Data Selection Panel

These are the main input data that will be used in further analysis unless you specify otherwise. Use the drop down menus to select the data and information you want to use in your analysis.

<table>
<thead>
<tr>
<th>Name of Data Matrix</th>
<th>Name of Variable Matrix</th>
<th>Name of Filename Matrix</th>
<th>Name of Totalcounts Matrix</th>
<th>Name of Samplenames Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>Select Data</td>
<td>Select Variables</td>
<td>Select Filenames</td>
<td>Select Totalcounts</td>
<td>Select Samples</td>
</tr>
</tbody>
</table>

This tutorial contains navigation buttons that enable you to move throughout the tutorial.

Please use the navigation buttons and not the page up/page down or arrow keys to navigate through the tutorials.

This is the 'Next' button. It takes you to the next frame or stop point.

This is the 'Previous' button. It takes you to the previous frame or stop point.

This is the 'Go to frame' button. It takes you to a specified frame.

This is the 'Go to URL' button. It takes you to a website link.

Press the 'Next' button below to start this tutorial.
## Data Selection Panel

These are the main input data that will be used in further analysis unless you specify otherwise. Use the drop down menus to select the data and information you want to use in your analysis.

<table>
<thead>
<tr>
<th>Name of Data Matrix</th>
<th>Name of Variable Matrix</th>
<th>Name of Filename Matrix</th>
<th>Name of Totalcounts Matrix</th>
<th>Name of Samplenames Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>ndataset</td>
<td>exactmass</td>
<td>filenames</td>
<td>totalcounts</td>
<td>samplenames</td>
</tr>
</tbody>
</table>

---

This tutorial will cover how to run DPCA on your data using the spectragui.

This method is based on the paper: Yandle and Macfie in J. Chemometrics, v. 3, p. 589-600 (1989) and the code was adapted from code kindly provided by Matt Wagner PhD.

The data is automatically mean centered within this function.

No support of this function is provided.

It is included because I had the code for it.
From the 'MVA' menu choose -> Run DPCA
### Data Selection Panel

These are the main input data that will be used in further analysis unless you specify otherwise. Use the drop-down menus to select the data and information you want to use in your analysis.

<table>
<thead>
<tr>
<th>Name of Data Matrix</th>
<th>Name of Variable Matrix</th>
<th>Name of Filename Matrix</th>
<th>Name of Totalcounts Matrix</th>
<th>Name of Samplenames Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>ndataset</td>
<td>exactmass</td>
<td>filenames</td>
<td>totalcounts</td>
<td>samplenames</td>
</tr>
</tbody>
</table>

### Input Information

Make sure the data you want to use is selected in the menus of the 'Data Selection Panel'.

Load Selected Data

**Data:** None  
**Samples:** None  
**Variables:** None

### Run DPCA

Show Eigenvalue Scren Plot

### DPCA Summary

<table>
<thead>
<tr>
<th>PC#</th>
<th>%Var</th>
<th>%VarTotal</th>
</tr>
</thead>
</table>
Data Selection Panel

These are the main input data that will be used in further analysis unless you specify otherwise. Use the drop down menus to select the data and information you want to use in your analysis.

Name of Data Matrix: ndataset
Name of Variable Matrix: exactmass
Name of Filename Matrix: filenames
Name of Totalcounts Matrix: totalcounts
Name of Samplenames Matrix: sampelnames

Input Information

Make sure the data selected above is the data you want to analyze and then press the "Load Selected Data" button.

DPCA Panel

Load Selected Data

Then press the 'Load Selected Data' button.

Data:
Samples:
Variables:
None
None
None

Run DPCA

Show Eigenvalue Screen Plot

DPCA Summary

PC#  %Var  %VarTotal

Close Panel
Data Selection Panel

These are the main input data that will be used in further analysis unless you specify otherwise. Use the drop down menus to select the data and information you want to use in your analysis.

DPCA Panel

Input Information
Make sure the data selected above is the data you want to analyze and then press the "Load Selected Data" button.

Load Selected Data

Data: ndatess
Samples: samplenames
Variables: exactmass

Once the data is loaded, press the 'Run DPCA' button.

DPCA Summary

PC# %Var %VarTotal

Close Panel
Data Selection Panel

These are the main input data that will be used in further analysis unless you specify otherwise. Use the drop down menus to select the data and information you want to use in your analysis.

Name of Data Matrix: ndataset
Name of Variable Matrix: exactmass
Name of Filename Matrix: filenames
Name of Totalcounts Matrix: totalcounts
Name of Samplenames Matrix: samplenames

Input Information

Make sure the data selected above is the data you want to analyze and then press the "Load Selected Data" button.

Load Selected Data

Data:
Samples:
Variables:

Run DPCA

DPCA Summary

PC# | %Var | %VarTotal
--- | ---: | :--:
1   | 52   | 52   
2   | 19.5 | 71.5 |
3   | 12.6 | 84   |
4   | 8    | 92   |
5   | 4.3  | 96.3 |
6   | 3.7  | 100  |
7   | 0    | 100  |
8   | 0    | 100  |
9   | 0    | 100  |
10  | 0    | 100  |

Plot DPCA Scores and Loadings

These plots are for a quick check of the results and cannot be saved.

Loading plots default to m/z vs PC#.

PC#:
Plot Scores
Plot Loads

Save DPCA Data to Workspace

Number of PCs to keep
Scores output name
Loadings output name
Percent variance output name
Model output name
Save To Workspace

To see a scree plot of the eigenvalues, press the 'Show Eigenvalue Scree Plot'
The plot appears here.

Some people use this plot to determine how many PCs to keep within a model. Typically, you look for points of inflection in the curve, or where the curve hits a minimum.
Data Selection Panel

These are the main input data that will be used in further analysis unless you specify otherwise. Use the drop down menus to select the data and information you want to use in your analysis.

Input Information

Make sure the data selected above is the data you want to analyze and then press the “Load Selected Data” button.

DPCA Panel

Plot DPCA Scores and Loadings

These plots are for a quick check of the results and cannot be saved. Loading plots default to m/z vs PC#.

Save DPCA Data to Workspace

You can create scores and loadings plots using this panel.

These plots are intended to allow you to browse the results and cannot be saved.

DPCA Summary

<table>
<thead>
<tr>
<th>PC#</th>
<th>%Var</th>
<th>%VarTotal</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>52</td>
<td>52</td>
</tr>
<tr>
<td>2</td>
<td>19.5</td>
<td>71.5</td>
</tr>
<tr>
<td>3</td>
<td>12.6</td>
<td>84</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>92</td>
</tr>
<tr>
<td>5</td>
<td>4.3</td>
<td>96.3</td>
</tr>
<tr>
<td>6</td>
<td>3.7</td>
<td>100</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>
Data Selection Panel

These are the main input data that will be used in further analysis unless you specify otherwise. Use the drop down menus to select the data and information you want to use in your analysis.

Input Information
Make sure the data selected above is the data you want to analyze and then press the "Load Selected Data" button.

DPCA Panel

Plot DPCA Scores and Loadings
These plots are for a quick check of the results and cannot be saved. Loading plots default to m/z vs PC#.

Choose what you want to plot from the drop down menus.

The menu on the left is the x-axis, the menu on the right is the y-axis.

All loadings plots are plotted as the chosen PC number from the y-axis list vs m/z. The spectragui does not support plotting a cross plot of two PC loadings.
Data Selection Panel

These are the main input data that will be used in further analysis unless you specify otherwise.
Use the drop-down menus to select the data and information you want to use in your analysis.

Name of Data Matrix: ndataset
Name of Variable Matrix: exactmass
Name of Filename Matrix: filenames
Name of Totalcounts Matrix: totalcounts
Name of Samplenames Matrix: samplenames

Input Information

Make sure the data selected above is the data you want to analyze and then press the "Load Selected Data" button.

Load Selected Data

Data: ndataset
Samples: samplenames
Variables: exactmass

Run DPCA

DPCA Summary

<table>
<thead>
<tr>
<th>PC#</th>
<th>%Var</th>
<th>%VarTotal</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>52</td>
<td>71.5</td>
</tr>
<tr>
<td>2</td>
<td>19.5</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>12.6</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5.3</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>4.3</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>3.7</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>

Plot DPCA Scores and Loadings

These plots are for a quick check of the results and cannot be saved.
Loading plots default to m/z vs PCs.

Plot Scores
Plot Loads

Save DPCA Data to Workspace

Number of PCs to keep

Percent variance output name

Model output name

Save To Workspace

After selecting what you want to plot, press the 'Plot Scores'
Data Selection Panel
These are the main input data that will be used in further analysis unless you specify otherwise.
Use the drop-down menus to select the data and information you want to use in your analysis.

Name of Data Matrix: ndataset
Name of Variable Matrix: exactmass
Name of Filename Matrix: filenames
Name of Totalcounts Matrix: totalcounts
Name of Samplenames Matrix: samplenames

Input Information
Make sure the data selected above is the data you want to analyze and then press the "Load Selected Data" button.

Load Selected Data

Data: ndataset
Samples: samplenames
Variables: exactmass

Run DPCA

DPCA Summary

<table>
<thead>
<tr>
<th>PC#</th>
<th>%Var</th>
<th>%VarTotal</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>52</td>
<td>52</td>
</tr>
<tr>
<td>2</td>
<td>19.5</td>
<td>71.5</td>
</tr>
<tr>
<td>3</td>
<td>12.6</td>
<td>84</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>92</td>
</tr>
<tr>
<td>5</td>
<td>4.3</td>
<td>96.3</td>
</tr>
<tr>
<td>6</td>
<td>3.7</td>
<td>100</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>

Plot DPCA Scores and Loadings
These plots are for a quick check of the results and cannot be saved.
Loading plots default to m/z vs PCs.

samp/var: 1

And then 'Plot Loads' to create the plots.

Save DPCA Data to Workspace
Number of PCs to keep:

Percent variance output name:
Model output name:
Save To Workspace

Close Panel
Data Selection Panel

These are the main input data that will be used in further analysis unless you specify otherwise. Use the drop down menus to select the data and information you want to use in your analysis.

Name of Data Matrix
Name of Variable Matrix
Name of Filename Matrix
Name of Totalcounts Matrix
Name of Samplenames Matrix

Input Information
Make sure the data selected above is the data you want to analyze and then press the “Load Selected Data” button.

Load Selected Data

Data: ndataset
Samples: samplenames
Variables: exactmass

DPCA Panel

Plot DPCA Scores and Loadings
These plots are for a quick check of the results and cannot be saved.
Loading plots default to m/z vs PCs.

samp/var:
1:

Plot Scores
Plot Loads

Save DPCA Data to Workspace
Number of PCs to keep:

Scores output name:

Loadings output name:

Percent variance output name:

Model output name:

Next enter names for the scores, loadings, percent variance, and model in their respective boxes.

These names should be unique names to prevent overwriting data in the Matlab workspace.

Show Eigenvalue Scre Plot

DPCA Summary

<table>
<thead>
<tr>
<th>PC#</th>
<th>%Var</th>
<th>%VarTotal</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>52</td>
<td>52</td>
</tr>
<tr>
<td>2</td>
<td>19.5</td>
<td>71.5</td>
</tr>
<tr>
<td>3</td>
<td>12.6</td>
<td>84</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>92</td>
</tr>
<tr>
<td>5</td>
<td>4.3</td>
<td>96.3</td>
</tr>
<tr>
<td>6</td>
<td>3.7</td>
<td>100</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>

Close Panel
Data Selection Panel

These are the main input data that will be used in further analysis unless you specify otherwise. Use the drop down menus to select the data and information you want to use in your analysis.

Name of Data Matrix: name = ndataset
Name of Variable Matrix: name = exactmass
Name of Filename Matrix: name = filenames
Name of Totalcounts Matrix: name = totalcounts
Name of Samplenames Matrix: name = samplenames

Input Information

Make sure the data selected above is the data you want to analyze and then press the “Load Selected Data” button.

Data: ndataset
Samples: samplenames
Variables: exactmass

DPCA Panel

Plot DPCA Scores and Loadings

These plots are for a quick check of the results and cannot be saved.
Loading plots default to m/z vs PCs.

1. samp/var = 1
2. Plot Scores
3. Plot Loads

Press the ‘Save To Workspace’ button.

DPCA Summary

PC#  %Var  %VarTotal
1   52    52
2   19.5  71.5
3   12.6  84
4   8     92
5   4.3   96.3
6   3.7   99
7   0     100
8   0     100
9   0     100
10  0     100

Press the ‘Save To Workspace’ button.
Data Selection Panel

These are the main input data that will be used in further analysis unless you specify otherwise. Use the drop down menus to select the data and information you want to use in your analysis.

- Name of Data Matrix: ndataset
- Name of Variable Matrix: exactmass
- Name of Filename Matrix: filenames
- Name of Totalcounts Matrix: totalcounts
- Name of Samplenames Matrix: samplenames

DPCA Panel

- The data is saved to the Matlab workspace and entered into the respective drop down menus of the 'MVA Data Selection Panel'.

DPCA Summary

<table>
<thead>
<tr>
<th>PC#</th>
<th>%Var</th>
<th>%VarTotal</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>52</td>
<td>52</td>
</tr>
<tr>
<td>2</td>
<td>19.5</td>
<td>71.5</td>
</tr>
<tr>
<td>3</td>
<td>12.6</td>
<td>84</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>92</td>
</tr>
<tr>
<td>5</td>
<td>4.3</td>
<td>96.3</td>
</tr>
<tr>
<td>6</td>
<td>3.7</td>
<td>100</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>

DPCA Scores and Loadings

Scores and loadings plots are for a quick check of the results and cannot be saved. Plots default to m/z vs PC.

- PC# of interest (var): 1
- PC# of interest (var): 2

Save DPCA to Workspace

- Number of PCs to keep: 6
- Scores output name: dpcascores
- Loadings output name: dpcaloads
- Percent variance output name: dpc_variance
- Model output name: dpcamodel

Close Panel
Data Selection Panel

These are the main input data that will be used in further analysis unless you specify otherwise. Use the drop down menus to select the data and information you want to use in your analysis.

Name of Data Matrix: ndataset
Name of Variable Matrix: exactmass
Name of Filename Matrix: filenames
Name of Totalcounts Matrix: totalcounts
Name of Samplenames Matrix: samplenames

That's it for this tutorial.

Press the green button on the left to go back to the previous step. Press the button the right to go back to the beginning of the tutorial.