bridge

April 19, 2009

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

Description

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

Usage

```
bridge.2samples(sample1,sample2,B=1000,min.iter=0,batch=10,mcmc.obj=NULL,all.out
```

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.
`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` is 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.
`affy`	A logical value indicating if the data correspond to affy data or cDNA microarray data. If affy=FALSE, a bivariate distribution is used.
`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==TRUE`) 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. More details can be found in the references below.

Value

An object of type `bridge2` 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.
`lambda.gamma1`	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 weigths.
`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 weigths. 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,m
```

---

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
```

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==TRUE) 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 weigths.
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

```
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=N
```

`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

http://expression.microslu.washington.edu/expression/vantwoutjvi2002.html

References

van't Wout, A. B. ,Lehrma, G. K., Mikheeva, S. A., O'Keeffe, G. C., Katze, M. G., Bumgarner, R. E., Geiss, G. K. and Mullins, J. I. Cellular gene expression upon human immunodeficiency virus type 1 infection of CD4+-T-Cell lines Journal of Virology, 2003. 77(2):1392-1402.

Index