Robust Principal Component Analysis

EMMANUEL J. CANDES and XIAODONG LI,Stanford University
YIMA, University of Illinois at Urbana-Champaign,Microsoft Research Asia
JOHN WRIGHT,Microsoft Research Asia

***Abstract*— This article demonstrates that for a given data matrix, which is a superimposition of a lower-rank component and a sparse component, it is conceivable to recuperate both the low-rank and the sparse components precisely by comprehending an exceptionally useful convex program called Principal Component Pursuit; under some reasonable suspicions; among every plausible decompositions, essentially minimizing a weighted mix of the nuclear norm and of the l_1 norm.**

I. INTRODUCTION

A. Motivation

Decomposition of a large matrix M can be done as follows:

$$M = L_0 + S_0,$$

where, L_0 has a low rank and S_0 is sparse.The least complex and most useful assumption here is that the data all lie near some low-dimensional subspace.Mathematically,

$$M = L_0 + N_0,$$

where, N_0 is a small perturbation matrix. Classical PCA looks for the best rank-k estimate of L_0 by solving

$$minimize||M - L|| \quad and \quad subject \quad to \quad rank(L) \leq k$$

...(1)

B. Applications

- Video Surveillance
- Face Recognition
- Latent Semantic Indexing
- Ranking and Collaborative Filtering

C. Separation of Components

Singular Value Decomposition of $L_0 \in R^{n_1 \times n_2}$ is given as :

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

where, r = rank of the matrix, $\sigma_1, ..., \sigma_r$ are 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 :

$$max_i||U_* e_i||_2 \leq \frac{\mu r}{n_1}, \ max_i||V_* e_i||_2 \leq \frac{\mu r}{n_2}$$

and

$$||UV^*||_\infty \leq \sqrt{\frac{\mu r}{n_1 \times n_2}}$$

D. Main Result

The matrices L_0 whose principal components are sensibly spread can be recuperated with probability nearly one from arbitrary and completely known corruption patterns.In fact, this works for extensive estimations of the rank,i.e.,on the order of n/$(logn)^2$ when μ is not too large, where n = order of the matrix and μ = incoherence condition parameter.

E. Implications for Matrix Completion from Grossly Corrupted Data

A matrix with significant fraction of its entries corrupted as well as some of them missing can be recovered by the proposed algorithm. Assumption : Each observed entry is corrupted with probability τ independently of others.Then, there exists a constant c such that with probability at least $1 - cn^{-10}$, *Principal Component Pursuit* with $\lambda = 1/\sqrt{0.1n}$ is exact,i.e., L'=L_0 provided that,

$$rank(L_0) \leq \rho_r n \mu^{-1} (logn)^{-2} \qquad ...(2)$$

Hence, perfect recovery from incomplete and corrupted entries is possible by convex optimization. And the above result proves that matrix completion is stable versus gross errors.

II. ARCHITECTURE OF THE PROOF

This section presents the key steps underlying the evidence of the main result, equation (1). The result is proved for square matrices and is extended for rectangular matrices.

A. An elimination theorem

Definition : S' is a trimmed version of S if supp(S')⊂supp(S) and $S'_{ij} = S_{ij}$ whenever $S'_{ij} \neq 0$ This definition in turn asserts that is Principal Component Pursuit correctly recovers the low-rank and sparse components of $M_0 = L_0 + S_0$, it also correctly recuperates the components of a matrix $M'_0 = L_0 + S'_0$ Hence, if the solution to equation (1) with input data $M_0 = L_0 + S_0$ is unique and exact, then the solution to the same equation with input M'_0 is exact as well.

B. Derandomization

THEOREM : Suppose L_0 obeys the conditions of equation (2) and that the locations of the nonzero entries of S_0 follow the Bernoulli model with parameter $2\rho s$, and the signs of S_0 are i.i.d. ± 1. Then if the PCP solution is exact with high probability, then it is also exact with at least the same probability for the model in which the signs are xed and the locations are sampled from the Bernoulli model with parameter ρ_s.

C. Key Lemmas

- Suppose ω_0 is sampled from the Bernoulli model with parameter ρ_0. Then, with high probability,

$$||P_T - \rho_0^{-1} P_T P_{\Omega_0} P_T|| \leq \epsilon$$

  provided that $\rho_0 \geq C_0 \epsilon^{-2}(\mu r logn)/n$ for some numerical constant $C_0 > 0$. For rectangular matrices, we need $\rho_0 \geq C_0 \epsilon^{-2}(\mu r logn_1)/n_2$
- *LEMMA :* Assume that $\Omega\ Ber(\rho)$ with parameter $\rho \leq \rho_s$ for some $\rho_s > 0$. Set $j_0 = 2[logn]$. Then, under other assumptions of equation (2), the matrix W^L obeys :
  (a)$||W^L|| < 1/4$
  (b)$||P_\Omega(UV^* + W^L)||_F < \lambda/4$
  (c)$||P_{\Omega^\perp}(UV^* + W^L)|| < \lambda/4$
  Since $||P_\Omega P_T|| < 1$ with large probability, W^S is well defined and the following holds:
- *LEMMA :* Assume that S_0 is supported on a set Ω sampled and that the signs of S_0 are independent and identically distributed symmetric. Then, under the other assumptions of equation (2), the matrix W^S obeys :
  (a)$||W^S|| < 1/4$
  (a)$||P_{\Omega^\perp} W^S|| < \lambda/4$

III. INPUT & OUTPUT IMAGES

Fig. 1. Input Image

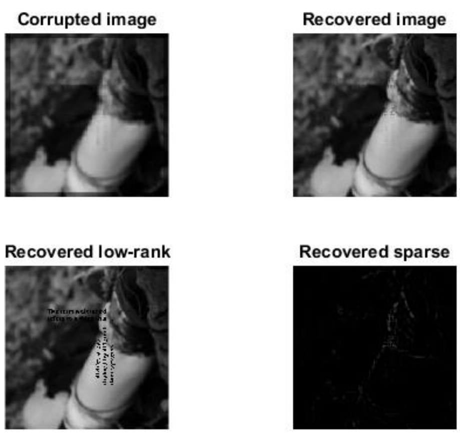

Fig. 2. Output Images

IV. CONCLUSIONS

The article comes to a conclusion that one can unravel the low-rank and sparse components precisely by convex programming and this provably works under exceptionally broad conditions that are much more extensive than the best known outcomes.The results even generalize to the case when there are both incomplete and corrupted entries and the problem can be solved with remarkable efficiency and accuracy. Even more generally, the issues of sparse signal recovery, low-rank matrix completion, classical PCA and robust PCA can all be considered as special instances of a general measurement model of the form :

$$M = A(L_0) + B(S_0) + C(N_0)$$

where, A,B and C are known linear maps.

V. FUTURE SCOPE

The striking capacity of convex optimization in recoupling low-rank matrices and sparse signals in high dimensional spaces recommend that they will be an effective apparatus for handling gigantic data sets, that emerge in image/video processing, web data analysis and bioinformatics.

REFERENCES

[1] V. Chandrasekaran, S. Sanghavi, P. Parrilo, and A. Willsky. Rank-sparsity incoherence for matrix decomposition. preprint, 2009.
[2] E. J. Cand‘es and Y. Plan. Matrix completion with noise. Proceedings of the IEEE (to appear), 2009.
[3] S. Becker, J. Bobin, and E. J. Cand‘es. NESTA: A fast and accuract rst-order method for sparse recovery. preprint, 2009.
[4] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming (web page and software). http://stanford.edu/boyd/cvx, June 2009.
[5] D. Gross. Recovering low-rank matrices from few coecients in any basis. CoRR, abs/0910.1879, 2009.