

Parameter Space Exploration for VisRSeq

Joseph Pober

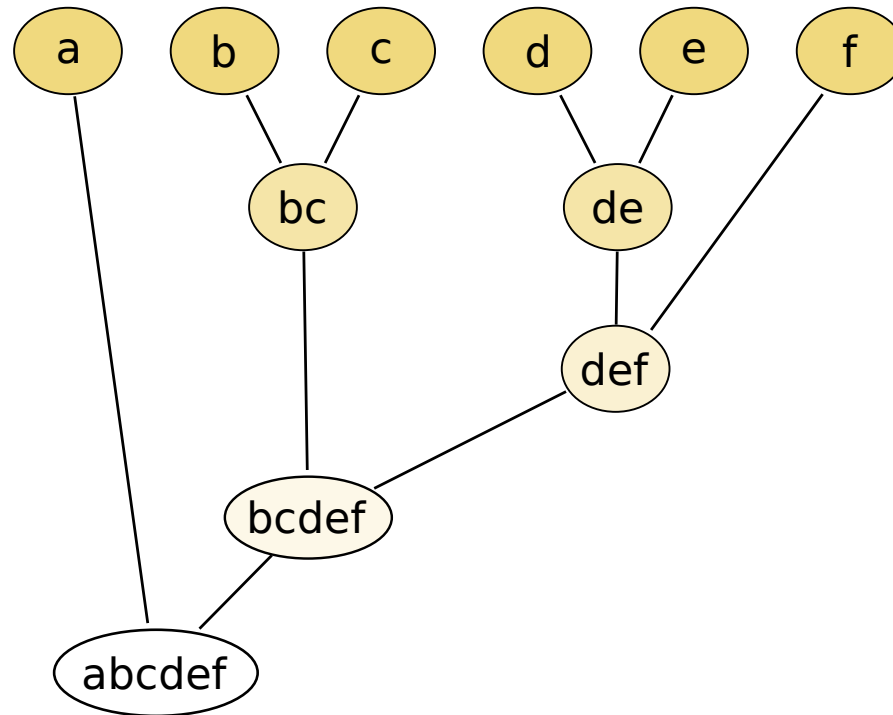
Content

- Recap of project
- Visualization design
- Prototype

Project recap

- VisRSeq lets non-programmers use R
- VisRSeq allows only one plot at a time
- Parameter space exploration
- Clustering

Hierarchical clustering



Results

Cluster



Using the minus button zooms out one hierarchy-level and displays the parent cluster again.

The scroll wheel can also be used to zoom out again. The zoom-out action will be performed after the wheel is turned a certain amount of degrees.

While zooming in, the cluster shortly becomes transparent, before disappearing. Underneath it the clusters it is composed of will be rendered.

The user can pan the viewport while zoomed in and will see the child-clusters of the other clusters.

Normal View

The screenshot displays the VisRSeq (v0.72.9) software interface. The main workspace is titled "workspace" and contains a single application icon labeled "k-means". The interface is divided into several panels:

- File:** Located at the top left, containing standard file management icons.
- Data:** A large empty area on the left side of the workspace.
- Apps:** A search panel on the right side, showing a search bar and a list of clustering applications: "Hierarchical Clustering (hclust)", "K-Means Clustering", and "Partitioning Around Medoids".
- Parameters:** A panel below the Apps panel, currently showing the configuration for the "k-means" application. It includes sections for "variables", "Plot Options", and "Output Column Name".
- Console:** A panel at the bottom left, currently empty.

The "Parameters" panel for "k-means" is configured as follows:

- variables:** standardize variables
- centers:** 3
- maximum number of c...:** 15
- Plot Options:**
 - choice of plot: K-means clustering
 - iterations: 10
 - start: 1
 - algorithm: Hartigan-Wong
- Output Column Name:** cluster ID: clusterID

At the bottom right of the interface, there is a "Normal" dropdown menu and a "Run" button with a circular arrow icon.

Ranges

The screenshot displays the VisRSeq (v0.72.9) software interface. The main workspace is titled "workspace" and contains a single application icon labeled "k-means". The interface is divided into several panels:

- Apps Panel:** Located on the right, it features a search bar and a list of clustering applications: "Hierarchical Clustering (hclust)", "K-Means Clustering", and "Partitioning Around Medoids".
- Parameters Panel:** Below the Apps panel, it shows the configuration for the selected "k-means" app. The "Input" tab is active, displaying the following settings:
 - variables:** (empty field)
 - standardize variables:** Yes No
 - centers:** 3 (range 3 to 3)
 - maximum number of c...:** 15 (range 15 to 15)
 - Plot Options:**
 - Plot the within groups sums of squar...
 - choice of plot:** Recommended number of clusters u... K-means clustering
 - iterations:** 10 (range 10 to 10)
 - start:** 1 (range 1 to 1)
 - Hartigan-Wong
 - Lloyd
 - Forgy
 - MacQueen
 - Output Column Name:** cluster ID (range clusterID to clusterID)
- Console Panel:** Located at the bottom left, it is currently empty.
- Runs Panel:** Located at the bottom right, it shows "Runs: 2" and a "Run" button.

Distribution

VisRSeq (v0.72.9)

File

Data workspace k-means [k]

Apps

Search

[Clustering]

- Partitioning Around Medoids
- Hierarchical Clustering (hclust)
- K-Means Clustering

Parameters

Input Code

k-means

variables:

standardize variables: Yes No

centers: 3 | 10

maximum number of c... 11 | 14

Plot Options

Runs Runs: 7 Run

K-means clustering

K-means clustering

wt

mpg

Console

* Among all indices:

Hovering

VisRSeq (v0.72.9)

File

Data workspace k-means [b]

Apps

Search

[Clustering]

- Partitioning Around Medoids
- Hierarchical Clustering (hclust)
- K-Means Clustering

Parameters

Input Code

k-means

variables:

standardize variables: Yes No

centers: 3 | 10

maximum number of c... 11 | 14

Plot Options

Runs Runs: 7 Run

K-means clustering

K-means clustering

mpg

mpg

Console

* Among all indices:

Selecting

VisRSeq (v0.72.9)

File

Data workspace k-means [k]

Apps

Search

[Clustering]

- Partitioning Around Medoids
- Hierarchical Clustering (hclust)
- K-Means Clustering

Parameters

Input Code

k-means

variables:

standardize variables: Yes No

centers: 3 | 8

maximum number of c... 11 | 14

Plot Options

Runs Runs: 7 Run

Console

* Among all indices:

Inspecting

The screenshot displays the VisRSeq (v0.72.9) software interface. The main workspace is titled "workspace k-means [k]" and contains four data visualization plots: two scatter plots on the left and two plots on the right, one of which is a 3x3 grid of smaller plots. The right-hand panel is divided into three sections: "Apps" with a search bar and icons for "Hierarchical Clustering (hclust)", "K-Means Clustering", and "Partitioning Around Medoids"; "Parameters" with an input tab and settings for "k-means" including "standardize variables" (checked), "centers" (set to 3), and "maximum number of c..." (set to 11); and a "Console" at the bottom with the text "* Among all indices:". The bottom right corner features a "Runs" section with a "Run" button and a counter showing "7".

Zooming & Panning

The screenshot displays the VisRSeq (v0.72.9) software interface. The main workspace is titled "workspace k-means [k]" and contains a "K-means clustering" plot. This plot is a grid of scatter plots for variables "mpg", "cyl", and "disp", showing data points colored by cluster. To the left of the main plot is a "centering" plot. To the right is a vertical plot with a y-axis ranging from -1.0 to 1.0. The interface includes a "Data" panel on the left, a "Search" panel on the right with icons for "Hierarchical Clustering (hclust)", "K-Means Clustering", and "Partitioning Around Medoids", and a "Parameters" panel at the bottom right. The "Parameters" panel is set to "Input" mode and shows settings for "k-means", including "standardize variables" (Yes/No), "centers" (3/8), and "maximum number of c..." (11/14). A "Plot Options" section is also visible. At the bottom, a "Console" panel shows the message "* Among all indices:". The bottom right corner features a "Runs" panel with a "Run" button and a counter showing "7".

Zooming out

The screenshot displays the VisRSeq (v0.72.9) software interface. The main workspace shows a k-means clustering plot with data points colored by cluster. The plot is titled "K-means clustering" and has axes labeled "mpg" (x-axis, ranging from -1 to 2) and "mpg" (y-axis, ranging from -1.0 to 1.0). A smaller version of the same plot is visible in the top-left corner of the workspace.

On the right side, the "Apps" panel is open, showing a search bar and several clustering algorithms: Hierarchical Clustering (hclust), K-Means Clustering, and Partitioning Around Medoids. The "Parameters" panel is also open, showing the configuration for the "k-means" algorithm. The parameters are:

- variables: (empty)
- standardize variables: Yes No
- centers: 3 (left) | 8 (right)
- maximum number of c...: 11 (left) | 14 (right)

At the bottom, the "Console" panel shows the message: "* Among all indices:". The "Runs" panel at the bottom right shows "Runs: 7" and a "Run" button.

Challenges

- Visual comparison of plots
- Parallelization

Thank you for your attention!