Package 'nonparametricSummaryPSM'

July 3, 2019

Type Package

Title Nonparametric methods to find optimal weights to combine posterior similarity matrices

Version 0.1.0

Author Magdalena Strauss

Maintainer Magdalena Strauss <ms58@sanger.ac.uk>

Description The purpose of this code is to compute summary PSMs (posterior similarity matrices) from a set of multiple PSMs obtained for instance by means of subsampling. This implements the Dirichlet process and Pitman-Yor process based methods for combining PSMs proposed in Strauss et al. (2019). GPseudoClust: deconvolution of shared pseudo-trajectories at single-cell resolution.

License GPL-3.0

Encoding UTF-8

LazyData true

RoxygenNote 6.1.1

R topics documented:

`checkConvergence` *checkConvergence*

Description

checkConvergence

Usage

```
checkConvergence(PSMs)
```

Arguments

PSMs	3-dimensional array of PSMs, for each j PSMs[,,j] is the PSM of subsampled chain j, j = 1, ..., m

Value

results: list of length m-1 of results of processPSM function applied to the following subsets of PSMs: 1,2; 1,2,3; 1,2,3,4; ... 1,2,...,m

coph: vector of length m-1 of cophenetic correlation coefficients measuring how well the summary PSM obtained using the PY+PEAR method on the first k submatrices (k=1, ..., l) is described by a hierarchcial clustering tree for a summary PSM based on the first l subsampled chains

coph_DP: as coph, but for summary PSMs obtained by the DPM+PEAR method

distPY: vector of length m-2 of Frobenius (Euclidean) norm of distances between consecutive summary PSMs, that is summary PSMs obtained from chains 1,2,..,k,k+1 and 1,2,...k (PY+PEAR)

distDP: as distPY, but for DPM+PEAR method

Author(s)

Magdalena Strauss

---

computeSumClustPEAR *computeSumClustPear*

Description

Compute summary PSMs (posterior similarity matrices) from a set of multiple PSMs obtained for instance by means of subsampling. This implements the Dirichlet process and Pitman-Yor process based methods for combining PSMs proposed in Strauss et al. Unravelling shared pseudo-trajectories at single-cell resolution. Internal function

Usage

```
computeSumClustPEAR(PSM, maxCl = 10)
```

Arguments

PSM	posterior similarity matrix

Value

Summary clustering computed using the PEAR criterion (Fritsch and Ickstadt, 2009, using the mcclust package (Fritsch, 2012))

Author(s)

Magdalena Strauss

computeWeightsSumClust *computeWeightsSumClust*

Description

Internal function

Usage

```
computeWeightsSumClust(allocs)
```

Value

PSM and summary clustering obtained from Pitman-Yor process with allocs as input, weights used to compute the summary PSM from the individual PSMs

Author(s)

Magdalena Strauss

computeWeightsSumClustDPM *computeWeightsSumClustDPM*

Description

Internal function

Usage

```
computeWeightsSumClustDPM(allocs)
```

Value

PSM and summary clustering obtained from Dirichlet process with allocs as input, weights used to compute the summary PSM from the individual PSMs

Author(s)

Magdalena Strauss

`processPSMs` *processPSMs*

Description

processPSMs

Usage

```
processPSMs(PSMs)
```

Arguments

`PSMs`	3-dimensional array of PSMs, for each j PSMs[,,j] is the PSM of subsampled chain j

Value

weightedPSM: weighted summary PSM obtained using a Pitman-Yor process mixture model with variable selection

sumClustPEAR: final summary clustering obtained from weightedPSM using the PEAR criterion

weightedPSM_DP weighted summary PSM obtained using a Dirichlet process mixture model with variable selection

sumClustPEAR_DP: final summary clustering obtained from weightedPSM_DP using the PEAR criterion

weights: weights which were used for the computation of the summary PSM (Pitman-Yor based model)

weights_DP: weights which were used for the computation of the summary PSM (Dirichlet based model)

Author(s)

Magdalena Strauss

`RhatConcentration` *RhatConcentration*

Description

RhatConcentration

Usage

```
RhatConcentration(concentrationSamples, nChainsTest)
```

Arguments

concentrationSamples:	nIterations x nChains matrix of nIterations samples of the concentration parameter alpha for each of nChains subsampled chains
nChainsTest:	number of subsamples for which we want to test if the number of sufficient for convergence

Value

GR-statistics across subsampled chains (see Strauss et al. 2019) for groups of

Author(s)

Magdalena Strauss

Index