Is an Affine Constraint Needed for Affine Subspace Clustering?

Chong You — Electrical Engineering and Computer Sciences, University of California, Berkeley

Chun-Guang Li — Information and Communication Engineering, Beijing University of Posts and Telecommunications

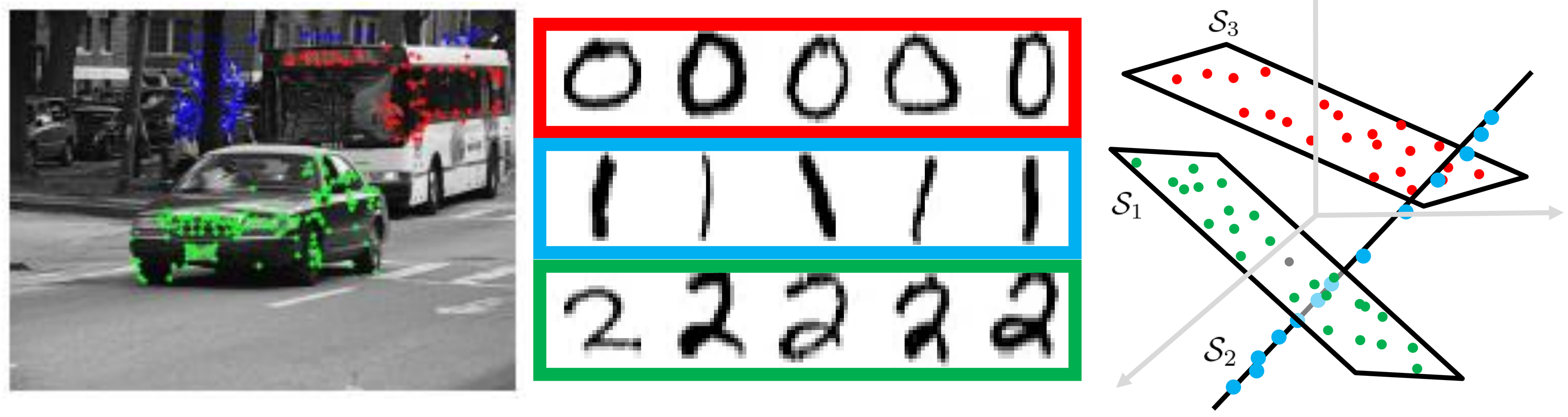
Daniel P. Robinson — Applied Mathematics and Statistics, Johns Hopkins University, Baltimore

René Vidal — Mathematical Institute for Data Science, Johns Hopkins University, Baltimore

ICCV 2019 Seoul, Korea

Introduction

- Vision datasets often contain multiple classes, each lying in a low-dimensional subspace
- In many cases, the subspaces do not pass through the origin, i.e., they are affine
- Affine subspace clustering (ASC): discover affine subspaces in an unsupervised manner

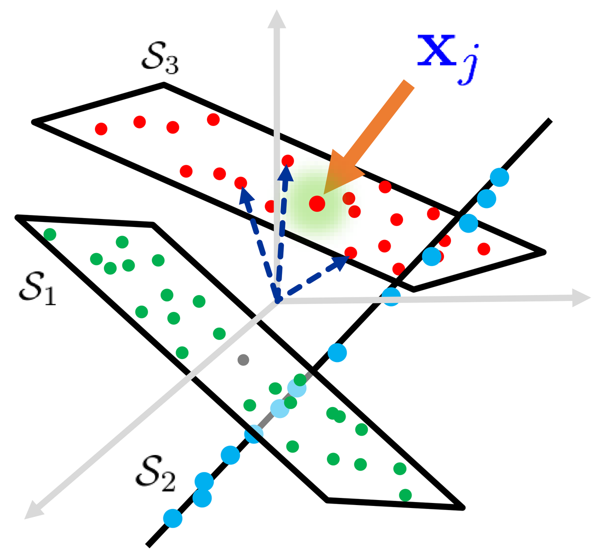


Prior Work

- Data from multiple affine subspaces is self-expressive, i.e., $\mathbf{x}_j = X\mathbf{c}_j$, $X = [\mathbf{x}_1, \cdots, \mathbf{x}_N]$
  - find the self-expression $\mathbf{c}_j$ via solving

$$\min_{\mathbf{c}_j} f(\mathbf{c}_j) \;\; \text{s.t.} \;\; \mathbf{x}_j = X\mathbf{c}_j,\; c_{jj} = 0,\; \mathbf{c}_j \in \mathcal{C} \quad (1)$$

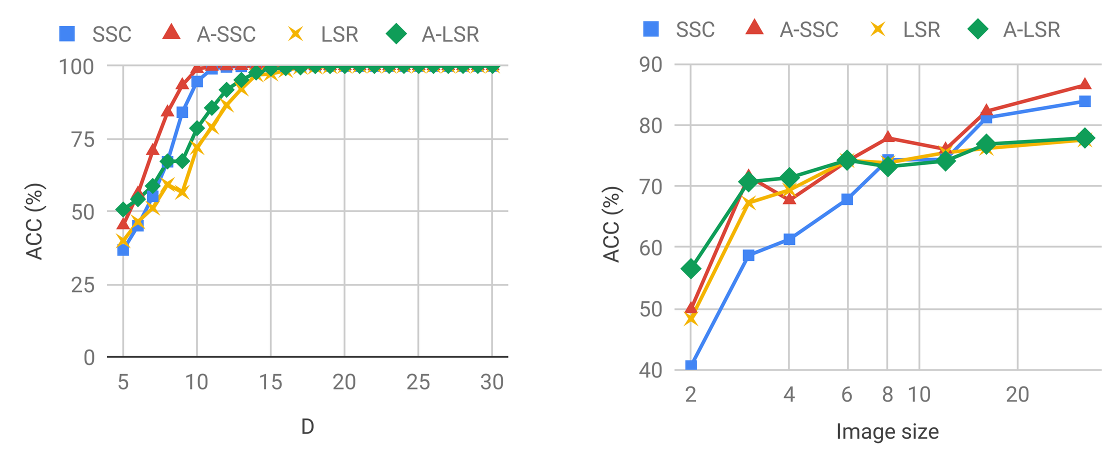
  - apply spectral clustering to affinity $|c_{ij}| + |c_{ji}|$



- ASC without an affine constraint: $\mathcal{C} = \mathbb{R}^N$
  - ✗ No explicit modeling of the affine structure
  - ✓ More often applied, many scalable algorithms
- ASC with an affine constraint: $\mathcal{C} = \{\mathbf{1}^\top \mathbf{c}_j = 1\}$
  - ✓ With explicit modeling of the affine structure
  - ✗ Rarely applied, no scalable algorithms (noisy case)

Contributions

- When the ambient dimension is high enough, the affine constraint is not needed
  - Randomly generated subspaces: both ASC with/without an affine constraint are guaranteed to produce correct affinity
  - Computer vision datasets: difference in performance between ASC with/without an affine constraint is small or negligible

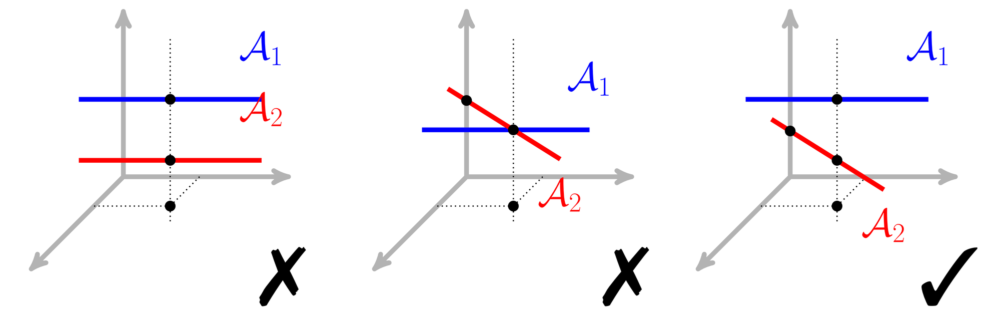


Varying ambient dimension for randomly generated subspaces (left) and for the Coil-100 image database (right)

Theoretical Analysis for Affine Subspace Clustering

Preliminaries

- **Definition:** A function $f : \mathbb{R}^{N\times N} \to \mathbb{R}$ is said to satisfy the Enforced Block Diagonal (EBD) conditions if $f(\mathbf{C}) = f(\mathbf{P}^\top \mathbf{C}\mathbf{P})$ for any permutation $\mathbf{P}$ and $f(\mathbf{C}) \geq f(\mathbf{C}_0)$ for any $\mathbf{C}_0$ that contains only the diagonal blocks of $\mathbf{C}$
  - The EBD conditions are satisfied for $f(\cdot) = \|\cdot\|_1$, $\|\cdot\|_F^2$, $\|\cdot\|_*$, and so on
- **Definition:** A collection of affine subspaces $\{\mathcal{A}_\ell\}_{\ell=1}^n$ is said to be affinely independent if $\dim(\mathrm{aff}(\cup_{\ell=1}^n \mathcal{A}_\ell)) + 1 = \sum_{\ell=1}^n \dim(\mathcal{A}_\ell) + n$



The collection of two lines (i.e., 1D affine subspaces) in $\mathbb{R}^3$ is affinely independent if they are skew lines

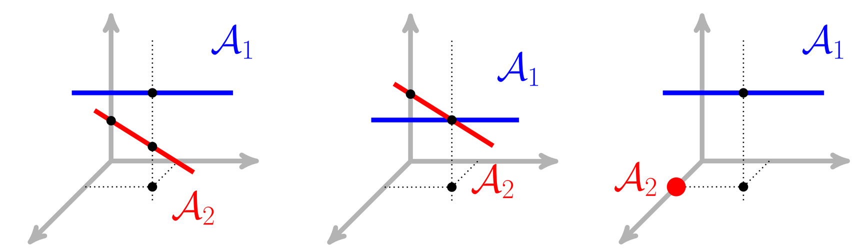
- **Definition:** A collection of affine subspaces $\{\mathcal{A}_\ell \subseteq \mathbb{R}^D\}_{\ell=1}^n$ is said to be drawn from the random affine subspace model if they are drawn independently and uniformly from the space of affine $\{d_\ell\}$-dimensional subspaces of $\mathbb{R}^D$

Geometric Conditions

Given data $\{\mathbf{x}_j\}_{j=1}^N$ drawn from $\{\mathcal{A}_\ell\}_{\ell=1}^n$, assume that f satisfies the EBD conditions.

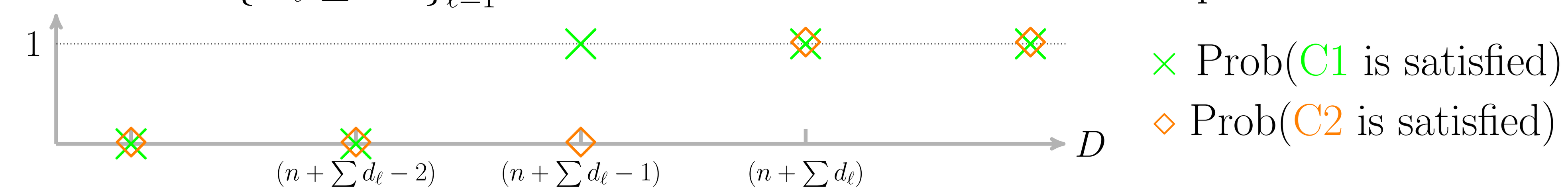
Theorem: (ASC without affine constraint) The solution to (1) gives a correct affinity if
- C1: $\{\mathcal{A}_\ell\}_{\ell=1}^n$ is affinely independent, and
- C2: $\mathbf{0} \notin \mathrm{aff}(\cup_{\ell=1}^n \mathcal{A}_\ell)$

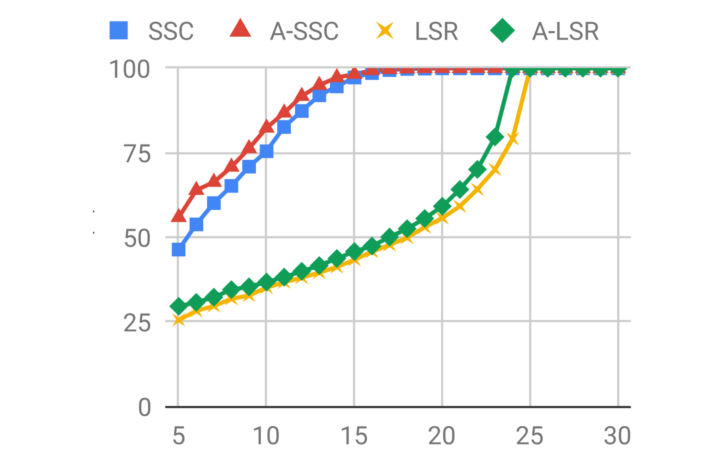
Theorem: (ASC with affine constraint) The solution to (1) gives a correct affinity if
- C1: $\{\mathcal{A}_\ell\}_{\ell=1}^n$ is affinely independent



Conditions Under Random Affine Subspace Model

Theorem: Let $\{\mathcal{A}_\ell \subseteq \mathbb{R}^D\}_{\ell=1}^n$ be drawn from the random affine subspace model.





Percentage of correct affinities (y-axis) vs. D (x-axis) for $n = 5$ and $d_\ell = 4$

Conclusion: For affine subspaces drawn from the random model, both ASC with/without an affine constraint produce correct affinities with probability 1 if $D \geq n + \sum_{\ell=1}^n d_\ell$

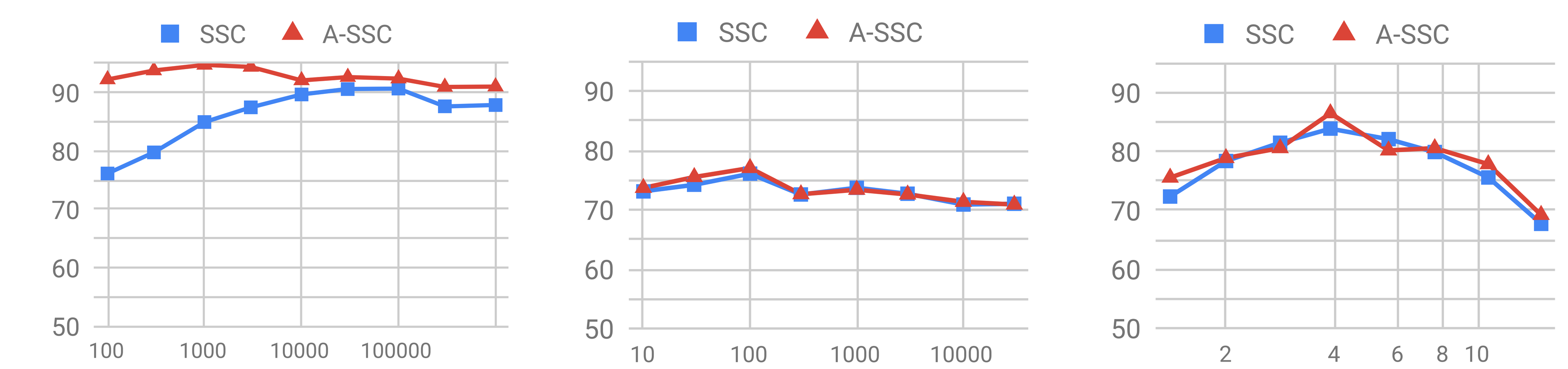
Empirical Evaluation for Affine Subspace Clustering

- We study the following methods

Affine constraint?	$f(\cdot) = \|\cdot\|_1$	$f(\cdot) = \|\cdot\|_F^2$
Without	SSC	LSR
With	A-SSC	A-LSR

- We test on the following real datasets

	Hopkins 155	MNIST	Coil-100
Data type	Motion	Digit image	Object image
Avg. D	57	3,472	1,024



Clustering accuracy (y-axis) vs. model parameter (x-axis) on Hopkins 155, MNIST and Coil-100

Conclusion: For real applications, difference between ASC with/without an affine constraint is small for high-dimensional data

[1] C.-Y. Lu, H. Min, Z.-Q. Zhao, L. Zhu, D.-S. Huang, S. Yan, Robust and Efficient Subspace Segmentation via Least Squares Regression, *European Conference on Computer Vision*, 2012.

[2] C.-G. Li, C. You, R. Vidal On Geometric Analysis of Affine Sparse Subspace Clustering, *IEEE Journal on Selected Topics in Signal Processing*, 2018.

[3] E. Elhamifar, R. Vidal, Sparse Subspace Clustering, *IEEE Conference on Computer Vision and Pattern Recognition*, 2009.

This work was supported by NSF under grant #1618637 and #1704458, Northrop Grumman Corporation's REALM Program, NSFC under grant #61876022 and MOE at Peking University's Open Project Fund.