Learning Structures and Time Series Homework 3

- Akankshi Mody, am92786

```
Question 1
              a) X1:0.85, + 0.4 5, + 2,

X2:0.65, + 0.65, + 82

N3:0.45, + 0.85, + 83
                                bonding matrix: 3

$1, $2

$1, $2

$2, $5, $2

$2, $5, $2

$3, $4, $5, $2

$3, $4, $5, $2

$4, $5, $5, $2

$4, $5, $5, $2

$5, $6, $5, $5, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6, $6

$5, $6, $6

$5, $6, $6

$5, $6, $6

$5, $6, $6

$5, $6, $6

$5, $6, $6

$5, $6, $6

$5, $6, $6

$5, $6, $6

$5, $6, $6

$5, $6, $6

$5, $6, $6

$5, $6, $6

$5, $6, $6

$5, $6, $6

$5, $6, $6

$5, $6, $6

$5, $6, $6

$5, $6

$5, $6, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5, $6

$5,
                           we know, cors(1, x) = cor(Y, x)
                         COMLX, , x2) = cov(0.8 &, + 0.4 &, + 21, 0.64, +06 &, +22)
                     = 0.48 + 0.24 = 0.72

Similarly, corr(x2, x3) = 0.24+ 0.48 = 0.72
                                          wr ( X3, X1) = 0.32+032 = 0.64
```

(ROLLOWS)	
Estimated loadings matrix	
7) (2 (3)	
X1 1 0.72 0.6 4 38 0 + 3 N O 3 A	
K2 0.72 1 0.72	
X3 0.64 0.72 1 (xister graderal)	
b) communality of x, = (0.8) + (0.4) = 0.8	
X2-(N:()+f0:6)-0.72	
X3 = (0·4) + (0·8) = 0·8	
(x x) 100 = (xx) 78700, want no	6
(1)	
a) Var (xi) = 9112+ 212+ 212+ Var (ei)=1	
a) Var (xi) = Ai12+ Ai2+ rip + Var (e;) = 1	
This matrix from (point wise) square of FA loadings matrix gives var (x;) explained 2; and 2;	
loadings matin gives your x:) explained	
2 and 2'	by
	0
X, 0.64 0.16 O.8 reptained	
Vo	
X3 0.16 0.64 0.8	
Total 1:16 1:16 2-32	
Var 2 3 4.0 2 4.1 (8) + 0 - 2 4.1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
120= 18:0+ 8:0 = (18: x) = 0:00 H	
	4

b) No, it cannot be equivalent computationally to PCA - Unlike PCA, where number of PCs = munber of variables and 100% variance is explained by all PCs, FA here has only 2 factors for 3 variables and only 2.32 out of 3 variance is explained. Questión 3.

M= [0.7071 0.7071] F= factor pattern metrix
[0.7071 ~ 0.7071] a) F= [0.8 0.4] M= [0.7071 0.7071]

0.6 0.6 0.8 0.7071 - 0.7071 FM 2 F = [0.85 0.28] 0.64 0.72 0.8 $PM \times M_{1} \times t_{1} = \begin{bmatrix} 0.82 & 0.58 \\ 0.82 & 0 \end{bmatrix} \begin{bmatrix} 0.501 & 0.501 \\ 0.62 & 0 \end{bmatrix} \begin{bmatrix} 0.501 & 0.501 \\ 0.62 & 0.602 \end{bmatrix}$

0-8 0.72 0'64 0.72 0.72 0.72 Thus FXFT = FXMXMX ET Also, FXMXM'AF has coro matrix of 1 0.72 0.64 (Same as (a) 0.72 10.64 0.72 b) from F*, 10000 5M communality of X1 (0.85)2 +(0.283=0.8 $X_2 = (0.85)^2 + 0 = 0.12$ $X_3 = (0.85)^2 + (-0.28)^2 = 0.8$ F* is orthornormal rotation of F, communality is same. Information of factors is still same after the rotation

Question 4 X1 = a2 + E1 X22 b& + &2 X3 = C & + 83 a) for correlation to be same. corr(x, x2)=0.72=ab corr(2, 4): ab (1) corr (x2, x3) = p.72 = bc. corr(2, 4) = bc (2) corr (x3, x1) = 0.64 2 ac corr (8, 6) 2 ac 3 From above equations, b=0.72; 0.72 = a=C ac = 0.64 . a = C=0.8 b=0.72 = 0.9 model to produce same cost matri as two factor model b) 3(a) and 4(a), we can see that for a combination of namifest variables there can be many sets of factors, and thus many sets of correlation matrices.

5. Run a "principal components" style of factor analysis to extract 6 factors.

a) Submit your SAS code.

Principal Components Style Factor Analysis The FACTOR Procedure Initial Factor Method: Principal Components **Prior Communality Estimates: ONE** Eigenvalues of the Correlation Matrix: Total = 6 Average = 1 Difference Proportion Eigenvalue Cumulative 3.16922321 2.16287646 0.5282 0.5282 1.00634675 0.24343802 0.1677 0.6959 0.1272 0.8231 0.76290873 0.21039227 0.55251646 0.23526997 0.0921 0.9152 0.31724648 0.12548811 0.0529 0.9680 0.19175838 0.0320 1.0000 6 factors will be retained by the NFACTOR criterion.

Factor Pattern									
Factor1 Factor2 Factor3 Factor4 Factor5 Fac									
BEEFS	BEEFS	0.78219	-0.31363	0.38883	-0.23490	-0.10787	0.26797		
PRIVILEGE	PRIVILEGE	0.70268	-0.30973	0.18990	0.60569	-0.02123	-0.08333		
NEWLEARN	NEWLEARN	0.82140	-0.21777	-0.23756	-0.16709	0.43688	-0.05153		
RAISES	RAISES	0.87704	0.11590	0.00490	-0.27139	-0.25930	-0.27649		
CRITICAL	CRITICAL	0.40022	0.80479	0.39938	0.07429	0.16271	0.02533		
ADVANCE	ADVANCE	0.67791	0.32172	-0.59975	0.15293	-0.14347	0.18237		

b) How many factors would you retain? Why?

The eigenvalue for the first two factors is more than 1. By the Kaiser rule, I would retain 2 factors.

c) Try to interpret the first two factors.

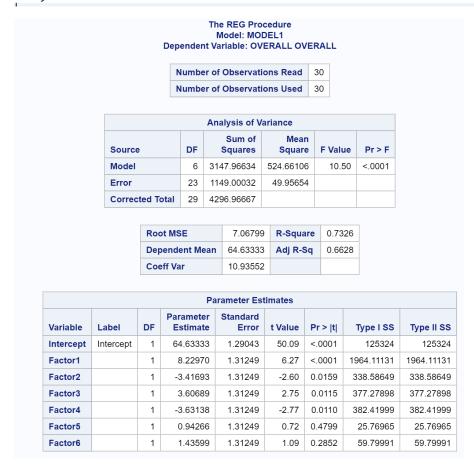
Factor 1 exhibits high weightage from all the variables except for "Critical". This means that an employee who has a high score on Factor 1 likely has high scores on all other variables in the dataset – i.e. on handling complaints well, special privileges, opportunity to learn new things, raises based on

merit, and advances employees to better jobs. But this employee is not critical. Such an employee would be a valuable asset to a company and would be liked well by employees working under him/her.

Factor 2 has a high score only for "Critical". Such a supervisor will be considered critical and will probably be disliked by employees working under him/her.

6. How well do the six factors of question 5 explain the OVERALL supervisor rating – individually and collectively?

PROC REG DATA=WORK.prinfactors;
 model OVERALL = Factor1 Factor2 Factor3 Factor4 Factor5 Factor6 /ss1 ss2;
RUN;



Collective: We run a regression with all the factors to determine the overall rating. The six factors then explain 73.26% of the OVERALL rating. From the output we see that Factors 5 and 6 are insignificant.

Individual: Variance explained is be given by R-square * Type I SS / SSE.

Factors	Variance explained (%)
Factor 1	45.70912743
Factor 2	7.879641514
Factor 3	8.780099623

Factor 4	8.899742069
Factor 5	0.599715612
Factor 6	1.391673523
Total	73.25999977

7. Run a principal factor analysis, with R-square type initial estimates of communalities, to extract six factors.

a) Submit your SAS code.

PROC FACTOR DATA = WORK.EvaluateSupervisors METHOD = principal PRIORS=SMC NFACTORS=6
MINEIGEN=0
OUT=WORK.prinfactors;

VAR REES DETAILED NEWLEARN RAISES CRITICAL ADVANCE.

VAR BEEFS PRIVILEGE <u>NEWLEARN</u> RAISES CRITICAL ADVANCE; RUN;

The FACTOR Procedure Initial Factor Method: Principal Factors									
Prior Communality Estimates: SMC									
	BEEF	FS PRIVILEGE NEWLEARN RAISES CRITICAL ADVANCE							
	0.6250553	5 0.3753480	0.55967371	0.6751	3759 0.1	8574001	0.48759750		
Ξiς	genvalues	of the Reduce	d Correlation Ma		= 2.90855	217 Avera		_	
		Eigenvalue	Differer	ice	Propor	tion	Cumulati	ve	
_		2.71771612	2.316781			9344	0.93	_	
-		0.40093481	0.222028	808	0.1	378	8 1.072		
3	3	0.17890672			0.0615		1.1337		
4	1	00403456	0.161507		-0.0014		1.1324		
_	5	16554195	0.053887	702					
•	6	21942897			-0.0	754	1.000		
		3 factors	will be retained	by the Mil	NEIGEN c		2		
		BEEFS	BEEFS	0.74755	-0.36273	1			
		PRIVILEGE	PRIVILEGE	0.61091	-0.30273				
		NEWLEARN		0.76629	-0.05146				
		RAISES	RAISES	0.84947	0.11042				
		CRITICAL	CRITICAL	0.32091	0.25308	0.2676	0		
		ADVANCE ADVANCE 0.61147 0.39882 -0.15045					_		

Only the first 3 factors will be retained by since the eigenvalues for the last 3 factors are negative.

b) Try to interpret the first two factors.

Factor 1 exhibits high weightage from all the variables except for "Critical". This means that an employee who has a high score on Factor 1 likely has high scores on all other variables in the dataset –

i.e. on handling complaints well, special privileges, opportunity to learn new things, raises based on merit, and advances employees to better jobs. But this employee is not critical. Such an employee would be a valuable asset to a company and would be liked well by employees working under him/her.

Factor 2 exhibits low weightage on all variables. Also, it has negative loadings on the first 3 variables. Such a supervisor probably does not handle employee complaints well, does not give them privilege and opportunity to learn. He/she is critical and has little ability to advance employee to better jobs.

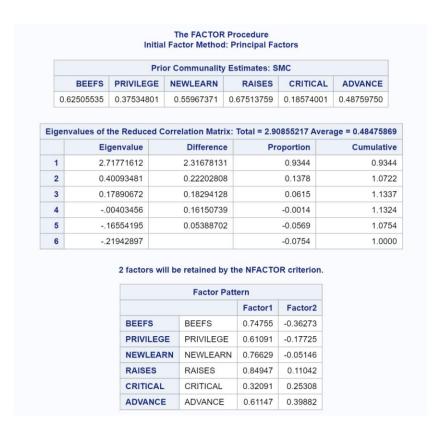
c) Succinctly compare the factor analysis of Q7 with the factor analysis of Q5.

In Q7, the total number of retained factors is 3 whereas in Q5, all the factors had with positive eigenvalues. The variance being explained by has also reduced.

8. Run a principal factor analysis, with R-square type initial estimates of communalities, to extract two factors, apply a varimax rotation to the initial factor pattern, and add factor scores for all 30 supervisors to a dataset called WORK. Evaluate Supervisors_scores (including all original data as well).

a) Submit your SAS code.

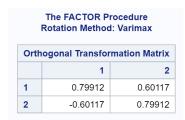
```
PROC FACTOR DATA = WORK.EvaluateSupervisors METHOD = principal PRIORS=SMC NFACTORS=2
ROTATE=varimax
OUT=WORK.EvaluateSupervisors_scores;
    VAR BEEFS PRIVILEGE NEWLEARN RAISES CRITICAL ADVANCE;
RUN;
```



b) Are the first two initial factors the same as the first two factors in the preceding question?

Yes, they are the same.

c) Verify that the varimax factor rotation matrix is orthonormal.



The magnitude of each vector:

$$(0.79912)^2 + (0.60117)^2 = 0.638592774 + 0.361405369 = 1$$

$$(0.60117)^2 + (0.79912)^2 = 0.361405369 + 0.638592774 = 1$$

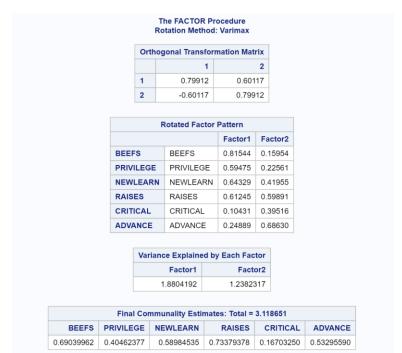
Thus, the vectors are normalized.

Dot Product:

$$0.79912 * 0.60117 + 0.79912* -0.60117 = 0$$

Thus, the vectors are orthogonal.

Thus, the varimax factor rotation matrix is orthonormal.



d) Try to interpret the varimax rotated factor pattern.

Factor exhibits high loading on BEEFS, PRIVILEGE, NEWLEARN, RAISES and low loadings on CRITICAL AND ADVANCE. Such a supervisor is not critical and does not have the ability to advance employee to better jobs.

Factor 2 exhibits low but positive loadings on all variables with the highest loadings are on ADVANCE AND RAISES. Such a supervisor would have the ability to advance employee to better jobs as well as giving raises.

- 9. Calculate means, standard deviations, and correlation for the factor scores of the 30 supervisors of the varimax rotated factors that you stored in WORK.EvaluateSupervisors_scores in the preceding question. [You may use PROC CORR, which includes PROC MEANS output by default.]
- a) Given the assumptions of factor analysis, do your summary statistics surprise you? Why or why not?

PROC CORR DATA=WORK.EvaluateSupervisors_scores; RUN;

Simple Statistics									
Variable	N	Mean	Std Dev	Sum	Minimum	Maximum	Label		
OVERALL	30	64.63333	12.17256	1939	40.00000	85.00000	OVERALL		
BEEFS	30	66.60000	13.31476	1998	37.00000	90.00000	BEEFS		
PRIVILEGE	30	53.13333	12.23543	1594	30.00000	83.00000	PRIVILEGE		
NEWLEARN	30	56.36667	11.73701	1691	34.00000	75.00000	NEWLEARN		
RAISES	30	64.63333	10.39723	1939	43.00000	88.00000	RAISES		
CRITICAL	30	74.76667	9.89491	2243	49.00000	92.00000	CRITICAL		
ADVANCE	30	42.93333	10.28871	1288	25.00000	72.00000	ADVANCE		
Factor1	30	0	0.85285	0	-1.86118	1.36525			
Factor2	30	0	0.77645	0	-1.31002	2.01667			

Pearson Correlation Coefficients, N = 30 Prob > r under H0: Rho=0									
	OVERALL	BEEFS	PRIVILEGE	NEWLEARN	RAISES	CRITICAL	ADVANCE	Factor1	Factor2
OVERALL OVERALL	1.00000	0.82542 <.0001	0.42612 0.0189	0.62368 0.0002	0.59014 0.0006	0.15644 0.4091	0.15509 0.4132	0.82349 <.0001	0.19094 0.3122
BEEFS BEEFS	0.82542 <.0001	1.00000	0.55829 0.0013	0.59674 0.0005	0.66920 <.0001	0.18771 0.3205	0.22458 0.2328	0.95614 <.0001	0.20547
PRIVILEGE PRIVILEGE	0.42612 0.0189	0.55829 0.0013	1.00000	0.49333 0.0056	0.44548 0.0136	0.14723 0.4375	0.34329 0.0633	0.69736 <.0001	0.29057 0.1193
NEWLEARN NEWLEARN	0.62368 0.0002	0.59674 0.0005	0.49333 0.0056	1.00000	0.64031 0.0001	0.11597 0.5417	0.53162 0.0025	0.75429 <.0001	0.54034
RAISES RAISES	0.59014 0.0006	0.66920 <.0001	0.44548 0.0136	0.64031 0.0001	1.00000	0.37688 0.0401	0.57419 0.0009	0.71812 <.0001	0.77135
CRITICAL CRITICAL	0.15644 0.4091	0.18771 0.3205	0.14723 0.4375	0.11597 0.5417	0.37688 0.0401	1.00000	0.28334 0.1292	0.12230 0.5197	0.50894
ADVANCE ADVANCE	0.15509 0.4132	0.22458 0.2328	0.34329 0.0633	0.53162 0.0025	0.57419 0.0009	0.28334 0.1292	1.00000	0.29183 0.1176	0.88390
Factor1	0.82349 <.0001	0.95614 <.0001	0.69736 <.0001	0.75429 <.0001	0.71812 <.0001	0.12230 0.5197	0.29183 0.1176	1.00000	0.2904° 0.1195
Factor2	0.19094 0.3122	0.20547 0.2760	0.29057 0.1193	0.54034 0.0021	0.77135 <.0001	0.50894 0.0041	0.88390 <.0001	0.29041 0.1195	1.00000

The means of Factors 1 and 2 do not surprise me since they match our assumptions. But, the standard deviations of Factors should be 1, since the data is standardized but they are not. Even the correlations of the factors are supposed to be different in the basic Factor Analysis model.

10. Extract the default number of factors by the maximum likelihood method with R-square type initial estimates of communalities. You may need to use the ULTRAHEYWOOD or HEYWOOD option to deal with estimated communalities that exceed 1. You may assume that the assumptions of the maximum likelihood model apply.

a) Does the output provide support for the hypothesis that common factors exist?

Significance Tests Based on 30 Observations								
Test	DF	Chi-Square	Pr > ChiSq					
H0: No common factors	15	65.5127	<.0001					
HA: At least one common factor								
H0: 2 Factors are sufficient	4	2.8155	0.5892					
HA: More factors are needed								

We can see that the p-value for no common factors is less than 0.05. Hence, we can reject this null hypothesis indicating that the data is has at least one common factor.

b) Does the output provide support for the hypothesis that the default number of extracted factors is adequate?

We can see that the p-value is 0.5892. Hence, we fail to reject the null hypothesis can extract the sufficient number of factors which is 2.