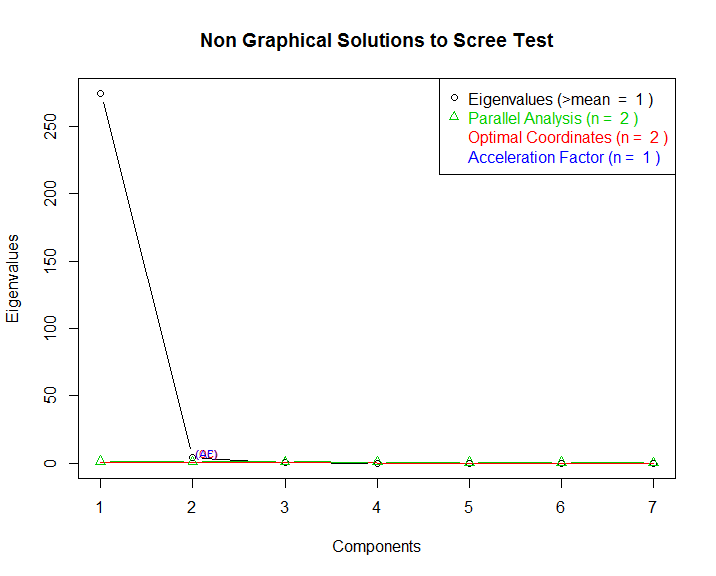
## STA 135 Homework 6 Haozhe Gu 999200555

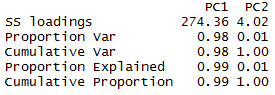
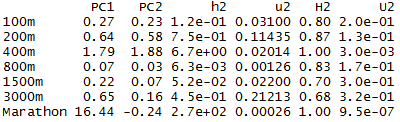
1. (Problem 9.28)

**Using Covariance Matrix S**

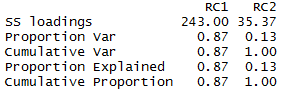
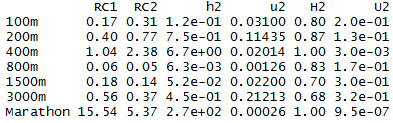
**-**determine Number of Factors, m should be 2



**Unrotated**



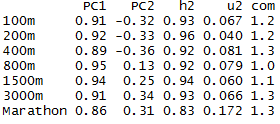
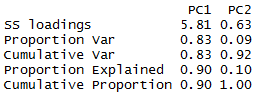
**Varimax Rotation**



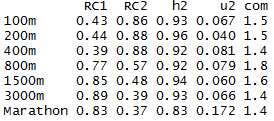
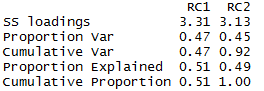
Using covariance matrix, most of the variance are explained by the first factor and the largest factor loading is associate with Marathon. This set up give us very little extra information about the dataset, so we choose to use R, the correlation matrix.

**Using Correlation Matrix R**

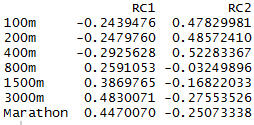
**Unrotated**

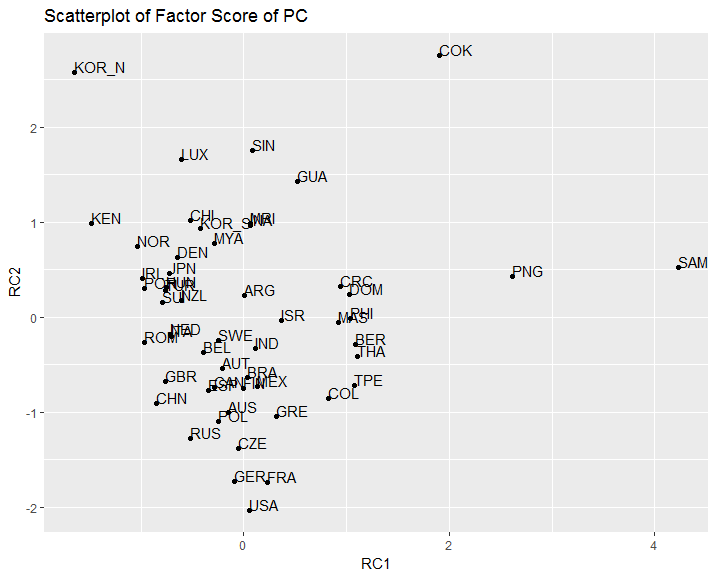
**Varimax Rotation**

**Factor score:**

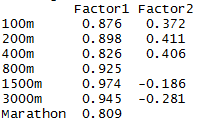
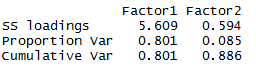


**Outlier Plot:**

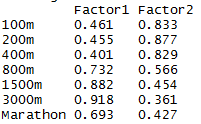
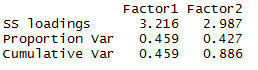


**ML Factor Analysis**

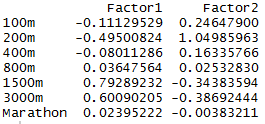
**Unrotated**

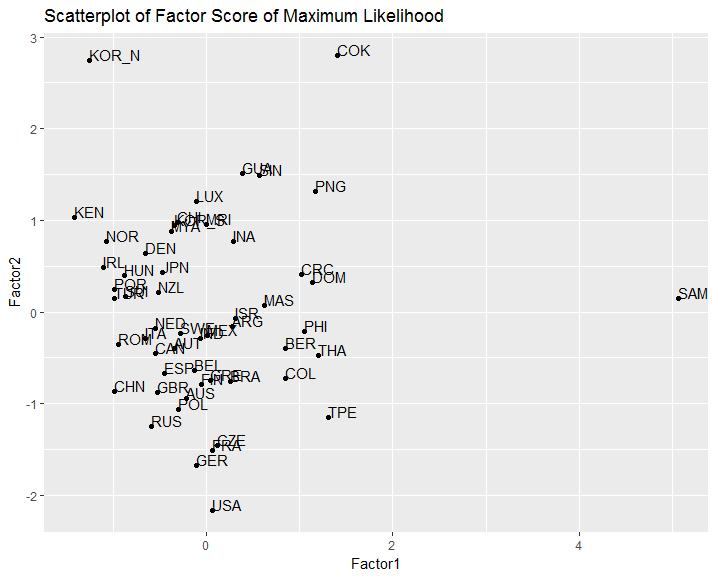
**Varimax Rotation**

**Factor score:**



**Outlier Plot:**



For both method, the unrotated loading for factor one seems like the “overall score” factor which loading of all events are similar. The Second factor seems to contrast short range performance and long range performance. However, the unrotated loading still rely heavily on factor one.

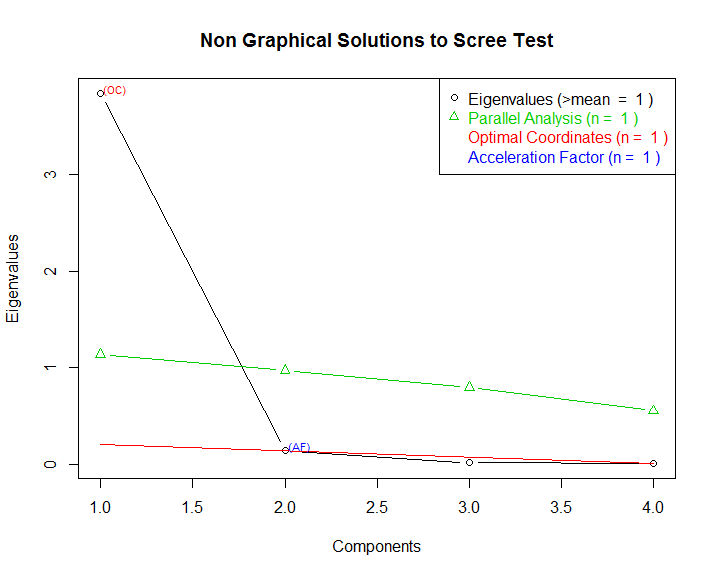
The analysis after varimax rotation makes both factors explain approximately equal amount variance, and they sum up to explain 88.6% to 92% total variance. The first factor load heavily on long-run performance which might represent “Endurance”, the second factor load heavily on short-run performance which might represent “Speed”. For m=2 model, SAM, COK and KOR\_N might be Outliers.

1. (Problem 9.34)

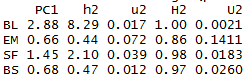
BL = Breaking Length, EM = Elastic Modulus, SF = Stress at Failure, BS = Burst Strength

**Using Covariance Matrix S**

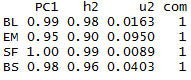
**-**determine Number of Factors, we will use m = 1



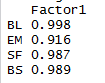
**Using Covariance Matrix S**

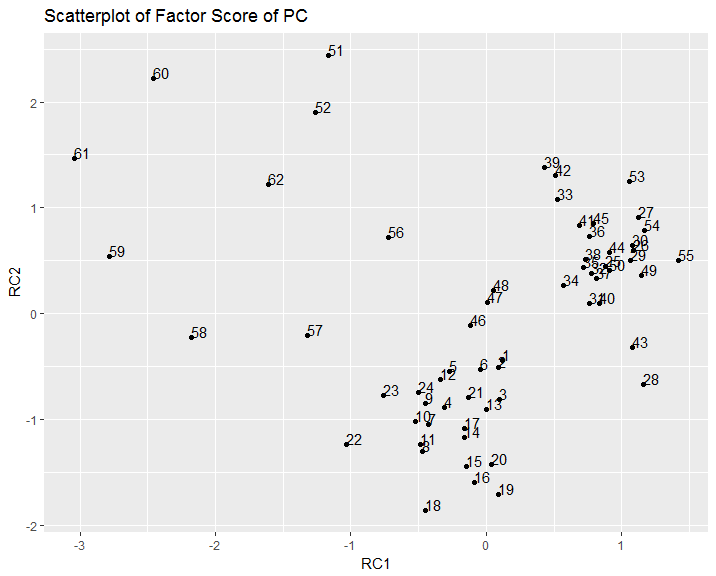
**Using Correlation Matrix R**

**ML Factor Analysis**

**Outlier Plot:**

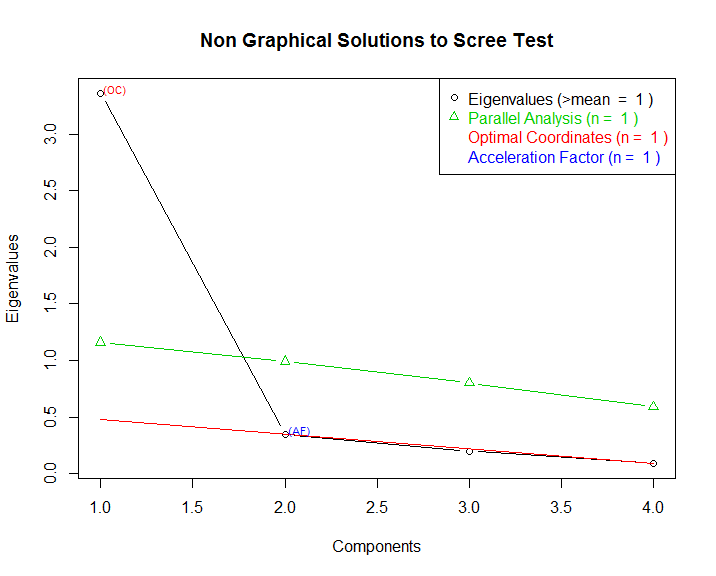


From the above three methods, the first factor explains about 94.7% to 99% variance of the total variance. So, it is possible to summarize the data using this single factor. Using R correlation matrix and Maximum Likelihood methods, this single factor have loading equally distributed for each variable. So this factor might represent “paper property”. The result using S covariance matrix have highest loading on BL, slightly high loading on SF and low loading on EM and BS. For m=2 model, observation 60,61,51,59 might be Outliers.

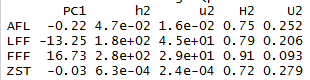
1. (Problem 9.35)

AFL = Arithmetic Fiber Length, LFF = Long Fiber Fraction, FFF = Fine Fiber Fraction, ZST = Zero Span tensile

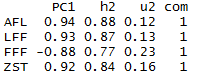
**-**determine Number of Factors, we will use m = 1



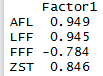
**Using Covariance Matrix S**

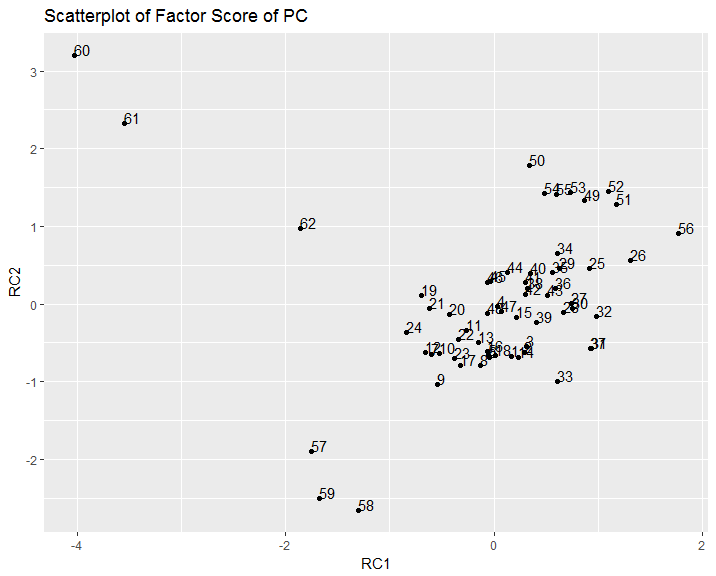
**Using Correlation Matrix R**

**ML Factor Analysis**

**Outlier Plot:**



Three method with single factor explains 78.1% to 86% of total variance. So a single factor model might be enough to summarize the data. For Maximum Likelihood and PC using Correlation Matrix, the loading seems to contrast FFF with three other variables with roughly similar weight. Factor Analysis using S covariance matrix, the single factor seems to contrast LFF and FFF only. Plot of factor scores for m=2 suggest observation 60,61and 62 might be Outliers

## APPENDIX

