

# Package ‘slfm’

August 10, 2015

**Type** Package

**Title** Tools for Fitting Sparse Latent Factor Model

**Version** 0.2.2

**Author** Joao Duarte and Vinicius Mayrink

**Maintainer** Joao Duarte <jdanielnd@gmail.com>

**Description** Set of tools to find coherent patterns in microarray data using a Bayesian sparse latent factor model (Duarte and Mayrink 2015 - [http://link.springer.com/chapter/10.1007%2F978-3-319-12454-4\\_15](http://link.springer.com/chapter/10.1007%2F978-3-319-12454-4_15)). Considerable effort has been put into making slfm fast and memory efficient, turning it an interesting alternative to simpler methods in terms of execution time. It implements versions of the SLFM using both type of mixtures: using a degenerate distribution or a very concentrated normal distribution for the spike part of the mixture. It also implements additional functions to help pre-process the data and fit the model for a large number of arrays.

**URL** <https://github.com/jdanielnd/slfm>

**BugReports** <https://github.com/jdanielnd/slfm/issues>

**Depends** R (>= 3.1.0)

**Imports** Rcpp (>= 0.11.0), coda, lattice

**LinkingTo** Rcpp, RcppArmadillo

**License** GPL-2

**NeedsCompilation** yes

**Repository** CRAN

**Date/Publication** 2015-08-10 09:09:59

## R topics documented:

slfm-package	2
plot_matrix	2
process_matrix	3
slfm	3
slfm_list	4

---

slfm-package	<i>slfm: the sparse latent factor model package for R.</i>
--------------	------------------------------------------------------------

---

### Description

Set of tools to find coherent patterns in microarray data using a Bayesian sparse latent factor model (Duarte and Mayrink 2015 - <http://link.springer.com/chapter/10.1007> Considerable effort has been put into making slfm fast and memory efficient, turning it an interesting alternative to simpler methods in terms of execution time. It implements versions of the SLFM using both type of mixtures: using a degenerate distribution or a very concentrated normal distribution for the spike part of the mixture. It also implements additional functions to help pre-process the data and fit the model for a large number of arrays. Includes functions to:

### Details

\* pre-process a set of matrices \* fit models to a set of matrices \* detailed summary of model fit

---

plot_matrix	<i>Plot matrix</i>
-------------	--------------------

---

### Description

This function plots a data matrix

### Usage

```
plot_matrix(y, standardize.rows = TRUE, reorder.rows = TRUE,
            reorder.cols = TRUE, high.contrast = TRUE)
```

### Arguments

y	matrix to be plotted
standardize.rows	standardize matrix rows for plot
reorder.rows	reorder matrix rows based on pattern
reorder.cols	reorder matrix cols based on pattern
high.contrast	apply transformation to matrix to increase contrast

---

process_matrix	<i>Pre-process data for SLFM</i>
----------------	----------------------------------

---

**Description**

This function pre-process the data to be used for fitting a sparse latent factor model.

**Usage**

```
process_matrix(path, output_path, sample_size)
```

**Arguments**

path	path containing the set of matrices to be processed
output_path	path to save the processed matrices
sample_size	number of matrices to be used on the principal component analysis

---

slfm	<i>Sparse Latent Factor Model</i>
------	-----------------------------------

---

**Description**

This function is used to fit a Bayesian sparse latent factor model.

**Usage**

```
slfm(x, a = 2.1, b = 1.1, gamma_a = 1, gamma_b = 1, omega_0 = 0.01,
     omega_1 = 10, sample = 1000, burnin = round(0.25 * sample), lag = 1,
     degenerate = FALSE)
```

**Arguments**

x	matrix with the pre-processed data
a	prior shape parameter for Gamma distribution
b	prior scale parameter for Gamma distribution
gamma_a	prior parameter for Beta distribution
gamma_b	prior parameter for Beta distribution
omega_0	prior variance of the spike component
omega_1	prior variance of the slab component
sample	sample size after burn-in
burnin	burn-in size
lag	lag for MCMC
degenerate	use the degenerate version of mixture

**Value**

x: data matrix  
 q\_star: matrix of MCMC chains for q\_star parameter  
 alpha: summary table of MCMC chains for alpha parameter  
 lambda: summary table of MCMC chains for lambda parameter  
 sigma: summary table of MCMC chains for sigma parameter  
 classification: classification of each alpha ('present', 'marginal', 'absent')

**References**

1. Duarte, J. D. N. and Mayrink, V. D. (2015). Factor analysis with mixture modeling to evaluate coherent patterns in microarray data. In *Interdisciplinary Bayesian Statistics*, volume 118 of Springer Proceedings in Mathematics & Statistics, pages 185-195. Springer International Publishing.

**Examples**

```
mat <- matrix(rnorm(2000), nrow = 20)
slfm(mat, sample = 1000)
```

---

 slfm\_list

---

*Fit SLFM to the matrices inside a directory*


---

**Description**

This function is used to fit a Bayesian sparse latent factor model to a directory of numeric matrices.

**Usage**

```
slfm_list(path = ".", recursive = TRUE, a = 2.1, b = 1.1, gamma_a = 1,
  gamma_b = 1, omega_0 = 0.01, omega_1 = 10, sample = 1000,
  burnin = round(0.25 * sample), lag = 1, degenerate = FALSE)
```

**Arguments**

path	path to read the matrices from
recursive	if the function should look recursively inside folders
a	prior shape parameter for Gamma distribution
b	prior scale parameter for Gamma distribution
gamma_a	prior parameter for Beta distribution
gamma_b	prior parameter for Beta distribution
omega_0	prior variance of the spike component
omega_1	prior variance of the slab component
sample	sample size after burn-in

burnin	burn-in size
lag	lag for MCMC
degenerate	use the degenerate version of mixture

# Index

[plot\\_matrix](#), 2  
[process\\_matrix](#), 3  
  
[slfm](#), 3  
[slfm-package](#), 2  
[slfm\\_list](#), 4