Package ‘bridge’

March 22, 2017

Title  Bayesian Robust Inference for Differential Gene Expression

Version  1.38.0

Author  Raphael Gottardo

Description  Test for differentially expressed genes with microarray data. This package can be used with both cDNA microarrays or Affymetrix chip. The package fits a robust Bayesian hierarchical model for testing for differential expression. Outliers are modeled explicitly using a $t$-distribution. The model includes an exchangeable prior for the variances which allow different variances for the genes but still shrink extreme empirical variances. Our model can be used for testing for differentially expressed genes among multiple samples, and can distinguish between the different possible patterns of differential expression when there are three or more samples. Parameter estimation is carried out using a novel version of Markov Chain Monte Carlo that is appropriate when the model puts mass on subspaces of the full parameter space.

Maintainer  Raphael Gottardo <raph@stat.ubc.ca>

Depends  R (>= 1.9.0), rama

License  GPL (>= 2)

biocViews  Microarray,OneChannel,TwoChannel,DifferentialExpression

NeedsCompilation  yes

R topics documented:

bridge.2samples ................................................................. 1
bridge.3samples ............................................................... 4
hiv ................................................................. 6

Index  8

bridge.2samples  Bayesian Robust Inference for Differential Gene Expression (BRIDGE) with two Samples
桥接器

**Description**

测试不同表达的基因在两样本设置下。此代码可以用于cDNA微阵列或Affymetrix芯片。

**Usage**

```r
des.2samples(sample1, sample2, B=1000, min.iter=0, batch=10, mcmc.obj=NULL, all.out=TRUE, affy=FALSE, verbose=FALSE, log=FALSE, robust=TRUE)
```

**Arguments**

- `sample1`: 样本1的矩阵。每行对应一个不同的基因。
- `sample2`: 样本2的矩阵。每行对应一个不同的基因。
- `B`: 迭代使用的MCMC算法的次数。
- `min.iter`: MCMC算法的热化期长度。`min.iter` 应小于 `B`。
- `batch`: 稀疏值，用于存储MCMC。每隔`batch`迭代将存储一次。
- `mcmc.obj`: 作为 `des.2samples` 返回的类型为 `bridge2` 的对象。使用 `mcmc.obj` 初始化MCMC。如果没有 `mcmc.obj`，MCMC将初始化为最小平方估计。
- `all.out`: 一个逻辑值，指示是否应输出所有参数。如果`all.out`是FALSE，仅输出后验均值。这可能用于节省内存。
- `affy`: 一个逻辑值，指示数据是否对应于Affy数据或cDNA微阵列数据。如果affy=FALSE，使用双变量分布。
- `verbose`: 一个逻辑值，指示是否应打印当前MCMC迭代次数。
- `log`: 一个逻辑值，指示数据是否对数变换。
- `robust`: 一个逻辑值，指示是否使用`t`模型（`robust=TRUE`）或高斯模型（`robust=FALSE`）。在`t`模型的情况下，自由度将被估计。

**Details**

此代码为 robust Bayesian hierarchical model 用于测试差异表达。可以通过使用 `bridge.2samples`。每个变异系数都相互独立。

**Value**

- A matrix, each row contains the sampled values from the corresponding gene effect in sample 1.
- A matrix, each row contains the sampled values from the corresponding gene effect in sample 2.
- A vector containing the sampled values for the precision of the gene effect prior in sample 1.
lambda.gamma2 A vector containing the sampled values for the precision of the gene effect prior in sample 2.

rho A vector containing the sampled values from between sample correlation coefficient rho. If affy=TRUE, rho is identically zero.

lambda_eps1 A matrix, each row contains the sampled values from the corresponding gene specific error precision in sample 1.

lambda_eps2 A matrix, each row contains the sampled values from the corresponding gene specific error precision in sample 2.

a.eps1 A vector containing the sampled values for the mean of the prior of the error precisions in sample 1.

b.eps1 A vector containing the sampled values for the variance of the prior of the error precisions in sample 1.

a.eps2 A vector containing the sampled values for the mean of the prior of the error precisions in sample 2.

b.eps2 A vector containing the sampled values for the variance of the prior of the error precisions in sample 2.

w1 A matrix, each element (i,j) correspond to the posterior mean of the sampled weights of replicate j in gene i and sample 1. To save memory, we only store the posterior mean of the weights.

nu1 A matrix containing the sampled degrees of freedom in sample 1.

w2 A matrix, each element (i,j) correspond to the posterior mean of the sampled weights of replicate j in gene i and sample 2. To save memory, we only store the posterior means of the weights. If affy=FALSE, this is identically equal to w2 as we have a bivariate t-distribution.

nu2 A matrix containing the sampled degrees of freedom in sample 2. If affy=FALSE, this is identically equal to nu2 as we have a bivariate t-distribution.

p The mixing proportion in the two component mixture.

post.p The posterior probabilities of differential expression.

move The proportion of moves between components. This should be used as a diagnostic tool.

Author(s)
Raphael Gottardo

References
Robust Estimation of cDNA Microarray Intensities with Replicates Raphael Gottardo, Adrian E. Raftery, Ka Yee Yeung, and Roger Bumgarner Department of Statistics, University of Washington, Box 354322, Seattle, WA 98195-4322

See Also
bridge.3samples

Examples
data(hiv)

bridge.hiv<-bridge.2samples(hiv[1:10,c(1:4)],hiv[1:10,c(5:8)],B=2000,min.iter=0,batch=1,mcmc.obj=NULL,verbose=FALSE,robust=TRUE)
bridge.3samples

Bayesian Robust Inference for Differential Gene Expression (BRIDGE) with three Samples

Description

Test for differentially expressed genes in a three sample set-up. This code can be used with both cDNA microarrays or Affymetrix chip.

Usage

bridge.3samples(sample1,sample2,sample3,B=1000,min.iter=0,batch=10,mcmc.obj=NULL,all.out=TRUE,verbose=FALSE,log=FALSE,robust=TRUE)

Arguments

- **sample1**: The matrix of intensity from the sample 1. Each row corresponds to a different gene.
- **sample2**: The matrix of intensity from the sample 2. Each row corresponds to a different gene.
- **sample3**: The matrix of intensity from the sample 3. Each row corresponds to a different gene.
- **B**: The number of iteration used in the MCMC algorithm.
- **min.iter**: The length of the burn-in period in the MCMC algorithm. min.iter should be less than B.
- **batch**: The thinning value to be used in the MCMC. Only every batch-th iteration will be stored.
- **mcmc.obj**: An object of type bridge2, as returned by bridge.2samples. mcmc.obj can be used to initialized the MCMC. If no mcmc.obj, the MCMC is initialized to the least squares estimates.
- **all.out**: A logical value indicating if all the parameters should be output. If all.out is FALSE, only the posterior mean is output. This could be used to save memory.
- **verbose**: A logical value indicating if the current MCMC iteration number should be printed out.
- **log**: A logical value indicating if the data are log transformed.
- **robust**: A logical value indicating if a t model (robust==TRUE) or a Gaussian model (robust==FALSE) should be used. In the case of the t-model, the degrees of freedoms are estimated.

Details

This code fits a robust Bayesian hierarchical model for testing for differential expression. Outliers are modeled explicitly using a $t$-distribution. The model includes an exchangeable prior for the variances which allow different variances for the genes but still shrink extreme empirical variances. This function DO NOT perform normalization. The data should be normalized before hands such as centering the mean expression of each experiment.

More details can be found in the references below.
**Value**

An object of type `bridge3` containing the sampled values from the posterior distribution.

- **gamma1**: A matrix, each row contains the sampled values from the corresponding gene effect in sample 1.
- **gamma2**: A matrix, each row contains the sampled values from the corresponding gene effect in sample 2.
- **gamma3**: A matrix, each row contains the sampled values from the corresponding gene effect in sample 3.
- **lambda.gamma1**: A vector containing the sampled values for the precision of the gene effect prior for the component corresponding to sample 1.
- **lambda.gamma2**: A vector containing the sampled values for the precision of the gene effect prior for the component corresponding to sample 2.
- **lambda.gamma3**: A vector containing the sampled values for the precision of the gene effect prior for the component corresponding to sample 3.
- **lambda.gamma12**: A vector containing the sampled values for the precision of the gene effect prior for the component where sample 1 and sample 2 are combined.
- **lambda.gamma13**: A vector containing the sampled values for the precision of the gene effect prior for the component where sample 1 and sample 3 are combined.
- **lambda.gamma23**: A vector containing the sampled values for the precision of the gene effect prior for the component where sample 2 and sample 3 are combined.
- **lambda.gamma123**: A vector containing the sampled values for the precision of the gene effect prior for the component where all the samples are combined.
- **lambda_eps1**: A matrix, each row contains the sampled values from the corresponding gene precision in sample 1.
- **lambda_eps2**: A matrix, each row contains the sampled values from the corresponding gene precision in sample 2.
- **lambda_eps3**: A matrix, each row contains the sampled values from the corresponding gene precision in sample 3.
- **a.eps1**: A vector containing the sampled values for the mean of the prior of the genes precision in sample 1.
- **b.eps1**: A vector containing the sampled values for the variance of the prior of the genes precision in sample 1.
- **a.eps2**: A vector containing the sampled values for the mean of the prior of the genes precision in sample 2.
- **b.eps2**: A vector containing the sampled values for the variance of the prior of the genes precision in sample 2.
- **a.eps3**: A vector containing the sampled values for the mean of the prior of the genes precision in sample 3.
- **b.eps3**: A vector containing the sampled values for the variance of the prior of the genes precision in sample 3.
- **w1**: A matrix, each element \((i,j)\) correspond to the posterior mean of the sampled weights of replicate \(j\) in gene \(i\) and sample 1. To save memory, we only store the posterior means of the weights.
w2  A matrix, each element (i,j) correspond to the posterior mean of the sampled weights of replicate j in gene i and sample 2. To save memory, we only store the posterior means of the weigths.

w3  A matrix, each element (i,j) correspond to the posterior mean of the sampled weights of replicate j in gene i and sample 3. To save memory, we only store the posterior means of the weigths.

w1  A matrix, each element (i,j) correspond to the posterior mean of the sampled weights of replicate j in gene i and sample 1. To save memory, we only store the posterior means of the weigths.

nu1  A matrix containing the sampled degrees of freedom in sample 1.

nu2  A matrix containing the sampled degrees of freedom in sample 2.

nu3  A matrix containing the sampled degrees of freedom in sample 3.

w.mix  The posterior mixing proportions in the mixture component.

prop.model  The posterior proportions of each component for each each gene.

move  The proportion of moves between components. This should be used as a diagnostic tool.

Author(s)
Raphael Gottardo

References
Robust Estimation of cDNA Microarray Intensities with Replicates Raphael Gottardo, Adrian E. Raftery, Ka Yee Yeung, and Roger Bumgarner Department of Statistics, University of Washington, Box 354322, Seattle, WA 98195-4322

See Also
bridge.2samples

Examples
```r
sample1<-matrix(exp(rnorm(150)),50,3)
sample2<-matrix(exp(rnorm(200)),50,4)
sample3<-matrix(exp(rnorm(150)),50,3)
mcmc.bridge3<-bridge.3samples(sample1,sample2,sample3,B=10,min.iter=0,batch=10,mcmc.obj=NULL,all.out=TRUE,verbose=FALSE,robust=TRUE)
```

hiv  Cellular gene expression upon human immunodeficiency virus type 1 infection of CD4+-T-Cell lines

Description
This data set consists of 4 experiments using the same RNA preparation on 4 different slides. The expression levels of ~7000 cellular RNA transcripts were assessed in CD4-T-cell lines at time $t=24$ hour after infection with HIV virus type 1. The first 4 columns correspond to the first treatment state (hiv infected). The second four represent the control state. The experiment is a balanced dye swap experiment. Finally, the last two columns contain the row and column positions of each gene on the array (slide).
Usage

data(hiv)

Source


References

Index

*Topic **datasets**
  hiv. 6

*Topic **models**
  bridge.2samples.1
  bridge.3samples.4

*Topic **robust**
  bridge.2samples.1
  bridge.3samples.4

bridge.2samples.1, 6
bridge.3samples.3, 4

hiv. 6