Auto-Weighted Two-Dimensional Principal Component Analysis with Robust Outliers

Rui Zhang, Feiping Nie, and Xuelong Li, *Fellow, IEEE*
ruizhang8633@gmail.com, feipingnie@gmail.com, xuelong_li@ieee.org

Problem

Two-dimensional principal component analysis (2DPCA) serves as an efficient approach for both dimensionality reduction and high-quality reconstruction. However, conventional 2DPCA method do have some limitations:

1. It is sensitive to the outliers such that associated results could be compromised
2. The mean is preset as arithmetic average value of all the data points, which is irrational.

Method

Robust 2DPCA with optimal mean (R2DPCA-OM) problem can be proposed as

$$\min_{\mathbf{M},\mathbf{U},\mathbf{V}} \sum_{i=1}^{l} \|\mathbf{A_i} - \mathbf{M} - \mathbf{U}\mathbf{U}^T(\mathbf{A_i} - \mathbf{M})\mathbf{V}\mathbf{V}^T\|_F$$

$$\text{s.t.} \quad \mathbf{U}^T\mathbf{U} = \mathbf{I}_{k_1} \ \text{ and } \ \mathbf{V}^T\mathbf{V} = \mathbf{I}_{k_2}. \qquad (3)$$

Based on (3), re-weighted form of R2DPCA-OM can be illustrated as

$$\min_{\mathbf{M},\mathbf{U},\mathbf{V}} \sum_{i=1}^{l} \mathbf{w_i}\|\mathbf{A_i} - \mathbf{M} - \mathbf{U}\mathbf{U}^T(\mathbf{A_i} - \mathbf{M})\mathbf{V}\mathbf{V}^T\|_F^2$$

$$\text{s.t.} \quad \mathbf{U}^T\mathbf{U} = \mathbf{I}_{k_1} \ \text{ and } \ \mathbf{V}^T\mathbf{V} = \mathbf{I}_{k_2} \qquad (4)$$

where $\mathbf{w_i} \leftarrow \frac{1}{2\|\mathbf{A_i}-\mathbf{M}-\mathbf{U}\mathbf{U}^T(\mathbf{A_i}-\mathbf{M})\mathbf{V}\mathbf{V}^T\|_F}$ is to be updated iteratively.

Eq. (4) can be further rewritten as

$$\max_{\mathbf{U},\mathbf{V}} \sum_{i=1}^{l} \mathbf{w_i}\mathbf{Tr}(\mathbf{U}^T(\mathbf{A_i} - \mathbf{M})\mathbf{V}\mathbf{V}^T(\mathbf{A_i} - \mathbf{M})^T\mathbf{U})$$

$$\text{s.t.} \quad \mathbf{U}^T\mathbf{U} = \mathbf{I}_{k_1} \ \text{ and } \ \mathbf{V}^T\mathbf{V} = \mathbf{I}_{k_2} \qquad (5)$$

where optimal mean $\mathbf{M} = \frac{\sum_{i=1}^{l} \mathbf{w_i}\mathbf{A_i}}{\sum_{i=1}^{l} \mathbf{w_i}}$.

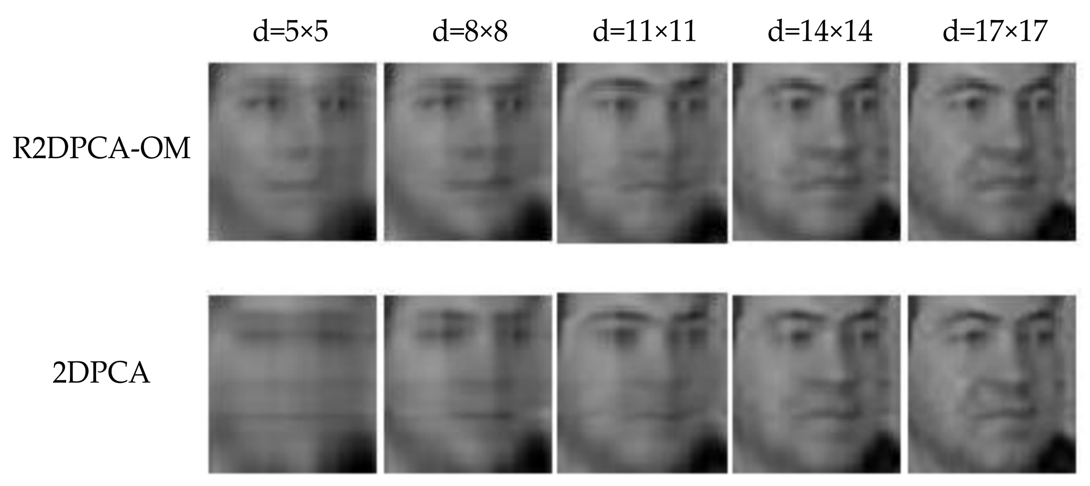
- **The major superiority of the proposed re-weighted form as R2DPCA-OM in (4) is connected with the self-adaptive weight.**

Contributions

Different from traditional 2DPCA method, the proposed method utilizes the robust 2DPCA with optimal mean (R2DPCA-OM) as the objective function, which is robust to the outliers.

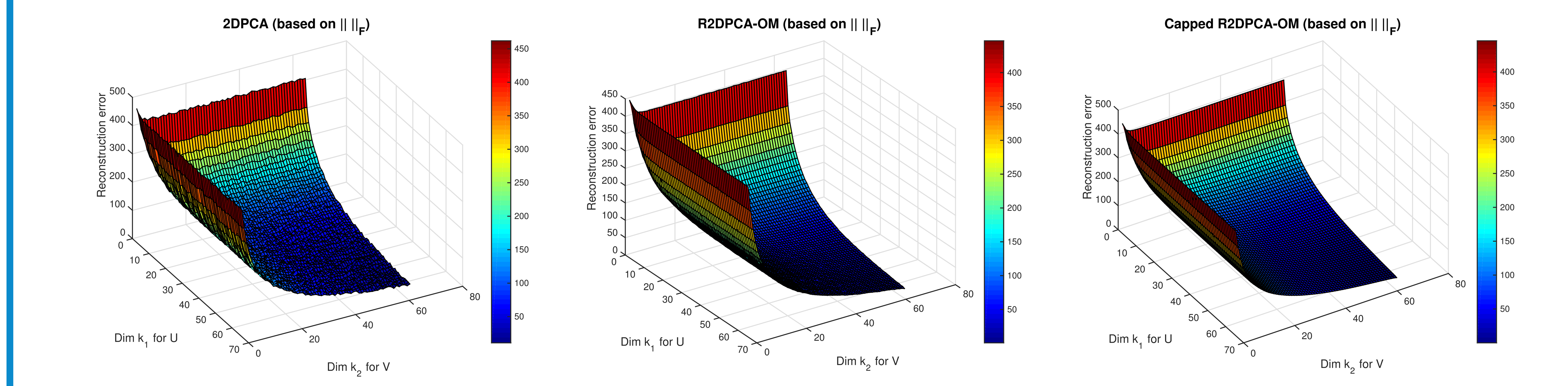
1. Associated algorithm seeks the optimal mean in each iteration instead of traditional data preprocessing.
2. The proposed R2DPCA-OM method has a self-adaptive weight, which assigns the smaller weight to the term with larger outliers automatically.

Core algorithm and samples



Input: $\mathbf{A_i} \in \mathbb{R}^{m\times n}, (\mathbf{i} = \mathbf{1}, \mathbf{2}, \ldots, \mathbf{l})$.
Output: $\mathbf{Y} = \mathbf{U}^T\mathbf{C}\mathbf{V}$ represents the projection.
Initialize $\mathbf{w_i} = \mathbf{1}, (\mathbf{i} = \mathbf{1}, \mathbf{2}, \ldots, \mathbf{l})$ and $\mathbf{V}\mathbf{V}^T = \mathbf{I_n}$;
while *not converge* **do**
  Update $\mathbf{M} \leftarrow \frac{\sum_{i=1}^{l} \mathbf{w_i}\mathbf{A_i}}{\sum_{i=1}^{l} \mathbf{w_i}}$;
  Update
  $\mathbf{P_1} \leftarrow \sum_{i=1}^{l} \mathbf{w_i}(\mathbf{A_i} - \mathbf{M})\mathbf{V}\mathbf{V}^T(\mathbf{A_i} - \mathbf{M})^T$;
  Update $\mathbf{U} \leftarrow \arg\max_{\mathbf{U}^T\mathbf{U}=\mathbf{I}_{k_1}} \mathbf{Tr}(\mathbf{U}^T\mathbf{P_1}\mathbf{U})$;
  Update
  $\mathbf{P_2} \leftarrow \sum_{i=1}^{l} \mathbf{w_i}(\mathbf{A_i} - \mathbf{M})^T\mathbf{U}\mathbf{U}^T(\mathbf{A_i} - \mathbf{M})$;
  Update $\mathbf{V} \leftarrow \arg\max_{\mathbf{V}^T\mathbf{V}=\mathbf{I}_{k_2}} \mathbf{Tr}(\mathbf{V}^T\mathbf{P_2}\mathbf{V})$;
  for $\mathbf{i} = \mathbf{1} : \mathbf{n}$ **do**
   Update
   $\mathbf{w_i} \leftarrow \frac{1}{2\|\mathbf{A_i}-\mathbf{M}-\mathbf{U}\mathbf{U}^T(\mathbf{A_i}-\mathbf{M})\mathbf{V}\mathbf{V}^T\|_F}$;
return $\mathbf{U} \in \mathbb{R}^{m\times k_1}$ and $\mathbf{V} \in \mathbb{R}^{n\times k_2}$;

Reconstruction error comparison



The comparisons of reconstruction error are performed for 2DPCA, R2DPCA-OM and capped R2DPCA-OM under the noised data of dataset FEI.

We compare the proposed robust 2DPCA with optimal mean (R2DPCA-OM) and capped R2DPCA-OM with the 2DPCA approach on FEI and YALE via the reconstruction error. We randomly select 25% of each dataset and set 20% size of the selected images with Gaussian noise to compare the reconstruction error represented by $\sum_i \mathbf{m}(\mathbf{X_i^o} - \mathbf{X_i^r})$, where $\mathbf{X_i^o}$ is the original image and $\mathbf{X_i^r}$ is the reconstructed data. Moreover, the measure $\mathbf{m}$ is chosen as both $\|\cdot\|_F$ and $\|\cdot\|_*$ to ensure a just comparison.

- **We could observe that surfaces of the proposed R2DPCA-OM and capped R2DPCA-OM methods are more smooth than that of the 2DPCA method, which represent stronger robustness to the outliers for the proposed approaches.**

Further extensions of capped ℓ_2 norm

- **Extension of the robust 2DPCA problem in (3).** Given the possible situation that the outliers might be extraordinarily huge for certain **i**-th term in (3), the superiority of the proposed robust problem in (3) and (4) might be largely compromised. Actually, we could address this situation by introducing the capped form of the problem (3) as

$$\min_{\mathbf{M},\mathbf{U},\mathbf{V}} \left(\sum_{i=1}^{l} \min(\|\mathbf{A_i} - \mathbf{M} - \mathbf{U}\mathbf{U}^T(\mathbf{A_i} - \mathbf{M})\mathbf{V}\mathbf{V}^T\|_F, \varepsilon)\right)$$

$$\text{s.t.} \quad \mathbf{U}^T\mathbf{U} = \mathbf{I}_{k_1} \ \text{ and } \ \mathbf{V}^T\mathbf{V} = \mathbf{I}_{k_2} \qquad (1)$$

where ε is the threshold parameter. We could observe that if the outliers of certain i-th term in (1) is very large, Eq. (1) would automatically replace the related term by the threshold ε. In other words, the capped R2DPCA-OM in (1) could avoid the ill-defined situation mentioned above. Besides, solving the problem (1) is basically the same as solving the problem (3) as previously mentioned. The only difference reflects on a novel threshold-sensitive weight, which is introduced to the proposed R2DPCA-OM in (4) as

$$\mathbf{w_i} = \frac{\mathbf{Ind}}{2\|\mathbf{A_i} - \mathbf{M} - \mathbf{U}\mathbf{U}^T(\mathbf{A_i} - \mathbf{M})\mathbf{V}\mathbf{V}^T\|_F} \qquad (2)$$

where the indicative function *Ind* is defined as

$$\mathbf{Ind} = \begin{cases} \mathbf{1}, & \|\mathbf{A_i} - \mathbf{M} - \mathbf{U}\mathbf{U}^T(\mathbf{A_i} - \mathbf{M})\mathbf{V}\mathbf{V}^T\| \le \varepsilon \\ \mathbf{0}, & \textit{Otherwise} \end{cases}$$

Equipped with the weight defined in (2), the R2DPCA could be extended to unraveling the capped robust 2DPCA in (1) correspondingly.

YALE($\times 10^3$)	$\mathbf{m} = \|\cdot\|_*$	
$\mathbf{k_1 \times k_2}$	8×60	80×40
capped R2DPCA-OM	**6.3413**	2.8597
R2DPCA-OM	6.3638	**2.8561**
2DPCA	6.3653	3.3154