Robust Principal Component Analysis

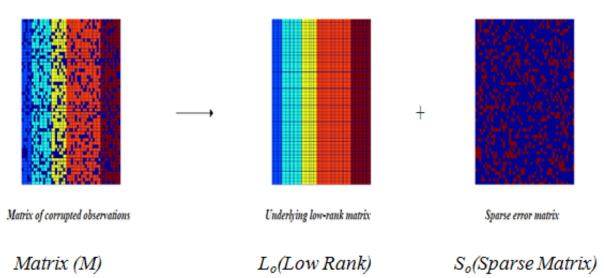
Report by Pinak Divecha
Roll No: 1401098

***Abstract*—Principal Component Analysis (PCA) has been widely used for the representation of shape, appearance, and motion. One drawback of typical PCA methods is that they are least squares estimation techniques and hence fail to account for "outliers" which are common in realistic training sets. To overcome the problem of outliers, robust principle component analysis is used. In Robust PCA the goal is to decompose the data matrix A in two components such that A=L+S, where L is a low-rank matrix and S is a sparse noise matrix. To decomposition the data matrix they have used technique named as Principal Component Pursuit method. In this paper, they have applied the method of augmented Lagrange multipliers (ALM) for optimization.**

Keywords. Principal components, Low rank, Sparse matrix, nuclear-norm minimization,l1-norm minimization, Convex optimization.

I. Introduction

The proposed robust principal component analysis (robust PCA) has the ability to recover the low rank model from sparse noise. This extends to the situation where a fraction of the entries are missing as well. There are many important applications in which the data under study can naturally be modeled as a low-rank plus a sparse contribution. Depending on the applications, either the low-rank component or the sparse component could be the object of interest like Video Surveillance, Face Recognition, Latent Semantic Indexing, detection of objects in a cluttered background.



II. Separation

Suppose, we are given a large matrix $M \in R^{n_1,n_2}$, and know that it may be decomposed as,

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

Where L_0 is a Low rank Matrix and S_0 is a sparse matrix. Here, both components are of arbitrary magnitude. We don't know the low-dimensional column and row space of L_0, not even their dimension. Similarly, we don't know the location of the nonzero entries of S_0 and not even how many there are. One state-of-the-art approach is to decompose M into a low rank and a sparse component. This approach has been shown to be advantageous for several reasons. First, it overcomes the weakness of standard PCA where the solution could be extremely skewed even if a single entry is corrupted, and second, exact recovery of L_0 can be guaranteed under certain assumptions. By using principle component pursuit (PCP) estimate solving one can exactly recover the low-rank matrix L_0 from M

$$\begin{aligned} minimize \quad & ||L||_* + \lambda ||S||_1 \\ subject\ to \quad & L + S = M \end{aligned}$$

Here, $||L||_* := \sum_i \sigma_i(L)$ denote the nuclear norm of low rank matrix L that is the sum of the singular value of L, and let $||S||_1 := \sum_{ij} |S_{ij}|$ denote the l_1 norm of matrix S seen as sum of absolute values of element S.

A. Low rank could not be sparse

Every algorithm works under some constraints, similarly to disentangle the low-rank and the sparse component we need to impose that the low-rank component (L_0) is not sparse (S_0). If the data matrix is both sparse and low rank then we won't be able to decide whether it is low-rank or sparse. Suppose the matrix M has one in the top left corner and zeros everywhere else. Then M is equal to the $e_1 e_1^*$ which both sparse and low-rank.

$$\boldsymbol{M} = \boldsymbol{e}_1 \boldsymbol{e}_n^* = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

Here e_1 is the basis vector and e_1^* is the transpose of basis vector. If we project the basis vector on the column vector of L_0 and if the projection on the column vector is very very large (μ is large)then we can say that low-rank component is sparse. To avoid such meaningless situations, we will assume that the sparsity pattern of the sparse component is selected uniformly at random.

B. Main result

The surprise is that under these minimal assumptions, the simple PCP solution perfectly recovers the low-rank and the sparse components. Which is given as following:

Theorem 1.1: Suppose L_0 is nxn obeying the incoherence condition (for n = n1 = n2) and that the support set of S_0 is uniformly distributed among all sets of cardinality m. 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 (PCP) with $\lambda = \frac{1}{\sqrt{n}}$ is exact, i.e.,$L = L_0$ And $S = S_0$, provided that the rank

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

In the general rectangular case where L_0 is $n1$ x $n2$, PCP with $\lambda = \frac{1}{\sqrt{n}}$ succeeds with probability at least 1 - cn^{-10}.Under the assumption of the theorem,

$$minimize \qquad ||L||_* + \frac{1}{\sqrt{n}} ||S||_1 \quad , n_1 = max(n_1, n_2)$$

One would have to choose the right scalar λ to balance the two terms $||L||_* + \lambda ||S||_1$ appropriately . This is, however, clearly not the case. In this sense, the choice =1n 1 is universal. Further, it is not a priori very clear why $\lambda = \frac{1}{\sqrt{n_1}}$ is a correct choice no matter what L_0 and S_0 are. It is the mathematical analysis which reveals the correctness of this value.

We have seen that theorem 1.1 asserts that it is possible to recover a low-rank matrix even though a significant fraction of its entries are corrupted. In some applications, however, some of the entries may be missing as well. And this is a significant extension of the matrix completion problem, which seeks to recover L_0. The technique developed in this paper is established is given as:

Theorem 1.2: Suppose L_0 is $n \times n$, obeying the incoherence condition (for n = n1 = n2), and that Ω_{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 τ independently of the others. Then, there is a numerical constant c such that with probability at least 1 - cn^{-10}, Principal Component Pursuit with $\lambda = \frac{1}{\sqrt{0.1n}}$ is exact, that is, $\hat{L}$= L_0, provided that

$$rank(L_0) \leq \rho_r n \mu^{-1} (logn)^{-2} \; and \; \tau \leq \tau_s. \qquad (3)$$

In this equation, ρ_r and τ_s are positive numerical constants. For general n_1 n_2 rectangular matrices, PCP with $\lambda = \frac{1}{\sqrt{0.1n_1}}$ succeeds from $m = 0.1 n_1 n_2$ corrupted entries with probability at least 1-cn^{-10} , provided that $rank(L_0) \leq \rho_r n_2 \mu^{-1} (logn_1)^{-2}$. In short, perfect recovery from incomplete and corrupted entries is possible by convex optimization.

III. Conclusion

This paper delivers some rather surprising news: one can disentangle the low-rank and sparse components exactly by convex programming, and this provably works under very broad conditions. Analysis has revealed rather close relationships between matrix completion and matrix recovery (from sparse errors) and our results even generalize to the case when there are both incomplete and corrupted entries. In addition, Principal Component Pursuit does not have any free parameter and can be solved by simple optimization algorithms with remarkable efficiency and accuracy. And this can be used in computer vision, signal processing, data analysis and many others. So far robustPCA is limited to the low-rank component being exactly low-rank, and the sparse component being exactly sparse. Hope to see the solution of this problem in future.

IV. Algorithm

Algorithm 1 given below is a special case of a more general class of augmented Lagrange multiplier algorithms known as alternating directions methods

Algorithm 1 (Principal Component Pursuit by Alternating Directions [33, 51])

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

V. References

[1]//http://ieeexplore.ieee.org/abstract/document/937541/

[2]http://papers.nips.cc/paper/3704-robust-principal-component-analysis-exact-recovery-of-corrupted-low-rank-matrices-via-convex-optimization

[3]Robust Principal Component Analysis: Exact Recovery of Corrupted Low-Rank Matrices by Convex Optimization. John Wright , Yigang Peng, Yi Ma Visual Computing Group Microsoft Research Asia

[4]Analysis of Robust PCA via Local Incoherence. Huishuai Zhang, Yi Zhou, Yingbin Liang.