**lm**

lm(formula, data,...)

----------------

for (i in 1:24)

B = sigMat[,2]

CD4T = sigMat[,3]

CD8T = sigMat[,4]

G = sigMat[,5]

Mo =sigMat[,6]

NK = sigMat[,7]

RBC = sigMat[,8]

*lmCellType* = lm( mixData[ ,i+1] ~ Mo + G + RBC + NK + B + CD4T + CD8T)

----------------

**nnls** *The Lawson-Hanson NNLS implemention of non-negative least squares*

**Usage**

nnls(A, b)

**Arguments**

A numeric matrix with m rows and n columns (like our Signature Matrix)

b numeric vector of length m (like each sample data from Mix Data matrix)

----------------

for (i in 1:24)

*nnlmCellType* = nnls(as.matrix(sigMat[ ,2:8] ), mixData[ ,i+1])

----------------

Testing the code:

> sigMat <- read.csv("sigMat.csv")

> head(sigMat)

X B CD4T CD8T G Mo NK RBC

1 cg00008647 0.072352 0.074320 0.075871 0.072724 0.076460 0.299599 0.098206

2 cg00025044 0.065995 0.070191 0.068315 0.068832 0.068543 0.068232 0.237987

3 cg00035969 0.046206 0.051926 0.058779 0.121245 0.056430 0.054227 0.298819

4 cg00058449 0.955530 0.331838 0.242658 0.210894 0.257774 0.074019 0.615413

5 cg00067758 0.160810 0.123704 0.117611 0.108495 0.109115 0.311912 0.125920

6 cg00080081 0.206366 0.404653 0.763532 0.123728 0.092218 0.391698 0.185771

> mixData <- read.csv("mix.csv")

> head(mixData)

F1 TS222 TS223 TS224 TS225 TS226 TS227 TS228 TS230 TS231 TS232 TS233 TS234 TS235 TS236 TS237 TS239 TS240 TS241

1 cg00008647 0.09215215 0.1004936 0.09963833 0.08647534 0.09683944 0.09012591 0.09003169 0.08404113 0.09807222 0.09503048 0.07873398 0.09948778 0.08319033 0.13356119 0.12322904 0.1065307 0.08416226 0.09947758

2 cg00025044 0.07875750 0.0718306 0.05976316 0.06629688 0.06779879 0.06505567 0.05125000 0.05690012 0.06422897 0.08014669 0.05671956 0.07764061 0.07467158 0.07225627 0.07424998 0.0855494 0.06070764 0.05203387

3 cg00375132 0.66036724 0.6184495 0.64738109 0.62751629 0.61732089 0.68072326 0.68866222 0.63161782 0.61419489 0.62655935 0.65432543 0.67761328 0.68379685 0.65617892 0.64056193 0.6763053 0.61492740 0.59953871

4 cg00058449 0.38287571 0.3935523 0.32232215 0.31977209 0.29101751 0.30704108 0.39895734 0.34230193 0.30787707 0.39424970 0.27924748 0.31809167 0.35398068 0.26503311 0.29868792 0.3457603 0.27385215 0.37261756

5 cg00067758 0.17042666 0.1822439 0.15842878 0.15692897 0.17434268 0.14846462 0.16829814 0.14742457 0.17809051 0.15723072 0.13447059 0.16744498 0.16442464 0.20055635 0.18771386 0.1839165 0.16250987 0.16155390

6 cg00080081 0.29916706 0.3164539 0.27483600 0.25019743 0.30984334 0.22790136 0.31671306 0.31443528 0.33581875 0.32164902 0.21688311 0.27521445 0.20584383 0.32667290 0.30337944 0.3256258 0.28333774 0.41941575

TS242 TS243 TS244 TS245 TS246 TS247

1 0.08654056 0.10468277 0.08597740 0.12363592 0.08502952 0.09755766

2 0.06564620 0.06694654 0.06950524 0.06382271 0.06080372 0.08605473

3 0.66255279 0.60962961 0.67351120 0.62742431 0.60526543 0.62977694

4 0.36220627 0.41521958 0.38230979 0.32389201 0.34415185 0.35190162

5 0.15555901 0.17173229 0.16585459 0.18816591 0.16449188 0.15842138

6 0.20418662 0.39988519 0.22089247 0.33580029 0.38962618 0.32068938

> i= 1

> no = sprintf("%01d", i)

> no

[1] "1"

> no2 = sprintf("%01d", i+221)

> no2

[1] "222"

> if (i+221 > 228){ no2 = sprintf("%01d", i+222)}

> no2

[1] "222"

> if (i+221 >238){no2 = sprintf("%01d", i+223)}

> no2

[1] "222"

> Mix = mixData[,i+1]

> print(head(Mix))

[1] 0.09215215 0.07875750 0.66036724 0.38287571 0.17042666 0.29916706

> model = nnls(as.matrix(sigMat[,2:8]), Mix)

> model

Nonnegative least squares model

x estimates: 0.1328745 0.2235356 0.1202559 0.2581776 0.08752862 0.1000892 0.1198779

residual sum-of-squares: 1.701

reason terminated: The solution has been computed successfully.

> coefs <- coef(model)

> coefs

[1] 0.13287449 0.22353560 0.12025592 0.25817762 0.08752862 0.10008918 0.11987785

> coefs[1:7]

[1] 0.13287449 0.22353560 0.12025592 0.25817762 0.08752862 0.10008918 0.11987785

> sum(coefs)

[1] 1.042339

> coefsNorm <- coefs\*100/ sum(coefs)

> coefsNorm

[1] 12.747720 21.445570 11.537118 24.769058 8.397325 9.602361 11.500848

> sampleName = paste("TS", no2, sep = "")

> sampleName

[1] "TS222"

> class(coefsNorm)

[1] "numeric"

> as.data.frame(coefsNorm)

coefsNorm

1 12.747720

2 21.445570

3 11.537118

4 24.769058

5 8.397325

6 9.602361

7 11.500848

> t(as.data.frame(coefsNorm))

[,1] [,2] [,3] [,4] [,5] [,6] [,7]

coefsNorm 12.74772 21.44557 11.53712 24.76906 8.397325 9.602361 11.50085

> if (i ==1){

+ write.table( t(as.data.frame(coefsNorm)), file = "testnnlm.csv", append=T, sep="," , row.names = sampleName, col.names=c("\",\"Mo","G","RBC", "NK", "B" , "CD4T" , "CD8T"))

+ }

Warning message:

In write.table(t(as.data.frame(coefsNorm)), file = "testnnlm.csv", :

appending column names to file

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| \ | \"Mo" | G | RBC | NK | B | CD4T | CD8T |
| TS222 | 12.74771986 | 21.44557034 | 11.53711823 | 24.76905783 | 8.397325414 | 9.602360586 | 11.50084774 |

--------------------

> i= 8

> no = sprintf("%01d", i)

> no

[1] "8"

> no2 = sprintf("%01d", i+221)

> no2

[1] "229"

> if (i+221 > 228){ no2 = sprintf("%01d", i+222)}

> no2

[1] "230"

> if (i+221 >238){no2 = sprintf("%01d", i+223)}

> no2

[1] "230"

> Mix = mixData[,i+1]

> print(head(Mix))

[1] 0.08404113 0.05690012 0.63161782 0.34230193 0.14742457 0.31443528

> head(sigMat[,2:8])

B CD4T CD8T G Mo NK RBC

1 0.072352 0.074320 0.075871 0.072724 0.076460 0.299599 0.098206

2 0.065995 0.070191 0.068315 0.068832 0.068543 0.068232 0.237987

3 0.046206 0.051926 0.058779 0.121245 0.056430 0.054227 0.298819

4 0.955530 0.331838 0.242658 0.210894 0.257774 0.074019 0.615413

5 0.160810 0.123704 0.117611 0.108495 0.109115 0.311912 0.125920

6 0.206366 0.404653 0.763532 0.123728 0.092218 0.391698 0.185771

> model = nnls(as.matrix(sigMat[,2:8]), Mix)

> model

Nonnegative least squares model

x estimates: 0.09740562 0.2220413 0.1409176 0.3797247 0.07873423 0.05733232 0.09314326

residual sum-of-squares: 1.737

reason terminated: The solution has been computed successfully.

> coefs <- coef(model)

> coefs

[1] 0.09740562 0.22204126 0.14091758 0.37972471 0.07873423 0.05733232 0.09314326

> coefsNorm <- coefs\*100/ sum(coefs)

> coefsNorm

[1] 9.109297 20.765124 13.178502 35.511556 7.363163 5.361674 8.710685

> sampleName = paste("TS", no2, sep = "")

> sampleName

[1] "TS230"

> as.data.frame(coefsNorm)

coefsNorm

1 9.109297

2 20.765124

3 13.178502

4 35.511556

5 7.363163

6 5.361674

7 8.710685

> t(as.data.frame(coefsNorm))

[,1] [,2] [,3] [,4] [,5] [,6] [,7]

coefsNorm 9.109297 20.76512 13.1785 35.51156 7.363163 5.361674 8.710685

> if (i ==1){

+ write.table( t(as.data.frame(coefsNorm)), file = "testnnlm.csv", append=T, sep="," , row.names = sampleName, col.names=c("\",\"Mo","G","RBC", "NK", "B" , "CD4T" , "CD8T"))

+ }else{

+ write.table( t(as.data.frame(coefsNorm)), file = "testnnlm.csv", append=T, sep="," , row.names = sampleName, col.names= F)

+ }

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| \ | \"Mo" | G | RBC | NK | B | CD4T | CD8T |
| TS222 | 12.74771986 | 21.44557034 | 11.53711823 | 24.76905783 | 8.397325414 | 9.602360586 | 11.50084774 |
| TS230 | 9.109297058 | 20.7651239 | 13.17850168 | 35.51155589 | 7.363163136 | 5.361673584 | 8.710684752 |