Multivariate Analysis

Assignment #3

Due: 5/1/2018

Perform the calculations in calculator or matrix software.

1. (20%) Textbook 9.12

Hint: As the factor analysis is performed by maximum likelihood method, please use S_n instead of S.

The following maximum likelihood estimates of the factor loadings for an $m = 1$ model were obtained:

Variable	Estimated factor loadings F_1
1. ln(length)	.1022
2. ln(width)	.0752
3. ln(height)	.0765

Using the estimated factor loadings, obtain the maximum likelihood estimates of each of the following.

(a) Specific variances.

(b) Communalities.

(c) Proportion of variance explained by the factor.

(d) The residual matrix $\mathbf{S}_n - \hat{\mathbf{L}}\hat{\mathbf{L}}' - \hat{\Psi}$.

Hint: Convert **S** to $\mathbf{S}_n$.

Perform the calculations in SAS or SPSS.

2. (40%) Textbook 9.32

9.32. Perform a factor analysis of the data on bulls given in Table 1.10. Use the seven variables YrHgt, FtFrBody, PrctFFB, Frame, BkFat, SaleHt, and SaleWt. Factor the sample covariance matrix **S** and interpret the factors. Compute factor scores, and check for outliers. Repeat the analysis with the sample correlation matrix **R**. Compare the results obtained from **S** with the results from **R**. Does it make a difference if **R**, rather than **S**, is factored? Explain.

(a) (5%) Factor **S** using principal components (rotated)

(b) (5%) Factor **S** using maximum likelihood method (rotated)

(c) (5%) Factor **R** using principal components (rotated)

(d) (5%) Factor **R** using maximum likelihood method (rotated)

(e) (10%) Compare the results obtained from **S** with the results from **R**. Interpret the factors.

(f) (10%) Plot the scatter plots of factor2 vs factor1 in (a) and (c). Any outlier?

3. (40%) A survey is undertaken to determine consumer perceptions of six competing brands of soft drinks. The data collected are given in file drinks.DAT. The brands rated were as follows: (1) Pepsi; (2) Coke; (3) Gatorade; (4) Allsport; (5) Lipton tea; (6) Nestea. Respondents used a 7-point scale (1= strongly disagree to 7= strongly agree) to indicate their level of agreement/disagreement with the following 10 statements (in each of the statements substitute "Brand X" with the brands listed above). The ten statements are given below:

X1 Brand X has a refreshing taste.
X2 I prefer Brand X because it has fewer calories than other drinks.
X3 Brand X quenches my thirst immediately.
X4 I like the sweet taste of Brand X.
X5 I prefer drinking Brand X after workouts and sports because it gives me energy.
X6 I prefer Brand X because it comes in environment friendly packaging.
X7 Brand X has minerals and vitamins that help quench my deep down body thirst.
X8 Brand X has a unique flavor of its own.
X9 Brand X has the right mix of minerals and vitamins that are healthy for my body.
X10 I prefer to drink Brand X when I am really thirsty.

Use principal components factor analysis to analyze the data (correlation matrix).
(a) (10%) Determine the appropriate number of factors to effectively account for the variance in the data. Show the Scree plot and rotated factor loadings.
(b) (9%) Label the factors and explain their meanings.
(c) (9%) Calculate the "average" factor scores of each brand.
Then make all scatter plots of any two factor scores.
Note that you will have only six points (brands) in each scatter plot.
(d) (12%) Use the scatter plots to interpret the positions of the six brands.