# Low-Rank-Sparse Subspace Representation for Robust Regression

Yongqiang[1], Daming[1,2], Junbin[3], Dansong[1]

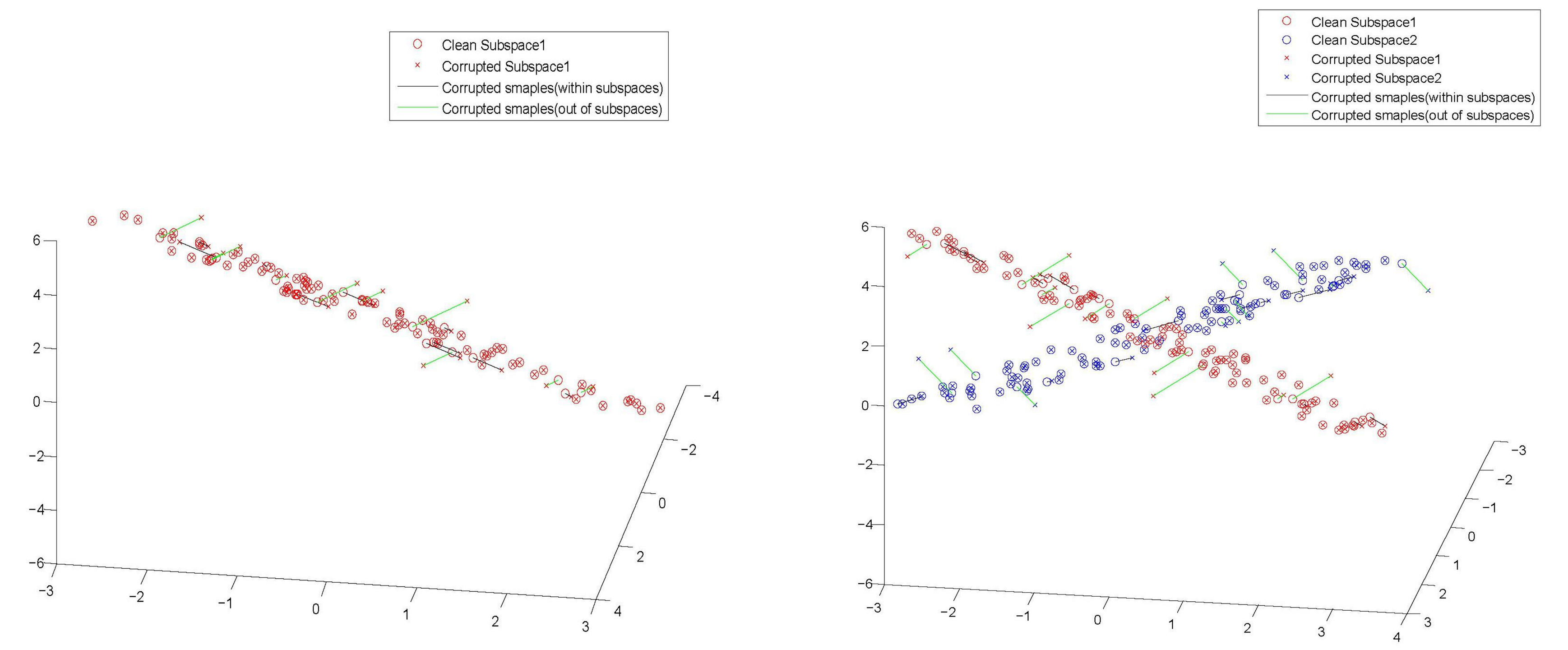
1 Harbin Institute of Technology; 2 Shenzhen University; 3 The University of Sydney

IEEE 2017 Conference on Computer Vision and Pattern Recognition

CVPR July 21-26 HONOLULU 2017

## Problem:

- The current methods have studied low-rank regression models that are robust against typical noises (like Gaussian noise and out-sample sparse noise) or outliers. However, few of them can handle the outliers/noise lying on the sparsely corrupted disjoint subspaces.




(a) An independent corrupted subspace (b) Corrupted disjoint subspaces

Figure 1. Different Corrupted Subspaces

- As shown in Fig. 1, In case (a), current low-rank robust regression (LR-RR) works well; In case (b), it tends to fail.

## Contribution:

- Unlike most of the existing regression methods, we propose an approach with two phases of low-rank-sparse subspace recovery and regression optimization being carried out simultaneously.
- We also apply the linearized alternating direction method with adaptive penalty to solved the formulated LRS-RR problem and prove the convergence of the algorithm and analyze its complexity.
- We demonstrate the efficiency of our method for the high-dimensional corrupted data on both synthetic data and two benchmark datasets against several state-of-the-art robust methods.

## Methods:

- We propose a low-rank-sparse subspace representation for robust regression by learning a clean dictionary-the basis for subspaces-that satisfies the condition of sparse noise:

$$\hat{\mathscr{J}}_{\text{LRS-RR}} = \min_{\mathbf{T},\hat{\mathbf{D}},\mathbf{A},\mathbf{Z},\mathbf{J},\mathbf{E}} \frac{\eta}{2}\left\|\mathbf{W}(\mathbf{Y}-\mathbf{T}\hat{\mathbf{D}})\right\|_F^2 + \|\mathbf{A}\|_* + \|\mathbf{Z}\|_* + \lambda_2\|\mathbf{J}\|_1 + \lambda_1\|\mathbf{E}\|_1$$

$$\text{s.t. } \mathbf{X} = \mathbf{A}\mathbf{Z} + \mathbf{E},\ \hat{\mathbf{D}} = \left[\mathbf{A}\mathbf{Z}; \mathbf{1}^{\mathbf{T}}\right],\ \mathbf{Z} = \mathbf{J}, \mathbf{J} \geq \mathbf{0}.$$

**Y**: dependent variables; **X**: independent variables; **T**: the mapping
**A**: dictionary; **Z**: low-rank-sparse coefficients; **D** : cleaned samples
**E**: sparse noise; **J**: non-negative sparse constraint for **Z**

- We solve it by the LADMAP method.

## Convergence and Complexity:

- It is solved iteratively via the four subproblems:
  - 2 least square(**T**; **D**)
  - 2 low-rank (**A**, **E**; **Z**, **J**)
- Low-rank subproblems converge to a KKT point.
- In the iterations of subproblems, the computational costs are mainly matrix inversion and SVT.

Fig. 2 gives the plots of relative Frobenius norm errors of **D** and **E** varying with iteration number.

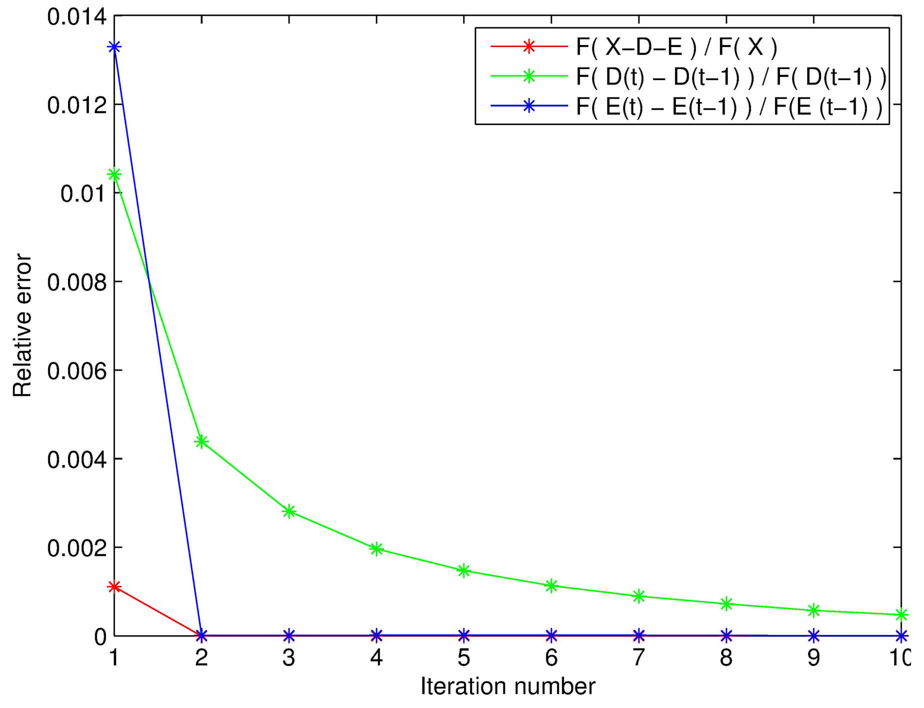



Figure 2. Convergence result on Synthetic Data.

## Experiments:

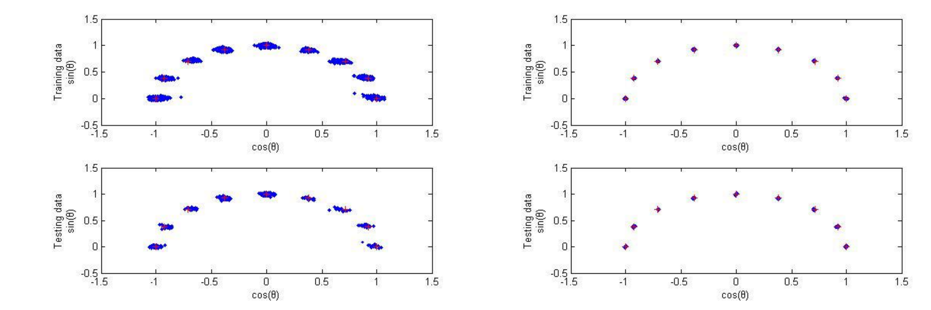
- Synthetic Data

*RAE and its standard deviation on synthetic data*

| Method | $RAE_{\mathbf{T}}$ | $RAE_{\mathbf{Y}}$ |
|---|---|---|
| LSR | $0.269 \pm 0.121$ | $0.035 \pm 0.012$ |
| RANSAC | $0.256 \pm 0.133$ | $0.036 \pm 0.013$ |
| RPCA+LSR | $0.464 \pm 0.030$ | $0.051 \pm 0.006$ |
| LR-RR | $0.035 \pm 0.015$ | $0.015 \pm 0.006$ |
| LRS-RR | $\mathbf{0.005 \pm 0.0005}$ | $\mathbf{0.011 \pm 0.003}$ |

- CMU PIE Database for Pose Estimation

*Comparison of yaw angle error and standard deviation*



LR-RR LRS-RR

| Method | Pose Angle Err | Time(s) |
|---|---|---|
| LSR | $27.56^o \pm 23.60^o$ | 0.05 |
| RANSAC | $23.20^o \pm 20.39^o$ | 0.22 |
| RPCA+LSR | $20.45^o \pm 19.51^o$ | 0.25 |
| LR-RR | $1.97^o \pm 5.77^o$ | 3.03 |
| LRS-RR | $\mathbf{1.03^o \pm 5.65^o}$ | 10.02 |

- YaleB Database for Reconstruction of Corrupted Faces

*Models errors & Fitting errors*

| Method | Model Err | Fitting Err |
|---|---|---|
| RANSAC | $1.058 \pm 0.040$ | $0.185 \pm 0.007$ |
| RPCA+LSR | $1.075 \pm 0.051$ | $0.187 \pm 0.007$ |
| LR-RR | $1.069 \pm 0.044$ | $0.185 \pm 0.006$ |
| LRS-RR | $\mathbf{1.045 \pm 0.049}$ | $\mathbf{0.164 \pm 0.006}$ |

## Conclusions and Future Work:

- Our method can deal with outliers/noise inside or outside the disjoint subspaces, and can obtain much more exact regression model in this complex situation.
- The current LRS-RR is time-consuming, the optimization algorithm will be parallelized in the future.