Fast Subspace Clustering Based on the Kronecker Product

Lei Zhou, Xiao Bai, Liang Zhang
School of Computer Science and Engineering
Beihang University, China

Jun Zhou
School of Information and Communication Technology
Griffith University, Australia

Edwin Hancock
Department of Computer Science
University of York, York, U.K.

Outline

- Introduction
- Traditional subspace clustering methods
- The proposed method
- Experimental results
- Conclusion

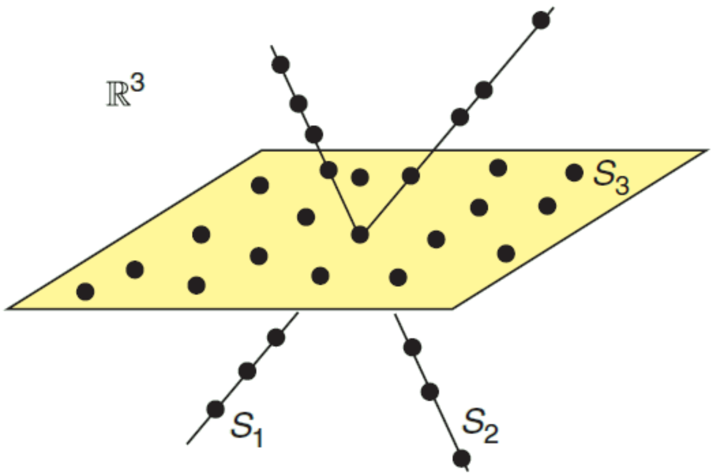
Outline

- Introduction
- Traditional subspace clustering methods
- The proposed method
- Experimental results
- Conclusion

- For many high dimensional data, their intrinsic dimension is often much smaller than the dimension of the ambient space.

- **(Subspace Clustering)**

Let $X = [x_1, x_2, \dots, x_N] \in R^{D \times N}$ be a given set of points drawn from k linear or affine subspaces $\{S_i\}_{i=1}^{k}$. The goal of subspace clustering is to find the segmentation of the points according to the subspaces.



Outline

- Introduction
- Traditional subspace clustering methods
- The proposed method
- Experimental results
- Conclusion

Self-representation based Subspace Clustering

These methods are based on the self-expressiveness property of data lying in a union of subspaces, which states that each point in a union of subspaces can be written as a linear combination of other data points in the subspaces.

$$x_i = \sum_{j \neq i} Z_{ij} x_j$$

where $Z_{ij} = 0$ if the i-th and j-th data points are from different subspaces.

Therefore, in the case of k subspaces, Z has the following block diagonal form through some transformations

$$Z = \begin{bmatrix} Z_1 & 0 & \cdots & 0 \\ 0 & Z_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & Z_k \end{bmatrix}$$

Sparse Subspace Clustering (SSC)

SSC calculates the similarity among data points by solving the following optimization problem:

$$\min_{Z} \ \|Z\|_1$$
$$\text{s.t.} \ X = XZ, \ Z_{ii} = 0$$

Low-Rank Representation (LRR)

LRR uses the lowest rank representation rather than the sparsest representation to build the similarity graph. The objective function of LRR is:

$$\min_{Z} \ \|Z\|_*$$
$$\text{s.t.} \ X = XZ$$

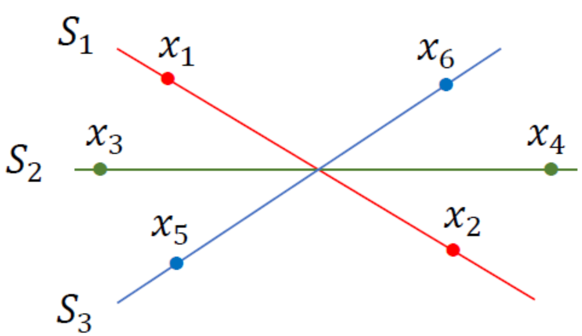
Sparse subspace clustering. CVPR, 2009
Robust Subspace Segmentation by Low-Rank Representation. ICML, 2010

Outline

- Introduction
- Traditional subspace clustering methods
- The proposed method
- Experimental results
- Conclusion

Motivation

Kronecker product: $$A \otimes B = \begin{bmatrix} a_{11} \times B & \cdots & a_{1n} \times B \\ \vdots & \ddots & \vdots \\ a_{m1} \times B & \cdots & a_{mn} \times B \end{bmatrix}$$



$$X = [x_1, x_2, x_3, x_4, x_5, x_6]$$

Conventional sparse subspace clustering:

$$X = XC \longrightarrow C = \left(\begin{array}{cc|cc|cc} 0 & -1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & -1 & 0 \end{array}\right)$$

Kronecker product based model:

$$X = X(C_1 \otimes C_2) \longrightarrow C_1 \otimes C_2 = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

北京航空航天大學
BEIHANG UNIVERSITY

- **Formulation**:

$$\min_{C_i} \|X - X(C_1 \otimes C_2)\|_F^2 + \lambda \|C_1 \otimes C_2\|_F^2$$

- **Optimization**:

$$\begin{aligned}
&\|X - X(C_1 \otimes C_2)\|_F^2 \\
=&tr((X - X(C_1 \otimes C_2))^T (X - X(C_1 \otimes C_2))) \\
=&\|X\|_F^2 - 2tr(X(C_1 \otimes C_2)X^T) \\
&+ tr(X(C_1 \otimes C_2)(X(C_1 \otimes C_2))^T)
\end{aligned}$$

$$\Phi = -2tr(X(C_1 \otimes C_2)X^T) + tr(X(C_1 \otimes C_2)(X(C_1 \otimes C_2))^T)$$

Algorithm:

Algorithm 1: Subspace Clustering Based on Kronecker Product.

Input: A set of data points $X = \{x_i\}_{i=1}^N$, the number of subspaces n, the number of small matrices k and the balance parameter λ.

Steps:

1. Learn the small matrices $C_1, C_2, \cdots, C_k$.

for $i = 1, ..., k$ **do**

Fix $C_1, \cdots, C_{i-1}, C_{i+1}, \cdots, C_k$, update C_i.
Optimize Eq. (8), estimate C_i by ridge regression solution.

end

2. Calculate the self-representation coefficient matrix C by the Kronecker product of small matrices, $C = \otimes_{i=1}^k C_i$.
3. Construct an affinity matrix by $W = |C| + |C|^T$.
4. Calculate the Laplacian matrix L of W.
5. Calculate the eigenvector matrix V of L corresponding to its n smallest nonzero eigenvalues.
6. Perform k-means clustering algorithm on the rows of V.

Output: The clustering result of X.

Outline

- Introduction
- Traditional subspace clustering methods
- The proposed method
- Experimental results
- Conclusion

The subspace clustering performance on the CMU PIE dataset.

No. Objects	5 Objects		10 Objects		20 Objects		40 Objects		60 Objects	
	Time	Acc.	Time	Acc.	Time	Acc.	Time	Acc.	Time	Acc.
SSC	243.6	92.47	1182	89.25	3618	84.31	14502	82.37	-	-
KrSSC	**12.7**	91.28	**26.8**	88.27	**61.4**	83.86	**150.2**	81.75	**274.3**	79.48
LRR	216.4	94.53	852.5	92.14	2743	89.21	11463	85.47	-	-
KrLRR	**9.7**	92.51	**20.4**	90.72	**57.2**	88.13	**145.8**	85.21	**254.8**	83.65
TRR	152.7	97.35	548.2	96.05	2167	94.54	8427	91.74	-	-
KrTRR	**7.5**	95.21	**18.3**	94.52	**52.8**	93.84	**143.5**	90.23	**260.1**	87.26
NVR3	190.5	98.51	624.6	97.51	2536	95.75	11826	93.15	-	-
KrNVR3	**11.3**	97.14	**25.7**	96.26	**72.4**	93.96	**180.4**	91.57	**312.5**	89.15

The subspace clustering performance on the MNIST dataset.

No. Points	500		1000		10000		30000		70000	
	Time	Acc.	Time	Acc.	Time	Acc.	Time	Acc.	Time	Acc.
SSC	152.4	83.36	638.2	82.45	-	-	-	-	-	-
KrSSC	**7.3**	81.25	**18.7**	81.17	**192.4**	79.42	**411.5**	76.15	**683.2**	73.34
LRR	145.5	85.75	614.8	85.14	-	-	-	-	-	-
KrLRR	**7.1**	83.24	**16.4**	83.20	**160.8**	81.52	**384.5**	79.21	**641.5**	76.53
TRR	113.2	90.28	476.4	89.78	-	-	-	-	-	-
KrTRR	**6.5**	88.95	**15.8**	88.65	**168.2**	85.76	**403.8**	83.26	**795.6**	81.53
NVR3	118.5	91.85	531.1	91.28	-	-	-	-	-	-
KrNVR3	**8.3**	90.08	**22.5**	90.14	**243.6**	86.27	**627.5**	83.87	**968.4**	82.41

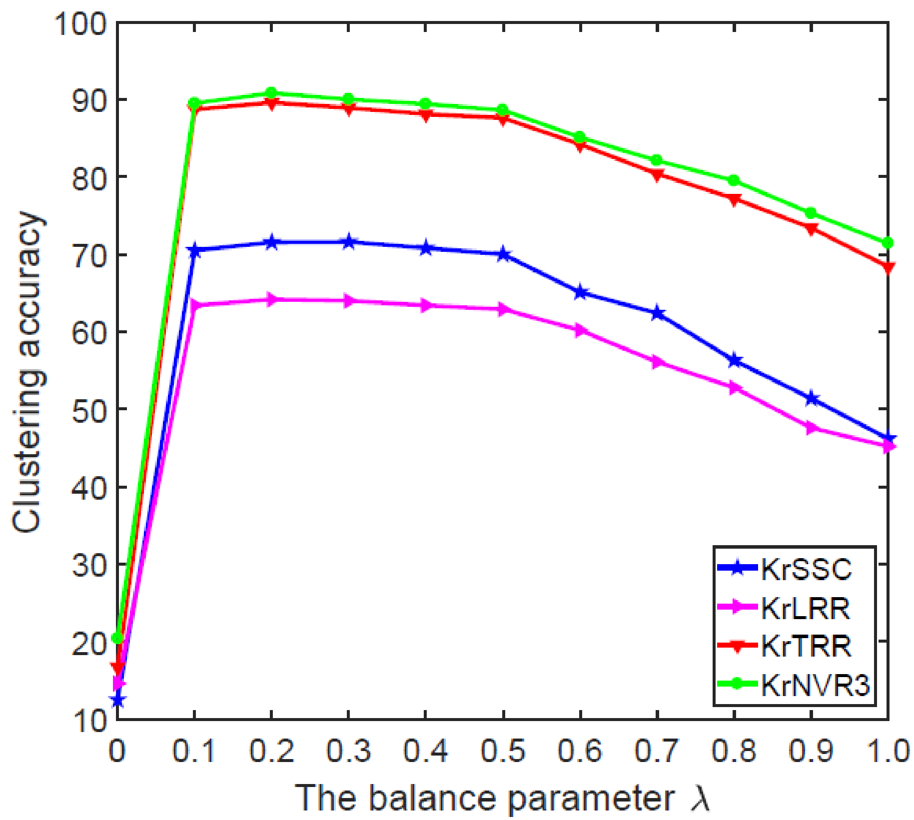
北京航空航天大學 BEIHANG UNIVERSITY

The subspace clustering performance on the synthetic dataset.

No. Points	500		5000		10000		50000		100000	
	Time	Acc.	Time	Acc.	Time	Acc.	Time	Acc.	Time	Acc.
SSC	135.4	94.15	1824	93.86	5413	91.05	-	-	-	-
KrSSC	**6.2**	92.12	**53.4**	91.18	**164.2**	89.73	**231.5**	85.04	**285.7**	81.85
LRR	118.6	95.27	1645	94.57	4853	92.14	-	-	-	-
KrLRR	**6.0**	93.24	**49.3**	92.21	**152.7**	89.49	**216.2**	86.03	**274.3**	82.20
TRR	89.5	98.85	1627	97.15	5825	95.69	-	-	-	-
KrTRR	**5.9**	98.06	**46.7**	96.53	**185.3**	95.05	**250.3**	93.16	**314.2**	89.06
NVR3	96.4	99.91	1752	98.61	6024	97.10	-	-	-	-
KrNVR3	**6.0**	99.07	**52.8**	98.11	**207.5**	96.24	**260.1**	93.89	**321.5**	90.62

The average running time and clustering accuracy of our methods with different k.

k	2	3	4	5
average running time (seconds):				
KrSSC	715.6	285.7	61.2	25.4
KrLRR	682.5	274.3	52.7	20.6
KrTRR	755.1	314.2	84.3	31.5
KrNVR3	794.3	321.5	91.6	36.2
average clustering accuracy:				
KrSSC	83.14	81.85	75.42	67.25
KrLRR	84.43	82.20	77.16	68.17
KrTRR	90.75	89.06	84.27	73.41
KrNVR3	92.54	90.62	85.34	75.24



The average clustering accuracy with different balance parameter λ.

Outline

- **Introduction**
- **Traditional subspace clustering methods**
- **The proposed method**
- **Experimental results**
- **Conclusion**

Conclusions

- **Main Contributions**

  – We have presented a fast subspace clustering model based on the Kronecker product. We learn the representation matrix of spectral clustering using the Kronecker product of a set of smaller matrices.

  – The memory space and computational complexity of our methods achieve significant efficiency gain compared with several baseline approaches.

  – Moreover, we have presented results on synthetic data which has verified the scalability of our methods on large scale datasets.

Thanks for watching！