Measurement error model for correlation coefficient estimation

Beiying Ding
Robert Gentleman

October 29, 2024

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 Ding and Gentleman (2003) 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            s2
   -0.8608527    -0.1231664   1007.4823404   1009.4220486     0.4702038     0.6422618
```

```
$counts
function gradient
      36       33

$convergence
[1] 0

> cor.me.matrix(exp,se)

$corr.true
             [,1]       [,2]        [,3]        [,4]        [,5]        [,6]
 [1,]   1.00000000 -0.1231664 -0.06338327 -0.21090361 -0.52659070  0.72724729
 [2,]  -0.12316639  1.0000000  0.67778519 -0.90718679  0.66835432  0.57871992
 [3,]  -0.06338327  0.6777852  1.00000000  0.26717687 -0.62668850  0.59225624
 [4,]  -0.21090361 -0.9071868  0.26717687  1.00000000 -0.74886860 -0.79313558
 [5,]  -0.52659070  0.6683543 -0.62668850 -0.74886860  1.00000000 -0.02696545
 [6,]   0.72724729  0.5787199  0.59225624 -0.79313558 -0.02696545  1.00000000
 [7,]   0.92313520 -0.8198625 -0.73306354 -0.22187059 -0.98729964 -0.82198578
 [8,]  -0.19623152 -0.3572551  0.83941955  0.01188849  0.11953143  0.79534299
 [9,]   0.77049407  0.2523858 -0.60589763 -0.76807130  0.74189770  0.76891369
[10,]  -0.27731460 -0.8588763  0.62284995  0.84967670  0.78888431 -0.95713168
             [,7]       [,8]       [,9]      [,10]
 [1,]   0.9231352 -0.19623152  0.7704941 -0.2773146
 [2,]  -0.8198625 -0.35725511  0.2523858 -0.8588763
 [3,]  -0.7330635  0.83941955 -0.6058976  0.6228500
 [4,]  -0.2218706  0.01188849 -0.7680713  0.8496767
 [5,]  -0.9872996  0.11953143  0.7418977  0.7888843
 [6,]  -0.8219858  0.79534299  0.7689137 -0.9571317
 [7,]   1.0000000  0.80944950 -0.3515832 -0.8701093
 [8,]   0.8094495  1.00000000  0.3319837  0.3363241
 [9,]  -0.3515832  0.33198374  1.0000000 -0.6581544
[10,]  -0.8701093  0.33632411 -0.6581544  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 es-
timate 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.

References

Beiying Ding and Robert Gentleman. Measurement error model for correlation coefficient estimation and its application in microarray analysis. 2003.