Stockholms universitet

STATISTISKA INSTITUTIONEN HT-2016

Written exam in Multivariate Methods, 7.5 ECTS credits
Thursday, 27th October 2016, 15:00 – 20:00
Time allowed: FIVE hours
Examination Hall: Brunsvikssalen

You are required to answer all **6 (six)** questions as well as motivate your solutions. The total amount of points is 80. In order to pass this part, you need to get at least 40 points. Points from this exam will be added to your results from the computer lab assignment. The final grades are assigned as follows: **A** (91+), **B** (81-90), **C** (71-80), **D** (61-70), **E** (51-60), **Fx** (30-49), and **F** (0-29).

You are **allowed** to use a pocket calculator, a language dictionary, and a list of formulas (attached).

The teacher reserves the right to examine the students **orally** on the questions in this examination.

1. (15 points) Let us analyse the following 3-variate dataset with 6 observations. Each observation consists of 3 measurements and recorded in the following matrix

$$\begin{matrix} 7 & 4 & 3 \\ 4 & 1 & 8 \\ 6 & 3 & 5 \\ 8 & 6 & 1 \\ 8 & 5 & 7 \\ 7 & 2 & 9 \end{matrix}$$

What portion of total variance each variable accounts for? Compute the correlation matrix. Next, find eigenvalues of the correlation matrix and interpret them in style of PCA. How much of total variance the first two principal components "explain"? Have you mean-adjusted and/or standardized the original data set before the analysis: why yes/no?

2. (12 points)

(a) Points A and B have the following coordinates with respect to orthogonal axes X_1 and X_2: $A=(3,-3)$; $B=(7,1)$. If the axes X_1 and X_2 are rotated 300° clockwise to produce a new set of orthogonal axes X_1^* and X_2^*, find the coordinates of A and B with respect to X_1^* and X_2^*.

(b) Coordinates of a point A with respect to an orthogonal set of axes X_1 and X_2 are $(2,2)$. The axes X_1 and X_2 are rotated counter-clockwise by an angle θ. If the new coordinates of the point A with respect to the rotated axes are $(2.8284, 0)$, find θ. Provide geometric motivation.

3. (12 points)

a) What problems cluster analysis is intended to solve? What types of cluster analysis you know: name them. What assumptions on data are imposed in order to apply "cluster analysis"? How strict you should be with those assumptions: speculate and exemplify.

b) Cluster the following hypothetical data set into two groups using average linkage method and the associated similarity matrixes. Moreover, cluster the same data into 4 clusters using Ward's method. Analyse and discuss your findings.

Subject ID	Income in tEUR	Education (in years)
S1	17	10
S2	23	12
S3	25	14
S4	28	15
S5	30	20
S6	35	18

4. (15 points) The correlation matrix for a hypothetical data set is given in the following table:

	X_1	X_2	X_3	X_4
X_1	1.000			
X_2	0.7	1.000		
X_3	0.3	0.25	1.000	
X_4	0.35	0.2	0.4	1.000

The following estimated factor loadings were extracted by the principal axis factoring procedure:

Variable	F_1	F_2
X_1	0.90	0.20
X_2	0.70	0.15
X_3	0.20	0.90
X_4	0.20	0.70

Compute and discuss the following: (a) specific variances; what high specific variance indicates? Explain using data above; (b) communalities and % of shared variance; interpret both; (c) proportion of variance explained by each factor, what can you say about chosen factors? (d) Estimated or reproduced correlation matrix; how good is the estimate? Discuss; and (e) residual matrix, compute RMSR and interpret.

5. (14 points) Do the following for the data given below:
   a) Assume that data is transformed into mean corrected form. Will the results of the statistical techniques (e.g. factor analysis, principal component analysis) be affected by the transformation? Why or why not? (2p)
   b) Assume that data is transformed into standardized form. Will the results of the statistical techniques (e.g. factor analysis, principal component analysis) be affected by standardizing the data? Why or why not? (2p)
   c) Compute the total, between-group, and within-group SSCP matrices. What conclusions can you draw from these matrices? Are there all assumptions fulfilled to apply discriminant

analysis? (out of 10 points)

Financial Data for Failed and non-Failed firms

Observations (Failed Firms)	EBITASS	ROTC	Observations (non-failed)	EBITASS	ROTC
1	0.1	0.2	1	-0.01	-0.03
2	0.2	0.2	2	-0.05	-0.11
3	0.2	0.3	3	0.09	0.12
4	0.1	0.2	4	0.03	0.05
5	0.3	0.2	5	0.04	0.06
6	0.2	0.1			

6. (12 points) Consider the two-indicator two-factor model represented by the following equations:

$$X_1 = 0.104F_1 + 0.824F_2 + U_1$$
$$X_2 = 0.065F_1 + 0.959F_2 + U_2$$
$$X_3 = 0.065F_1 + 0.725F_2 + U_3$$
$$X_4 = 0.906F_1 + 0.134F_2 + U_4$$
$$X_5 = 0.977F_1 + 0.116F_2 + U_5$$
$$X_6 = 0.827F_1 + 0.016F_2 + U_6$$

The usual assumptions hold for the above model. Answer the following questions assuming that the correlation between the common factors F_1 and F_2 is given by Corr(F_1, F_2)= ϕ_{12} = - 0.4. Repeat all your calculations in assumption that correlation changed to Corr(F_1, F_2)= ϕ_{12} = 0.4 and discuss the differences in detail. Try to provide intuition for at least some of your answers: without calculating, what you would expect in case correlation is 0.9, 0, -0.9?

(a) What are the pattern loadings of indicators X_1, X_4 and X_6 on the factors F_1 and F_2?

(b) Compute the correlation between the indicators X_1 and X_2.

(c) What percentage of the variance of indicators X_1 and X_2 is not accounted for by the common factors F_1 and F_2?

Formula Sheet, Multivariate Methods

Matrices

Transpose – exchange rows and columns
Identity (I) – diag (1,1...) of order n*n
Inverse of A (A^{-1}): $AA^{-1} = A^{-1}A = I$
$A + B = B + A$; $x(A + B) = xA + xB$; $AB \neq BA$ (*in general*);
If order (A)=m*n, order (B)=n*p, then C=AB is of order m*p

D=det A=$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$

detA=$a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}$ where cofactor $A_{ij} = (-1)^{i+j} D_{ij}$ (i-row, j-column of D)
Cramer's rule: $x_j = D_j / D$ where D=detA and D_j is the determinant that arises when the j column of D is replaced by the column elements $b_1, \ldots, b_n$. (**Ax=b**)

Vectors

$\boldsymbol{a} = (a_1 a_2 \ldots a_p)$
A right-angle triangle: α - angle between a and c; c – hypotenuse; $\cos\alpha = \frac{a}{c}$, $\sin\alpha = \frac{b}{c}$
Length of vector a=$\|\boldsymbol{a}\| = \sqrt{a_1^2 + a_2^2}$
Basis vectors $\boldsymbol{e_1} = (1\ 0), \boldsymbol{e_2} = (0\ 1)$
$\boldsymbol{a} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2$
Scalar product $\boldsymbol{ab} = a_1b_1 + a_2b_2 + \cdots + a_pb_p$; $\boldsymbol{ab} = \|\boldsymbol{a}\|\|\boldsymbol{b}\|\cos\alpha$
Length of the projection: $\|\boldsymbol{a_p}\| = \|\boldsymbol{a}\|\cos\alpha$
Variance of x_i: $s_1^2 = \frac{\|x_i\|^2}{n-1}$; Generalized variance: $GV = \left(\frac{\|x_1\|\cdot\|x_2\|}{n-1} \cdot \sin\alpha\right)^2$

Distances

Euclidean: $D_{AB} = \sqrt{\sum_{j=1}^{p}(a_j - b_j)^2}$
Statistical: $SD_{ij}^2 = \left(\frac{x_i - x_j}{s}\right)^2$, s-standard deviation
Mahalanobis: $MD_{ik}^2 = \frac{1}{1-r^2}\left[\frac{(x_{i1}-x_{k1})^2}{s_1^2} + \frac{(x_{i2}-x_{k2})^2}{s_2^2} - \frac{2r(x_{i1}-x_{k1})(x_{i2}-x_{k2})}{s_1 s_2}\right]$

Variance, Sum of Squares, and Cross Products

Variance: $s_j^2 = \frac{\sum_{i=1}^{n} x_{ij}^2}{n-1} = \frac{SS}{df}$ (sum of squares/degrees of freedom)
Covariance: $s_{jk} = \frac{\sum_{i=1}^{n} x_{ij}x_{ik}}{n-1} = \frac{SCP}{df}$ (sum of the cross products/degrees of freedom)
SSCP – sum of squares and cross products matrix $\begin{pmatrix} SSX_1 & SCP \\ SCP & SSX_2 \end{pmatrix}$
S – covariance matrix $S_t = \frac{SSCP_t}{df}$
Within-Group Analysis: $SSCP_w = SSCP_1 + SSCP_2$ (pooled SSCP matrix) $S_w = \frac{SSCP_w}{n_1+n_2-2}$ (pooled cov m)
Between-Group Analysis: $SS_j = \sum_{g=1}^{G} n_g(\bar{x}_{jg} - \bar{x}_j)^2$; $SCP_{jk} = \sum_{g=1}^{G} n_g(\bar{x}_{jg} - \bar{x}_j)(\bar{x}_{kg} - \bar{x}_k)$
$SSCP_t = SSCP_w + SSCP_b$

Principal Components Analysis

$x_1^* = \cos\theta * x_1 + \sin\theta * x_2$; $x_2^* = -\sin\theta * x_1 + \cos\theta * x_2$
$\boldsymbol{\Sigma}$ covariance matrix; λ-eigenvalues; $|\boldsymbol{\Sigma} - \boldsymbol{\lambda I}| = \boldsymbol{0}$; γ-eigenvector; $(\boldsymbol{\Sigma} - \boldsymbol{\lambda I})\boldsymbol{\gamma} = \boldsymbol{0}$; $\boldsymbol{\gamma}'\boldsymbol{\gamma} = \boldsymbol{1}$;

Factor Analysis

Assumptions: 1. Means of indicators, common factor, unique factors are zero.
2. Variances of indicators and common factors are one. 3.$E(\xi_i\varepsilon_i) = 0$ and $E(\varepsilon_i\varepsilon_j) = 0$

Two-Factor Model: $x_1 = \lambda_{11}\xi_1 + \lambda_{12}\xi_2 + \varepsilon_1$

$$x_2 = \lambda_{21}\xi_1 + \lambda_{22}\xi_2 + \varepsilon_2$$
$$\vdots$$
$$x_p = \lambda_{p1}\xi_1 + \lambda_{p2}\xi_2 + \varepsilon_p$$

The variance of x: $E(x^2) = E(\lambda_1\xi_1 + \lambda_1\xi_2 + \varepsilon_1)^2$; $Var(x) = \lambda_1^2 + \lambda_2^2 + Var(\varepsilon) + 2\lambda_1\lambda_2\phi$
The correlation between any indicator and any factor (the structure loading):
$E(x\xi_1) = E[(\lambda_1\xi_1 + \lambda_1\xi_2 + \varepsilon_1)\xi_1]$; $Corr(x\xi_1) = \lambda_1 + \lambda_2\phi$
The shared variance between the factor and an indicator: $Shared\ variance = (\lambda_1 + \lambda_2\phi)^2$
The correlation between two indicators:

$$E(x_j x_k) = E[(\lambda_{j1}\xi_1 + \lambda_{j2}\xi_2 + \varepsilon_j)(\lambda_{k1}\xi_1 + \lambda_{k2}\xi_2 + \varepsilon_k)]$$
$$Corr(x_j x_k) = \lambda_{j1}\lambda_{k1} + \lambda_{j2}\lambda_{k2} + (\lambda_{j1}\lambda_{k2} + \lambda_{j2}\lambda_{k1})\phi$$

Confirmatory Factor Analysis

The covariance matrix (one-factor model, two indicators): $\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix}$

Evaluating model fit: χ^2-test H_0: $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\theta})$ H_a: $\boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}(\boldsymbol{\theta})$ (test whether the difference between the sample and the estimated covariance matrix is a zero matrix)
$\chi^2 = \sum_{i=1}^{k} \frac{[n_i - E(n_i)]^2}{E(n_i)}$

Cluster Analysis
Measure of similarity – squared Euclidean distance between two points
Hierarchical clustering:
Centroid method – each group is replaced by centroid
Nearest-neighbor or single-linkage method – the distance between two clusters is represented by the minimum of the distance between all possible pair of subjects in the two clusters
Farthest-neighbor or complete-linkage method - … the maximum of the distances…
Average-linkage method - … the average distance…
Ward's method – does not compute distances between clusters. Method tries to minimize the total within-group sums of squares.

Discriminant Analysis
Assumptions: multivariate normality, equality of covariance matrices
Discriminant function: $Z = w_1x_1 + w_2x_2$
$\lambda = \frac{between-group\ sum\ of\ squares}{within-group\ sum\ of\ squares}$
$\boldsymbol{\Sigma}$-variance-covariance matrix, **T**-total SSCP matrix. $\boldsymbol{\gamma}$-vector of weights.
Discriminant function $\boldsymbol{\xi} = \boldsymbol{X}'\boldsymbol{\gamma}$. **B** and **W** are between-groups and within-group SSCP matrices.
Maximize $\lambda = \frac{\gamma' B \gamma}{\gamma' W \gamma}$
$|\boldsymbol{W}^{-1}\boldsymbol{B} - \lambda\boldsymbol{I}| = 0$; $\boldsymbol{\gamma} = \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)$ - Fisher's discriminant function

Logistic regression
$odds = \frac{p}{1-p}$
$\ln odds = \beta_0 + \beta_1X_1 + \cdots + \beta_kX_k$
$p = \frac{1}{1+e^{-(\beta_0+\beta_1X_1+\cdots+\beta_kX_k)}}$
Maximum likelihood estimation: $P(Y = 1) = p = \frac{e^{\beta X}}{1+e^{\beta X}}$
$L = \prod_{i=1}^{n} p_i^{y_i}(1-p_i)^{1-y_i}$
Quadratic equations: $ax^2 + bx + c = 0; x = \frac{-b\pm\sqrt{b^2-4ac}}{2a}$
Cubic equations:

$$y^3 + ay^2 + by + c = 0; y = x - \frac{a}{3};\ x^3 + px + q = 0; x_1 = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}$$

Added formulas:

RMSR $= \sqrt{\frac{\sum_{i=1}^{p}\sum_{j=i}^{p} res_{ij}^2}{p(p-1)/2}}$, where res_{ij} is the correlation matrix between the ith and jth var, p is number of variables.

Cubic equations:

There is an analogous formula for polynomials of degree three:
$ax^3 + bx^2 + cx + d = 0$ is:

$$x = \sqrt[3]{\left(\frac{-b^3}{27a^3} + \frac{bc}{6a^2} - \frac{d}{2a}\right) + \sqrt{\left(\frac{-b^3}{27a^3} + \frac{bc}{6a^2} - \frac{d}{2a}\right)^2 + \left(\frac{c}{3a} - \frac{b^2}{9a^2}\right)^3}} + \sqrt[3]{\left(\frac{-b^3}{27a^3} + \frac{bc}{6a^2} - \frac{d}{2a}\right) - \sqrt{\left(\frac{-b^3}{27a^3} + \frac{bc}{6a^2} - \frac{d}{2a}\right)^2 + \left(\frac{c}{3a} - \frac{b^2}{9a^2}\right)^3}} - \frac{b}{3a}.$$

Department of Statistics

Stockholms universitet

Correction sheet

Date: 27/10 - 2016

Room: Brunnsvikssalen

Exam: Multivariate Methods

Course: Multivariate Methods

Anonymous code: MME-0004

☒ I authorise the anonymous posting of my exam, in whole or in part, on the department homepage as a sample student answer.

NOTE! ALSO WRITE ON THE BACK OF THE ANSWER SHEET

Mark answered questions

1	2	3	4	5	6	7	8	9	Total number of pages
X	X	X	X	X	X	~~X~~	~~X~~	~~X~~	9
Teacher's notes 13	12	11	11	14	7				

Points	Grade	Teacher's sign.
68		AA

SU, DEPARTMENT OF STATISTICS

Room: Brunns v. **Anonymous code:** MME-CCH **Sheet number:** 1

2. a. formulas: $x_1^* = \cos\theta \cdot x_1 + \sin\theta \cdot x_2$,

$x_2^* = -\sin\theta \cdot x_1 + \cos\theta \cdot x_2$

300° clockwise (given info) implies

$\theta = 360 - 300 = 60°$

Info: $A = (3, -3)$, $B = (7, 1)$

$\Rightarrow$ $x_1^* = \cos 60 \cdot 3 + \sin 60 \cdot (-3) = -1,1$

$x_2^* = -\sin 60 \cdot 3 + \cos 60 \cdot (-3) = -4,1$

for A, $\Rightarrow A^* = (-1,1 \; , \; -4,1)$

For B: $x_1^* = \cos 60 \cdot 7 + \sin 60 \cdot 1 = 4,4$

$x_2^* = -\sin 60 \cdot 7 + \cos 60 \cdot 1 = -5,6$

$\Rightarrow B^* = (4,4 \; , \; -5,6)$

$\Rightarrow$ $\boxed{A^* = (-1,1 \; , \; -4,1), \quad B^* = (4,4 \; , \; -5,6)}$ ok

1 pu

Answers in boxes.

2. b. Info: $A = (2, 2)$, $A^* = (2,8284 \; , \; 0)$

$\Rightarrow$ $x_1 = 2$, $x_2 = 2$, $x_1^* = 2,8284$, $x_2^* = 0$

$\Rightarrow$ {with formula for PCA}

$2,8284 = \cos\theta \cdot 2 + \sin\theta \cdot 2$

$0 = -\sin\theta \cdot 2 + \cos\theta \cdot 2$

ptp.

2.6. (continuation)

$\Rightarrow \cos\theta = (2{,}8284 - 2\sin\theta)/2$

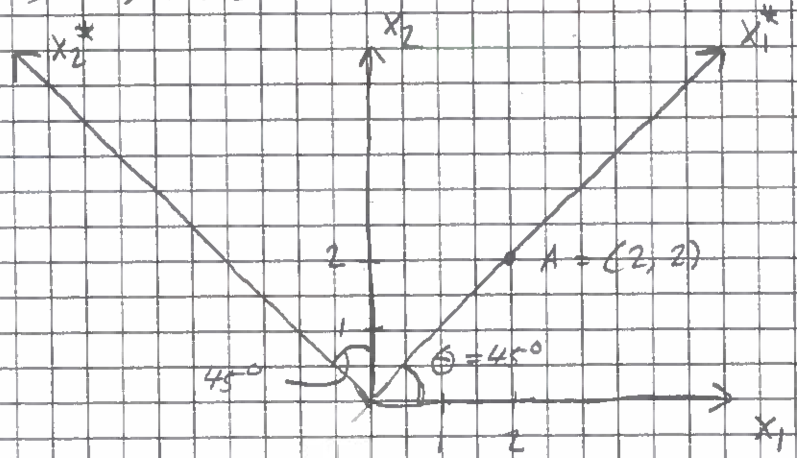
$\Rightarrow 0 = -\sin\theta \cdot 2 + (2{,}8284 - 2\sin\theta)/2 \cdot 2$

$0 = -2\sin\theta + 2{,}8284 - 2\sin\theta$

$4\sin\theta = 2{,}8284 \Rightarrow \sin\theta = 2{,}8284/4$

$\theta = \sin^{-1}(2{,}8284/4) = 45°$

Illustration:



Point $A = (2, 2)$ lies on the new axis x_1^*. Therefore, $x_2^* = 0$, as given in info.

ØL

SU, DEPARTMENT OF STATISTICS

Room: Brunnsv. Anonymous code: MME-0004 Sheet number: 2

3 a. Cluster analysis intends to group similar observations and thereby make classifation of observations easier.

Some types of cluster analysis follow:

- Hierarchical clustering
  - Centroid method clustering
  - Nearest-neighbor/single linkage method clustering
  - Farthest-neighbor/complete linkage method clustering
  - Average linkage method clustering
  - Ward's method clustering
- Non-hierarchical clustering

Cluster analysis assumes that data is homogenous. I believe assumption should be strictly followed regarding outliers since they can vastly affect the clustering process. Except for outliers I think the assumption can be relaxed (if for instance theres great distance between two rather homogenous groups of observations.

su

3. 8. Similarity matrix given by squared euclid. distances

For S_1 and S_2:

$ED_{12}^2 = (23-17)^2 + (12-10)^2 = 40$

(repeat for all observations)

$\Rightarrow ED_{12}^2 = 40$, $ED_{13}^2 = 80$, $ED_{14}^2 = 146$, $ED_{15}^2 = 269$, $ED_{16}^2 = 388$,
$ED_{23}^2 = 8$, $ED_{24}^2 = 34$, $ED_{25}^2 = 113$, $ED_{26}^2 = 180$,
$ED_{34}^2 = 10$, $ED_{35}^2 = 61$, $ED_{36}^2 = 116$,
$ED_{45}^2 = 29$, $ED_{46}^2 = 58$, $ED_{56}^2 = 29$

$\Rightarrow$ Simil. matrix

	S1	2	3	4	5	6
S1	0					
2	40	0				
3	80	8	0			
4	146	34	10	0		
5	269	113	61	29	0	
6	388	180	116	58	29	0

$\Rightarrow CL_1 = \{S2, S3\}$

Simil. matrix now given by average sq. euclid. dist. of possible paired obs.

For CL_1 and S1:

$(ED_{12}^2 + ED_{13}^2)/2 = (40+80)/2 = 60$

(repeat) see next page

SU, DEPARTMENT OF STATISTICS

Room: Brunnsv. Anonymous code: MMG-0004 Sheet number: 3

3. 6. (continuation)

$\Rightarrow$ Sim. matrix

	CL1	S1	4	5	6
CL1	0				
S1	60	0			
4	22	146	0		
5	87	269	29	0	
6	148	388	58	29	0

$\Rightarrow$ $CL1 = \{S2, S3, S4\}$

Average linkage for CL1 and S1:

$(ED_{12}^2 + ED_{13}^2 + ED_{14}^2)/3 = (40+80+146)/3 = 88{,}7$

(repeat)

$\Rightarrow$ Sim. matrix

	CL_1	S1	5	6
CL1	0			
S1	88,7	0		
5	67,7	269	0	
6	118	388	29	0

$\Rightarrow$ $CL_2 = \{S5, S6\}$

pto.

3.6. (continuation)

Average linkage for CL1 and CL2:

$$(ED^2_{25} + ED^2_{35} + ED^2_{45} + ED^2_{26} + ED^2_{36} + ED^2_{46}) / 6$$
$$= (113 + 61 + 29 + 180 + 116 + 58) / 6 = 92,8$$

(repeat)

$\Rightarrow$ Sim. matrix

	CL1	CL2	S1
CL1	0		
CL2	92,8	0	
S1	60	328,5	0

$\Rightarrow$ $\boxed{CL_1 = \{S1, S2, S3, S4\},\ CL_2 = \{S5, S6\}}$

Ward's method:

ESS	CL1	2	3	4	5
20	S1,S2	3	4	5	6
40	13	2	4	5	6
increasing	⋮				
4	23				
increasing	24				
↓	⋮				
4,75	34				
increasing	35				
↓	36				
14,5	45				
increas.	46				
8,25	56				

Example:
ESS for
$CL1 = \{S2, S3\}$

$$= \left(25 - \frac{25+23}{2}\right)^2 + \left(23 - \frac{25+23}{2}\right)^2$$
$$+ \left(14 - \frac{14+12}{2}\right)^2 + \left(12 - \frac{14+12}{2}\right)^2$$
$$+ 4 \cdot 0 = 4$$

ESS for cluster with one observ. is zero.

(repeat calculation for all combinations

next page

SU, DEPARTMENT OF STATISTICS

Room: Brunnsv. Anonymous code: MME-004 Sheet number: 4

3.6. (continues)

$\Rightarrow CL1 = \{S2, S3\}$

ESS for $CL1 = \{S2, S3\}$ and $CL2 = \{5, 6\}$:
4 + 8,25 = 12,25

(repeat for all comb.)

$\Rightarrow$

ESS	CL 1	2	3	4
42,7	S1,2,3	4	5	6
17,3	234			
increas.	235			
	236			
12,25	23	56	1	4

$CL2 = \{S5, S6\}$ chosen in list because it had second lowest ESS in previous step.

$\Rightarrow \boxed{CL1 = \{S2, S3\},\ CL2 = \{S5, S6\},\ CL3 = \{S1\},\ CL4 = \{S4\}}$

Both methods group $\{S2, S3\}$ and $\{S5, S6\}$ together. However, using Ward's method, $\{S5, S6\}$ gets grouped in an earlier step, compared to average linkage.

ok

4. a. Assum no correlation between constructs and $V(X_1) = V(X_2) = V(X_3) = V(X_4) = 1$

Specific variance $V(U)$ is given by (formula implies)

$$V(U) = V(X) - \lambda_1^2 - \lambda_2^2 = 1 - \lambda_1^2 - \lambda_2^2$$

(with info) for $V(U_1)$:

$V(U_1) = 1 - 0{,}9^2 - 0{,}2^2 = 0{,}15$, (repeat)
$V(U_2) = 0{,}4875$, $V(U_3) = 0{,}15$, $V(U_4) = 0{,}47$.

ok

Specific variance is variance not explained by factors for an indicator.

4. b. Communalities are variance shared between a factor and an indicator, which is given by squared factor loadings.

$\text{Comm.}(X_1, F_1) = 0{,}90^2 = 0{,}81$ (repeat)

$\Rightarrow$

Indic.	Communality		% SV
	F_1	F_2	
X_1	0,81	0,04	0,81+0,04=0,85=85%
X_2	0,49	0,0225	51,25%
X_3	0,04	0,81	85%
X_4	0,04	0,49	53%

Next page

SU, DEPARTMENT OF STATISTICS

Room: Brunns v. **Anonymous code:** MME-CC04 **Sheet number:** 5

4.b (contin.)

Shared variance (SV in previous table) is given by (formula sheet)

$SV = (\lambda_1 + \lambda_2 \phi)^2 = \lambda_1^2 + 2\lambda_1\lambda_2\phi + \lambda_2^2 = \lambda_1^2 + \lambda_2^2$

since we assume $\phi = 0$,

Consequentially, SV is sum of communalities to both factors for an indicator.

OK

c. Variance explained by F_1:

$0{,}81 + 0{,}49 + 0{,}64 + 0{,}64 = 1{,}38$,

By F_2: $1{,}3625$

$\Rightarrow$ Prop. explained by F_1

$1{,}38/(1{,}38 + 1{,}3625) = 0{,}503$,

By F_2: $0{,}497$

OK

$\Rightarrow$ Prop are roughly equal, implying that factors work equally well in the chosen model.

But how many % of the total? Both can be equally "good/bad"

OK

4. d. Correlations in reproduced corr. matrix are given by

$$Corr(X_1, X_2) = \lambda_{11}\lambda_{21} + \lambda_{12}\lambda_{22} + (\lambda_{11}\lambda_{22} + \lambda_{12}\lambda_{21})\phi$$
$$= \lambda_{11}\lambda_{21} + \lambda_{12}\lambda_{22} \quad \text{since} \quad \phi = 0$$

$$\Rightarrow Corr(X_1, X_2) = 0{,}9 \cdot 0{,}7 + 0{,}20 \cdot 0{,}15 = 0{,}66$$
(repeat)

$\Rightarrow$ Reprod. corr. matrix

	X_1	2	3	4
X_1	1			
2	0,66	1		
3	0,36	0,275	1	
4	0,32	0,245	0,67	1

? OK OK

4. e. Set $\mathbf{B}$ = "reprod. corr matrix above".

Set $\mathbf{A}$ = "given corr. matrix in info".

Residual matrix $\mathbf{R}$ is now given by

$$\mathbf{R} = \mathbf{A} - \mathbf{B}$$

$$= \begin{bmatrix} 0 & 0{,}04 & -0{,}06 & 0{,}03 \\ 0{,}04 & 0 & 0{,}025 & -0{,}045 \\ -0{,}06 & -0{,}025 & 0 & -0{,}27 \\ 0{,}03 & -0{,}045 & -0{,}27 & 0 \end{bmatrix}$$

next page.

SU, DEPARTMENT OF STATISTICS

Room: Brunnsv. Anonymous code: MMMG-0004 Sheet number: 6

4. e. (cont.)

$$RMSR = \{formula\}$$

$$= \sqrt{\frac{0^2 + 0,04^2 + (-0,06)^2 + \ldots (-0,27)^2 + 0^2}{4(4-1)/2}}$$

$$= \sqrt{0,1633/6} = 0,16$$

The mean difference between correlations of hypothetical data and reprod. correlations is 0,16 which could be consider decent measure (although lower is desired) of RMSR.

5. a) no, since mean-correction results in the same covariance matrix FA and PCA are applied to.

5b. yes, since standardization leads to corresponding correlation matrix being used instead of covariance matrix when FA / PCA is applied.

ptp.

(5) a. Set $X = EBITASS$, $Y = ROTC$,

$\Rightarrow \bar{x} = (0{,}1 + 0{,}2 + \ldots + 0{,}04)/11 = 0{,}109$, $\bar{y} = 0{,}117$

$SS(X) = (0{,}1 - 0{,}109)^2 + (0{,}2 - 0{,}109)^2 + \ldots + (0{,}04 - 0{,}109)^2$

$= 0{,}1123$, $SS(Y) = 0{,}1422$, $CP(XY) = 0{,}0998$

$$\Rightarrow \boxed{SSCP_T = \begin{bmatrix} 0{,}1123 & 0{,}0998 \\ 0{,}0998 & 0{,}1422 \end{bmatrix}}$$

Repeat for group 1 (failed firms):

$$SSCP_1 = \begin{bmatrix} 0{,}0283 & 0 \\ 0 & 0{,}02 \end{bmatrix}$$

For group 2 (non-failed firms):

$$SSCP_2 = \begin{bmatrix} 0{,}0112 & 0{,}0187 \\ 0{,}0187 & 0{,}0319 \end{bmatrix}$$

$$\boxed{SSCP_W = SSCP_1 + SSCP_2 = \begin{bmatrix} 0{,}0395 & 0{,}0187 \\ 0{,}0187 & 0{,}0519 \end{bmatrix}}$$

$$\boxed{SSCP_B = SSCP_T - SSCP_W = \begin{bmatrix} 0{,}0728 & 0{,}0811 \\ 0{,}0811 & 0{,}0903 \end{bmatrix}}$$

m

ok.

SS is relatively small within groups compared to between. Assumption that groups are homogenous could be considered fulfilled for DA.

SU, DEPARTMENT OF STATISTICS

Room: Brunnsv. Anonymous code: ______ Sheet number: 7

6 (a) Pattern loadings given by coefficients in equations.

Indicator	Pattern loading	
	F_1	F_2
X_1	0,164	0,824
X_4	0,906	0,134
X_6	0,827	0,016.

ok

Answer in table above.

6 (b) Formula for $Corr(x_j; x_k)$ goves with info.

$$Corr(X_1, X_2) = 0{,}164 \cdot 0{,}065 + 0{,}824 \cdot 0{,}959 + (0{,}164 \cdot 0{,}959 + 0{,}824 \cdot 0{,}065) \cdot (-0{,}4)$$

$= 0{,}736$ ok for $\phi_{12} = -0{,}4$

For $\phi_{12} = 0{,}4$: $Corr(X_1, X_2) = 0{,}858$

ok

ptp.

6 (c) Variance accounted for (shared variance, SV) is given by

$SV = (\lambda_1 + \lambda_2 \phi)^2 \Rightarrow SV_1 = \left(0{,}164 + 0{,}824(-0{,}4)\right)^2 = 0{,}051$ for $\phi_{12} = -0{,}4$, $SV_1 = 0{,}188$ for $\phi_{12} = 0{,}4$,

$SV_2 = 0{,}102$ for $\phi_{12} = -0{,}4$ and
$SV_2 = 0{,}201$ for $\phi_{12} = 0{,}4$

$\Rightarrow$

Indicator	SV	
	$\phi_{12} = -0{,}4$	$\phi_{12} = 0{,}4$
X_1	0,051	0,188
X_2	0,102	0,201
Total	0,153	0,389

Total variance: $V(X_1) + V(X_2) = \{\text{assumption: } V(X_1) = V(X_2) = 1\} = 1 + 1 = 2$

$\Rightarrow$ Variance not acc. for:
2 − "Total SV" = 2 − 0,153 = 1,847
$\Rightarrow$ Perc. not acc. for: 1,847 / 2 = 92,4% for $\phi_{12} = -0{,}4$ and 80,6% for $\phi_{12} = 0{,}4$.

1x

$\boxed{\begin{array}{l} 92{,}4\% \text{ for } \phi_{12} = 0{,}4 \text{ and} \\ 80{,}6\% \text{ for } \phi_{12} = 0{,}4 \end{array}}$

next page

SU, DEPARTMENT OF STATISTICS

Room: Brunnsv. Anonymous code: MME-0004 Sheet number: 8

6. c. (cont.)

Looking at formula for SV:
$(\lambda_1 + \lambda_2 \phi)^2$, increasing correlation ϕ between constructs increases the value of term $\lambda_2 \phi$ and thereby SV, which in turn decreases percentage of variance not accounted for.

$\Rightarrow$ We should find following for percentage of variance not acc. for (PVNAF) for given correlations:

$$PVNAF(\phi = 0{,}9) < PVNAF(\phi = 0) < PVNAF(\phi = -0{,}9).$$

Our calculus aggrees with this statement (see previous results)

Looking at formula for $Corr(X_j, X_k)$, correlation between indicators increase when ϕ increase

$\Rightarrow Corr(X_1, X_2, \phi = 0{,}9) > Corr(X_1, X_2, \phi = 0)$
$> Corr(X_1, X_2, \phi = -0{,}9)$.

This is also backed by given calculus results.

(1) Let each column represent variables X, Y, Q

$\Rightarrow \bar{x} = (7+4+\ldots+7)/6 = 6{,}67$, $\bar{y} = 3{,}5$, $\bar{q} = 5{,}5$

$$\Rightarrow s_X = \sqrt{\frac{(7-6{,}67)^2 + (4-6{,}67)^2 + \ldots + (7-6{,}67)^2}{6-1}} = 1{,}506$$

where $6-1=5$ is degrees of freedom, (6 observations),

$s_Y = 1{,}871$, $s_Q = 3{,}082$,

Standardize observations:

$$z_X = \frac{x-\bar{x}}{s_X}$$

$\Rightarrow$

z_X	z_Y	z_Q
0,221	0,267	−0,811
−1,771	−1,336	0,811
⋮		
0,221	−0,802	1,136

$\Rightarrow SS(X) = 0{,}221^2 + (-1{,}771)^2 + \ldots + 0{,}221^2 = 5$

since mean is zero for standardized variables, $SS(Y) = 5$, $SS(Q) = 5$,

$CP(X,Y) = 4{,}260$, $CP(X,Q) = -2{,}155$, $CP(Y,Q) = -3{,}72$

$$\Rightarrow SSCP = \begin{bmatrix} 5 & 4{,}260 & -2{,}155 \\ 4{,}260 & 5 & -3{,}729 \\ -2{,}155 & -3{,}729 & 5 \end{bmatrix}$$

next page

SU, DEPARTMENT OF STATISTICS

Room: Brunnsv. Anonymous code: MME-CCC4 Sheet number: 9

Corr. matrix given by SSCP / (d.f.)
where d.f. = 6 − 1 = 5

$\Rightarrow$ Corr. matrix R,

$$R = \begin{bmatrix} 1 & 0,85 & -0,43 \\ 0,85 & 1 & -0,75 \\ -0,43 & -0,75 & 1 \end{bmatrix}$$

answer: (correlation matrix).

VG +

Eigenvalues given when

$|R - \lambda I| = 0$ where λ is eigenvalue, I is identity matrix.

$$|R - \lambda I| = (1-\lambda)^3 + 0,85(-0,75)(-0,43) + (-0,43)\cdot 0,85(-0,75) - (-0,43)(-0,43)(1-\lambda) - 0,85\cdot 0,85(1-\lambda) - (-0,75)(-0,75)(1-\lambda)$$

$$= (1-\lambda)^3 + 2\cdot 0,85(-0,75)(-0,43) - (-0,43)^2(1-\lambda) - 0,85^2(1-\lambda) - (0,75)^2(1-\lambda)$$

$$= (1-\lambda)(1-\lambda)^2 + 0,548 - 0,185(1-\lambda) - 0,723(1-\lambda) - 0,563(1-\lambda)$$

$$= (1-\lambda)(1-2\lambda+\lambda^2) - 1,471 + 1,471\lambda + 0,548$$

$$= 1 - 2\lambda + \lambda^2 - \lambda + 2\lambda^2 - \lambda^3 + 1,471\lambda - 0,923$$

$$= -\lambda^3 + 3\lambda^2 - 1,529\lambda + 0,077$$

ptp.

1. (cont.)

$\Rightarrow -\lambda^3 + 3\lambda^2 - 1{,}529\lambda + 0{,}077 = 0$

$\lambda^3 - 3\lambda^2 + 1{,}529\lambda - 0{,}077 = 0$

$\lambda_1 \approx 2.37$
$\lambda_2 \approx 0.58$

Set $\lambda = x - (-3)/3 = x + 1$

$\Rightarrow (x+1)^3 - 3(x+1)^2 + 1{,}529(x+1) - 0{,}077 = 0$

$(x+1)(x^2+2x+1) - 3(x^2+2x+1) + 1{,}529x + 1{,}529 - 0{,}077 = 0$

$(x-2)(x^2+2x+1) + 1{,}529x + 1{,}452 = 0$

$\underline{x^3} + \underline{2x^2} + \underline{x} - \underline{2x^2} - \underline{4x} - \underline{2} + \underline{1{,}529x} + \underline{1{,}452} = 0$

$x^3 - 1{,}471x - 0{,}548 = 0$

$\Rightarrow$ {with formula}

Comments: I'm getting no real solutions using formula and unfortunately not enough time to recalculate further.

The third degree polynomial should have three roots and thereby 3 eigenvalues where $\lambda_1 > \lambda_2 > \lambda_3$.

Total variance is given by $(\lambda_1 + \lambda_2 + \lambda_3)$, sum of all eigenvalues. How much variance first two principal components explain is then given by $(\lambda_1 + \lambda_2)/(\lambda_1 + \lambda_2 + \lambda_3)$. I used both mean corr. and stand. data to get correlation matrix. Perhaps better result would be achieved using covariance matrix instead.

Logic ok