Random Consensus Robust PCA

Daniel Pimentel-Alarcón & Robert Nowak

Wisconsin Institute for Discovery
UNIVERSITY of WISCONSIN-MADISON
Department of Electrical and Computer Engineering

AISTATS 2017
$X = L + S$

Robust PCA
What is this good for?
Lots of Applications
Lots of Applications
Lots of Applications
Lots of Applications
Lots of Applications
Lots of Applications
Lots of Applications
Background segmentation
Background segmentation
Background segmentation

$$\begin{align*}
\text{minimize} & \quad \|L\|_* + \lambda \|S\|_1 \\
\text{subject to} & \quad X = L + S
\end{align*}$$

Existing theory
Uniform Sampling + Incoherence

Existing theory


Uniform Sampling + Incoherence

Existing theory


Existing theory

Uniform Sampling

+ Incoherence

References


Existing theory

Uniform Sampling

+ Incoherence


Existing theory

Uniform Sampling + Incoherence

References
Existing theory

Uniform Sampling

+ Incoherence


In general

$$X = L + S$$
In general

To answer this:
Totally different way to think about the problem

- Incoherence
- Uniform
- With high probability
- Optimization

- Arbitrary
- Deterministic
- With probability 1
- Algebraic/Geometric

\( X = L + S \)
THE FOLLOWING PREVIEW HAS BEEN APPROVED FOR ALL AUDIENCES BY THE MOTION PICTURE ASSOCIATION OF AMERICA INC.

THE FILM ADVERTISED HAS BEEN RATED

RESTRICTED
UNDER 17 REQUIRES ACCOMPANYING PARENT OR GUARDIAN

GEOMETRY

www.filmratings.com   www.mpaa.org
**PCA**: Finds Subspace that Explains Data.
**PCA:** Finds *Subspace* that Explains Data.
• **PCA**: Finds **Subspace** that Explains Data.
• **Complication**: corrupted entries in **EVERY** column!
• **PCA**: Finds **Subspace** that Explains Data.
• **Complication**: corrupted entries in **EVERY** column!
• **PCA**: Finds *Subspace* that Explains Data.
• **Complication**: corrupted entries in *EVERY* column!
• **PCA**: Finds *Subspace* that Explains Data.
• **Complication**: corrupted entries in *EVERY* column!
• **PCA**: Finds **Subspace** that Explains Data.
• **Complication**: corrupted entries in **EVERY** column!
• **PCA**: Finds Subspace that Explains Data.
• **Complication**: corrupted entries in EVERY column!
• **ALL** columns are outliers!
Our main idea

- Take a few columns at a time (as RANSAC)
Our main idea

- Take a few columns at a time (as RANSAC)
Our main idea

• Take a few columns at a time (as RANSAC)
• Take a few coordinates at a time (projection)
Our main idea

- Take a few columns at a time (as RANSAC)
- Take a few coordinates at a time (projection)
Our main idea

- Take a few columns at a time (as RANSAC)
- Take a few coordinates at a time (projection)
Our main idea

• Take a few columns at a time (as RANSAC)
• Take a few coordinates at a time (projection)
Our main idea

- Take a few columns at a time (as RANSAC)
- Take a few coordinates at a time (projection)
Our main idea

• Take a few columns at a time (as RANSAC)
• Take a few coordinates at a time (projection)
Our main idea

- Take a few columns at a time (as RANSAC)
- Take a few coordinates at a time (projection)
Our main idea

- Take a few columns at a time (as RANSAC)
- Take a few coordinates at a time (projection)
Our main idea

• Take a few columns at a time (as RANSAC)
• Take a few coordinates at a time (projection)
Our main idea

- Take a few columns at a time (as RANSAC)
- Take a few coordinates at a time (projection)
Our main idea

- Take a few columns at a time (as RANSAC)
- Take a few coordinates at a time (projection)
Our main idea

• Take a few columns at a time (as RANSAC)
• Take a few coordinates at a time (projection)
Our main idea

- Take a few columns at a time (as RANSAC)
- Take a few coordinates at a time (projection)
Now the question is:
Given canonical projections, can I find the subspace?
Now the question is:

Given canonical projections, can I find the subspace?
Now the question is:
Given canonical projections, can I find the subspace?
Now the question is:
Given canonical projections, can I find the subspace?
Now the question is:
Given canonical projections, can I find the subspace?
A subspace can be recovered from $N = d - r$ canonical projections if and only if every subset of $n$ projections involves at least $n + r$ coordinates.
**Theorem** (Pimentel, Boston, Nowak, ISIT, 2015)

A subspace can be recovered from $N = d - r$ canonical projections if and only if every subset of $n$ projections involves at least $n + r$ coordinates.

$N = d - r$

$d=9$

$r=3$
This tells me
Which projections I need to reconstruct the subspace.
Our Algorithm: R2PCA
(Ransac Robust PCA)
Our Algorithm: R2PCA
(Ransac Robust PCA)

rank = r

uncorrupted
Our Algorithm: R2PCA
(Ransac Robust PCA)

rank > r

corrupted
Our Algorithm: R2PCA
(Ransac Robust PCA)
Our Algorithm: R2PCA
(Ransac Robust PCA)

Keep finding uncorrupted projections
Our Algorithm: R2PCA
(Ransac Robust PCA)

Keep finding uncorrupted projections
Our Algorithm: R2PCA
(Ransac Robust PCA)

Keep finding uncorrupted projections
Our Algorithm: R2PCA

(Ransac Robust PCA)

Keep finding uncorrupted projections
Our Algorithm: R2PCA
(Ransac Robust PCA)

If we find the right projections, we can find the subspace. Keep finding uncorrupted projections.
Background segmentation
In many cases, similar results
In other cases, better
In other cases, better
Original Video

This Work
(Pimentel, Nowak, 2017)

RPCA-ALM
(RPCA-ALM (Lin et al, 2011-2016))

Coherent Subspaces!!
Performance Analysis
Performance Analysis

- Few errors
- Coherent (bad)
- Incoherent (good)
Performance Analysis

- Coherent (bad)
- Incoherent (good)

Few errors

Many errors
Performance Analysis

RPCA-ALM (Lin et. al, 2011-2016)

Coherent (bad)

Incoherent (good)

Few errors

Many errors

(the lighter the better)
This work (Pimentel, Nowak, 2017)

Performance Analysis

Coherent (bad)

Incoherent (good)

Few errors

Many errors

(the lighter the better)
Our main result in a nutshell

We can find the subspace, \textit{and fast, unless} there are too many errors

Pimentel, Nowak, AISTATS, 2017
WOW, AMAZING
PLEASE TELL ME MORE
$S^* = r$-dimensional subspace in general position.
$S^* = r$-dimensional subspace in general position.

Assumptions
$S^* = r$-dimensional subspace in general position.
\( S^* = r\)-dimensional subspace in general position.

Columns lie in \( S^* \) generically.

**Assumptions**
\( S^* = r \)-dimensional subspace in general position.

Columns lie in \( S^* \) generically.

Assumptions
$S^* = r$-dimensional subspace in general position.

Columns lie in $S^*$ generically.

Assumptions
Take-home Message

• New (algebraic) method for Robust PCA.

• Arbitrary (non-uniform, even adversarial) sparsity patterns.

• No coherence assumptions.
Joint work with:
Thank you!