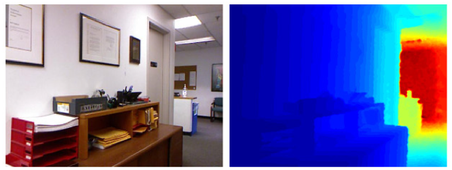
Dual Principal Component Pursuit: Improved Analysis and Efficient Algorithms

Zhihui Zhu[1], Yifan Wang[2], Daniel Robinson[1], Daniel Naiman[1], René Vidal[1], Manolis Tsakiris[2]

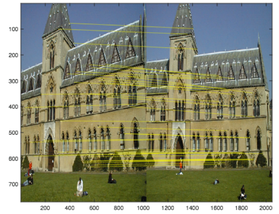
[1]Johns Hopkins University [2]ShanghaiTech University

Motivation

- Problem: Fit a subspace to data contaminated with outliers
- Prior work: Sparse & low-rank methods require subspace to be low dimensional
- Challenge: Many applications require subspace to be of high relative dimension



Estimating the geometry of a room from its depth map by fitting multiple planes

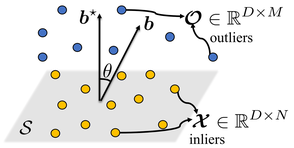


Estimating relative motions between two images by fitting a hyperplane

Dual Principal Component Pursuit (DPCP)

- Inliers $\boldsymbol{\mathcal{X}}$ span a d-dimensional subspace $\mathcal{S}$
- Given $\widetilde{\boldsymbol{\mathcal{X}}} = [\boldsymbol{\mathcal{X}}\ \boldsymbol{\mathcal{O}}]$, DPCP computes $\boldsymbol{b}^\star$ by

$$(1) \qquad \min_{\boldsymbol{b}\in\mathbb{S}^{D-1}:\ \text{unit sphere}} \|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}\|_1$$



- $\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}\|_1$ is a surrogate for $\|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}\|_0$, which counts the number of outliers to $\mathcal{S}$.
- Tsakiris & Vidal derived geometric conditions under which the solution to (1) is orthogonal to $\mathcal{S}$ and proposed two algorithms for solving (1): Alternating Linear Projection (ALP) and Iterative Reweighted Least Squares (IRLS).
- This paper derives tighter conditions and proves that the Projected SubGradient Method (PSGM) has a piecewise linear convergence rate for (1).

	Theory		Algorithms		
	Geometric analysis	Probabilistic analysis	Convergence guarantee		Scalable (# points)
Tsakiris & Vidal 2015 & 2018	✓	✗	ALP	✓	✗ ($\leq 10^3$)
			IRLS	✗	✓ ($\leq 10^4$)
This paper	✓	✓	PSGM	✓	✓ ($< 10^5$)

Global Optimality Analysis for DPCP

- Geometric measures:

$$c_{\boldsymbol{\mathcal{X}},\min} := \frac{1}{N} \min_{\boldsymbol{b}\in\mathbb{S}^{D-1}\cap\mathcal{S}} \|\boldsymbol{\mathcal{X}}^\top \boldsymbol{b}\|_1$$

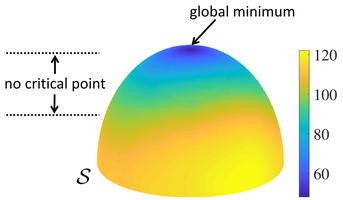
$$\eta_{\boldsymbol{\mathcal{O}}} := \frac{1}{M} \max_{\boldsymbol{b}\in\mathbb{S}^{D-1}} \Big\| \underbrace{(\mathbf{I} - \boldsymbol{b}\boldsymbol{b}^\top)\boldsymbol{\mathcal{O}}\operatorname{sign}(\boldsymbol{\mathcal{O}}^\top \boldsymbol{b})}_{\text{Riemannian subgradient}} \Big\|_2$$

$$\kappa_{\boldsymbol{\mathcal{O}}} := \frac{1}{M}\left(\max_{\boldsymbol{b}\in\mathbb{S}^{D-1}} \|\boldsymbol{\mathcal{O}}^\top \boldsymbol{b}\|_1 - \min_{\boldsymbol{b}\in\mathbb{S}^{D-1}} \|\boldsymbol{\mathcal{O}}^\top \boldsymbol{b}\|_1 \right)$$

	Data more well-distributed	Data well-distributed, $N, M \to \infty$
$c_{\boldsymbol{\mathcal{X}},\min}$	$\uparrow$	c_d
$\eta_{\boldsymbol{\mathcal{O}}}$	$\downarrow$	0
$\kappa_{\boldsymbol{\mathcal{O}}}$	$\downarrow$	0

- **Lemma**: Any critical point of (1) must either lie in $\mathcal{S}^\perp$, or have a principal angle

$$\theta \geq \theta^\natural := \arccos\left(\frac{M\eta_{\boldsymbol{\mathcal{O}}} + D - 1}{N c_{\boldsymbol{\mathcal{X}},\min}}\right).$$

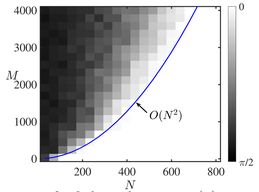


- **Theorem**: Any global solution to (1) must lie in $\mathcal{S}^\perp$ if $\cos^2(\theta^\natural) + \left(\frac{M\kappa_{\boldsymbol{\mathcal{O}}}}{N c_{\boldsymbol{\mathcal{X}},\min}}\right)^2 < 1$.
- This conditon holds for any fixed $\frac{M}{N}$ when $N \to \infty$ if data well distributed.

Probabilistic Analysis

- Outliers drawn uniformly from $\mathbb{S}^{D-1}$
- Inliers drawn uniformly from $\mathbb{S}^{D-1} \cap \mathcal{S}$
- **Theorem**: Any global solution to (1) must lie in $\mathcal{S}^\perp$ with probability exceeding $1 - O(\frac{1}{N^2})$ if

$$M \lesssim \frac{1}{d \max\left(D \log^2 D, \log N\right)} N^2.$$



θ of the solution to (1)

- Proof sketch: $c_{\boldsymbol{\mathcal{X}},\min} \gtrsim 1/\sqrt{d}$, $\eta_{\boldsymbol{\mathcal{O}}} \lesssim \frac{\sqrt{D}\log D}{\sqrt{M}}$, $\kappa_{\boldsymbol{\mathcal{O}}} \lesssim \frac{1}{\sqrt{M}}$
- Comparison with other methods: (Lerman and Maunu, Proceedings of IEEE, 2018)

Geodesic Gradient Descent	REAPER	$\ell_{2,1}$-RPCA	Coherence Pursuit
$M \lesssim \frac{\sqrt{D(D-d)}}{d} N$	$M \lesssim \frac{D}{d} N, d \leq \frac{D-1}{2}$	$M \lesssim \frac{1}{d \max(1, \frac{\log(M+N)}{d})} N$	$M \lesssim \frac{D-d^2}{d} N, d < \sqrt{D}$

Projected SubGradient Method (PSGM)

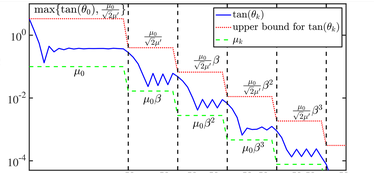
Spectral initialization: set $\widehat{\boldsymbol{b}}_0 = \arg\min_{\boldsymbol{b}\in\mathbb{S}^{D-1}} \|\widetilde{\boldsymbol{\mathcal{X}}}^\top \boldsymbol{b}\|_2$, K_0, K, μ_0, $\beta < 1$.

Piecewise geometrically diminishing step size: $\mu_k = \mu_0 \beta^{\lfloor (k-K_0)/K \rfloor}$.

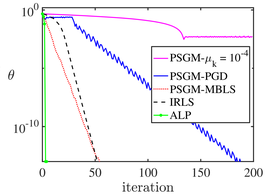
PSGM update: $\boldsymbol{b}_k \leftarrow \widehat{\boldsymbol{b}}_{k-1} - \mu_k \widetilde{\boldsymbol{\mathcal{X}}} \operatorname{sign}(\widetilde{\boldsymbol{\mathcal{X}}}^\top \widehat{\boldsymbol{b}}_{k-1})$; $\widehat{\boldsymbol{b}}_k \leftarrow \boldsymbol{b}_k / \|\boldsymbol{b}_k\|$.

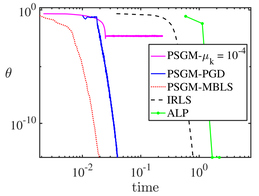
Theorem: If $\theta_0 < \theta^\natural$, PSGM has a piecewise linear convergence rate for $\tan(\theta_k)$, i.e.,

$$\tan(\theta_k) \lesssim \beta^{\lfloor (k-K_0)/K \rfloor}.$$



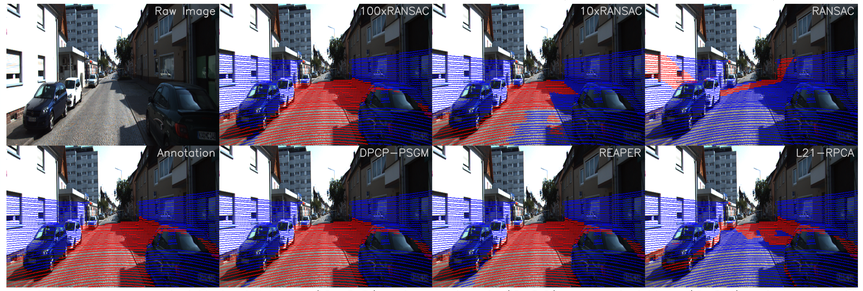
- PGD: Piecewise Geometrically Diminishing
- MBLS: Modified Backtracking Line Search
- $\mu_k = 10^{-4}$: constant step size
- $D = 30, d = 29, N = 500$, $\frac{M}{M+N} = 0.7$.





Experiments on 3D Point Cloud Road Data

- 360° 3D point clouds for Frame 21 of dataset KITTI-CITY-48
- determine 3D points that lie on the road plane (inliers) and off that plane (outliers)



- F1 measures: DPCP-PSGM (**0.93**), REAPER (0.89), ℓ_{21}-RPCA (0.25), RANSAC (0.02), 10xRANSAC (0.62), 100xRANSAC (0.82).

Supported by NSF grant 1704458. We thank Tianyu Ding, Yunchen Yang, Tianjiao Ding, and Ziyu Liu.