STUDY clustering overview

STEP 1
Build a STUDY

STEP 2
Precompute the data

STEP 3
Precluster the data

STEP 4
Cluster the data
Memory options should change when using STUDY vs single dataset
Build a STUDY
Build a STUDY, cont'd

Create a new STUDY set -- pop_study()

<table>
<thead>
<tr>
<th>dataset filename</th>
<th>browse</th>
<th>subject</th>
<th>session</th>
<th>condition</th>
<th>group</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>2</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>3</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>4</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>5</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>6</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>7</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>8</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>9</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>10</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Important note: Removed datasets will not be saved before being deleted from EEGLAB memory.

Update dataset info - datasets stored on disk will be overwritten (unset = Keep study info set).

Delete cluster information (to allow loading new datasets, set new components for clustering).

Choose dataset to add to STUDY -- pop_study()

Look in: S01

<table>
<thead>
<tr>
<th>Name</th>
<th>Date modified</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ignore.set</td>
<td>11/8/2009 7:06 PM</td>
<td>SET File</td>
</tr>
<tr>
<td>Memorize.set</td>
<td>11/8/2009 7:06 PM</td>
<td>SET File</td>
</tr>
<tr>
<td>Probe.set</td>
<td>11/12/2009 10:02 PM</td>
<td>SET File</td>
</tr>
</tbody>
</table>
Edit dataset info

**pop_study(): Pre-select components**

Enter maximum residual (topo map - dipole proj.) var. (in %)
NOTE: This will delete any existing component clusters!

- [ ] Keep only in-brain dipoles.

**Cancel**  **Help**  **Ok**

**Create a new STUDY set -- pop_study()**

**Edit STUDY set information - remember to save changes**

- STUDY set name: [ ]
  - [ ] Sternberg
- STUDY set task name: [ ]
  - [ ] Sternberg
- STUDY set notes:

**dataset filename**

<table>
<thead>
<tr>
<th></th>
<th>browse</th>
<th>subject</th>
<th>session</th>
<th>condition</th>
<th>group</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>...</td>
<td>S01</td>
<td></td>
<td>memorize</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>...</td>
<td>S01</td>
<td></td>
<td>ignore</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>...</td>
<td>S01</td>
<td></td>
<td>probe</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>...</td>
<td>S02</td>
<td></td>
<td>memorize</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>...</td>
<td>S02</td>
<td></td>
<td>ignore</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>...</td>
<td>S03</td>
<td></td>
<td>probe</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>...</td>
<td>S03</td>
<td></td>
<td>memorize</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>...</td>
<td>S03</td>
<td></td>
<td>ignore</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>...</td>
<td>S04</td>
<td></td>
<td>probe</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>...</td>
<td>S04</td>
<td></td>
<td>memorize</td>
<td></td>
</tr>
</tbody>
</table>

**Important note:** Removed datasets will not be saved before being deleted from EEGLAB memory

- [ ] Dataset info (condition, group, ...) differs from study info. [set] = Overwrite dataset info.
- [ ] Delete cluster information (to allow loading new datasets, set new components for clustering, etc.)

**Help**  **Cancel**  **Ok**
ICs to cluster
STUDY clustering overview

STEP 1
Build a STUDY

STEP 2
Precompute the data

STEP 3
Precluster the data

STEP 4
Cluster the data
Precompute data measures
Precompute data measures

TIP: Compute all measures so you can test different combinations for clustering

Recommend: 'alpha', .01 (time-consuming)
STUDY clustering overview

STEP 1
Build a STUDY

STEP 2
Precompute the data

STEP 3
Precluster the data

STEP 4
Cluster the data
Precluster the data
Precluster the data

Select and compute component measures for later clustering -- pop_precl...

Build pre-clustering matrix for STUDY set: Sternberg
Select the cluster to refine by sub-clustering (any existing sub-hierarchy will be overwritten)

ParentCluster 1 (336 ICs)

Note: Only measures that have been precomputed may be used for clustering.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>spectra</td>
<td>10</td>
<td>√</td>
<td>1</td>
</tr>
<tr>
<td>ERPs</td>
<td>10</td>
<td>√</td>
<td>1</td>
</tr>
<tr>
<td>dipoles</td>
<td>3</td>
<td>√</td>
<td>10</td>
</tr>
<tr>
<td>scalp maps</td>
<td>10</td>
<td>√</td>
<td>1</td>
</tr>
<tr>
<td>ERSPs</td>
<td>10</td>
<td>√</td>
<td>1</td>
</tr>
<tr>
<td>ITCs</td>
<td>10</td>
<td>√</td>
<td>1</td>
</tr>
<tr>
<td>Final dimensions</td>
<td>10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measures</th>
<th>Freq. range [Hz]</th>
<th>3.25</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time range [ms]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measures</th>
<th>Use channel values</th>
<th>Absolute values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time range [ms]</td>
<td>0.600</td>
<td>3.30</td>
</tr>
<tr>
<td>Freq. range [Hz]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Choosing data measures

What measure(s) should you use?

• It depends on your final cluster criteria…
  - If for example, your priority is dipole location,
    then cluster only based on dipole location…

But consider:

  - What is the difference between these two components?
Choosing data measures

Similar dipole location, very different orientation.

Obvious dramatic effect on scalp map topography:

But, do they perform the same functions?
Choosing data measures

ERPs seem different…
Choosing data measures

Spectra are similar, but they have variable responses to different conditions…
Choosing data measures

ERSPs have some similar features...
Choosing data measures

What data measures should you use?

It depends…

• broadly-matched ICs: use many/all of the measures.

• specifically-matched ICs: use one/few of the measures.
What should clusters look like?

ICs clustered by dipole location and ERSP activity.
Precluster: Use singular values from PCA

%% Do it yourself:
%% Load all ERSP data
%% decompose with PCA
%% plot singular values
(See code in ‘Tutorial_7_BuildSTUDY.m’)
STUDY clustering overview

STEP 1
Build a STUDY

STEP 2
Precompute the data

STEP 3
Precluster the data

STEP 4
Cluster the data
nclusts = 25; % choose # of clusters to create

[STUDY] = pop_clust(STUDY, ALLEEG,'algorithm','kmeans','clus_num',nclusts);
1. k initial "means" (in this case k=3, (shown in color)) are randomly selected from the data set (shown in grey).

2. k clusters are created by associating every observation with the nearest mean.

3. The centroid of each of the k clusters becomes the new means

4. Steps 2 and 3 are repeated until convergence has been reached.
Plot/edit clusters
Plot cluster data

- Plot mean scalp maps for easy reference.