Necessary and Sufficient Conditions for Sketched Subspace Clustering

Daniel Pimentel-Alarcón, Laura Balzano & Robert Nowak

University of Wisconsin-Madison

Allerton 2016
Subspace Clustering
Subspace Clustering
Subspace Clustering

We are given: Columns in a union of subspaces.
Subspace Clustering

We are given: Columns in a union of subspaces.
Goal: Cluster the columns, or find the subspaces.
**We are given**: Columns in a union of subspaces.

**Goal**: Cluster the columns, or find the subspaces.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>4</th>
<th>1</th>
<th>3</th>
<th>3</th>
<th>1</th>
<th>2</th>
<th>1</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>9</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>8</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>8</td>
<td>2</td>
<td>6</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>8</td>
<td>3</td>
<td>9</td>
<td>6</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>6</td>
<td>2</td>
</tr>
</tbody>
</table>

(predictions)

**Sketched Subspace Clustering**

Projections of $S_1^*$ and $S_2^*$
Sketched Subspace Clustering

We are given: Columns in a union of subspaces.
Goal: Cluster the columns, or find the subspaces.
Applications
Applications
Applications
Example: Network Topology Estimation
Example: Network Topology Estimation
Example: Network Topology Estimation
Example: Network Topology Estimation
Complication

Two subspaces (even orthogonal) can appear identical if they are only observed on a subset of coordinates.
Complication

Two subspaces (even orthogonal) can appear identical if they are only observed on a subset of coordinates.
Complication

Two subspaces (even orthogonal) can appear identical if they are only observed on a subset of coordinates.
Complication

Two subspaces (even orthogonal) can appear identical if they are only observed on a subset of coordinates.
Complication

Two subspaces (even orthogonal) can appear identical if they are only observed on a subset of coordinates.
Fortunately
Not all subsets of coordinates are bad
Fortunately
Not all subsets of coordinates are bad
If we pick *the right* subsets of coordinates, we will be fine.

**The catch:** How do we know which are *the right* subsets?
First thing to ask

How many subsets of coordinates are good?
(This depends on the subspaces)
New measure of similarity

# of subsets of $r+1$ coordinates where two subspaces differ
New measure of similarity

# of subsets of $r+1$ coordinates where two subspaces differ
Definition 1. Given $S, S' \in \text{Gr}(r, \mathbb{R}^d)$, define the partial coordinate discrepancy between $S$ and $S'$ as:

$$\delta(S, S') := \frac{1}{\binom{d}{r+1}} \sum_{\omega \in [d]^{r+1}} \mathbb{1}\{S_\omega \neq S'_\omega\}.$$
Definition 1. Given $S, S' \in \text{Gr}(r, \mathbb{R}^d)$, define the partial coordinate discrepancy between $S$ and $S'$ as:

$$\delta(S, S') := \frac{1}{\binom{d}{r+1}} \sum_{\omega \in [d]^{r+1}} \mathbb{1}_{\{S_\omega \neq S'_\omega\}}.$$
Definition 1. Given $S, S' \in \text{Gr}(r, \mathbb{R}^d)$, define the partial coordinate discrepancy between $S$ and $S'$ as:

$$
\delta(S, S') := \frac{1}{\binom{d}{r+1}} \sum_{\omega \in [d]^{r+1}} \mathbb{1}\{S_\omega \neq S'_\omega\}.
$$

New measure of similarity

# of subsets of $r+1$ coordinates where two subspaces differ
Definition 1. Given $S, S' \in \text{Gr}(r, \mathbb{R}^d)$, define the partial coordinate discrepancy between $S$ and $S'$ as:

$$\delta(S, S') := \frac{1}{\binom{d}{r+1}} \sum_{\omega \in [d]^{r+1}} 1\{S_\omega \neq S'_\omega\}.$$
New measure of similarity

# of subsets of \( r+1 \) coordinates where two subspaces differ

**Definition 1.** Given \( S, S' \in \text{Gr}(r, \mathbb{R}^d) \), define the partial coordinate discrepancy between \( S \) and \( S' \) as:

\[
\delta(S, S') := \frac{1}{\binom{d}{r+1}} \sum_{\omega \in [d]^{r+1}} \mathbb{1}_{\{S_\omega \neq S'_\omega\}}.
\]
Example

\[
S = \begin{bmatrix}
1 \\
1 \\
1 \\
1
\end{bmatrix} \quad S' = \begin{bmatrix}
1 \\
1 \\
-1 \\
-1
\end{bmatrix}
\]
Example

\[
\begin{bmatrix}
S \\
1 \\
1 \\
1 \\
1
\end{bmatrix}
\begin{bmatrix}
S' \\
1 \\
1 \\
-1 \\
-1
\end{bmatrix}
\]

\[
\delta(S, S') = \frac{4}{6}
\]
Example

\[ S = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad S' = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} \]

\[ \delta(S, S') = \frac{4}{6} \]

# of good combinations of \( r+1 \) coordinates
Example

$$\begin{bmatrix}
1 & 1 \\
1 & -1 \\
1 & -1 \\
\end{bmatrix}$$

$$\begin{bmatrix}
S \\
S' \\
\end{bmatrix}$$

$$\delta(S, S') = \frac{4}{6}$$

# of good combinations of r+1 coordinates
Example

\[
\begin{bmatrix}
S \\
1 \\
1 \\
1 \\
1
\end{bmatrix}
\begin{bmatrix}
S' \\
1 \\
1 \\
-1 \\
-1
\end{bmatrix}
\]

\[\delta(S, S') = \frac{4}{6}\]

# of good combinations of \(r+1\) coordinates
Example

\[
\delta(S, S') = \frac{4}{6}
\]

# of good combinations of \(r+1\) coordinates

# of total combinations of \(r+1\) coordinates
Example

\[
\begin{bmatrix}
1 & 1 \\
1 & 1 \\
1 & -1 \\
1 & -1 \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 1 \\
1 & 1 \\
1 & -1 \\
1 & -1 \\
\end{bmatrix}
\]

\[\delta(S, S') = \frac{4}{6}\]

# of good combinations of \(r+1\) coordinates

# of total combinations of \(r+1\) coordinates

Probability that 2 subspaces are different on \(r+1\) coordinates chosen randomly
What does this mean?

Depending on the subspaces, there may be way too many bad subsets!
What does this mean?

Depending on the subspaces, there may be way too many bad subsets!

Lucky break!

If we rotate subspaces randomly, all subsets will be good!
What does this mean?

Depending on the subspaces, there may be way too many bad subsets!

Lucky break!

If we rotate subspaces randomly, all subsets will be good!
Lucky break!
If we rotate subspaces randomly, all subsets will be good!
Lucky break!
If we rotate subspaces randomly, all subsets will be good!
What does this mean?

We can rotate, subsample and cluster
What does this mean?

We can rotate, subsample and cluster
What does this mean?

We can rotate, subsample and cluster
What does this mean?

We can rotate, subsample and cluster
A little more about $\delta$
A little more about \( \delta \)

Long story short: none implies the other.
A little more about $\delta$

Long story short: none implies the other.
What’s next?
δ is an *all or nothing* metric.

What’s next?
δ is an all or nothing metric.

\[
\delta(S, S') := \frac{1}{\binom{d}{r+1}} \sum_{\omega \in [d]^{r+1}} \mathbb{1}\{S_\omega \neq S'_\omega\}.
\]

kind of an \( \ell_0 \) norm.

What’s next?
• \( \delta \) is an *all or nothing* metric.

\[
\delta(S, S') := \frac{1}{d \choose r+1} \sum_{\omega \in [d]^{r+1}} \mathbb{1}_{\{S_\omega \neq S'_\omega\}}.
\]

kind of an \( \ell_0 \) norm.

• Can we come up with more practical metrics? kind of an \( \ell_1 \) norm.

What’s next?
Thank you.