A Nonconvex Projection Method for Robust PCA

Arika Dutta*, Filip Hanzely*, Peter Richtárik*,Ω,α
King Abdullah University of Science and Technology*, University of EdinburghΩ, MIPTα

Robust Principal Component Pursuit

Robust principal component pursuit is a robust matrix decomposition model in which we wish to decompose A into the sum of a low-rank matrix L and an error matrix S: A = L + S.

The celebrated principal component pursuit uses the ℓ2 norm to address the sparsity and one obtains

\[
\min_{L,S} \frac{1}{2} ||A - L - S||_{F}^{2} \text{subject to } L \geq 0, S \geq 0.
\]

Different RPCA formulations by using surrogate constraints and optimization functions.

\[\min_{L,S} \frac{1}{2} ||A - L - S||_{F}^{2} \text{subject to } \lambda \leq \text{rank}(L) \leq r \text{ and } \lambda \leq \text{rank}(S) \leq r.\]

An extended model of (Cherapanamjeri, Gupta, and Jain 2017; Gupta, and Jain 2017) is the restriction operator defined by

\[
\text{rank}(A) \leq r \text{ and } ||A - L - S||_{F} \leq \varepsilon
\]

Moving the constraint to the objective as a penalty, together with adding explicit constraints on the target rank and target sparsity leads to the following formulation (Zhou and Tao 2011) as:

\[
\min_{L,S} \frac{1}{2} ||A - L - S||_{F}^{2} \text{subject to } \lambda \leq \text{rank}(L) \leq r \text{ and } \lambda \leq \text{rank}(S) \leq r.
\]

Cherapanamjeri, Gupta, and Jain 2017; Chen et al. 2011, Tao and Yang 2011 proposed the robust matrix completion (RMC) problem as:

\[
\min_{L,S} \frac{1}{2} ||A - L - S||_{F}^{2} \text{subject to } \lambda \leq \text{rank}(L) \leq r \text{ and } \lambda \leq \text{rank}(S) \leq r.
\]

\[
\text{rank}(A) \leq r \text{ and } ||A - L - S||_{F} \leq \varepsilon
\]

Applications to Real-World Problems

Nonconvex Feasibility and Alternating Projections

Set feasibility:
Find point in the intersection of closed sets

\[
\text{Find } x \in X, \quad X = \bigcap_{i=1}^{n} X_{i}
\]

A natural alternating projection algorithm to solve set feasibility is Algorithm 1.

Algorithm 1: Alternating projection for set feasibility

\[
\text{Input: } i \in \{1, 2, ..., n\}, x_{0} \in X_{i}, \epsilon > 0
\]

\[
\text{Output: } x
\]

\[
\text{for } i = 1, n \text{ do }
\]

\[
\text{if } x_{i-1} \neq x_{i-1} \text{ then }
\]

\[
\text{if } x_{i-1} \in X_{i} \text{ then }
\]

\[
\text{end}
\]

\[
\text{end}
\]

Projection of \( L_{0} \) onto rank r constraint

Operator \( H_{r} = \text{rank } r \text{ SVD of } L_{0} \)

Projection of \( S_{0} \) on the sparsity constraint (approximate)

Keep largest \( \alpha \) fraction of values in each row and each column

The next Algorithm is proposed to solve the RMC problem:

Algorithm 2: Alternating projection method for RMC

\[
\text{Input: } L_{0}, S_{0}
\]

\[
\text{Output: } L_{\infty}, S_{\infty}
\]

Now we consider the following reformulations of RPCA:

\[
\min_{L,S} \frac{1}{2} ||A - L - S||_{F}^{2} \text{subject to } \lambda \leq \text{rank}(L) \leq r \text{ and } \lambda \leq \text{rank}(S) \leq r.
\]

Convergence of Algorithm 2

Local linear convergence:

\[
d_{X_{k}\cap X_{l}}([L_{0}, S_{0}]) < \varepsilon d_{X_{k}\cap X_{l}}([L_{0}, S_{0}])
\]

Cosine of angle between tangent space of \( X_{k} \) and \( X_{l} \)

Define \( X_{k} = X_{k} \cap X_{l} \)

Global linear rate

Numerical Results

We demonstrate our results on both synthetic and real data.

Sensitivity of Algorithm 3 to the initialization.

Figure 1: Three synthetic examples for RCPA (RAH, APG, and SVD) and GRASTA with respect to rank and error sparsity. Here \( r = \text{rank}(L) \) and \( \alpha = \text{the sparsity parameter.} \)

Figure 2: Three synthetic examples for Relative error for RPCA algorithms: (a) \( \delta_{r}^{\text{GRISTA}} = 0.03, \delta_{r}^{\text{APG}} = 0.09, \delta_{r}^{\text{SVD}} = 0.05 \), (b) \( \delta_{r}^{\text{GRISTA}} = 0.03, \delta_{r}^{\text{APG}} = 0.02, \delta_{r}^{\text{SVD}} = 0.05 \). We observe that APG has more success rate than both algorithms.

Figure 3: Background and foreground estimation on Stargate dataset [1000 frames]. Except RPCA GD and our method, all the methods fail to recover the exact foreground signal.

Figure 4: Background and foreground estimation on Stargate dataset [1000 frames]. We used 90% sample. GRASTA forms a fragmentary background and clutter around 150 frames to form a static video. We note that RPCA GD has more data points in the foreground.

Figure 5: Shadow and specular regions missed from two images captured under varying illumination and camera positions. Our feasibility approach provides a consistent reconstruction to that of dALM and APG.