Robust Principal Component Analysis?

Rajvi Prajapati,1401033
School of Engineering and Applied Sciences
Ahmedabad University
Email: rajvi.p.btechi14@ahduni.edu.in

***Abstract*—The article delivers the problem of recovering the low-rank and sparse components from a data matrix. This problem captures a broad spectrum of applications in the area of video surveillance, face recognition,Latent Semantic Indexing,Ranking and collaborative Filtering etc. It is NP-hard problem and thus not tractable in general but still under certain suitable assumptions like Incoherence and Random support,it is possible to recover both the low-rank and the sparse components exactly by solving a very convenient convex program called Principle component pursuit where the l_1-norm and the nuclear norm are used to induce sparse and low-rank structures, respectively. This enables the possibility of a principled approach to Robust principal component analysis even in the presence of gross errors or arbitrarily corrupted entries. This proposed model also benefits in the scenario where a fraction of the entries are missing. A special case of augmented Lagrange multiplier(ALM) algorithm known as alternating directions methods (ADM) is discussed to accomplish the recovery of Low-rank and Sparse matrix which is easily implementable and computationally efficient as per numeric results.**

Key Words: PCA,Sparcity,Nuclear norm Minimization,Outliers

I. Introduction

We are given a large data matrix M which may b decomposed as

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

We neither know the low-dimensional column and row space of Low-rank matrix L_0 nor even their dimensions. Similarly,We dont know the locations and number of the nonzero entries of sparse matrix S_0.

II. Matrix seperation methods

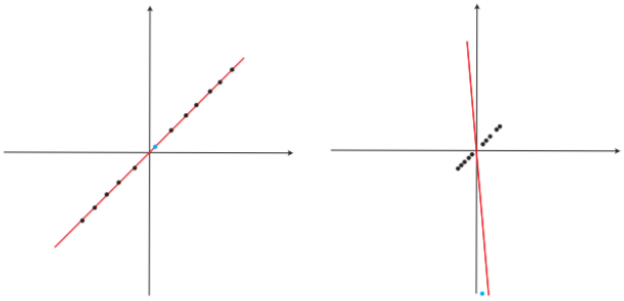
1) classical Principal Component Analysis: It is the most widely used tool for linear dimensionality reduction and clustering. PCA simply seeks the best rank-k estimate of L_0 in the l_2 sense, which can be solved efficiently via singular value decomposition (SVD) and thresholding. [1]

$$M = L_0 + N_0 \tag{2}$$

$$minimize \quad \| M - L \|$$

$$Subject\ to \quad rank(L) \leq k \tag{3}$$

Where N_0 is small and independent and identically distributed Guassian noise component. Classical PCA is highly sensitive to outliers and does not scale well with respect to the number of data samples and thus it suffers brittle faiure on grossly corrupted observations.



[Figure 1]

2) Robust Principal Component Analysis: Existing methods fail to provide a polynomial time algorithm with strong performance guarentees. RPCA suggests the possibilty to correctly recover underlying low-rank structure in the data, even in the presence of gross errors or outlying observations. Gross errors frequently occur in many applications like Image processing,Web data analysis,Bioinformatics,Occlusions,Sensor failures etc.The natural approach to incoprorate arbitrarily large sparse errors:

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

$$Subject\ to \quad M = L_0 + S_0 \tag{4}$$

Where, $\| L \| * = \Sigma_i \sigma_i(L)$(Sum of singular values) denotes the nuclear norm of matrix L and $\| S \|_1 = \Sigma_{ij} |M_{ij}|$(Sum of absolute values) of the matrix S denotes the l_1 norm.

The recently proposed Principal Component Pursuit (PCP) method utilizes a convex program that guarantees the recoverability by solving above equation even if rank of L grows linearly with the dimensions of M and S_0 has constant fractions of entries.

III. Prior Assumptions

Some prior assumptions are taken to avoid identifiability issue.

A. Incoherence

Objective of this assumption is to impose that low-rank component L_0 is not sparse.[2].Suppose that L_0 is a rank-r matrix with SVD

$$L_0 = U \Sigma V^* = \sum_{i=1}^{r} \sigma_i u_i v_i^* \tag{5}$$

where $\sigma_1, ..., \sigma_r$ are the positive singular values and $U = [u_1, ..., u_r], V = [v_1, ..., v_r]$ are the matrices of left and right singular vectors. Then,the incoherence condition with parameter μ states that

$$max_i||U * e_i||^2 \leq \frac{\mu r}{n_1}, \;\; max_i||V * e_i||^2 \leq \frac{\mu r}{n_2} \tag{6}$$

and

$$||UV^*||_\infty \leq \sqrt{\frac{\mu r}{n1n2}} \tag{7}$$

Here,$||M||_\infty max_{i,j}|M_{ij}|$,that is the l_∞norm of M seen as a long vector.Since the orthogonal projection P_U onto the column space of U is given by $P_U = UU^*$ is equivalent to $max_i||P_U e_i||^2 \leq \mu r/n_1$,and similarly for P_V. So,for small values of , the singular vectors are reasonably spread out not sparse.

B. Random Support

Objective of this assumption is to check whether sparse matrix has Low-rank. This issue will occur if all nonzero entries of S occur in a column or in a few columns. For instance,if the first column of S_0 is the opposite of that of L_0 and that all the other columns of S_0 vanish.Then it would be clear that we would not be able to recover L_0 and S_0 by any method. So, to avoid such meaningless situations, the sparsity pattern of S_0 is selected uniformly at random.

IV. Theorems and Algorithm

A. Theorem 1.1

Suppose L_0 is $n_1 \times n_2$, obeys (6) and (7) then there is a numerical constant c such that with probability at least $1 - cn_1^{-10}$,PCP with $\lambda = 1/\sqrt{n_1}$ is exact,that is, $\hat{L} = L_0$ and $\hat{S} = S_0$ provided that

$$rank(L_0) \leq \rho_r n_2 \mu^{-1} (logn_1)^{-2} \; and \; m \leq \rho_s n_1 n_2 \tag{8}$$

Under the assumption, minimizing

$$||L||_* + \frac{1}{n_1}||S||_1, \;\; n_1 = max(n_1, n_2) \tag{9}$$

always return the correct answer.

B. Theorem 1.2

Suppose L_0 in $n \times n$,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,

$$rank(L_0) \leq \rho_r n \mu^{-1} (logn)^{-2} \;\; , and \;\; \tau \leq \tau_s \tag{10}$$

C. Algorithm

Our algorithm for solving the convex Principal Component Pursuit problem works on augmented Lagrange multiplier (ALM) introduced in Lin et al. [2009a] and Yuan and Yang [2009].This works across a wide range of problem settings with no tuning parameters.The rank iterates often remains bounded by $rank(L_0)$ throughout the optimization.The augmented Lagrangian for this problem can be defined as

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

The solution is obtained by repeatedly setting $(L_k, S_k) = \arg min_{L,S} l(L, S, Y_k)$ and updating the Lagrange multiplier by $Y_{k+1} = Y_k + \mu(M - L_k - S_k)$. Now,

$$\arg \; min_S l(L, S, Y) = S_{\lambda/\mu}(M - L - \tfrac{1}{\mu}Y)$$
$$\arg \; min_L l(L, S, Y) = D_{1/\mu}(M - S - \tfrac{1}{\mu}Y)$$

Here,formally we are minimizing l w.r.t L by fixing S,then minimizing l w.r.t S fixing L and finally updating the Lagrange Multiplier Y based on residual $M - L - S$.

Algorithm 1 Principal Component Pursuit by Alternating Directions [Lin et al. 2009a;Yuan and Yang 2009]

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

Here, $\mu = n_1 n_2/4 \| M \|_1$.Algorithm is terminated when $\| M - L - S \|_F \leq \delta \| M \|_F$ with $\delta = 10^{-7}$.

V. Future implemantation

The above implematation is limited to low-rank component being exactly low and the sparse component being exactly sparse.The algorithms discussed above can be explored more if any one of the assumptions is relaxed.

References

[1] Zhou, Z., Li, X., Wright, J., Candes, E., Ma, Y. (2010, June). Stable principal component pursuit. In Information Theory Proceedings (ISIT), 2010 IEEE International Symposium on (pp. 1518-1522). IEEE.

[2] Zhang, H., Zhou, Y., Liang, Y. (2015). Analysis of robust PCA via local incoherence. In Advances in Neural Information Processing Systems (pp. 1819-1827).

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