Epitome of Robust Principal Component Analysis

Maunil Vyas
SEAS, Ahmedabad University

***Abstract*—This review article is an epitome of a famous Robust principal component analysis article published on may 2011 by EMMANUEL J. CANDE'S, XIAODONG LI, YI MA, JOHN WRIGHT. The phenomena that the authors are suggesting is a fundamental matrix problem. they are asking a simple question to themselves that suppose if they have a data matrix which is a combination of low rank and sparse component then is there a way to decompose the matrix into the low rank and sparse component? and in answer to that, they have proposed a method under suitable conditions which is capable of extracting both components by solving very convenient convex program called *Principal Component Pursuit*. The idea is to find all the feasible decompositions and then simply minimize the weighted combination of the nuclear norm and of the ℓ_1 norm. The proposed methodology is robust because it doesn't care about the quality of the data which a matrix have means it allows corruption and missing values in a matrix and still guarantees perfect separation under certain assumptions. RPCA has many real life applications like video surveillance, face recognition, Latent Semantic Indexing, Ranking and Collaborative Filterin, Image recovery etc.**

***Keywords*—Principal components,nuclear-norm minimization, $\ell_1$norm minimization, low-rank matrices, sparsity.**

I. Introduction

A. Motivation

Finding low dimensional structure in a high dimensional data is in a high demand. Areas like audio-video processing, web search, bioinformatics etc finding new data of thousands and millions of dimensional observation on daily basis. Dimensionality curse is there and it requires a heuristic approach to fits the meaningful observation from the dataset to some low dimensional subspace. In other words, if we stack all the observations as column vectors of a matrix $M \in R^{m\times n}$ the matrix should be (approximately) low rank. Principal component analysis (PCA) seeks the best for such low-rank representation of the given data matrix when the data set has a slight effect of noise, also PCA can be easily and efficiently computed by SVD but when a data set has some sort of corruption in values. PCA will lead a significant deviation from original. Unfortunately, corruption in data values are omniparous in mostly all the applications nowadays and thus Robustifying PCA is in demand.

There are number of Robustifying PCA approaches are there in the literature but no one guarantees the polynomial time algorithm for the broad condition.

Manuscript received February 9, 2017

B. Problem statement & result formulation

The problem that authors are trying to solve is about to compute the low rank component L_0 from highly corrupted measurement $M = L_0 + S_0$ unlike the small noise term $M = L_0 + N_0$ in PCA, the entries in S_0 can have arbitrary large magnitudes and their support assume to be sparse but unknown.

Problem formulation: Given $M = L_0 + S_0$, where L_0 and S_0 are unknown, but L_0 is known to be low rank and S_0 is known to be sparse, recover L_0.

The above statement immediately suggests a conceptual solution: seek the lowest rank L_0 that could have generated the data, subject to the constraint that the errors are sparse: $\|E\|_0 \leq k$. The Lagrangian reformulation of this optimization problem is

$$min_{L_0,S_0} Rank(L_0) \leq k \longrightarrow Subj : M = L_0 + S_0 \quad (1)$$

As per the authors claim (1) is a highly non convex optimization problem, and no efficient solution is known. so one approach is to find a tractable optimization problem by relaxing (1), replacing the ℓ_0 norm with the ℓ_1 norm, and the rank with the nuclear norm $\|L_0\|_* = \sum_i \sigma_i(L_0)$

$$min_{L_0,S_0} \|L_0\|_* + \lambda\|S_0\|_1 \longrightarrow Subj : M = L_0 + S_0 \quad (2)$$

Theoretically this guarantees to work, Surprisingly this leads to a perfect extraction of L_0 and S_0 in simulations too.

C. Critical Conditions to perform RPCA for the separation

- M cannot be sparse and low-rank.
- L cannot be sparse.
- Sparsity pattern should be uniformly random.

II. Main Results Suggested by the Authors

It's amazing to see that taking care of above mentioned assumptions will lead to a perfect separation of L_0 and S_0.

A. Only Corruption in data matrix

Theoram 1.1: Let L_0 is $n \times n$ of rank(L_0)$\leq$ $p_r n$ $\mu^{-1}\log n^{-2}$, S_0 is $n \times n$ randomly sparsity pattern of cardinality $m \leq p_s n^2$.then with probability $1 - O(n^{-10})$, PCP with $\lambda = 1/\sqrt{n}$ exactly separates L_0 and S_0, for rectangular $\lambda = 1/\sqrt{(max(dim))}$

B. Missing values with corruption in data matrix

$$min_{L_0,S_0}\|L_0\|_*+\lambda\|S_0\|_1 \longrightarrow Subj : M_{ij} = L_{ij}+S_{ij}(i,j) \in \Omega_{obs} \tag{3}$$

Theoram 1.2: Let L_0 is $n \times n$ of rank(L_0)$\leq$ $p_r n$ $\mu^{-1} \log n^{-2}$, Ω_{obs} random set of size m=0.1 n^2, each observed entry is corrupted with probability of $\tau \leq \tau_s$ then with probability $1 - O(n^{-10})$, PCP with $\lambda = 1/\sqrt{0.1n}$ exactly separates L_0, for rectangular $\lambda = 1/\sqrt{0.1(max(dim))}$.

C. Missing values in data matrix

$$\begin{gathered} min_{L_0,S_0}\|L_0\|_* + 1/\sqrt(n)\|S_0\|_1 \\ Subj : L^0_{ij} = L_{ij} + S_{ij}(i,j) \in \Omega_{obs} \end{gathered} \tag{4}$$

Results in S_0 = 0 with complete low rank component.

III. ALGORITHM

A. Augmented Lagrange multiplier

The authors have proposed **Augmented Lagrange multiplier** algorithm to compute PCP which works on augmented Lagrangian.

Algorithm 1 Augmented Lagrange multiplier

1: **procedure** INITIALIZE:(S_0 = $Y_0 = 0$, $\mu > 0$)
2: **while** not converged **do**
3: compute L_{k+1} = $D_{1/\mu}(M - S_k + \mu^{-1}Y_k)$;
4: compute S_{k+1} = $S_{\lambda/\mu}(M - L_{k+1} + \mu^{-1}Y_k)$;
5: compute Y_{k+1} = $Y_k + \mu(M - L_{k+1} - S_{k+1})$;
6: **return** L,S

A generic Lagrange multiplier algorithm would solve PCP by repeatedly setting (L_k, S_k) = arg $min_{L,S}$ $l(L, S, Y_k)$, and then updating the Lagrange multiplier matrix via $Y_k + 1$ = Y_k + $\mu(M - L_k - S_k)$-(Step 5),Step 4 is an extension of shrinkage operator to the whole matrix. and Step 3 is related to the singular value thresholding. As per the authors claim, The algorithm will terminate when $\|M - L - S\|_F \leq \delta\|M - L - S\|_F$ with $\delta = 10^{-7}$

B. Simulation

Here I have tried to implement the proposed method given by authors on Matlab. I have taken an image (256 x 256) and then corrupted it by imposing another image. After that, I have applied RPCA methodology using Augmented Lagrange multiplier algorithm and extracted L and S from the corrupted image and by my surprise, I got almost the same original image when the corruption is not at it's significant level,I have also analyzed that as the corruption increases the chances of getting perfect image becomes less. fig.1 and fig.2 are simulation results.

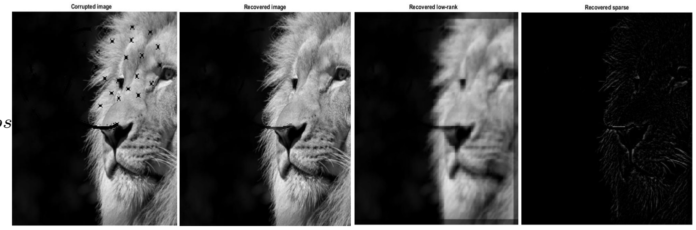

Fig. 1. Corruption:2.98 percentage , Corrupted values:1956

Fig. 2. Corruption:11.89 percentage, Corrupted values:7797

IV. CONCLUSION

As I have been part of this epitome preparation I got a chance to analyze such an amazing art of work done by the authors on Robustifying the principal component analysis. According to me, the paper has a significant potential to impact on diverse areas. The surprising fact about this paper is its generalized result which is applicable in almost all diverse areas and that sort of remarkable achievement from the authors leads it to become one the revolutionary change in the field of computer science and engineering, this kind of research work motivates the importance of mathematical fundamentals, by visualizing the importance of RPCA in different diverse fields actually makes the problem more beautiful and more challenging why challenging ? answer to that The world is filled with full of uncertainty.

ACKNOWLEDGMENT

I would like to express my special thanks of gratitude to the authors of RPCA as well as my professor Dr. Ratnik Gandhi who gave me the golden opportunity to do this wonderful epitome on the topic of RPCA, which also helped me in doing a lot of Research and I came to know about so many new things I am really thankful to them.

REFERENCES

[1] EMMANUEL. J. CANDE'S, XIAODONG. LI, YI. MA, JOHN. WRIGHT *Robust Principal Component Analysis*, Journal of the ACM, Vol. 58, No. 3, Article 11, Publication date: May 2011

[2] John. Wright, Yigang. Peng, YI. MA, *Robust Principal Component Analysis: Exact Recovery of Corrupted Low-Rank Matrices by Convex Optimization*,

[3] https://www.youtube.com/watch?v=DK8RTamIoB8