Subspace Clustering Reloaded: Sparse vs. Dense Representations

Working Paper

Author(s):
Dyler, Eva L.; Studer, Christoph; Baraniuk, Richard G.

Publication date:
2013

Permanent link:
https://doi.org/10.3929/ethz-b-000460355

Subspace Clustering Reloaded: Sparse vs. Dense Representations

Eva L. Dyer, Christoph Studer, and Richard G. Baraniuk

Rice University, Houston, TX, USA; e-mail: {e.dyer, studer, richb}@rice.edu

***Abstract*—State-of-the-art methods for learning unions of subspaces from a collection of data leverage sparsity to form representations of each vector in the dataset with respect to the remaining vectors in the dataset. The resulting sparse representations can be used to form a subspace affinity matrix to cluster the data into their respective subspaces. While sparsity-driven methods for subspace clustering provide advantages over traditional nearest neighbor-based approaches, sparse representations often produce affinity matrices with weakly connected components that are difficult to cluster, even with state-of-the-art clustering methods. In this work, we propose a new algorithm that employs dense (least-squares) representations to extract the subspace affinity between vectors in the dataset. We demonstrate the advantages of the proposed dense subspace clustering algorithm over state-of-the-art sparsity-driven methods on real and synthetic data.**

I. Sparse Subspace Clustering

Unions of linear subspaces are a widely used signal model for representing collections of high-dimensional data, such as images of faces acquired under varying illumination conditions, motion trajectories from different objects, or local field potentials in brain-machine-interface applications [1], [2]. In order to use this signal model, the subspaces that the collection of data live upon must be learned from the data by performing *subspace clustering*—learning subspaces present in the data and clustering the data based upon the subspace membership of each vector. Subspace clustering is challenging due to the fact that subspace estimation and segmentation must be performed simultaneously.

Sparse subspace clustering (SSC) has been shown to yield state-of-the-art performance on both synthetic and numerous real-world image datasets [1]. In SSC, a sparse representation of a signal in a collection of data is formed with respect to the remaining signals in the same dataset. The idea underlying this approach is that signals from the same subspace cluster will use one another in their sparse representations, thus revealing which points belong to the same subspace. SSC provides a powerful alternative to nearest neighbor (NN)-based approaches to subspace clustering and enables the derivation of guarantees that describe when convex methods [1], [3] or greedy algorithms [2] yield representations that only contain points from the same subspace.

Formally, SSC computes sparse representations $\mathbf{c}_i \in \mathbb{R}^d$ for each point $\mathbf{y}_i \in \mathbb{R}^n$, $i = 1, \ldots, d$, via the following optimization problem:

$$\mathbf{c}_i = \operatorname{argmin}_{\tilde{\mathbf{c}} \in \mathbb{R}^d} \|\tilde{\mathbf{c}}\|_1 \quad \text{subject to} \quad \mathbf{y}_i = \textstyle\sum_{j\,:\,j \neq i} \mathbf{y}_j \tilde{c}_j, \qquad (1)$$

which are used to construct an *affinity matrix* $\mathbf{C}^T = [\,\mathbf{c}_1 \cdots \mathbf{c}_d\,]$. Spectral clustering is then performed on the graph Laplacian of $\mathbf{W} = |\mathbf{C}| + |\mathbf{C}^T|$ to segment the data into subspace clusters [2].

II. Dense Subspace Clustering

While sparse representations result in affinity matrices that contain a small number of edges in the graph linking signals from *different* subspaces, recovering subspace clusters from the affinity matrices obtained via SSC is challenging due to the fact that sparse representations often produce weakly connected components between signals in the *same* subspace. To circumvent this issue, we propose a novel method for subspace clustering that is based on forming *dense* least-squares representations[1] from the data. This method, which we refer to as *dense subspace clustering* (DSC), produces affinity matrices that have more tightly connected components than those obtained via SSC. For this reason, spectral clustering algorithms operating on the affinity matrices obtained via DSC can recover the subspace clusters more reliably than SSC.

Specifically, we compute minimum ℓ_2-norm representations of each vector in the dataset via the pseudo-inverse $\bar{\mathbf{c}}_i = \mathbf{Y}_{(i)}^{\dagger}\mathbf{y}_i$. After forming dense representations for all d vectors in the dataset, we form an affinity matrix $\mathbf{W}$ according to $W_{k,\ell} = |\langle \bar{\mathbf{c}}_k, \bar{\mathbf{c}}_\ell \rangle| / (\|\bar{\mathbf{c}}_k\|_2 \|\bar{\mathbf{c}}_\ell\|_2)$ and perform spectral clustering on the graph Laplacian of $\mathbf{W}$. In the noisy case, we compute a thresholded version of the pseudoinverse, where small singular values are set to zero.

[1]We replace the ℓ_1 norm in (1) with the ℓ_2 norm resulting in (typically dense) least-squares representations of the data.

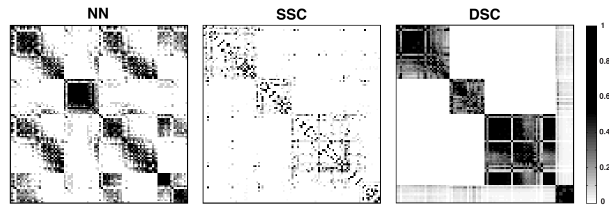

Fig. 1. *Affinity matrices from illumination subspaces.* Left: NN; middle: SSC via OMP; right: DSC. Classification errors for NN: $P_m = 34.4\%$, $P_a = 20.9\%$; SSC: $P_m = 8.8\%$, $P_a = 1\%$; DSC: $P_m = 5.6\%$, $P_a = 0.3\%$, where P_m is the miss rate (points that are not included in their correct subspace clusters) and P_a is the false alarm rate (points that are incorrectly included in a subspace cluster). DSC clearly outperforms NN and SSC.

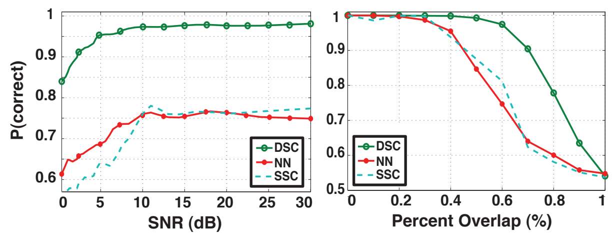

Fig. 2. *Probability of correct classification from noisy subspaces.* Left: 10-dimensional subspace intersection in $\mathbb{R}^{50}$ for varying SNR; right: 15 dB SNR with varying subspace intersection. DSC clearly outperforms NN and SSC.

III. Comparisons and Conclusions

In Figure 1, we display the affinity matrices computed from a set of images collected from four different faces under various illumination conditions; here, we compare nearest neighbor (NN) selection, SSC via orthogonal matching pursuit (OMP), and DSC as proposed here. In Figure 2, we test the classification performance of NN, OMP-based SSC, and DSC on noisy synthetic data. For these experiments, we generate vectors living on a pair of noisy and intersecting 10-dimensional subspaces in $\mathbb{R}^{50}$. Both results suggest that dense representations provide significant advantages over existing sparsity-based subspace clustering methods. In particular, the proposed DSC method requires low computational complexity and outperforms state-of-the-art subspace clustering methods.

References

[1] E. Elhamifar and R. Vidal, "Clustering disjoint subspaces via sparse representation," in *Proc. ICASSP*, Mar. 2010, pp. 1926–1929.

[2] E. L. Dyer, A. C. Sankaranarayanan, and R. G. Baraniuk, "Greedy feature selection for subspace clustering," *Preprint*, Apr. 2012.

[3] M. Soltanolkotabi and E. J. Candès, "A geometric analysis of subspace clustering with outliers," *arXiv:1112.4258v2 [cs.IT]*, Jan. 2012.