Package ‘anchoredDistr’

June 20, 2017

Type Package
Title Post-Processing for the Method of Anchored Distributions
Version 1.0.3
Author Heather Savoy
Maintainer Heather Savoy <frystacka@berkeley.edu>
Depends R (>= 3.1.0)
License GPL
LazyData TRUE
Imports np, RSQLite, DBI, methods, stats, reshape2, Rmisc, ggplot2, plyr, dplyr
Collate 'anchoredDistr.R' 'MADproject.R' 'calcLikelihood.R'
  'calcPosterior.R' 'data.R' 'plot.R' 'print.R' 'readMAD.R'
  'reduceData.R' 'test_convergence.R'
RoxygenNote 6.0.1
Suggests knitr, rmarkdown
VignetteBuilder knitr
NeedsCompilation no
Repository CRAN
Date/Publication 2017-06-20 06:35:54 UTC

R topics documented:

anchoredDistr ......................................................... 2
anchoredDistr  

**anchoredDistr**: A package for post-processing of the Method of Anchored Distributions.

---

**Description**

The anchoredDistr package provides functions for post-processing applications of the Method of Anchored Distributions (MAD) such as viewing data, calculating likelihood and posterior distributions, testing convergence of likelihood values, and applying dimension reduction.

---

**calcLikelihood**  

*Calculate the likelihood for the samples in a MADproject object.*

---

**Description**

calcLikelihood returns an updated MADproject with the likelihood values based on the observation and realization data in the MADproject or, optionally, a subset thereof.

**Usage**

```r
calcLikelihood(proj, dsubset, num_realz = max(proj@realizations$rid),
               samples = 1:proj@numSamples)
```

```r
## S4 method for signature 'MADproject,numeric'
calcLikelihood(proj, dsubset,
               num_realz = max(proj@realizations$rid), samples = 1:proj@numSamples)
```

```r
## S4 method for signature 'MADproject,ANY'
calcLikelihood(proj, num_realz, samples)
```
calcPosterior

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>proj</td>
<td>The MADproject object with data read from the MAD# databases.</td>
</tr>
<tr>
<td>dsubset</td>
<td>The subset of inversion data to use for the likelihood calculations.</td>
</tr>
<tr>
<td>num_realz</td>
<td>The number of realizations to use in the likelihood calculation (defaults to all in the realizations slot)</td>
</tr>
<tr>
<td>samples</td>
<td>A vector of sample IDs for which to calculate likelihood values (defaults to all available in the realizations slot)</td>
</tr>
</tbody>
</table>

Details

The likelihood calculation utilizes the **np** package for non-parametric density estimation with all inversion data as dependent (i.e. multivariate likelihood distributions are estimated).

Value

proj The updated MADproject object with a filled likelihood slot.

Methods (by class)

- proj = MADproject, dsubset = numeric: Calculates the likelihood using a subset dsubset of inversion data zid
- proj = MADproject, dsubset = ANY: Calculates the likelihood using all inversion data zid

Examples

data(pumping)
pumping <- calcLikelihood(pumping, 100) #Inversion data as time step 100

calcPosterior Return

Calculate the posterior for a MADproject object.

Description

calcPosterior returns a MADproject object with the posterior values for each sample

Usage

calcPosterior(proj)

### S4 method for signature 'MADproject'
calcPosterior(proj)

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>proj</td>
<td>The MADproject object with the likelihood slot filled.</td>
</tr>
</tbody>
</table>
Value

proj An updated MADproject object with the posterior slot filled.

Methods (by class)

- MADproject: Calculate the posterior value for each sample

Examples

data(pumping)
pumping.100 <- calcLikelihood(pumping, 100)
pumping.100 <- calcPosterior(pumping.100)

MADproject-class
An S4 class to represent a MAD project

Description

An S4 class to represent a MAD project

Slots

madname  Character string, the name of the MAD# project
resultname  Character string, the name of the MAD# project’s result
xpath  Character string, the path where to find the MAD# project’s result
numLocations  A numeric value for the number of measurement locations
numTimesteps  A numeric value for the number of time steps in series
numSamples  A numeric value for the number of samples
numAnchors  A numeric value for the number of local parameters to be inferred
numTheta  A numeric value for the number of global parameters to be inferred
truevalues  A numeric vector of length numAnchors*numTheta that contains the true values of the parameters being inferred, if known (i.e. for validation)
observations  A numeric vector containing the observed values of inversion data
realizations  A data.frame containing the the ensemble of simulated inversion data for each sample
priors  A data.frame containing the values sampled from the priors for each parameter to be inferred plus estimated marginal density
likelihoods  A data.frame of likelihood values calculated for each sample
posterior  A data.frame containing the posterior values estimated for each sample of each parameter
**Description**

plotMAD plots the data contained in various slots of the provided MADproject. If no specific slot is specified, all slots that can be plotted will be.

**Usage**

```r
plotMAD(x, y)
```

```r
## S4 method for signature 'MADproject,ANY'
plotMAD(x)
```

```r
## S4 method for signature 'MADproject,character'
plotMAD(x, y)
```

**Arguments**

- `x` The MADproject object
- `y` character string from the following options: "observations", "realizations", "priors", or "posteriors"

**Details**

All plots utilize `ggplot2` plotting functions. The slots that can be plotted are priors, observations, realizations, and posteriors.

**Methods (by class)**

- `x = MADproject, y = ANY`: Plots all available slots in the MADproject object
- `x = MADproject, y = character`: Plots the `y` slot in the MADproject object

**Examples**

```r
data(pumping)
plotMAD(pumping, "realizations")
plotMAD(pumping, "priors")
plotMAD(pumping, "posteriors")
plotMAD(pumping)
```
print.MADproject  
Print the data contained in MADproject object slots.

Description

print.MADproject plots information on the contents in the slots of the provided MADproject.

Usage

```r
## S3 method for class 'MADproject'
print(x, ...)
```

Arguments

- `x` The MADproject object
- `...` Not supported

Examples

```r
data(pumping)
print(pumping)
```

---

A dataset containing a MADproject object called 'pumping'

Usage

```r
data(pumping)
```

Format

A MADproject object with an observed time series, three prior samples, realizations for each sample, and true values for validation where the parameter to be inferred is the global mean of the natural log conductivity of a 2D aquifer.

- `@observations` a time series of drawdown, in meters
- `@numSamples` the value 3, [unitless]
- `@realizations` the ensemble of simulated drawdown time series
**readMAD**

Read the SQLite databases from MAD# into the MADproject object.

---

**Description**

`readMAD` returns an updated MADproject object with data from the MAD# databases.

**Usage**

```r
readMAD(proj, location)
```

// S4 method for signature 'MADproject,numeric'
readMAD(proj, location)

**Arguments**

- **proj**: The MADproject object with the slots madname, resultname, and xpath specified.
- **location**: The measurement ID(s) to be read from the MAD# databases

**Value**

- **proj**: An updated MADproject object with slots numTimesteps, numLocations, numSamples, numAnchors, numTheta, observations, priors, truevalues (if present), and realizations filled in from the MAD# databases.

**Methods (by class)**

- `proj = MADproject, location = numeric`: Reads the MAD# databases for information related to location

**Examples**

```r
## Not run:
example <- new("MADproject", madname="Example", resultname="results", xpath=getwd())
example <- readMAD(example, 1:3) #Observations 1-3
## End(Not run)
```
reduceData

Apply the requested dimension reduction technique to the inversion data for the given MADproject object

Description
reduceData returns a modified MADproject object where the observations and realizations slots are replaced by a lower-dimensional representation

Usage
reduceData(proj, method, params, ...)

## S4 method for signature 'MADproject, 'function', ANY'
reduceData(proj, method)

## S4 method for signature 'MADproject, 'function', 'function''
reduceData(proj, method, params, ...)

Arguments
- **proj**: MADproject object that has been filled by MAD# databases
- **method**: function to apply to data
- **params**: function for estimating the initial values of the parameters of method based on the data, if method has parameters
- **...**: additional arguments as needed for nls to fit method to the data

Value
proj MADproject object with reduced dimensions for inversion

Methods (by class)
- proj = MADproject, method = function, params = ANY: Fit method without any parameters (e.g., min)
- proj = MADproject, method = function, params = function: Fit method with parameters and initial guesses for the parameter values

Examples
data(pumping)
pumping.min <- reduceData(pumping, min)
plotMAD(pumping.min, "realizations")
testConvergence

Test (visually) the convergence of a MADproject object.

Description

testConvergence returns a plot to help the user visualize if there are enough realizations in the project for converged likelihood values.

Usage

testConvergence(proj, dsubset, samples = 1:proj@numSamples, NR = 10, NS = 7)

## S4 method for signature 'MADproject,numeric'
testConvergence(proj, dsubset,
  samples = 1:proj@numSamples, NR = 10, NS = 7)

## S4 method for signature 'MADproject,ANY'
testConvergence(proj, samples, NR, NS)

Arguments

- **proj**: The MADproject object to be tested.
- **ds subset**: The subset of inversion data to use for the likelihood calculations.
- **samples**: A vector of sample IDs to sample from to calculate likelihood values (defaults to all available in the MADproject object)
- **NR**: The number of different realization totals for which to calculate likelihood values (defaults to 10)
- **NS**: The number of randomly selected samples to test (defaults to 7) out of samples

Methods (by class)

- proj = MADproject, dsubset = numeric: Tests the convergence using a subset dsubset of inversion data zid
- proj = MADproject, dsubset = ANY: Tests the convergence using all inversion data zid

Examples

```r
## Not run:
data(pumping)
testConvergence(pumping, dsubset=1:3) #Inversion data as time steps 1-3

## End(Not run)
```
tutorial anchoredDistr example data

Description

A dataset containing a MADproject object called 'tutorial'

Usage

data(tutorial)

Format

A MADproject object with an observed head at three wells, 145 prior samples, realizations for each sample, and true values for validation where the parameters to be inferred are four anchors of the local decimal log transmissivity of a 1D aquifer

@observations three head measurements, in meters
@realizations the ensemble of simulated head measurements
@numAnchors the value 4
@priors the samples for each anchor’s prior distribution based on the kriging mean and variance at the anchor locations

Source

http://www.mad.codeplex.com/releases/
Index

*Topic datasets
  pumping, 6
  tutorial, 10

anchoredDistr, 2
anchoredDistr-package (anchoredDistr), 2

calclikelihood, 2
calclikelihood,MADproject,ANY-method
  (calclikelihood), 2
calclikelihood,MADproject,numeric-method
  (calclikelihood), 2
calcPosterior, 3
calcPosterior,MADproject-method
  (calcPosterior), 3

MADproject (MADproject-class), 4
MADproject-class, 4

plotMAD, 5
plotMAD,MADproject,ANY-method
  (plotMAD), 5
plotMAD,MADproject,character-method
  (plotMAD), 5

print.MADproject, 6
pumping, 6

readMAD, 7
readMAD,MADproject,numeric-method
  (readMAD), 7
reduceData, 8
reduceData,MADproject,function,ANY-method
  (reduceData), 8
reduceData,MADproject,function,function-method
  (reduceData), 8

testConvergence, 9
testConvergence,MADproject,ANY-method
  (testConvergence), 9
testConvergence,MADproject,numeric-method
  (testConvergence), 9

tutorial, 10