

Project Report: Robust Principle Component Analysis

Abstract—In this paper, we talk about a data matrix which is superposition of a low rank component and a sparse component. The question here is can we recover both low rank and sparse components individually from the data matrix? Under some suitable assumptions it is possible to recover both these components exactly by a very convenient convex program called Principal Component Pursuit; among all feasible decompositions, simply minimize a weighted combination of the nuclear norm and of the ℓ_1 norm. This paper also allows us to recover the principle components if some of its entries are arbitrarily corrupted. This also holds true if some of its entries are missing.

I. INTRODUCTION

Suppose, that we are given a matrix M which can be decomposed in two parts such that

$$M = L_0 + S_0 \quad (1)$$

where L_0 is a low rank matrix and S_0 is a sparse matrix. We do not know the dimensions of either of them. A solution to this problem might prove to be very important in today's data intensive process of scientific discovery. In current application domains of this problem, data is usually in millions or billions of dimensions. So we also need to reduce the dimensionality and bring the data in lesser dimensions. So we convert our large data matrix into low rank matrix L_0 . Classical Principle Component Analysis seeks the best rank- k estimate of L_0 by solving

$$\begin{aligned} \text{minimize} \quad & \|M - L\| \\ \text{subject to} \quad & \text{rank}(L) \leq k \end{aligned}$$

PCA is the most widely used statistical tool for data analysis and dimensionality reduction. Many approaches have been proposed but none of them yield a polynomial time algorithm with strong performance guarantees under broad conditions. The problem we study here can be considered as an idealized version of Robust PCA, in which our aim is to recover a low rank matrix L_0 from highly corrupted measurements $M = L_0 + S_0$. The entries in S_0 can have arbitrarily high magnitude but is sparse. Separation makes sense when a matrix M is either sparse or low rank and not both. Suppose M is a matrix $e_1 e_1^*$. Then M is both sparse and low rank. So to make problem meaningful we need to make sure matrix L_0 is low rank and not sparse. Using the general notion of incoherence introduced in Cands and Recht [2009] for the matrix completion problem; this is an assumption concerning the singular vectors of the low-rank component. The SVD of L_0 can be written as where L_0 is a $n_1 \times n_2$ matrix

$$L_0 = U \Sigma V^*$$

where Σ has positive singular values and U and V are matrices of left and right singular vectors. Thus the incoherence condition with parameter μ states that

$$\max \|U^* e_i\|^2 \leq \frac{\mu r}{n_1}, \max \|V^* e_i\|^2 \leq \frac{\mu r}{n_2} \quad (2)$$

$$\|UV^*\|_\infty \leq \sqrt{\frac{\mu r}{n_1 n_2}} \quad (3)$$

Here, $\|M\|_\infty = \max_{i,j} |M_{ij}|$, that is, is the l_∞ norm of M seen as a long vector. Note that since the orthogonal projection P_U onto the column space of U is given by $P_U = UU^*$, (2) is equivalent to $\max_i \|P_U e_i\|^2 \leq \frac{\mu r}{n_1}$, and similarly for P_V . Another issue arises if the sparse matrix is low rank. To avoid such situations we assume that the sparsity pattern of the sparse component is selected uniformly at random.

If we have a matrix M of dimensions $n_1 \times n_2$ then under some weak assumptions the Principal Component Pursuit estimate solving

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

exactly recovers matrix L_0 and S_0 . This is valid even if entries in low rank matrices increases linearly in the dimensions and the errors in sparse matrix are upto a constant fraction of all entries. Algorithmically, we will see that this problem can be solved by efficient and scalable algorithms, at a cost not so much higher than the classical PCA.

II. APPLICATIONS

There are many important applications in which the data can be modeled as a low rank and sparse matrix.

A. Video Surveillance

Suppose we are given a surveillance frames of a video, we need to identify activities that stand out from background. If we stack video as a matrix, then the low rank matrix corresponds to the stationary background and sparse matrix captures moving objects from the video.

B. Face Recognition

It is well known that images of a convex, Lambertian surface under varying illuminations span a low-dimensional subspace. This fact has been a main reason why low-dimensional models are mostly effective for imagery data. In particular, images of a humans face can be well-approximated by a low-dimensional subspace. Being able to correctly retrieve this subspace is crucial in many applications such as face recognition and alignment. However, realistic face images often suffer from self-shadowing, specularities, or saturations in brightness, which make this a difficult task and subsequently compromise the recognition performance.

C. Ranking and Collaborative Filtering

The problem of anticipating user tastes is gaining increasing importance in online commerce and advertisement. Companies now routinely collect user rankings for various products, for example, movies, books, games, or web tools, among which the Netflix Prize for movie ranking is the best known [Netflix, Inc.]. The problem is to use incomplete rankings provided by the users on some of the products to predict the preference of any given user on any of the products. This problem is typically cast as a low-rank matrix completion problem. However, as the data collection process often lacks control or is sometimes even ad hoc a small portion of the available rankings could be noisy and even tampered with. The problem is more challenging since we need to simultaneously complete the matrix and correct the errors.

Similar problems also arise in many other applications such as Ranking and collaborative Filtering, Latent Semantic Indexing, graphical model learning, linear system identification, and coherence decomposition in optical systems

III. MAIN RESULT

A. Theorem 1.1

Suppose L_0 is $n \times n$, obeys equations (2) and (3). Fix any $n \times n$ matrix Σ of signs. Suppose that the support set Ω of S_0 is uniformly distributed among all sets of cardinality m , and that $\text{sgn}([S_0]_{ij}) = \Sigma_{ij}$ for all $(i, j) \in \Omega$. Then, there is a numerical constant c such that with probability at least $1 - cn^{-10}$ (over the choice of support of S_0), Principal Component Pursuit (1) with $\lambda = 1/\sqrt{n}$ is exact, that is, $\hat{L} = L_0$ and $\hat{S} = S_0$, provided that

$$\text{rank}(L_0) \leq \rho_r n \mu^{-1} (\log n)^{-2} \text{ and } m \leq \rho_s n^2 \quad (4)$$

In this equation, ρ_r and ρ_s are positive numerical constants. In the general rectangular case, where L_0 is $n_1 \times n_2$, PCP with $\lambda = \frac{1}{\sqrt{n_{(1)}}}$ succeeds with probability at least $1 - cn_{(1)}^{-10}$, provided that $\text{rank}(L_0) \leq \rho_r n_{(2)} \mu^{-1} (\log n_{(1)})^{-2}$ and $m \leq \rho_s n_1 n_2$.

Matrix L_0 whose principal components are spreaded can be recovered with probability almost one from arbitrary and

completely unknown corruption patterns. It also works for higher ranks like $n/\log(n)^2$ when μ is not large. Minimizing

$$\|L\|_* + \frac{1}{\sqrt{n_{(1)}}} \|S\|_1$$

where,

$$n_{(1)} = \max(n_1, n_2)$$

under the assumption of theorem, this always gives correct answer. Here we chose $\lambda = \frac{1}{\sqrt{n_{(1)}}}$ but it is not clear why that has happened. It has been due to mathematical analysis why we are taking that value.

IV. CONCLUSION

From the above results we can say that if conditions in equation (2) and (3) are satisfied then we can get a low rank and sparse matrix from matrix M . The main phenomena is to select a uniformly random sparse matrix. So this approach can be used to decompose a given matrix in two components i.e. low rank and sparse matrix.

REFERENCES

- [1] Cands, E. J., Li, X., Ma, Y., and Wright, J. 2011. Robust principal component analysis? J. ACM 58, 3, Article 11 (May 2011), 37 pages.
- [2] Yi Ma, Compressive Principal Component Pursuit, 2012
- [3] C HANDRASEKARAN , V., S ANGHAVI , S., P ARRILO , P., AND W ILLSKY , A. 2009. Rank-sparsity incoherence for matrix decomposition. Siam J. Optim.