R documentation

of 'MethylTransition/man/ParameterEstimation.Rd'

December 20, 2018

ParameterEstimation   *"ParameterEstimation"*

Description

Estimated the parameters that represent the probabilities of three active DNA methylation change types during n cell cycle(s).

Usage

```
ParameterEstimation(observation_matrix, iter = 50, cell_cycle = 1)
```

Arguments

observation_matrix

The transition matrix(5×5) from the original state to the terminational state.

	original_class1	original_class2	original_class3	original_class4	original_class5
terminational_class1	a1	b1	c1	d1	e1
terminational_class2	a2	b2	c2	d2	e2
terminational_class3	a3	b3	c3	d3	e3
terminational_class4	a4	b4	c4	d4	e4
terminational_class5	a5	b5	c5	d5	e5

observation_matrix[1,1](a1) is the ratio of original_class1 to terminational_class1,observation_matr is the ratio of original_class2 to terminational_class1 and so on
observation_matrix[2,1](a2) is the ratio of original_class1 to terminational_class2,observation_matr is the ratio of original_class2 to terminational_class2 and so on
The sum of the ratio that original_class1 change to 5 classes should be 1. That is a1+a2+a3+a4+a5=1.

iter   The iteration times of the parameter estimation using the Newton-Raphson method with different initial guesses.

cell_cycle   The cell cycle times from the original state to the terminational state.

Details

The transition matrix of this model describes the changes of DNA methylation during one cell cycle in three steps: passive demethylation by DNA replication, active DNA methylation changes affected by DNA methylation-modifying enzymes and DNA methylation combinations during homologous recombination. For each CpG site in a chromsome, the methylation states are one of these four types: 0-0, 0-1, 1-0, 1-1. The transition matrix after DNA replication would be :

	original(0-0)	original(0-1)	original(1-0)	original(1-1)
after_replication(0-0)	1	a	$1-a$	0
after_replication(0-1)	0	$1-a$	0	$1-a$
after_replication(1-0)	0	0	0	a
after_replication(1-1)	0	0	0	0

among this matrix ,a is the methylation change probability with DNA replication and equal to 0.5. Then the transition matrix after active DNA methylation changes would be:

	after_replication(0-0)	after_replication(0-1)	after_replication(1-0)
after_enzymemodifying(0-0)	$(1-u) \times (1-u)$	$(1-u-p+u \times p) \times d$	$d \times (1-u-p+u \times p)$
after_enzymemodifying(0-1)	$u \times (1-u)$	$(1-u-p+u \times p) \times (1-d)$	$d \times (u+p-u \times p)$
after_enzymemodifying(1-0)	$u \times (1-u)$	$(u+p-u \times p) \times d$	$(1-d) \times (1-u-p+u \times p)$
after_enzymemodifying(1-1)	$u \times u$	$(u+p-u \times p) \times (1-d)$	$(1-d) \times (u+p-u \times p)$

The paramater u described the methylation probablity on CpG site. The paramater d described the de-methylation probablity on 5mCpG site. The paramater p described the methylation probablity on semi-CpG site. Thus the mathlation state change is $P = Penzymemodifying \cdot Preplication \cdot original_classes$ We using $t_{i,j}$ represent the i and j vector of the matrix $Penzymemodifying \cdot Preplication$ Then two chromsomes are combined during homologous recombination. The observed DNA methylation of each CpG site is the combination of the methylation types in both chromsomes. The observed transition matrix would be :

	original_class1(0)	original_class2(1/4)	original_class3(1/2)	original_class4(3/4)	orig
terminational_class1(0)	$x_{1,1}$	$x_{1,2}$	$x_{1,3}$	$x_{1,4}$	
terminational_class2(1/4)	$x_{2,1}$	$x_{2,2}$	$x_{2,3}$	$x_{2,4}$	
terminational_class3(1/2)	$x_{3,1}$	$x_{3,2}$	$x_{3,3}$	$x_{3,4}$	
terminational_class4(3/4)	$x_{4,1}$	$x_{4,2}$	$x_{4,3}$	$x_{4,4}$	
terminational_class5(1)	$x_{5,1}$	$x_{5,2}$	$x_{5,3}$	$x_{5,4}$	

and

$$x_{1,1} = t_{1,1} \times t_{1,1}$$

$$x_{1,1} = t_{,1,1} \times t_{,1,1}$$

$$x_{1,2} = 1/4 \times (t_{1,1} \times t_{1,2} + t_{1,1} \times t_{1,3} + t_{1,2} \times t_{1,1} + t_{1,3} \times t_{1,1})$$

$$x_{1,3} = 1/6 \times (t_{1,1} \times t_{1,4} + t_{1,2} \times t_{1,2} + t_{1,2} \times t_{1,3} + t_{1,3} \times t_{1,2} + t_{1,3} \times t_{1,3} + t_{1,4} \times t_{1,1})$$

$$x_{1,4} = 1/4 \times (t_{1,2} \times t_{1,4} + t_{1,3} \times t_{1,4} + t_{1,4} \times t_{1,2} + t_{1,4} \times t_{1,3})$$

$$x_{1,5} = t_{1,4} \times t_{1,4}$$

$$x_{2,1} = t_{1,1} \times t_{2,1} + t_{1,1} \times t_{3,1} + t_{2,1} \times t_{1,1} + t_{3,1} \times t_{1,1}$$

$$x_{2,2} = 1/4 \times (t_{1,1} \times t_{2,2} + t_{1,1} \times t_{2,3} + t_{1,2} \times t_{2,1} + t_{1,3} \times t_{2,1} + t_{1,1} \times t_{3,2} + t_{1,1} \times t_{3,3} + t_{1,2} \times t_{3,1} + t_{1,3} \times t_{3,1} + t_{2,1} \times t_{1,2} + t_{2,1} \times$$

$$x_{2,3} = 1/6 \times (t_{1,1} \times t_{2,4} + t_{1,2} \times t_{2,2} + t_{1,2} \times t_{2,3} + t_{1,3} \times t_{2,2} + t_{1,3} \times t_{2,3} + t_{1,4} \times t_{2,1} + t_{1,1} \times t_{3,4} + t_{1,2} \times t_{3,2} + t_{1,2} \times t_{3,3} + t_{1,3} \times$$

$$x_{2,4} = 1/4 \times (t_{1,2} \times t_{2,4} + t_{1,3} \times t_{2,4} + t_{1,4} \times t_{2,2} + t_{1,4} \times t_{2,3} + t_{1,2} \times t_{3,4} + t_{1,3} \times t_{3,4} + t_{1,4} \times t_{3,2} + t_{1,4} \times t_{3,3} + t_{2,2} \times t_{1,4} + t_{2,3} \times$$

$$x_{2,5} = t_{1,4} \times t_{2,4} + t_{1,4} \times t_{3,4} + t_{2,4} \times t_{1,4} + t_{3,4} \times t_{1,4}$$

$$x_{3,1} = t_{1,1} \times t_{4,1} + t_{2,1} \times t_{2,1} + t_{2,1} \times t_{3,1} + t_{3,1} \times t_{2,1} + t_{3,1} \times t_{3,1} + t_{4,1} \times t_{1,1}$$

$$x_{3,2} = 1/4 \times (t_{1,1} \times t_{4,2} + t_{1,1} \times t_{4,3} + t_{1,2} \times t_{4,1} + t_{1,3} \times t_{4,1} + t_{2,1} \times t_{2,2} + t_{2,1} \times t_{2,3} + t_{2,2} \times t_{2,1} + t_{2,3} \times t_{2,1} + t_{2,1} \times t_{3,2} + t_{2,1} \times$$

$$x_{3,3} = 1/6 \times (t_{1,1} \times t_{4,4} + t_{1,2} \times t_{4,2} + t_{1,2} \times t_{4,3} + t_{1,3} \times t_{4,2} + t_{1,3} \times t_{4,3} + t_{1,4} \times t_{4,1} + t_{2,1} \times t_{2,4} + t_{2,2} \times t_{2,2} + t_{2,2} \times t_{2,3} + t_{2,3} \times$$

$$x_{3,4} = 1/4 \times (t_{1,2} \times t_{4,4} + t_{1,3} \times t_{4,4} + t_{1,4} \times t_{4,2} + t_{1,4} \times t_{4,3} + t_{2,2} \times t_{2,4} + t_{2,3} \times t_{2,4} + t_{2,4} \times t_{2,2} + t_{2,4} \times t_{2,3} + t_{2,2} \times t_{3,4} + t_{2,3} \times$$

$$x_{3,5} = t_{1,4} \times t_{4,4} + t_{2,4} \times t_{2,4} + t_{2,4} \times t_{3,4} + t_{3,4} \times t_{2,4} + t_{3,4} \times t_{3,4} + t_{4,4} \times t_{1,4}$$

$$x_{4,1} = t_{2,1} \times t_{4,1} + t_{3,1} \times t_{4,1} + t_{4,1} \times t_{2,1} + t_{4,1} \times t_{3,1}$$

$$x_{4,2} = 1/4 \times (t_{2,1} \times t_{4,2} + t_{2,1} \times t_{4,3} + t_{2,2} \times t_{4,1} + t_{2,3} \times t_{4,1} + t_{3,1} \times t_{4,2} + t_{3,1} \times t_{4,3} + t_{3,2} \times t_{4,1} + t_{3,3} \times t_{4,1} + t_{4,1} \times t_{2,2} + t_{4,1} \times$$

$$x_{4,3} = 1/6 \times (t_{2,1} \times t_{4,4} + t_{2,2} \times t_{4,2} + t_{2,2} \times t_{4,3} + t_{2,3} \times t_{4,2} + t_{2,3} \times t_{4,3} + t_{2,4} \times t_{4,1} + t_{3,1} \times t_{4,4} + t_{3,2} \times t_{4,2} + t_{3,2} \times t_{4,3} + t_{3,3} \times$$

$$x_{4,4} = 1/4 \times (t_{2,2} \times t_{4,4} + t_{2,3} \times t_{4,4} + t_{2,4} \times t_{4,2} + t_{2,4} \times t_{4,3} + t_{3,2} \times t_{4,4} + t_{3,3} \times t_{4,4} + t_{3,4} \times t_{4,2} + t_{3,4} \times t_{4,3} + t_{4,2} \times t_{2,4} + t_{4,3} \times$$

$$x_{4,5} = t_{2,4} \times t_{4,4} + t_{3,4} \times t_{4,4} + t_{4,4} \times t_{2,4} + t_{4,4} \times t_{3,4}$$

$$x_{5,1} = t_{4,1} \times t_{4,1}$$

$$x_{5,2} = 1/4 \times (t_{4,1} \times t_{4,2} + t_{4,1} \times t_{4,3} + t_{4,2} \times t_{4,1} + t_{4,3} \times t_{4,1})$$

$$x_{5,3} = 1/6 \times (t_{4,1} \times t_{4,4} + t_{4,2} \times t_{4,2} + t_{4,2} \times t_{4,3} + t_{4,3} \times t_{4,2} + t_{4,3} \times t_{4,3} + t_{4,4} \times t_{4,1})$$

$$x_{5,4} = 1/4 \times (t_{4,2} \times t_{4,4} + t_{4,3} \times t_{4,4} + t_{4,4} \times t_{4,2} + t_{4,4} \times t_{4,3})$$

$$x_{5,5} = t_{4,4} \times t_{4,4}$$

The cost function was defined by

$$f_{cost} = \sum_{i=1,j=1}^{n=5} (o_{i,j} - x_{i,j})$$

and minimized using the Newton-Raphson method.

Value

estimated_parameters The estimated parameters using the maximum likelihood estimation and the Newton-Raphson method.

predicted_matrix The calculated transition matrix using the estimated parameters.

References

Zhao, C. et.al.(2018). A DNA methylation state transition model reveals the programmed epigenetic heterogeneity in pre-implantation embryos. Under revision.

Examples

```
# Let's start from a transtion matrix
observation_matrix <- matrix(c(0.9388,0.0952,0.0377,0,0.0001,
                               0.0497,0.5873,0.1887,0.0344,0.0149,
                               0.0096,0.2381,0.4151,0.0653,0.0876,
                               0.0019,0.0635,0.3396,0.6151,0.253,
                               0,0.0159,0.0189,0.2852,0.6444),5,5,byrow=T)

# The DNA methylation states change from original state to the terminational state after 1 cell cycle.
ParameterEstimation(observation_matrix,iter=30,cell_cycle=1)

# The DNA methylation states change from original state to the terminational state after 2 cell cycle.
# ParameterEstimation(observation_matrix,iter=1,cell_cycle=2)
# if this function was not successful to estimated the function, you may try more iterations with different in
ParameterEstimation(observation_matrix,iter=50,cell_cycle=2)

# The DNA methylation states change from original state to the terminational state after 30 cell cycle.
ParameterEstimation(observation_matrix,iter=50,cell_cycle=30)
```

Index

ParameterEstimation,