# Robust PCA Report

Lingyu Zhang

April 17, 2018

## Theoretical part

### 0.1 ADMM for robust PCA

For Robust PCA, the objective is

$$\min_{L,S} ||L||_* + \lambda||S||_1$$

$$\text{s.t. } L + S = X$$

Then the augmented Lagrangian is as below:

$$L_\rho(L, S, U) = \lambda||S||_1 + ||L||_* + < U, X - L - S > + \frac{\rho}{2}||L + S - X||_F^2$$

At each iteration, we firstly minimize the objective function with regard to L

$$\min_L \frac{1}{2}||L + S - X - \frac{U}{\rho}||_F^2 + \frac{1}{\rho}||L||_*$$

We then can obtain closed form solution using matrix shrinkage operation as below:

$$L^{k+1} = MatShrink(X - S^k + \frac{1}{\rho}U^k, \frac{1}{\rho})$$

where the matrix shrinkage operator is

$$MatShrink(Z, \frac{1}{\rho}) = UDiag(max\{\sigma - \frac{1}{\rho}, 0\})V^T$$

And $UDiag(\sigma)V^T$ is the SVD of Z. Then we minimize S

$$\min_S \frac{1}{2}||L + S - X - \frac{U}{\rho}||_F^2 + \frac{\lambda}{\rho}||S||_1$$

We obtained the closed-form solution using $l_1$ shrinkage operation

$$S^{k+1} = Shrink(X - L^{k+1} + \frac{1}{\rho}U^k, \frac{\lambda}{\rho})$$

where the $l_1$ shrinkage operator is

$$[Shrink(Z,\tau)]_{ij} = sgn(Z)(|Z|-\tau)_+ = \begin{cases} Z_{ij}-\tau & Z_{ij} > \tau \\ Z_{ij}+\tau & Z_{ij} < -\tau \\ 0 & -\tau \leq Z_{ij} \leq \tau \end{cases}$$

Finally we update U.

$$U^{k+1} = U^k + \beta\gamma(X - L^{k+1} - S^{k+1})$$

where $\gamma \in \{0, \frac{1+\sqrt{5}}{2}\}$ and we try $\beta \in \{0.1, 1, 5, 10\}$

## 0.2 Symmetric ADMM for robust PCA

For symmetric ADMM, At each iteration, we firstly minimize the objective function with regard to L

$$\min_L \frac{1}{2}||L+S-X-\frac{U}{\rho}||_F^2 + \frac{1}{\rho}||L||_*$$

We then can obtain closed form solution using matrix shrinkage operation as below:

$$L^{k+1} = MatShrink(X - S^k + \frac{1}{\rho}U^k, \frac{1}{\rho})$$

where the matrix shrinkage operator is

$$MatShrink(Z, \frac{1}{\rho}) = UDiag(max\{\sigma - \frac{1}{\rho}, 0\})V^T$$

And $UDiag(\sigma)V^T$ is the SVD of Z. Then we update U

$$U^{k+\frac{1}{2}} = U^k + \beta(X - L^{k+1} - S^k)$$

Then we minimize S

$$\min_S \frac{1}{2}||L+S-X-\frac{U}{\rho}||_F^2 + \frac{\lambda}{\rho}||S||_1$$

We obtained the closed-form solution using $l_1$ shrinkage operation

$$S^{k+1} = Shrink(X - L^{k+1} - \frac{1}{\rho}U^k, \frac{\lambda}{\rho})$$

where the $l_1$ shrinkage operator is

$$[Shrink(Z,\tau)]_{ij} = sgn(Z)(|Z|-\tau)_+ = \begin{cases} Z_{ij}-\tau & Z_{ij} > \tau \\ Z_{ij}+\tau & Z_{ij} < -\tau \\ 0 & -\tau \leq Z_{ij} \leq \tau \end{cases}$$

Finally we update U

$$U^{k+1} = U^k + \beta\gamma(X - L^{k+1} - S^{k+1})$$

### 0.3 Alternating minimization for penalized robust PCA

For each update step, we firstly update L as below:

$$L^{k+1} = \arg\min_{L} ||L||_* + \frac{\beta}{2}||L + S^k - X||_F^2$$

We then use soft-shrinkage to singular to solve the sub-problem.

$$L^{k+1} = MatShrink(X - S^k, \frac{1}{\beta})$$

We then update S as below

$$S^{k+1} = \arg\min_{S} \lambda||S||_1 + \frac{\beta}{2}||L^{K+1} + S - X||_F^2$$

We then use soft-shinkage to solve this sub-problem

$$S^{k+1} = Shrink(X - L^{k+1}, \frac{\lambda}{\beta})$$

## Experimental part

### Compare the performance of ADMM Symmetric ADMM and alternating minimization[1] for Robust PCA

For data preparation, we firstly reshape the data to $20800 * 200$ dimension.We initialize the parameters, the initial value of matrix L S and U are zeros. We set the hyperparameters as below:

$$maxiter = 1000;$$

$$\lambda = \frac{1}{\sqrt{20800}};$$

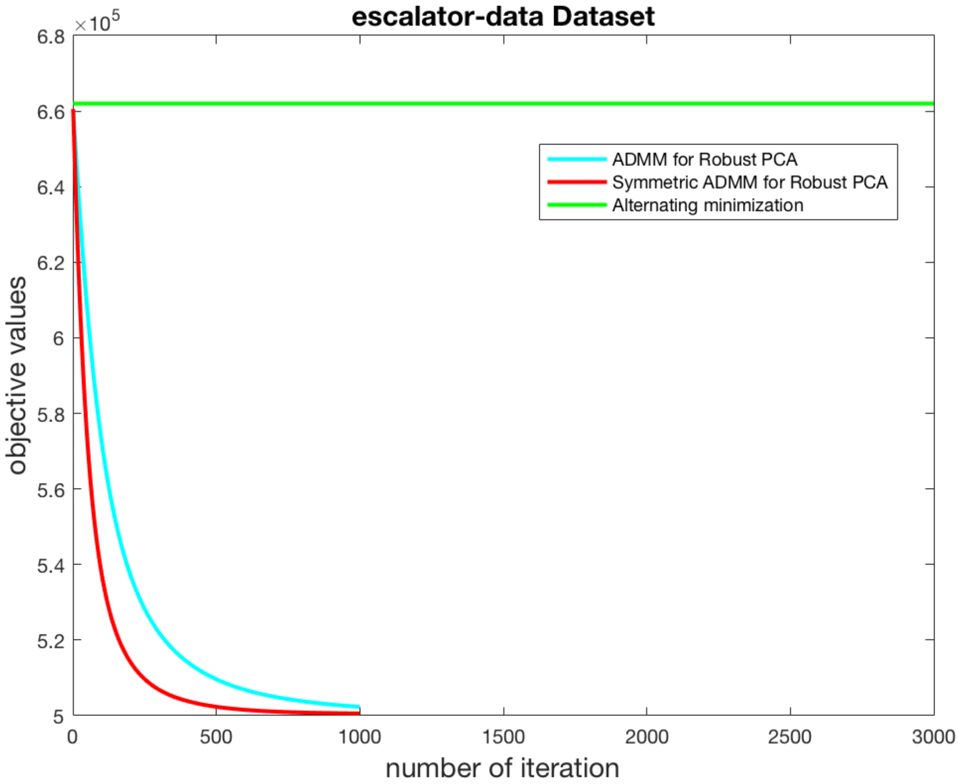
$$\rho = 10 * \lambda$$

$$\beta = 1e8;$$

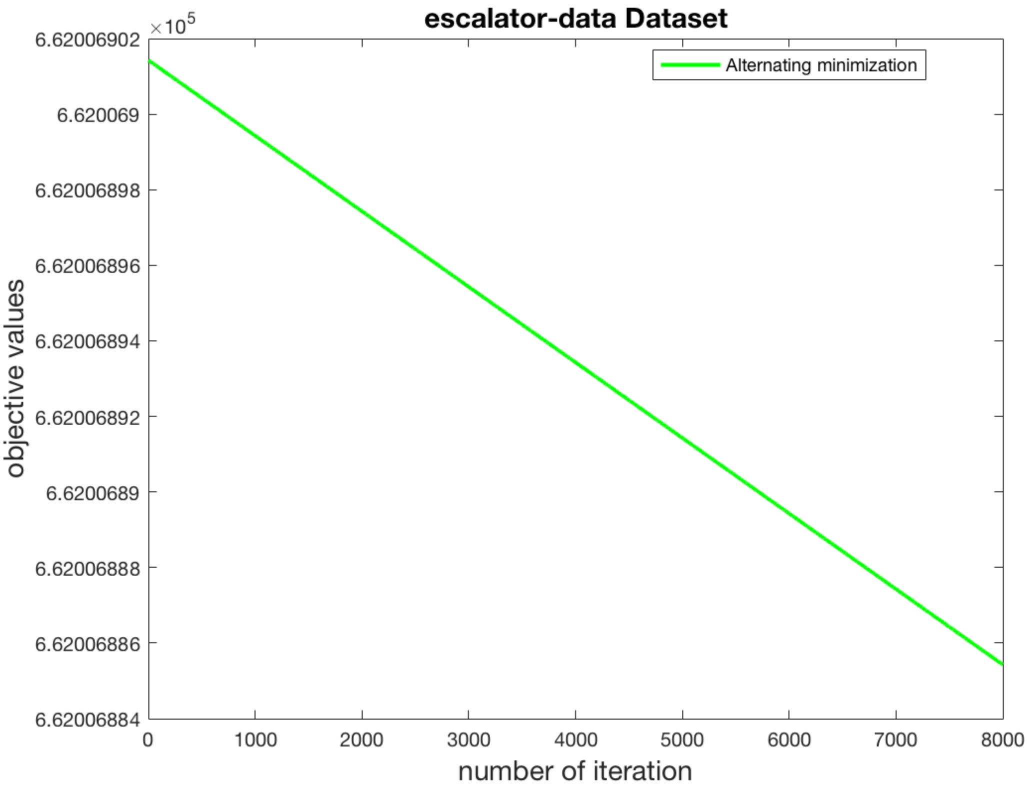
$$\gamma = 1;$$

Below is the objective values plot for stochastic gradient over iterations on both objective functions.

[1]For alternating minimization, I found there are some problems in my previous code for assignment 3, so I re-wrote my algorithm for alternating minimization this time, I attached my updated code in the appendix of this assignment 4




We can find that for ADMM and symmetric ADMM the objective value is going down with the optimization process iteration increases and finally converged. In order to check whether the objective function for alternating minimization is going down, we plot the alternating minimization training process solely for 8000 iterations and the plot is as below:




We can find that for Alternating minimization, the objective function is keeping going down while in a relatively samll slope compared with ADMM method.
Below is the time and objective value compairson for 1000 iterations.

```
>> rpca_test_lingyu
rank(L): 103
ADMM for Robust PCA time = 457.2522, objective value = 502331.472052

rank(L): 102
Symmetric ADMM for Robust PCA time = 467.4250, objective value = 500575.233114

rank(L): 200
Alternating minimization for Robust PCA time = 1361.4853, objective value = 662006.895428
```

We can find that after 1000 iterations, we obtain a low-rank L using both ADMM method and Symmetric ADMM method. For alternating minimization, we kept the iterations to 8000 and below is the record for the final objective value and time.

```
>> rpca_test_lingyu
rank(L): 200
Alternating minimization for Robust PCA time = 3642.0511, objective value = 662006.885428
```

In order to verify whether we have successfully separate the background with the ob-

ject in the given image, for each method, we visualized the some columns of L and S as image, in order to do it, we reshape each column of W as $130 * 160$.
Below is the visualized columns of L learnt by ADMM method.

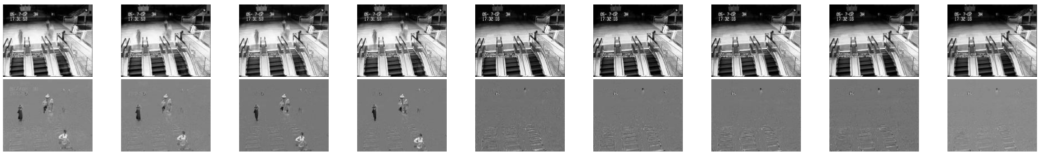


Figure 1: Visualization of column 1,2,3,4,196,197,198,199,200 of L (first row), column 1,2,3,4,196,197,198,199,200 of S (second row) by ADMM

Below is the visualized columns of L learnt by Symmetric ADMM method. Below is

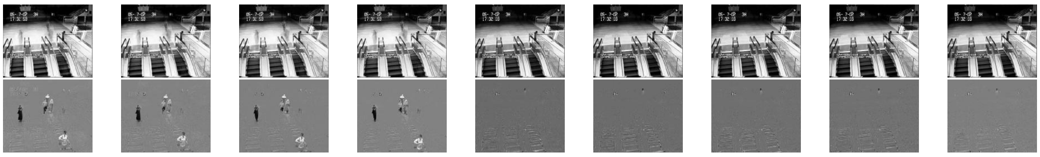


Figure 2: Visualization of column 1,2,3,4,196,197,198,199,200 of L (first row), column 1,2,3,4,196,197,198,199,200 of S (second row) by Symmetric ADMM

the visualized columns of L learnt by Alternating minimization method.

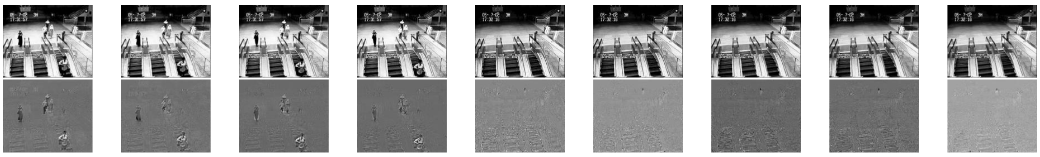


Figure 3: Visualization of column 1,2,3,4,196,197,198,199,200 of L (first row), column 1,2,3,4,196,197,198,199,200 of S (second row) by Alternating minimization