

Function Documentation

October 4, 2017

abundance_plot	<i>abundance table visualization</i>
----------------	--------------------------------------

Description

abundance table visualization

Usage

```
abundance_plot(count_matrix, title = "")
```

Arguments

count_matrix	a matrix of integers
title	an optional title

Value

a plot

Note

requires the ggplot2 package

association_plot	<i>a bipartite graph showing the association between X and Y</i>
------------------	--

Description

a bipartite graph showing the association between X and Y

Usage

```
association_plot(MPPI, mppi_threshold = 0.5, inc_legend = F, lwdx = 5,  
graph_layout = "circular", lab_dist = 2, ...)
```

Arguments

MPPI	a data.frame with n_vars x n_taxa rows and these four columns: 1) covariate: the names of the columns of X 2) taxa: the names of the columns of Y 3) mppi: the marginal posterior probability of inclusion for each taxa by covariate parameter 4) beta: a point estimate of for each taxa by covariate parameter
mppi_threshold	the threshold for inclusion in the plot
inc_legend	boolean to include the legend
lwdx	a scalar multiplier for growing or shrinking the widths of edges
graph_layout	either "circular" for a round layout or "bipartite" for a side-by-side layout
lab_dist	a scalar argument for the distance between node centers and the node labels
...	passthrough arguments

Value

a plot

Note

requires the igraph package

bfdr

calculate the Bayesian False Discovery Rate

Description

calculate the Bayesian False Discovery Rate

Usage

```
bfdr(mppi_vector, threshold = 0.1)
```

Arguments

mppi_vector	A vector of marginal posterior probabilities of inclusion.
threshold	The expected false discovery rate threshold

Value

selected: A boolean vector of selected (= T) and rejected (= F) variables
threshold: The BFDR threshold

References

Newton, M. A., Noueiry, A., Sarkar, D., & Ahlquist, P. (2004). Detecting differential gene expression with a semiparametric hierarchical mixture method. *Biostatistics*, 5(2), 155-76. doi:10.1093/biostatistics/5.2.155

dmbvs	<i>an R wrapper to C code for spike-and-slab Dirichlet–Multinomial Bayesian variable selection</i>
-------	--

Description

an R wrapper to C code for spike-and-slab Dirichlet–Multinomial Bayesian variable selection

Usage

```
dmbvs(XX, YY, intercept_variance, slab_variance, bb_alpha, bb_beta, GG, thin,
      burn, proposal_alpha = 0.5, proposal_beta = 0.5, init_alpha = 0,
      init_beta = 0, exec = file.path(".", "dmbvs.x"), output_location = NULL,
      r_seed = NULL, selection_type = "Gibbs")
```

Arguments

XX	covariate matrix (without intercept)
YY	count matrix
intercept_variance	a scalar for the prior variance on the intercept of the log-linear predictors
slab_variance	a scalar for the prior variance on the slab of the spike-and-slab
bb_alpha	a scalar for the alpha hyperparameter of the Beta-Bernoulli spike inclusion prior
bb_beta	a scalar for the beta hyperparameter of the Beta-Bernoulli spike inclusion prior
GG	the total number of MCMC iterations
thin	the MCMC thinning interval
burn	the number of MCMC iterations out of GG that will be discarded
proposal_alpha	initial value, either a scalar or a vector of length ncol(YY), if a scalar, that value is used for all proposals on alpha
proposal_beta	initial value, either a scalar or a matrix with ncol(XX) columns and ncol(YY) rows, if a scalar, that value is used for all proposals on beta
init_alpha	either a scalar or a vector of size ncol(YY)
init_beta	either a scalar or a matrix with ncol(YY) rows and ncol(XX) columns, inclusion initialization uses non-zero elements of init_beta
exec	the path to the C executable
output_location	if NULL, output goes to an the directory output/ created in the working directory
r_seed	an integer seed to pass to GSL's random number generator
selection_type	either "ss" (Stochastic Search) or "Gibbs", determines the MCMC mechanism for variable selection

Value

alpha: a matrix with iterations in the rows and the alphas in the columns

alpha_accept: the Metropolis-Hastings acceptance ratio for the alphas

beta: a matrix with iterations in the rows and the per-iteration beta matrix – flattened by rows – in the columns

beta_accept: the Metropolis-Hastings acceptance ratio for the betas

hyperparameters: a list containing the hyperparameters, the MCMC parameters, and the data from the original function call

simulate_dirichlet_multinomial_regression

simulate data from a Dirichlet-Multinomial regression model

Description

simulate data from a Dirichlet-Multinomial regression model

Usage

```
simulate_dirichlet_multinomial_regression(n_obs = 100, n_vars = 100,
  n_taxa = 40, n_relevant_vars = 4, n_relevant_taxa = 4, beta_min = 0.5,
  beta_max = 1, signoise = 1, n_reads_min = 1000, n_reads_max = 2000,
  theta0 = 0.01, rho = 0.4)
```

Arguments

n_obs:	the number of samples
n_vars:	number of covariates excluding the intercept
n_taxa:	number of species
n_relevant_vars:	number of relevant nutrients
n_relevant_taxa:	number of relevant species
beta_min:	minimum absolute value of the regression parameters
beta_max:	maximum absolute value of the regression parameters
signoise:	scalar multiplier on the regression parameters
n_reads_min:	lower bound on uniform distribution for number of reads in each sample
n_reads_max:	upper bound on uniform distribution for number of reads in each sample
theta0:	the dispersion parameter
rho:	the correlation between covariates

Value

XX: (design matrix) with intercept: $n_obs * (n_vars + 1)$

YY: (count matrix) rows: n_obs samples, columns: n_taxa species

alphas: simulated intercept vector

betas: simulated coefficient matrix $n_taxa * (n_vars + 1)$

n_reads_min , n_read_max : row sum parameters

θ_0 , ϕ , ρ , σ : simulation inputs

Note

Requires the dirmult and MASS packages

Index

abundance_plot, [1](#)
association_plot, [1](#)

bfdr, [2](#)

dmbvs, [3](#)

simulate_dirichlet_multinomial_regression,
[4](#)