Riemannian CUR Decompositions for Robust Principal Component Analysis

Keaton Hamm (UTA) Mohamed Meskini (UTA) HanQin Cai (UCF)

Highlights

- Combine Riemannian optimization and Robust CUR decompositions
- State-of-the-art computational complexity on Robust PCA
- Strong robustness to outliers

Robust PCA

Principal Component Analysis (PCA): fundamental technique for dimension reduction, but it is very sensitive to outliers.

Robust PCA (RPCA): aims to recover a low-rank matrix L and a sparse matrix S from their sum $D = L + S$.

Non-convex optimization problem:

$$\min_{L',S'} \|D - L' - S'\|_{\mathrm{F}} \quad \text{subject to } L' \text{ is low-rank} \quad \text{and} \quad S' \text{ is sparse.}$$

Assumptions:

A1 *The underlying $L \in \mathbb{R}^{n\times n}$ is a rank-r matrix with μ-incoherence, that is*

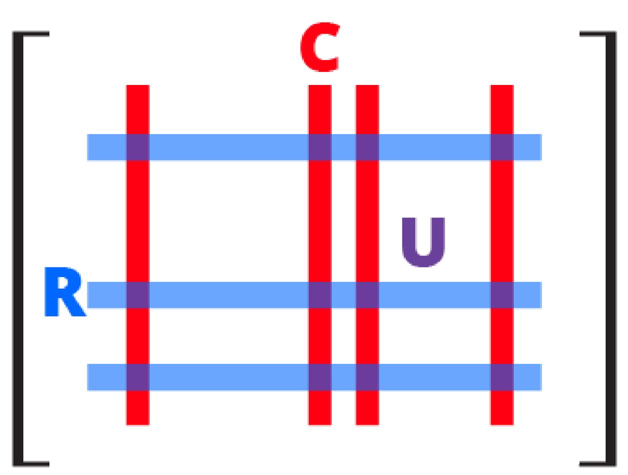
$$\|U\|_{2,\infty} \le \sqrt{\mu r/n}, \quad \text{and} \quad \|V\|_{2,\infty} \le \sqrt{\mu r/n}$$

hold for some numerical constant μ, where $L = U\Sigma V^\top$ is the compact SVD of L.

A2 *The underlying $S \in \mathbb{R}^{n\times n}$ is α-sparse. That is, S has at most αn non-zero entries in each row, and at most αn non-zero entries in each column.*

Preliminaries

CUR decompositions [1]:



Theorem: If the columns of $C = L_{:,J}$ span the column space of L, the rows of $R = L_{I,:}$ span the row space of L, and denote $U = L_{I,J}$, then $L = CU^\dagger R$.

Manifold and tangent space [2]: The set of low-rank matrices is indeed a Riemannian manifold. The tangent space of the manifold at L_k (the estimate of L at the k–th iteration of the algorithm below) is defined as

$$T_k := \{W_k A^\top + B V_k^\top : A, B \in \mathbb{R}^{n\times r}\}$$

where $U_k \Sigma_k V_k^\top$ is the SVD of L_k. The projection of an arbitrary matrix X onto the tangent space T_k has a close form

$$P_{T_k} X = W_k W_k^\top X + X V_k V_k^\top - W_k W_k^\top X V_k V_k^\top.$$

Iterated Robust CUR

IRCUR [3]: an iterative method for robust CUR decomposition.

- **Main idea:** Using CUR decompositions to replace the low-rank approximations in a classic iterated RPCA framework.
- **Pros:** improved computational complexity—$\mathcal{O}(r^2 n \log^2 n)$
- **Cons:** relatively low robustness

Accelerated Alternating Projections

AccAltProj [4]: a special case of Riemannian optimization for RPCA

- **Main idea:** In the framework of alternating projections, utilizing Riemannian gradient descent (fixed stepsize) for the updates of low-rank component.
- **Pros:** very robust against outliers
- **Cons:** standard computational complexity—$\mathcal{O}(n^2 r)$

Riemannian CUR

RieCur: To bridge this gap, we propose *Riemannian CUR* (RieCUR), which combines the ideas of robust CUR decomposition and Riemannian gradient descent.

- RieCUR has the improved complexity $\mathcal{O}(r^2 n \log^2 n)$, which is the same as IRCUR (although the constant appears to be slightly larger for RieCUR based on experiments in the sequel).
- RieCUR appears to tolerate outliers as well as AccActProj in terms of reconstruction vs. sparsity based on our numerical experiments, whereas IRCUR degrades as the amount of outliers increases.

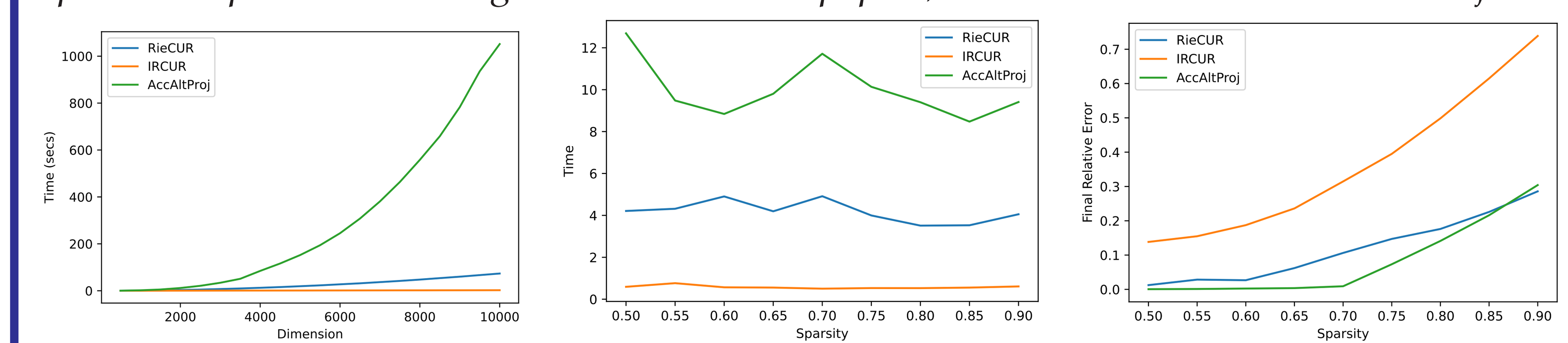
Algorithm: Riemannian CUR (RieCUR) for RPCA
Initialize L_0 and S_0
for $k = 0, 1, 2 \cdots$ **do**
$C_{k+1} = (P_{T_k}(D - S_k))_{:,J}$
$R_{k+1} = (P_{T_k}(D - S_k))_{I,:}$
$U_{k+1} = (P_{T_k}(D - S_k))_{I,J}$
$L_{k+1} = C_{k+1} U_{k+1}^\dagger R_{k+1}$
$\zeta_{k+1} = \gamma^k \zeta_0$
$(S_{k+1})_{:,J} = \mathcal{T}_{\zeta_{k+1}}(D - L_{k+1})_{:,J}$
$(S_{k+1})_{I,:} = \mathcal{T}_{\zeta_{k+1}}(D - L_{k+1})_{I,:}$
end for

The main difference of RieCUR compared with IRCUR and AccAltProj:

- RieCur utilizes submatrices of the tangent space projection $P_{T_k}(D - S_k)$ rather than the entire matrix as AccAltProj does, and IRCUR does not utilize the tangent space projection at all, but rather works with submatrices of $D - S_k$ directly.
- RieCur uses of the tangent space projection which adds a small amount of computation time to IRCUR but with the benefit of making the procedure more robust to outliers, as CUR decompositions are known to suffer from outliers.

Numerical Experiments

Synthetic data: An $n \times n$ rank-r matrix L is formed via $L = PQ^\top$, where $P, Q \in \mathbb{R}^{n\times r}$ have i.i.d. $\mathcal{N}(0,1)$ entries. The locations of the non-zero entries of the sparse matrix S are sampled uniformly and independently without replacement. The k-th iteration relative computing error is defined as $e_k = \|D - L_k - S_k\|_{\mathrm{F}} / \|D\|_{\mathrm{F}}$.
Speed comparisons among RieCUR (this paper), IRCUR [3] and AccAltProj [4]:



Left: Varying dimension n vs runtime, where $r = 5$, $\alpha = 0.3$, and n varies from 500 to 10000. Each algorithm stops once $e_k < 10^{-6}$ or 40 iterations has been reached Center: Varying sparsity factor α vs runtime, where $r = 5$, and $n = 2000$. Each algorithm stops once $e_k < 10^{-3}$ or 100 iterations. Right:Error vs. Sparsity for the three algorithms considered. In all trials L is 2000×2000 with rank 5. Each algorithm stops after 100 iterations.

Video background subtraction: The video we have used for this test is `restaurant`. The video is black and white with $3,055$ frames of size 120×160. A single data matrix of the video is obtained as follows: each frame is vectorized to form a column vector of size $19,200$ and columns are concatenated into a data matrix of size $19,200 \times 3,055$.

(Top row) One original frame from the `restaurant`. (Center row) Background of the same frame recovered from AccAltProj (left), RieCUR (center), and IRCUR (right). (Bottom row) Foreground of the same frame recovered from AccAltProj (left), RieCUR (center), and IRCUR (right).

References

[1] M. W. Mahoney and P. Drineas, "Cur matrix decompositions for improved data analysis," *Proceedings of the National Academy of Sciences*, 2009.

[2] B. Vandereycken, "Low-rank matrix completion by riemannian optimization," *SIAM Journal on Optimization*, 2013.

[3] H. Cai, K. Hamm, L. Huang, J. Li, and T. Wang, "Rapid robust principal component analysis: CUR accelerated inexact low rank estimation," *IEEE Signal Processing Letters*, 2020.

[4] H. Cai, J.-F. Cai, and K. Wei, "Accelerated alternating projections for robust principal component analysis," *Journal of Machine Learning Research*, 2019.