Robust Principal Component Analysis on Graphs

Nauman Shahid
Vassilis Kalofolias, Xavier Bresson, Michael Bronstein & Pierre Vandergheynst
EPF Lausanne, Switzerland

SPARS 2015, Cambridge UK
What is this talk about?

Sparsity?

NO!

True for most of the people in this auditorium.

Unfortunately “sparsity” in this talk represents “gross errors”

This talk is about Principal Component Analysis / Low-rank representation: The most widely used tool for linear dimensionality reduction & clustering!

SPARSITY!!!!
Outline

Principal Component Analysis
Robust Principal Component Analysis

Robust Principal Component Analysis on Graphs
*How and Why?*

Solution using ADMM

Results:
*Clustering in low-dimensional space*
*Low-rank recovery*
Robust Principal Component Analysis

Original data space $X \in \mathbb{R}^{p \times n}$

- Principal directions / basis
- PC space $Q$

Not robust to gross errors

Robust to Gaussian noise

$$\min_{U, Q} \| X - UQ \|_F^2 \quad \text{s.t.} \quad U^\top U = I$$

Solution: $SVD(X) = U \Sigma Q'$

$$U = [U_1 \mid U_2] \quad \Sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \quad Q' = \begin{bmatrix} Q_1' \top \\ Q_2' \top \end{bmatrix}$$

Example:
- Gross errors
- Low-rank ($L$)
- Sparse ($S$)

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

solved using ADMM
Robust Principal Component Analysis on Graphs

Original data space

\( X \in \mathbb{R}^{p \times n} \)

\[ \sigma_1 U_1 \]

\[ \sigma_2 U_2 \]

linear subspace

standard form

\[
\min_{U, Q} \|X - UQ\|_F^2 \\
\text{s.t. } U^T U = I
\]

Robust PCA

\[
\min_{L, S} \|L\|_* + \lambda \|S\|_1 \\
\text{s.t. } X = L + S
\]

\[ L = UQ \]

How & why?

Original data space

\( X \in \mathbb{R}^{p \times n} \)

non-linear smooth manifold

characterized via a discrete graph \( G = (V, E) \)

Edge weights: \( W_{i,j} = \exp\left(-\frac{\|x_i - x_j\|^2}{\sigma^2}\right) \)

\[ \Phi = D - W \]

unnormalized Laplacian

\[ \Phi = I - D^{-1/2}WD^{-1/2} \]

normalized
Robust Principal Component Analysis on Graphs: How?

\[ \min_{L,S} \|L\|_* + \lambda \|S\|_1 + \gamma \text{tr}(L\Phi L^T) \quad \text{s.t.} \quad X = L + S \]
Robust Principal Component Analysis on Graphs

**Why?**

\[
\min_{L,S} \|L\|_* + \lambda\|S\|_1 + \gamma \text{tr}(L\Phi L^T) \quad \text{s.t. } X = L + S
\]

Original data space

- Exploit local similarity information
- Enhanced low-rank representation, smooth on the manifold
- Better clusters in low-dimensional space
- Two important applications of PCA!

**IMPORTANT**: Smooth low-rank, NOT only principal components

non-linear smooth manifold

convexity
Solution using ADMM

\[
\min_{L, S} \|L\|_* + \lambda \|S\|_1 + \gamma \text{tr}(L\Phi L^T) \quad \text{s.t. } X = L + S
\]

Augmented Lagrangian

\[
(L, S, U)^{k+1} = \min_{L, S, U} \|L\|_* + \lambda \|S\|_1 + \gamma \text{tr}(U\Phi U^T) + \langle Z_1, X - L - S \rangle \\
+ \frac{r_1}{2} \|X - L - S\|_F^2 + \langle Z_2, U - L \rangle + \frac{r_2}{2} \|U - L\|_F^2
\]

\[
Z_1^{k+1} = Z_1^k + r_1 (X - L - S)^{k+1}, Z_2^{k+1} = Z_2^k + r_2 (U - L)^{k+1}
\]

\[
L^{k+1} = \text{prox}_{1/(r_1 + r_2)\|\cdot\|_*} \left( \frac{r_1 (X - S + Z_1/r_1) + r_2 (U + Z_2/r_2)}{r_1 + r_2} \right)
\]

SVDB : AΩB

singular value soft-thresholding

\[
S^{k+1} = \text{prox}_{\frac{1}{r_1} \|\cdot\|_1} \left( X - L^{k+1} + \frac{Z_1}{r_1} \right)
\]

constant

element wise soft-thresholding

\[
U^{k+1} = r_2 (\gamma H + r_2 I)^{-1} \left( L^{k+1} - \frac{Z_2}{r_2} \right)
\]

system of linear equations

\[
O(pn^2)
\]

\[
O(I|E|n)
\]

p < n
Results

• Clustering in low-dimensional space
• Static background separation from dynamic foreground
Results: Clustering

<table>
<thead>
<tr>
<th>Dataset</th>
<th>No Corruptions</th>
<th>Sample specific corruptions</th>
<th>Corruptions on the whole data set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Occlusions (25% of image size)</td>
<td>Missing pixels (25% of image size)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>15%</td>
<td>25%</td>
</tr>
<tr>
<td>CMU PIE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ORL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YALE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COIL20</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Bar charts showing clustering results for different datasets and corruptions levels.
Results: Principal Components for 3 classes of ORL dataset
Results: Background Separation from Videos
Please have a look at the full version of the paper on arXiv
Demo & code available at:
https://lts2.epfl.ch/blog/nauman/recent-projects/

Fast version and its code will be available in 2 weeks.

Future Work: Theoretical investigation