Summary on Robust PCA using Principal Component Pursuit

Aneet Kumar Dutta (18111401)

26 July 2019

Paper Reference: Robust Principal Component Analysis?by Candes,Li,Ma and Wright

Abstract

In this paper, the author tried to recover two matrices L(low-rank matrix) and S(Sparse matrix) given a large data matrix M. The problem is solved by a convex program called Principal Component Pursuit which minimizes a combination of the nuclear norm and l_1 norm. Solving this problem suggests that we will be able to recover the principal components of the data matrix even if some of the entries in the matrix are corrupt or outliers.

Optimization Problem

The optimization problem to solve RPCA problem is:

minimize $||L||_* + \lambda||S||_1$
subject to $L + S = M$

Solving the above optimization problem exactly recovers low-rank L and sparse S because:

- The nuclear norm of the L matrix is the sum of the singular values of L.
- Since the number of non-zero singular values is the rank of a matrix, therefore minimizing the nuclear norm will yield us a low-rank matrix.
- Minimizing the l_1 norm of the matrix S will ensure sparsity of S.

This paper solves the convex PCP problem using Augmented Lagrangian Multiplier(ALM). Now the optimization equation looks like:

$\text{argmin}_{L,S}\ l(L,S,Y) = ||L||_* + \lambda||S||_1 + < Y, M-L-S > + \frac{\mu}{2}||M-L-S||_F^2$.

Here, Y is the Lagrange Multiplier Matrix.

The term $< Y, M-L-S > + \frac{\mu}{2}||M-L-S||_F^2$ is added in the optimization problem due to the constraint $L + S = M$.

The above optimization problem is solved alternatively by updating L and S using singular value thresholding and soft thresholding.

$\text{argmin}_S\ l(L,S,Y) = S_{\lambda\mu^{-1}}(M - L + \mu^{-1}Y)$.

$S_\tau[x] = sign(x)max(|x| - \tau, 0)$, this is called soft thresholding.

$\text{argmin}_L\ l(L,S,Y) = D_{\mu^{-1}}(M - S + \mu^{-1}Y)$.

$D_\tau[X] = US_\tau(\Sigma)V^*$, this is called singular value thresholding.

Finally, Y is updated.

Algorithm

- initialize: $S_0 = Y_0 = 0$, $\mu = \frac{n1n2}{4||M||_1}$, $\delta = 10^{-7}$, λ is the hyperparameter
- while $||M - L - S||_F > \delta||M||_F$ do
  1) compute $U, \Sigma, V^T =$SVD$(M - S_k + \mu^{-1}Y_k)$.
  2) $L_{k+1} = US_{\mu^{-1}}(\Sigma)V^T$.
  3) $S_{k+1} = S_{\lambda\mu^{-1}}(M - L_{k+1} + \mu^{-1}Y_k)$
  4) $Y_{k+1} = Y_k + \mu(M - L_{k+1} - S_{k+1}$
- Return L, S