

# Package: FeatSeekR (via r-universe)

October 10, 2026

**Type** Package

**Title** FeatSeekR an R package for unsupervised feature selection

**Version** 1.13.0

**Description** FeatSeekR performs unsupervised feature selection using replicated measurements. It iteratively selects features with the highest reproducibility across replicates, after projecting out those dimensions from the data that are spanned by the previously selected features. The selected a set of features has a high replicate reproducibility and a high degree of uniqueness.

**License** GPL-3

**Encoding** UTF-8

**Imports** pheatmap, MASS, pracma, stats, SummarizedExperiment, methods

**RoxygenNote** 7.2.3

**Suggests** rmarkdown, knitr, BiocStyle, testthat (>= 3.0.0)

**VignetteBuilder** knitr

**BugReports** <https://github.com/tcapraz/FeatSeekR/issues>

**URL** <https://github.com/tcapraz/FeatSeekR>

**biocViews** Software, StatisticalMethod, FeatureExtraction, MassSpectrometry

**Config/testthat/edition** 3

**Config/pak/sysreqs** zlib1g-dev

**Repository** <https://bioc.r-universe.dev>

**Date/Publication** 2026-04-28 13:00:40 UTC

**RemoteUrl** <https://github.com/bioc/FeatSeekR>

**RemoteRef** HEAD

**RemoteSha** c28c249bcd22b2470dbc7bc09e11d600005116a6

## Contents

|                                 |   |
|---------------------------------|---|
| FeatSeek . . . . .              | 2 |
| FeatSeekR . . . . .             | 3 |
| plotSelectedFeatures . . . . .  | 3 |
| plotVarianceExplained . . . . . | 4 |
| simData . . . . .               | 5 |

|              |          |
|--------------|----------|
| <b>Index</b> | <b>7</b> |
|--------------|----------|

---

FeatSeek

*FeatSeek*

---

## Description

This function ranks features of a 2 dimensional array according to their reproducibility between conditions.

## Usage

```
FeatSeek(
  data,
  conditions = NULL,
  max_features = NULL,
  init = NULL,
  verbose = TRUE
)
```

## Arguments

|                           |                                                                                                                                                                                                                                                                                                                                                 |
|---------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <code>data</code>         | SummarizedExperiment with assay named <code>data</code> , where samples belongs to different conditions. Which sample belongs to which condition should be indicated in <code>colData</code> slot <code>conditions</code> . Or <code>matrix</code> with features x samples. Each conditions have multiple samples from replicated measurements. |
| <code>conditions</code>   | factor of length samples, indicating which sample belongs to which condition. Only required if <code>data</code> is provided as <code>matrix</code> .                                                                                                                                                                                           |
| <code>max_features</code> | integer number of features to rank                                                                                                                                                                                                                                                                                                              |
| <code>init</code>         | character vector with names of initial features. If <code>NULL</code> the feature with highest F-statistic will be used                                                                                                                                                                                                                         |
| <code>verbose</code>      | logical indicating whether messages should be printed                                                                                                                                                                                                                                                                                           |

## Value

SummarizedExperiment containing one assay with the selected features. `rowData` stores for each selected feature the F-statistic under `metric`, the cumulative explained variance under `explained_variance` and the feature names under `selected`

## Examples

```
# run FeatSeek to select the top 20 features
data <- array(rnorm(100*30), dim=c(30, 100),
dimnames <- list(paste("feature", seq_len(30)), NULL))
conds <- rep(seq_len(50), 2)
res <- FeatSeek(data, conds, max_features=20)

# res stores the 20 selected features ranked by their replicate
# reproducibility
```

---

FeatSeekR

*FeatSeekR an R package for unsupervised feature selection*


---

## Description

FeatSeekR performs unsupervised feature selection using replicated measurements. It iteratively selects features with the highest reproducibility across conditions, after projecting out those dimensions from the data that are spanned by the previously selected features. The selected a set of features has a high replicate reproducibility and a high degree of uniqueness.

## Details

For information on how to use this package please type `vignette("FeatSeekR-vignette")`.

Please post questions regarding the package to the Bioconductor Support Site:

<https://support.bioconductor.org>

## Author(s)

Tümay Capraz

---

plotSelectedFeatures    *plotSelectedFeatures*


---

## Description

plot correlation matrix of selected feature sets

## Usage

```
plotSelectedFeatures(res, n_features = NULL, assay = "selected")
```

**Arguments**

|                         |                                                                                                                      |
|-------------------------|----------------------------------------------------------------------------------------------------------------------|
| <code>res</code>        | result SummarizedExperiment from FeatSeek function                                                                   |
| <code>n_features</code> | top <code>n_features</code> to plot. if NULL then the maximum number of features in <code>res</code> will be plotted |
| <code>assay</code>      | assay slot to plot from result SummarizedExperiment object, default is the selected features slot                    |

**Value**

returns heatmap of selected features

**Examples**

```
# run FeatSeek to select the top 20 features
data <- array(rnorm(100*30), dim=c(30,100),
             dimnames = list(paste("feature", seq_len(30)), NULL))
conds <- rep(seq_len(50), 2)
res <- FeatSeek(data, conds, max_features=20)

# res stores the 20 selected features ranked by their replicate
# reproducibility
# plot the top 5 features
plotSelectedFeatures(res, n_features=5)
```

---

`plotVarianceExplained` *plotVarianceExplained*

---

**Description**

plot variance explained from 1 to `max_features` in `res`

**Usage**

```
plotVarianceExplained(res)
```

**Arguments**

|                  |                                                    |
|------------------|----------------------------------------------------|
| <code>res</code> | result SummarizedExperiment from FeatSeek function |
|------------------|----------------------------------------------------|

**Value**

returns plot of variance explained vs number of features

## Examples

```
# run FeatSeek to select the top 20 features
data <- array(rnorm(100*30), dim=c(30,100),
             dimnames = list(paste("feature", seq_len(30)), NULL))
conds <- rep(seq_len(50), 2)
res <- FeatSeek(data, conds, max_features=20)

# res stores the 20 selected features ranked by their replicate
# reproducibility
plotVarianceExplained(res)
```

---

simData

*simData*


---

## Description

simulate Data with orthogonal feature clusters and replicated samples. Each feature cluster corresponds to a different latent factor and contains 10 redundant features. E.g. choosing samples = 100, n\_latent\_factors = 5 and replicates = 2 will simulate a 50 x 200 data matrix, where the first 100 samples belong to replicate 1 and sample 101-200 belong to replicate 2.

## Usage

```
simData(conditions, n_latent_factors, replicates)
```

## Arguments

|                  |                                               |
|------------------|-----------------------------------------------|
| conditions       | number of conditions to generate samples from |
| n_latent_factors | number of latent factors to generate          |
| replicates       | number of replicates to generate              |

## Details

simData constructs n\_latent\_factors by generating a random matrix  $\mathbf{Q}$  whose row vectors  $\mathbf{Q}_i \sim \mathcal{N}(0, 1)$  with  $n$  samples and  $i \in \{1, \dots, n\_latent\_factors\}$  are orthonormal, each corresponding to a different latent factor. To simulate a set of redundant feature groups, it generates 10 features  $X_j$  for each latent factor  $\mathbf{Q}_i$  by scaling each latent factor by a random factor  $\delta_j \sim \mathcal{N}(0, 1)$  and adding replicate specific noise  $\epsilon_c \sim \mathcal{N}(0, 0.1)$  with  $c \in \{1, \dots, replicates\}$  preserving orthogonality.

## Value

SummarizedExperiment object carrying simulated data, with colData indicating which sample belongs to which replicate

**Examples**

```
# simulate data 100 samples from 100 conditions, 20 features generated by 2  
# latent factors and 2 replicates  
simData(conditions=100, n_latent_factors=2, replicates=2)
```

# Index

\* **package**

FeatSeekR, [3](#)

FeatSeek, [2](#)

FeatSeekR, [3](#)

plotSelectedFeatures, [3](#)

plotVarianceExplained, [4](#)

simData, [5](#)