# Summary and an implementation of RPCA (Candès, 2009)

Raj Derasari
School of Engineering and Applied Science,
Ahmedabad University

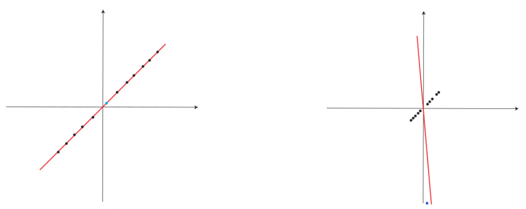
***Abstract*—Principal Component Analysis (PCA) is one of the most widely used tools for data optimization applications. However, it is sensitive to outlier samples. The paper proposes a scheme to implement a robust PCA, claiming that under suitable assumptions, it is possible to decompose a corrupted data matrix into a low rank component ($L$) and a sparse component ($S$). One of the implications of being able to utilize this algorithm is to be able to achieve foreground-background separation from a still image source, and in face recognition, as a way of removing peculiarities in face images.**

***Index Terms*—Principal components, robustness vis-a-vis outliers, nuclear-norm minimization, $\ell_1$-norm minimization, duality, low-rank matrices, sparsity, video surveillance, facial recognition, latent semantic indexing, ranking and collaborative filtering**

## I. Introduction

### A. The requirement of a robust PCA

Robust PCA has received attention recently due to its ability to recover a low rank model from sparse "noise". Robust PCA is required where PCA fails: PCA is sensitive to cases of extreme corruption in data samples. This can often be seen happening in image processing, computer vision, hardware failure, or data tampering.



(a) PCA with no outliers (b) PCA with gross outliers

Figure 1.1: Sensitive nature of classic PCA

### B. Agenda

Here, it is shown that, it is possible to decompose the data matrix M as a superposition of a lower rank component $L$ and a sparse component $S$, s.t. $M = L_0 + S_0$. Under rather weak assumptions, it can be shown that the Principal Component Pursuit estimate solves for $L$ and $S$ by maximizing the function $||L||_* + \lambda||S||_1$ subject to $L + S = M$. Theoretically, this is guaranteed to work if the rank of $L$ grows linearly and the errors in $S$ are a constant fraction of all entries.

## II. Organization of the report

Section III describes the approach taken to come up with robust PCA. III(A) defines the assumptions the RPCA model makes. Section III(B) is about the two primary theorems (1.1, 1.2) in [1]. Section IV mentions some advantages and challenges of RPCA. Section V contains the underlying algorithm of RPCA. Section VI mentions an application of RPCA, which has been implemented in [5], and observations. Finally, Section VII concludes the paper.

## III. Approach

### A. Assumptions

Topic 1.3 lists some minimal assumptions, in order to determine the situation in which RPCA would be suitable. On imposing that the low rank component $L_0$ is not sparse, citing [**?**] the incoherence condition with the parameter $\mu$ suggests that

$$max||U^T e_i||^2 \leq \frac{\mu r}{n_1}.........(1)$$

$$max||V^T e_i||^2 \leq \frac{\mu r}{n_2}.........(1)$$

$$||UV^*||_\infty \leq \sqrt{\frac{\mu r}{n_1 n_2}}.........(2)$$

Here U and V are SVD-based singular-vectors, $n_1$ and $n_2$ are the dimensions of $L_0$, and $r$ is the rank of $L_0$.
The incoherence condition asserts that for small values of $\mu$, singular vectors are reasonably spread out – hence, not sparse. Another issue arises if the sparse matrix has low rank; likely to happen if all entries in S are concentrated to very few columns; then $L_0$ and $S_0$ would not be recoverable, and hence it is assumed that the sparsity pattern of the sparse component is selected uniformly at random.

### B. Theorems

Theorem 1.1: *Suppose that we have $L_0$ that follows (1). and (2). Fix any $n$ x $n$ matrix $\Sigma$ of signs. Suppose that the support set $\Omega$ of $S_0$ which is uniformly distributed among all sets of cardinality $m$ and sgn($[S_0]_{ij}$) $= \Sigma_{ij}$ for all $(i, j) \in \Omega$ Then, there is a numerical constant c, such that, with probability at least $1 - cn^{-10}$, Principal Component Pursuit with $\lambda = \frac{1}{\sqrt{n}}$ is exact, provided that*

$$rank(L_0) \leq \rho_r n \mu^{-1}(logn)^{-2}$$

$$m \leq \rho_s n^2.........(3)$$

Here $\rho_r$ and $\rho_s$ are positive constants.
Under the assumption of the theorem, minimizing

$$minimize: \quad ||L||_* + \lambda||S||_1,$$

$$subject\ to: \quad P_{\Omega_{obs}}(L + S) = Y,$$

then $L_0$ with singular vectors reasonably spread out, can be recovered with probability of almost 1, from arbitrary corruption patterns

Theorem 1.2: *Suppose $L_0$ is $nxn$, that follows (1). and (2), and that $\Omega_{obs}$ is uniformly distributed among all sets of cardinality $m$ obeying $m = 0.1n^2$ . Suppose for simplicity, that each observed entry is corrupted with probability $\tau$ independently of the others. Then, there is a numerical constant c, such that, with probability at least $1 - cn^{-10}$ , PCP with $\lambda = \frac{1}{\sqrt{n}}$ is exact, that is, $\hat{L} = L_0$, provided that*

$$rank(L_0) \leq \rho_r n \mu^{-1} (logn)^{-2}, \quad and \quad \tau \leq \tau_s \text{ .........(4)}$$

Based on this, determined $\lambda = \frac{1}{\sqrt{n}}$ is a universal choice for $\lambda$

## IV. Advantages and Challenges

Advantages of the new method are: robustness to grossly corrupted samples of $M$, applicability even on unknown component of $S$, naturally extending to matrix completion
Challenges: (Input Size, Accuracy) $\longleftrightarrow$ Running Time; *Hard to identify* L and S if the rank of M is extremely low $\rightarrow L$ is sparse and not recoverable.

## V. Algorithm

A suitable value for $\mu$ is $\mu = \frac{n_1 n_2}{4} ||M||_1$

**Algorithm 1** PCA by ADDM (Yuan-Yang,[4])

**initialize**: $S_0 = Y_0 = 0, \mu > 0$
**while** not converged **do**
compute $L_{k+1} = D_\mu(M - S_k + \mu^{-1} Y_k)$;
compute $S_{k+1} = S_{\lambda\mu}(M - L_{k+1} + \mu^{-1} Y_k)$;
compute $Y_{k+1} = Y_k + \mu(M - L_{k+1} - S_{k+1})$;
**end while**
**Output:** *L,S*.

Exit Condition: $||M - L - S||_F \leq \delta ||M||_F$.
Select $\delta$ low, along $\delta = 10^{-7}$
The functions are based on the Augmented Lagrange Multiplier (ALM) function $l(L, S, Y)$

$$l(L, S, Y) = ||L||_* + \lambda ||S||_1 + \langle Y, M - L - S \rangle + \frac{\mu}{2} ||M - L - S||_F^2$$

Let $S_T : \mathbb{R} \rightarrow \mathbb{R}$ denote the shrinkage operator

$$S_\tau[x] = sgn(x) \ max(|x| - \tau, 0)$$

$$Then, \ argmin(S) \ \ l(L, S, Y) = S_{\lambda\mu}(L, S, Y)$$

Let $D_\tau(X)$ denote the singular value thresholding operator.

$$D_\tau(X) = U S_\tau(\Sigma) V^T, \quad X = U \Sigma V^T \ \ is \ an \ SVD$$

$$Then, \ argmin(L) \ \ l(L, S, Y) = D_\mu(L, S, Y)$$

## VI. Implementation

An implementation of RPCA is now available at [5] Foreground-background separation in video frames to process L and S is an The following has been observed on running foreground-background

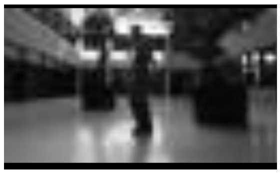


Figure 1.1 : Input Matrix M

Figure 1.2: Low Rank and Sparse Matrix

It can be observed that background-foreground separation is possible. The following statistics on RPCA for [7]

TABLE I
Observations in RPCA

| Input Size | 180 x 2304 |
|---|---|
| $\lambda$ | 0.0208 |
| $\mu$ | 1 |
| rank(L) | 4 |
| Iterations | 483 |
| Time Taken | 34.3 sec |

## VII. Conclusions

From this paper, we understand that under few assumptions and coherence conditions we can see that the low rank component $L$ and sparse component $S$ can be exactly recovered. RPCA can also be used to recover $L$ from grossly corrupted data, with a very high probability. RPCA offers some advantages such as robustness to grossly corrupted samples, having sparsity pattern of $S$ unknown ahead of time, and RPCA can be extended to matrix completion.

## Appendix

Nuclear norm: $||X||_* = \sum_i \sigma_i(X)$
$\ell_1$ norm: $||X||_1 = \sum_{i=1}^m \sum_{j=1}^n |X_{ij}|$
$\ell_\infty$ norm $||X||_\infty = max_{i,j} |X_{i,j}|$

Frobenius norm: $||X||_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |X_{ij}|^2}$

## References


[1] "Robust Principal Component Analysis", E. Candes, 2009
[2] "Principal Component Analysis",I. Jolliffe. Springer-Verlag, 1986.
[3] Candes, B. Recht. A simpler approach to matrix completion
[4] X. Yuan and J. Yang. Sparse and low-rank matrix decomposition via alternating direction methods, 2009.
[5] Github
[6] Conference on RPCA by Emanuel Candes, Youtube
[7] Video File