Measurement error model for correlation coefficient estimation

Beiying Ding
Robert Gentleman

October 17, 2016

Introduction

The `MeasurementError.cor` package fits a two-stage measurement error model for estimating correlation between two random variables under bivariate normality. It's application is perhaps most relevant for the gene expression data where both point and standard estimates are available. We have shown that the proposed measurement error corrected correlation estimate has lower bias compared with the usual sample pearson correlation. For details, refer to **?** as well as R help pages associated with each function.

The `cor.me.vector` and `cor.me.matrix` functions

The `cor.me.vector` calculates the measurement error model estimate of correlation between two observed vectors whereas `cor.me.matrix` calculates all pairwise measurement error model estimate of correlation in the matrix.

```
> library(MeasurementError.cor)
> exp <- matrix(abs(rnorm(100,1000,20)),ncol=10)
> se <- matrix(abs(rnorm(100,50,5)),ncol=10)
> cor.me.vector(exp[1,],se[1,],exp[2,],se[2,])

$estimate
      corr.me     corr.true           mu1           mu2            s1
  -0.97838743    0.53287537  994.23261742 1002.70590861    0.02269636
           s2
```

```
  24.12543892

$counts
function gradient
      69       66

$convergence
[1] 0
```

```
> cor.me.matrix(exp,se)

$corr.true
            [,1]       [,2]       [,3]       [,4]       [,5]       [,6]
 [1,]  1.0000000  0.5328754  0.7775642 -0.6440924 -0.3516747  0.5404121
 [2,]  0.5328754  1.0000000 -0.8986486  0.6358931  0.4239941  0.4023852
 [3,]  0.7775642 -0.8986486  1.0000000  0.3654855 -0.7681329 -0.7679597
 [4,] -0.6440924  0.6358931  0.3654855  1.0000000 -0.3822519 -0.2257236
 [5,] -0.3516747  0.4239941 -0.7681329 -0.3822519  1.0000000  0.6909843
 [6,]  0.5404121  0.4023852 -0.7679597 -0.2257236  0.6909843  1.0000000
 [7,]  0.6172037 -0.9014831  0.7864350  0.7572692 -0.9356390 -0.5894740
 [8,] -0.5365337  0.1502108  0.7819739  0.8255691  0.6574042 -0.4136084
 [9,]  0.3485890  0.9302826  0.7628944  0.8363437  0.8587852 -0.7952867
[10,]  0.7010484 -0.8762036 -0.8802381 -0.3382511  0.6547028 -0.4776167
            [,7]       [,8]       [,9]      [,10]
 [1,]  0.6172037 -0.5365337  0.3485890  0.7010484
 [2,] -0.9014831  0.1502108  0.9302826 -0.8762036
 [3,]  0.7864350  0.7819739  0.7628944 -0.8802381
 [4,]  0.7572692  0.8255691  0.8363437 -0.3382511
 [5,] -0.9356390  0.6574042  0.8587852  0.6547028
 [6,] -0.5894740 -0.4136084 -0.7952867 -0.4776167
 [7,]  1.0000000 -0.2775222  0.9691280 -0.8883205
 [8,] -0.2775222  1.0000000  0.7257698 -0.4794274
 [9,]  0.9691280  0.7257698  1.0000000  0.4005542
[10,] -0.8883205 -0.4794274  0.4005542  1.0000000

>
```

the quantity of interest, i.e. the model estimate of the correlation between the true value of two random variables whereas `cor.me` is the model estimate of correlation between the measurement errors of the two random variables. The second quantity may not be of interest. `mu1`,`mu2` and `s1`, `s2`

are the estimated mean and standard deviation of the two random variables. `cor.me.matrix` only returns the estimated correlation matrix.