Robust Principal Component Analysis and its application in Computer Vision

Yevhen Pozdniakov, Yaroslava Lochman

Linear Algebra
Interim Report

January 2019

Description of the problem

One of the most important problems in video surveillance is to distinguish background of the scene. This task is complicated by the presence of foreground objects: in busy scenes, every frame may contain some anomaly. Moreover, the background model needs to be flexible enough to accommodate changes in the scene, for example due to varying illumination or moving objects on the foreground.

The problem can be formulated in a way of finding decomposition of a given large matrix M, which is constructed by stacking video frames, as follows: $M = L + S$, where L has low-rank and corresponds to the background, S is sparse and represents foreground moving objects, illumination changes etc.

Sketch of the possible approaches

To recover a low-rank matrix from data one can use simple PCA method, but for highly noisy measurements we need to apply the so-called robust PCA algorithms. Natural approaches to robustifying PCA such as influence function techniques, multivariate trimming, alternating minimization, random sampling techniques had been explored before 2000s but these techniques were mostly heuristic, they didnt have guarantees to work. Later different convex and non-convex algorithms under different assumptions have been proposed for robust PCA. In short, the decomposition $M = L + S$ can be achieved by the methods such as Principal Component Pursuit method (PCP) [1], Stable PCP [2], Quantized PCP [3], Block based PCP [4] and Local PCP [5]. Then, among the optimization methods used the Augmented Lagrange Multiplier Method (ALM) [6], Alternating Direction Method (ADM) [7], Fast Alternating Minimization (FAM) [8], Iteratively Reweighted Least Squares (IRLS) [9,10]. Guaranteed methods for robust PCA were presented in [13, 1], where the authors showed recovery of an incoherent low rank matrix L^* through the following convex relaxation method:

$$\min_{L,S} \ \operatorname{rank}(L) + \lambda \|S\|_0 , \quad \text{s.t. } M = L + S$$

$$\downarrow$$

$$\min_{L,S} \ \|L\|_* + \lambda \|S\|_1 , \quad \text{s.t. } M = L + S$$

where $\|S\|_0$ is a total number of non-zero elements, $\|L\|_*$ denotes a nuclear norm (the sum of singular values), and $\|S\|_1$ denotes entry-wise l_1 norm, $\lambda > 0$ regularization parameter. Chandrasekaran et al.[13] considers a deterministic sparsity model for $S*$, where each row and column of the mn matrix S has at most fraction of non-zero entries. Candes et al. [1]

considers a different model with random sparsity and additional incoherence constraints, implements the in-exact augmented Lagrangian multipliers (IALM) method and provides guidelines for parameter tuning. Other related methods such as multi-block alternating directions method of multipliers (ADMM) have also been considered for robust PCA, e.g. in [11]. There exist works considering a non-convex method for robust PCA [20, 12]. The state-of-the-art results were achieved by the Non-convex Alternating Projections based Robust PCA [12], that have an alternative to the convex relaxation. Suppose we know the rank of the matrix L, and consider the following problem:

$$\min_{L,S} \ \|S\|_0\,, \quad \text{s.t. } M = L + S, \ \ \text{rank}\,(L) = r$$

The algorithm consists of the iterative alternating (non-convex) projections onto low-rank and sparse matrices

Pros and cons of the method

+ can recover a low-rank component of the data matrix and the sparse component, so the principal components of the matrix can easily be found even if the data is corrupted or if there some noise exist, thus it is robust to outliers that represent the sparse component;
+ under several suitable assumptions the methods guarantee the convergence;
+ our numerical experiments on fitting noisy data points by an affine subspace showed that the recovering low-rank component using RPCA works considerably better than PCA even for relatively big amount of noise (say one third of the data is corrupted or represents outliers);
+ can be implemented on the parallel and distributed computing infrastructures;
- it is reasonable to apply the method only for such data where the sparse component (corruptions, errors, noise etc.) is present for sure, otherwise there is no sense to use it instead of e.g. SVD for low-rank approximation. Robust PCA will give very similar result and sometimes a bit worse, and it will likely be slower, especially on big amounts of data;
- the computational complexity of the algorithm is relatively high. Suppose we have m n matrix ($m \leq n$). To reach a desired accuracy of eps (meaning Frobenius distance to the true low-rank component), the robust PCA via convex relaxation methods needs $O(m^2n/\varepsilon)$. However the state-of-the-art non-convex alternating minimization techniques for robust PCA need $O(mnr^2 \log(1/\varepsilon))$ where r is the rank of a low-rank componen. So it has lower computational complexity which in fact is close to PCA that needs $O(mnr \log(1/\varepsilon))$, so it is scalable to large matrices.

A numerical realization

The iteration process to get robust PCA is pretty simple. On every step we use SVD of difference between original matrix and sparse matrix to calculate low-rank mtrix. We shrink matrix S, using threshold. As the error term we use $\|M - L - S\|_2$. We stop iterative process if one of two conditions reaches:

1. error term less than tolerance expected accuracy (by default $\|M\|_2 \cdot 10^{-7}$)
2. number of iterations more than the maximum number of iterations (by default 10^3).

References

[1] E. J. Candes; X. Li; Y. Ma; J. Wright (2011). "Robust Principal Component Analysis?" J. of the ACM

[2] J. Wright; Y. Peng; Y. Ma; A. Ganesh; S. Rao (2009). "Robust Principal Component Analysis: Exact Recovery of Corrupted Low-Rank Matrices by Convex Optimization". Neural Information Processing Systems, NIPS 2009.

[3] S. Becker; E. Candes, M. Grant (2011). "TFOCS: Flexible First-order Methods for Rank Minimization". Low-rank Matrix Optimization Symposium, SIAM Conference on Optimization.

[4] G. Tang; A. Nehorai (2011). "Robust principal component analysis based on low-rank and block-sparse matrix decomposition". Annual Conference on Information Sciences and Systems, CISS 2011.

[5] B. Wohlberg; R. Chartrand; J. Theiler (2012). "Local Principal Component Pursuit for Nonlinear Datasets". International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2012.

[6] Z. Lin; M. Chen; L. Wu; Y. Ma (2013). "The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-Rank Matrices". Journal of Structural Biology. 181(2): 11627.

[7] X. Yuan; J. Yang (2009). "Sparse and Low-Rank Matrix Decomposition via Alternating Direction Methods". Optimization Online.

[8] P. Rodrguez; B. Wohlberg (2013). "Fast Principal Component Pursuit Via Alternating Minimization". IEEE International Conference on Image Processing, ICIP 2013.

[9] C. Guyon; T. Bouwmans; E. Zahzah (2012). "Foreground Detection via Robust Low Rank Matrix Factorization including Spatial Constraint with Iterative Reweighted Regression". International Conference on Pattern Recognition, ICPR 2012.

[10] C. Guyon; T. Bouwmans; E. Zahzah (2012). "Moving Object Detection via Robust Low Rank Matrix Decomposition with IRLS scheme". International Symposium on Visual Computing, ISVC 2012.

[11] A. Kyrillidis; V. Cevher (2012). "Matrix alps: Accelerated low rank and sparse matrix reconstruction". In SSP Workshop.

[12] P., Netrapalli; U., Niranjan; S., Sanghavi; A., Anandkumar; P., Jain (2014). "Nonconvex robust PCA". Advances in Neural Information Processing Systems. 1410: 11071115.

[13] V. Chandrasekaran; S. Sanghavi; P. A. Parrilo; A. S. Willsky (2011). "Rank-sparsity incoherence for matrix decomposition". SIAM Journal on Optimization, 21(2): 572596.