
- **Theorem:** Let $\Lambda : \mathbb{C}^{n_1} \times \mathbb{C}^{n_2} \to \mathbb{C}^{n_3}$ be a bilinear operator, with ℓ_p operator norm:

$$M_p = \sup_{||a||_p = ||b||_p = 1} ||\Lambda(a,b)||_p$$

  then

$$M_{p_\theta} \le M_{p_0}^{1-\theta} M_{p_1}^{\theta}, \quad \text{where} \quad \frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$$

- Only requires operator norms for $p = p_0, p_1$, then provides accurate bounds for every $p \in [p_0, p_1]$.



- First reduce the approximation analysis to operator norm analysis by careful construction of bi-linear forms, then apply Riesz-Thorin.
- When $p \in (1,2)$, choose $p_0 = 1, p_1 = 2$; when $p \in (2, +\infty)$, choose $p_0 = 2, p_1 = +\infty$.
- Only need to consider the simpler cases $p \in \{1, 2, +\infty\}$!

References

[1] Flavio Chierichetti, Sreenivas Gollapudi, Ravi Kumar, Silvio Lattanzi, Rina Panigrahy, and David P Woodruff. Algorithms for ℓ_p low rank approximation. ICML, 2017.

[2] Amit Deshpande, Luis Rademacher, Santosh Vempala, and Grant Wang. Matrix approximation and projective clustering via volume sampling. SODA, 2006