Robust Principal Component Analysis

Neha Aytoda-1644005
School of Engineering and Appied Science
Ahmadabad University
Email:neha.a.mtechcs16@ahduni.edu.in

***Abstract*—A ubiquitous problem in every domain working with high dimension datasets that can have a fraction of missing entries and fallacious data and the problem is to fill those missing entries and correct fallacious data.Suppose we have a data matrix,which is the superposition of a low-rank component and a sparse component and we want to recover each component individually this can be done using convex optimization techniques which are discussed in the paper.**

Keywords:Robust PCA, Low rank, Sparse Matrix,,Convex Optimization, Principal Component pursuit, Matrix Norms, Augmented Lagrange Multiplier.

I. Introduction

Principal Component Analysis is widely used in many areas of applied statistics. It is natural since interpretation and visualization in a fewer dimensional space is easier than in many dimensional space. But in real life situations, the data observed has some fractions of outliers. Principal Component Analysis is not prone to corrupted values. If we organize all the points and then apply SVD, hen SVD fails in such cases where we also have corrupted features. Robust PCA deals with the problem of making PCA robust to outliers and grossly corrupted observations. the Robust PCA problem can be formulated mathematically as the problem of decomposing a matrix consisting of sum of low-rank matrix and a sparse matrix.

II. THE MATRIX SEPARATION PROBLEM

Let the data matrix M

$$M = L_0 + S_0 \tag{1}$$

where,$L_0$is low-rank Matrix and S_0 is sparse matrix.here, both components are of arbitrary magnitude.

we aim to extract L0 and S0 exactly. Initially, as the number if unknowns are higher and we do not know the amount of errors in L0, it seems daunting to extract them. Hence, this seems to be a hard problem, as we do not know how to solve this non-convex, we try to relax the problem to a convex optimization problem as following:

$$\begin{aligned} minimize \quad & \parallel L \parallel_\Lambda + \lambda \parallel S \parallel_1 \\ Subjectto \quad & M = L_0 + S_0 \end{aligned} \tag{2}$$

where , $\parallel L \parallel_* = \Sigma_i \sigma_i(L)$ denote the nuclear norm of matrix L and $\parallel S \parallel_1 = \Sigma_{ij} j M_{ij}$of the matrix S denotes the l1 norm.

Under these relaxed assumptions, Principal Component Pursuit(PCP) estimates low-rank L_0 and sparse S_0 exactly. The solution holds true even if rank of L grows linearly with the dimensions of M and S_0 has constant fractions of entries.

III. PRIOR ASSUMPTIONS

While separating the data matrix M, what if M has both sparse and low-rank component? [1] has given an example of $amatrixMwhichisequaltoe_1 e_1^*$ which has 1 on top left left corner and remaining values as 0. Another issue arises certain direction in original data is poorly represented. In such cases, M is both low-rank and sparse. Hence we impose an incoherence condition which asserts that for small values of μ, S_0 is not sparse. The other condition is that the sparse matrix should not have low-rank i.e. we assume that the sparsity pattern of the sparse component is selected uniformly at random.

$$max_i \parallel U * e_i \parallel^2 \leq \frac{\mu r}{n_1} and max_i \parallel V * e_i \parallel^2 \leq \frac{\mu r}{n_2}$$

The incoherence parameter μ, which measures how column spaces and row spaces of L are aligned with previous basis and between themselves. In above discussed situations, value of μ is higher. But for smaller values of μ, the singular vectors are randomly spread out.

IV. THEOREMS

Under the above mentioned essential assumptions, the simple PCP approach perfectly recovers low rank and sparse component exactly with large probability.

A. *THEOREMS 1*

Suppose $L_0 isn * n$, obeys the above mentioned prior assumptions. Fix any$n * nmatrix|\Sigma$ of signs. Suppose that the support set$-\Omega$ of S_0 is uniformly distributed among all sets of cardinality m, and that $sgn([S_0]_{ij}) = \Sigma_{ij} forall(i;j) \in \Omega$. Then, there is a numerical constant c such that with probability at least$1 - cn^{-10}$ (over the choice of support of S_0), Principal Component Pursuit with $\lambda = \frac{1}{\sqrt{n}}$ is exact, that is,$\hat{L} = L_0$ and $\hat{S} = S_0$, provided that

$$rank(L_0) \leq \rho n \mu^{-1} (logn) and m \leq \rho_s n^2$$

B. Matrix completion from grossly corrupted data

In many situations, it is possible that some values can be corrupted. In some applications, some entries may be missing as well. Denoting P_Ω as an orthogonal projection on linear space of matrices having support on $\Omega \subset [n1] \times [n2]$, As we have only few entries of $L_0 + S_0$, we can write Y as,

$$Y = P_\Omega(L_0 + S_0) = P_{\Omega_{obs}} L_0 + S\prime_0$$

C. Theorem 2

Suppose $L_0 inn \times n$,obeys the incoherence conditions and that obeys the prior assumptions ans 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 τ independently of the others. Then, there is a numerical constant c such that with probability at least $1 - cn^{-10}$, Principle Component Pursuit with$\lambda = 1/\sqrt{0.1n}$

is exact, that is $L_0' = L_0$ provided that,

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

Here,ρ_τ and τ_s are positive numerical constants. For $n1 \times n2$ matrices, we take $\lambda = \frac{1}{\sqrt{0.1 n_{(1)}}}$ succeeds from m = 0.1 n1n2 corrupted entries with probability at least $1 - cn_{(1)}^{-10}$ (1) provided that $rank(L_0) \leq \rho_r \eta \mu^{-1} (logn_{-1})^{-2}$

V. ALGORITHM

We use the Augmented Langrange Multiplier method to solve this convex optimization problem. The algorithm is generalized to wide range of problems. The rank remains bounded by rank(L0) throughout the optimization. The augmented langrangian here is:

$$l[L, S, Y] = \parallel L \parallel_* + \lambda \parallel S \parallel_1 (Y, M - L - S) + \frac{\mu}{2} \parallel M - L - S \parallel_F^2$$

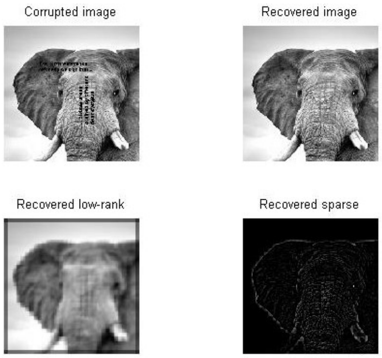
Below is the proposed Alternating Directions methods, which is a special case of augmented Lagrange multiplier(ALM).

Algorithm 1 Principal Component Pursuitby Alternating Directions

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

VI. APPLICATION AND IMPLEMENTATION

Robust PCA can be has numerous application such as Video Surveillance where it is often required to identify the activities that stand out from the background, face recognition where we effectively model low-dimensional for imagery data, Latent Semantic Indexing, Matrix Completion and Recommendation System.



Above image is showing the result of separation of a corrupted inage into low rank and sparce components. Rank of the low rank component comes out to 29 and cardinality of sparse matrix is 231387.

CONCLUSION

We can conclude from the above explanation and the results that one can disentangle the low-rank and the sparse components exactly by convex programming, this provably works under quite broad conditions. Also the above method can be used for matrix completion and matrix recovery from sparse errors and this also works in the case when there are both incomplete and corrupted entries.

REFERENCES

[1] E. Cand'es, E. J., Li, X., Ma, Y., Wright, J. (2011). Robust principal component analysis?. Journal of the ACM (JACM), 58(3), 11. Chicago

[2] E. Candes, X. Li, Y. Ma, and J. Wright. Robust principal component analysis. 2009. http://www-stat.stanford.edu/ candes/papers/RobustPCA.pdf

[3] https://github.com/dlaptev/RobustPCA