Robust Principal Component Analysis

Candès et al, *JACM*, 2011

Presented by Lu Meng

Jan 30, 2014

Motivation

- $$M = L_0 + N_0$$
  with L_0 low-rank and N_0 a small perturbation matrix
- Classical Principal Component Analysis (PCA) [1]

$$\begin{aligned} \min_{L} \quad & ||M - L||_2 \\ \text{subject to} \quad & \text{rank}(L) \leq k \end{aligned}$$

- sensitive to corrupted measurements
- none of current robustifying approaches yields a polynomial-time algorithm

[1] $||M||_2 = \max_i \sigma_i(M)$

Robust PCA

- $$M = L_0 + S_0$$
  with L_0 low-rank and S_0 sparse
- Principal Component Pursuit (PCP) [2] [3]

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

[2] $||M||_* = \sum_i \sigma_i(M)$
[3] $||M||_1 = \sum_{i,j} |M_{ij}|$

Applications

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

Identifiability and Uniqueness

- assumptions for identifiable PCP
  - L_0 is not sparse
  - sparsity pattern of the sparse component is selected uniformly random
- no tuning parameter: $\lambda = 1/\sqrt{max(n_1, n_2)}$ is universal
- conditions for unique optimal solution
- can be solved by tractable convex optimization and efficient algorithms

Algorithms

- For small problem sizes: off-the-shelf tools
- Large size: augmented Lagrange multiplier (ALM) method[4] [5]

$$l(L, S, Y) = ||L||_* + \lambda ||S||_1 + \langle Y, M - L - S\rangle + \frac{\mu}{2}||M - L - S||_F^2$$

repeatedly setting $(L_k, S_k) = \arg\min_{L,S} l(L, S, Y_k)$ and then updating the Lagrange multiplier matrix via $Y_{k+1} = Y_k + \mu(M - L_k - S_k)$.

  - [6] $\arg\min_S l(L, S, Y) = \mathcal{S}_{\lambda\mu^{-1}}(M - L + \mu^{-1} Y)$
  - [7] $\arg\min_L l(L, S, Y) = \mathcal{D}_{\mu^{-1}}(M - S + \mu^{-1} Y)$

---

[4] $\langle X, Y\rangle = \text{trace}(X^* Y)$
[5] $||X||_F^2 = \langle X, X\rangle$
[6] $\mathcal{S}_\tau[x] = \text{sgn}(x)\max(|x| - \tau, 0)$
[7] $\mathcal{D}_\tau(X) = U\mathcal{S}_\tau(\Sigma)V^*$, where $X = U\Sigma V^*$

Algorithm 1: PCP by Alternating Directions

initialize: $S_0 = Y_0 = 0, \mu > 0$

while *not not converaged* **do**

compute $L_{k+1} = \mathcal{D}_{\mu^{-1}}(M - S_k + \mu^{-1} Y_k)$;

compute $S_{k+1} = \mathcal{S}_{\lambda\mu^{-1}}(M - L_{k+1} + \mu^{-1} Y_k)$;

compute $Y_{k+1} = Y_k + \mu(M - L_{k+1} - S_{k+1})$

output: L, S

Examples

- surveillance video: PCP vs a state-of-the-art technique
- removing shadows from face images