Robust Principal Component Analysis

Reported by Shruti Agrawal - 1644007
School of Engineering and Applied Sciences,Ahmedabad University
Email: shruti.a.mtechcs16@ahduni.edu.in

***Abstract*—The paper shows an algorithm for solving the Robust Principle Component Analysis problem i.e. Recovering a low rank matrix with an unknown fraction of entries being arbitrary corrupted or missing .The data matrix is basically a super-position of low rank component and sparse component.It is shown that under some specific conditions and assumptions,it is possible to recover both low rank component as well as sparse components(errors).Recovering the Low-rank component and the Sparse component is possible by solving convex optimization program called Principle Component pursuit by minimizing a weighted combination of nuclear norm and of l1 norm.This suggests a possibility of principled approach to make robust Principle component Analysis.This can be useful in foreground background separation video surveillance,face recognition etc.**
Keywords: PCA,Matrix norms,low rank matrices, sparsity,outliers,video surveillance

I. INTRODUCTION

Suppose we are given a data matrix M such that M can be decomposed as $M = L_0 + S_0, where L_0$ has low rank and S_0 is sparse and both matrix are of arbitrary magnitude. We do not know the low-dimensional column and row space of L_0, similarly the non zero entries of S_0 are not known. There are many prior attempts to solve the above mentioned problem.

II. THEORETICAL ASPECTS

A. Classical Principal Component Analysis[1]

To solve the dimensionality and scale issue we must leverage on the fact that such data matrix are intrinsically lower in dimension, thus are indirectly sparse in some sense. Perhaps the simplest assumption is that the data in matrix all lie near some lower dimensional subspace, hence we can stack all the data points as column vector of a matrix M, and this column vector can be represented mathematically,

.

$$M = L_0 + S_0, where L_0$$

where L0 is essentially low rank and N0 is a small perturbation matrix.Classical Principal Component seeks the best rank-k estimate of L0 by solving

$$\textit{minimize } \|\text{M-L}\|$$
$$\textit{subject to rank}(\text{L}) \leq k.$$

Throughout this article $\|$M $\|$denotes l^2 norm. In classical PCA it is assumed that N_0 is small.

B. Robust Principal Component Analysis[2]

PCA is unequivocally the best statistical tool for data analysis and dimensionality reduction. However its brittleness to small corrupted data in data matrix puts its validity in jeopardy, as this small corruption could render the estimated $\hat{\text{L}}$ arbitrarily far from true L_0. Such problems are ubiquitous in modern applications such as image processing, web data analysis and many more.The problems mentioned above are the idealized version of robust PCA, where we recover low rank matrix from highly corrupted data matrix M such that $M = L_0 + S_0$ where unlike classical PCA S_0 can have arbitrarily large magnitude. and their support is assumed to be sparse.

III. ALGORITHM

The theorem above shows that incoherent low-rank matrices can be recovered from non-vanishing fractions of gross errors in polynomial time. For small problem sizes, PCP can be performed sing off-the-shelf tools such as interior pint methods. Despite the superior converge rates, interior point methods are typically limited to small problems, say $n < 100$, due to O^{n6} complexity of computing a step direction.

A. Principal component pursuit

1) Assumptions:: There is high possibility that the data matrix M has only the top left corner 1 and all other entries in the matrix are 0. Thus M is both sparse and low rank, thus to make the problem meaningful we assume that low rank matrix L_0 is not sparse. Also there is a possibility that the sparse matrix S_0 has all non-zero entries in few columns. To avoid such meaningless situations, we assume that the sparsity pattern of S_0 is uniformly random.

2) Claim:: Let the data matrix$M \epsilon R^{n1xn2}$. Also the low Prank matrix is L_0 and the sparse matrix is S_0.Let $\|$M$\|_*$ = $\Sigma_i \sigma_i$(P) denote the nuclear norm of any matrix M.Also $\|$M $\|_1$ denote the l_1 norm of any matrix P, then Principal component pursuit gives estimate,

$$\textit{minimize } \|\text{L }\|_* + \lambda k S k_1$$
$$\textit{subject to } \text{L + S = M}$$

The above estimate exactly recovers the Low-rank matrix L_0 and the sparse matrix S_0. Theoretically the claim is true even if the rank of matrix L_0 almost linearly and the errors in S_0 are upto a constant factors of all entries. Empirically we can solve this problem by efficient and scalable algorithms, at a cost not much higher than classical PCA.

B. *Alternating Directions Methods*

The below proposed Alternating Directions methods is a special case of augmented Lagrange multiplier(ALM).

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

IV. APPLICATIONS

There are some very important real life applications of Robust PCA wherein it is used to separate the low-rank and sparse distributions from a given data set. Some of them are as follows:

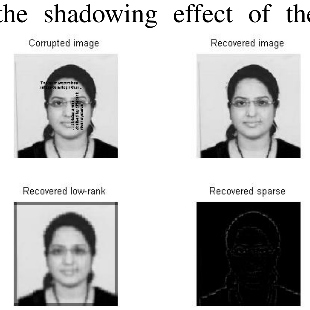
A. *Video Surveillance*

It is often required in a set of video frames to identify the activities that stand out from the background. If these various video frames are stacked upon on another to form a data matrix M then the low-rank component L_0 refers to the stationary background and the sparse components S_0 are the moving objects in the foreground as these occupy only a small fraction of the whole video frame. Thus, by separating the low-rank and the sparse components of the data matrix we can in essence separate the foreground from the background.

B. *Face Recognition*

It is established that images of a human's face can be approximated by a low-dimensional sub-space. Being able to correctly recognise a face is crucial to many applications such as face recognition and alignment. However, realistic faces and scenarios often suffer from self-shadowing, specularities, or saturations in brightness, which make it a difficult task and subsequently compromise the recognition performance. Thus, Robust PCA can be used to remove these shadows from faces for a more proper recognition.

In the application discussed below we have taken a corrupted data matrix M which is a corrupted image and from this corrupted data matrix we have achieved the L_0 low-rank matrix and S_0 completely.Here the speculation and the shadowing effect of the image are stored in S_0.



As shown above from the corrupted image M having dimensions 256x256, thus $M \epsilon R^{256x256}$ we have successfully recovered L_0 and have recorded speculations and image shadowing in S_0. The program took 16 seconds to converge with 330 iterations. The rank of low-rank matrix ;rank(L)= 38 which suggests it is indeed a low rank. The cardinality of set of the sparse matrix S_0 is card(S)= 207857. The error rate observed was 2.67.

V. CONCLUSION

In applications such as live video streaming ,where streaming of huge amount of data.Our aim is to process this data as quickly and efficiently as possible.Robust PCA algorithm in this paper ensures dimension reduction separating the low rank component as well as the sparse component.It corrects the errors in the data and also completes the matrix where the entries are missing.The cost of separation and recovering in not much higher than Classical Principle Component Analysis.

REFERENCES

[1] 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] E. Candes, X. Li, Y. Ma, and J. Wright. Robust principal component analysis. 2011. https://github.com/dlaptev/RobustPCA