

"Test after modifying functions to simulate data in HMTree package"

Heejung Shim

March 28, 2016

Contents

1	Modify simulate.data function	1
2	Add sim.data function	3

1 Modify simulate.data function

Let's check if modification doe not change outputs.

```
path.HMTree.package = "/home/hjshim/d/hjshim/software/HMTree/"
## read all functions
source(paste0(path.HMTree.package, "R/doppler.f.R"))
source(paste0(path.HMTree.package, "R/heaviside.f.R"))
source(paste0(path.HMTree.package, "R/spike.f.R"))
source(paste0(path.HMTree.package, "R/simulate_true_curve.R"))
source(paste0(path.HMTree.package, "R/simulate_noisy_curve.R"))
## read simulate.data function
source(paste0(path.HMTree.package, "R/simulate.data.R"))

curve.length = 1024
model.mode = 'Poisson'
num.groups=1
num.samples=10
WHMT.sigma=0.5
plotfn=FALSE
```

```

set.seed(666)
res.pois = simulate.data(curve.length=curve.length, model.mode=model.mode, num.
  groups=num.groups,num.samples=num.samples, WHMT.sigma=WHMT.sigma,plotfn=
  plotfn)
str(res.pois)
## List of 4
## $ lambda_vec : num [1:1024] 58.3 57.7 57.1 56.5 55.9 ...
## $ lambda_diff_vec: num [1:1024] 0 0 0 0 0 0 0 0 0 0 ...
## $ data : num [1:10, 1:1024] 63 66 63 62 50 52 55 71 57 55 ...
## $ data_mean : num [1:1024] 59.4 60.5 58.7 59.7 55.5 53.3 52.1 52.5 51.6 54.5
...

curve.length = 1024
model.mode = 'Normal'
num.groups=1
num.samples=10
WHMT.sigma=0.5
plotfn=FALSE

set.seed(666)
res.normal = simulate.data(curve.length=curve.length, model.mode=model.mode,
  num.groups=num.groups,num.samples=num.samples, WHMT.sigma=WHMT.sigma,plotfn
  =plotfn)
str(res.normal)
##List of 4
## $ lambda_vec : num [1:1024] 58.3 57.7 57.1 56.5 55.9 ...
## $ lambda_diff_vec: num [1:1024] 0 0 0 0 0 0 0 0 0 0 ...
## $ data : num [1:1024] 58.6 58.7 56.9 57.5 54.8 ...
## $ data_mean : num [1:1024] 58.6 58.7 56.9 57.5 54.8 ...

#####
## let's compare those outputs with files Or generated.
#####
### You need to modify the following two lines.
setwd('/home/hjshim/d/hjshim/projects/HMTprior/HMTTree_test/')
hmt.repodir.path = scan('.hmt.repodir.txt', what="")

pois.true = as.numeric(t(read.table(paste0(hmt.repodir.path, "data/sim/test_hmt
  _poiss.true.txt"))))

```

```

pois.mean = as.numeric(t(read.table(paste0(hmt.repodir.path, "data/sim/test_hmt_
_pois.noisy.txt"))))
pois.data = t(read.table(paste0(hmt.repodir.path, "data/sim/test_hmt_pois.
individual.noisy.txt")))

normal.true = as.numeric(t(read.table(paste0(hmt.repodir.path, "data/sim/test_
hmt_gaussian.true.txt"))))
normal.mean = as.numeric(t(read.table(paste0(hmt.repodir.path, "data/sim/test_
hmt_gaussian.noisy.txt"))))

sum(abs(res.pois$lambda_vec - pois.true) < 0.001)
sum(abs(res.pois$data_mean - pois.mean) < 0.001)
sum(abs(res.pois$data - pois.data) < 0.001)
## [1] 1024
## [1] 1024
## [1] 10240

sum(abs(res.normal$lambda_vec - normal.true) < 0.001)
sum(abs(res.normal$data_mean - normal.mean) < 0.001)
sum(abs(res.normal$data - normal.mean) < 0.001)
## [1] 1024
## [1] 1024
## [1] 1024

```

2 Add sim.data function

Here, I rename `simulate_noisy_curve` as `sim.data` and made multiple changes including an option which simulate multiple samples under normal model. Let's check if modification does not change outputs.

```

path.HMTree.package = "/home/hjshim/d/hjshim/software/HMTree/"
## read all functions
source(paste0(path.HMTree.package, "R/doppler.f.R"))
source(paste0(path.HMTree.package, "R/heaviside.f.R"))
source(paste0(path.HMTree.package, "R/spike.f.R"))
source(paste0(path.HMTree.package, "R/simulate_true_curve.R"))
##source(paste0(path.HMTree.package, "R/simulate_noisy_curve.R"))
## read simulate.data function

```

```

source(paste0(path.HMTree.package, "R/simulate.data.R"))

curve.length = 1024
model.mode = 'Poisson'
num.groups=1
num.samples=10
WHMT.sigma=0.5

set.seed(666)
res.pois = simulate.data(curve.length=curve.length, model.mode=model.mode, num.
  groups=num.groups,num.samples=num.samples, WHMT.sigma=WHMT.sigma)
str(res.pois)
## List of 2
## $ true.mean.curve: num [1:1024] 58.3 57.7 57.1 56.5 55.9 ...
## $ data : int [1:10, 1:1024] 63 66 63 62 50 52 55 71 57 55 ...

curve.length = 1024
model.mode = 'Normal'
num.groups=1
num.samples=1
WHMT.sigma=0.5

set.seed(666)
res.normal = simulate.data(curve.length=curve.length, model.mode=model.mode,
  num.groups=num.groups,num.samples=num.samples, WHMT.sigma=WHMT.sigma)
str(res.normal)
## List of 2
## $ true.mean.curve: num [1:1024] 58.3 57.7 57.1 56.5 55.9 ...
## $ data : num [1, 1:1024] 58.6 58.7 56.9 57.5 54.8 ...

curve.length = 1024
model.mode = 'Normal'
num.groups=1
num.samples=10
WHMT.sigma=0.5

set.seed(666)
res.normal.10 = simulate.data(curve.length=curve.length, model.mode=model.mode,
  num.groups=num.groups,num.samples=num.samples, WHMT.sigma=WHMT.sigma)
str(res.normal.10)

```

```

## List of 2
## $ true.mean.curve: num [1:1024] 58.3 57.7 57.1 56.5 55.9 ...
## $ data : num [1:10, 1:1024] 58.6 58.8 58.3 58.4 57.7 ...

#####
## let's compare those outputs with files Or generated.
#####
### You need to modify the following two lines.
setwd('/home/hjshim/d/hjshim/projects/HMTprior/HMTtree_test/')
hmt.repodir.path = scan('.hmt.repodir.txt', what="")

pois.true = as.numeric(t(read.table(paste0(hmt.repodir.path, "data/sim/test_hmt_
_pois.true.txt"))))
pois.mean = as.numeric(t(read.table(paste0(hmt.repodir.path, "data/sim/test_hmt_
_pois.noisy.txt"))))
pois.data = t(read.table(paste0(hmt.repodir.path, "data/sim/test_hmt_pois.
individual.noisy.txt")))

normal.true = as.numeric(t(read.table(paste0(hmt.repodir.path, "data/sim/test_
hmt_gaussian.true.txt"))))
normal.mean = as.numeric(t(read.table(paste0(hmt.repodir.path, "data/sim/test_
hmt_gaussian.noisy.txt"))))

sum(abs(res.pois$true.mean.curve - pois.true) < 0.001)
sum(abs(colMeans(res.pois$data) - pois.mean) < 0.001)
sum(abs(res.pois$data - pois.data) < 0.001)
## [1] 1024
## [1] 1024
## [1] 10240

sum(abs(res.normal$true.mean.curve - normal.true) < 0.001)
sum(abs(colMeans(res.normal$data) - normal.mean) < 0.001)
sum(abs(res.normal$data - normal.mean) < 0.001)
## [1] 1024
## [1] 1024
## [1] 1024

#####

```

```

## Check if function simulates proper data?
#####

curve.length = 1024
model.mode = 'Poisson'
num.groups=1
num.samples=30
WHMT.sigma=0.5
plotfn=FALSE

set.seed(666)
res.pois = simulate.data(curve.length=curve.length, model.mode=model.mode, num.
  groups=num.groups,num.samples=num.samples, WHMT.sigma=WHMT.sigma,plotfn=
  plotfn)

str(res.pois)
## List of 2
## $ true.mean.curve: num [1:1024] 58.3 57.7 57.1 56.5 55.9 ...
## $ data : int [1:10, 1:1024] 63 66 63 62 50 52 55 71 57 55 ...

curve.length = 1024
model.mode = 'Normal'
num.groups=1
num.samples=1
WHMT.sigma=0.5
plotfn=FALSE

set.seed(666)
res.normal = simulate.data(curve.length=curve.length, model.mode=model.mode,
  num.groups=num.groups,num.samples=num.samples, WHMT.sigma=WHMT.sigma,plotfn=
  =plotfn)
str(res.normal)
## List of 2
## $ true.mean.curve: num [1:1024] 58.3 57.7 57.1 56.5 55.9 ...
## $ data : num [1, 1:1024] 58.6 58.7 56.9 57.5 54.8 ...

```
