UNIVERSITI SAINS MALAYSIA

Peperiksaan Kursus Semasa Cuti Panjang
Sidang Akademik 1997/98

April 1998

MSG 462 - Analisis Multivariat

Masa: [3 jam]

ARAHAN KEPADA CALON:

Sila pastikan bahawa kertas peperiksaan ini mengandungi EMPAT soalan di dalam EMPAT halaman dan DUA PULUH SATU halaman lampiran yang bercetak sebelum anda memulakan peperiksaan ini.

Jawab **SEMUA** soalan.

1.(a) Diberikan matriks data

$$\underset{\sim}{X} = \begin{bmatrix} 1 & 4 & 3 \\ 2 & 1 & 0 \\ 5 & 6 & 4 \end{bmatrix}$$

dan gabungan-gabungan linear

$$\underset{\sim}{b}'\underset{\sim}{X} = [1\ 1\ 1]\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \quad \text{dan} \quad \underset{\sim}{c}'\underset{\sim}{X} = [1\ 2\ -3]\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}.$$

Nilaikan min, varians dan kovarians sampel bagi $\underset{\sim}{b}'\underset{\sim}{X}$ dan $\underset{\sim}{c}'\underset{\sim}{X}$.

(20/100)

(b) Katakan $\underset{\sim}{X}$ tertabur $N_3\left(\underset{\sim}{\mu}, \underset{\sim}{\sum}\right)$ dengan $\underset{\sim}{\mu}' = (2, -3, 1)$ dan $\underset{\sim}{\sum} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix}$. Cari taburan $3X_1 - 2X_2 + X_3$.

(20/100)

(c) Katakan $\underset{\sim}{A} = \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix}$.

(i) Tunjukkan bahawa $\underset{\sim}{A}$ adalah tentu positif.

(ii) Tentukan nilai-nilai eigen dan vektor-vektor eigen bagi $\underset{\sim}{A}$.

(iii) Tuliskan penghuraian spektrum bagi $\underset{\sim}{A}$.

(30/100)

...2/-

(d) (i) Dengan menggunakan data

$$\underset{\sim}{Y} = \begin{bmatrix} 10 & 8 & 8 & 9 \\ 1 & 6 & 4 & 6 \end{bmatrix}$$

nilaikan T^2 Hotelling untuk menguji $H_0 : \underset{\sim}{\mu}' = [5,\ 10]$.

(ii) Nyatakan taburan T^2 bagi keadaan dalam (i) di atas. Berikan anggapan-anggapan yang telah anda gunakan.

(iii) Dengan menggunakan (i) dan (ii) di atas, ujikan H_0 pada paras $\alpha = 0.01$. Apakah kesimpulan anda?

(30/100)

2.(a) Pertimbangkan dua set data:

$$\underset{\sim}{X}_1 = \begin{bmatrix} 3 & 2 & 4 \\ 7 & 4 & 7 \end{bmatrix} \text{ dan } \underset{\sim}{X}_2 = \begin{bmatrix} 6 & 5 & 4 \\ 9 & 7 & 8 \end{bmatrix}$$

di mana

$$\underset{\sim}{\bar{x}}_1 = \begin{bmatrix} 3 \\ 6 \end{bmatrix}, \quad \underset{\sim}{\bar{x}}_2 = \begin{bmatrix} 5 \\ 8 \end{bmatrix},$$

dan $\underset{\sim}{S}_p = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$.

(i) Hitungkan fungsi pembezalayan linear

$$y = (\underset{\sim}{\bar{x}}_1 - \underset{\sim}{\bar{x}}_2)' \ \underset{\sim}{S}_p^{-1} \ \underset{\sim}{x}.$$

(ii) Kelaskan cerapan $\underset{\sim}{x}_0' = [2,\ 7]$ sebagai dari populasi π_2 dengan menggunakan petua berikut:

Umpukkan $\underset{\sim}{x}_0$ kepada populasi π_1 jika

$$y_0 \geq \hat{m}$$

dan umpukkan $\underset{\sim}{x}_0$ kepada populasi π_2 jika

$$y_0 < \hat{m}$$

di mana

$$y_0 = (\underset{\sim}{\bar{x}}_1 - \underset{\sim}{\bar{x}}_2)' \ \underset{\sim}{S}_p^{-1} \ \underset{\sim}{x}_0$$

dan

$$\hat{m} = \frac{1}{2}(\underset{\sim}{\bar{x}}_1 - \underset{\sim}{\bar{x}}_2)' \ \underset{\sim}{S}_p^{-1} \ (\underset{\sim}{\bar{x}}_1 + \underset{\sim}{\bar{x}}_2)$$

(40/100)

...3/-

(b) Diberikan lima objek mempunyai matriks jarak berikut:

$$\underset{\sim}{D} = \{d_{ik}\} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{bmatrix} 0 & & & & \\ 3 & 0 & & & \\ 11 & 2 & 0 & & \\ 6 & 9 & 8 & 0 & \\ 9 & 7 & 10 & 5 & 0 \end{bmatrix} \end{array},$$

gunakan (i) kaedah pautan tunggal
(ii) kaedah pautan lengkap

daripada prosedur "*agglomerative hierachical*" untuk memperolehi gambarajah-gambarajah dendogram untuk lima objek itu.

(*60/100*)

3. Tulis nota pendek tentang tajuk-tajuk di bawah:

(a) Komponen prinsipal
(b) Analisis faktor
(c) Analisis kelompok
(d) Analisis pembezalayan

(*100/100*)

4. Bagi setiap bahagian yang berikut, huraikan kesimpulan-kesimpulan anda.

(a) Fisher (1936) mengkaji data daripada Anderson yang berkaitan dengan sampel-sampel rawak bunga iris dari turunan iris setosa (Species 1), versicolor (Species 2), dan virginica (Species 3). Balasan-balasan adalah empat sukatan: panjang "*sepal*", lebar "*sepal*", panjang "*petal*", dan lebar "*petal*". Bagi setiap jenis iris, cerapan-cerapan yang terdiri daripada empat sukatan itu diperhatikan untuk 50 pokok iris.

Data itu dianalisakan dengan menggunakan prosedur DISCRIMINANT dari pakej SPSS. Hasilnya diberikan di dalam Lampiran 1.

(*30/100*)

(b) Pertimbangkan bahagian (a). Komponen-komponen prinsipal bagi matriks kovarians pun telah diperolehi untuk setiap jenis iris melalui prosedur PRINCOMP dari pakej SAS. Output yang diperolehi adalah seperti yang diberikan di Lampiran 2.

(*30/100*)

...4/-

(c) Pengurus produk bagi sebuah firma yang membuat pembersih ingin camkan faktor-faktor major yang konsumer-konsumer menggunakan untuk menilai berbagai-bagai pembersih dalam pasaran. Faktor-faktor ini diandaikan tersembunyi; walau bagaimanapun, pengurusan percaya bahawa berbagai-bagai ciri pembersih adalah penanda-penanda bagi faktor-faktor ini. Analisis faktor dapat digunakan untuk mengenalpasti faktor-faktor ini. Suatu kajian dikendalikan yang mana 143 responden menilai tiga jenis pembersih pada 9 ciri produk dengan menggunakan skala pembezaan semantik lima-titik yang berikut:

Mencuci bersih kotoran - - - - - Gagal mencuci bersih kotoran

Jadual A menunjukkan suatu senarai bagi 9 ciri produk dan Jadual B menunjukkan matriks korelasi antara 9 ciri tersebut.

Lampiran 3 menunjukkan hasil analisis faktor dengan pakej SAS.

(40/100)

Jadual A. Senarai Ciri Pembersih

V1	:	Gentle to natural fabrics
V2	:	Won't harm colours
V3	:	Won't harm synthetics
V4	:	Safe for lingerie
V5	:	Strong, powerful
V6	:	Gets dirt out
V7	:	Makes colours bright
V8	:	Removes grease stains
V9	:	Good for greasy oil

Jadual B. Matriks Korelasi untuk Kajian Pembersih

	V1	V2	V3	V4	V5	V6	V7	V8	V9
V1	1.00000								
V2	0.41901	1.00000							
V3	0.51840	0.57599	1.00000						
V4	0.56641	0.49886	0.64325	1.00000					
V5	0.18122	0.18666	0.29080	0.38360	1.00000				
V6	0.17454	0.24648	0.34428	0.39637	0.57915	1.00000			
V7	0.23034	0.22907	0.41083	0.37699	0.59400	0.57756	1.00000		
V8	0.30647	0.22526	0.34028	0.40391	0.67623	0.70103	0.67682	1.00000	
V9	0.24051	0.21967	0.32854	0.42337	0.69269	0.62280	0.68445	0.69813	1.00000

-ooo0ooo-

- - - - - - - - D I S C R I M I N A N T A N A L Y S I S - - - - - - - -

On groups defined by SPECIES Population

150 (unweighted) cases were processed.
0 of these were excluded from the analysis.
150 (unweighted) cases will be used in the analysis.

Number of Cases by Group

	Number of Cases		
SPECIES	Unweighted	Weighted	Label
1	50	50.0	Setosa
2	50	50.0	Versicolor
3	50	50.0	Virginica
Total	150	150.0	

Group Means

SPECIES	X1	X2	X3	X4
1	5.00600	3.42800	1.46200	.24600
2	5.93600	2.77000	4.26000	1.32600
3	6.58800	2.97400	5.55200	2.02600
Total	5.84333	3.05733	3.75800	1.19933

Group Standard Deviations

SPECIES	X1	X2	X3	X4
1	.35249	.37906	.17366	.10539
2	.51617	.31380	.46991	.19775
3	.63588	.32250	.55189	.27465
Total	.82807	.43587	1.76530	.76224

Pooled Within-Groups Covariance Matrix with 147 degrees of freedom

	X1	X2	X3	X4
X1	.2650082			
X2	.9272109E-01	.1153878		
X3	.1675143	.5524354E-01	.1851878	
X4	.3840136E-01	.3271020E-01	.4266531E-01	.4188163E-01

Pooled Within-Groups Correlation Matrix

	X1	X2	X3	X4
X1	1.00000			
X2	.53024	1.00000		
X3	.75616	.37792	1.00000	
X4	.36451	.47053	.48446	1.00000

Correlations which cannot be computed are printed as '.'

Wilks' Lambda (U-statistic) and univariate F-ratio
with 2 and 147 degrees of freedom

Variable	Wilks' Lambda	F	Significance
X1	.38129	119.3	.0000
X2	.59922	49.16	.0000
X3	.05863	1180.	.0000
X4	.07112	960.0	.0000

Covariance Matrix for Group 1, Setosa

	X1	X2	X3	X4
X1	.1242490			
X2	.9921633E-01	.1436898		
X3	.1635510E-01	.1169796E-01	.3015918E-01	
X4	.1033061E-01	.9297959E-02	.6069388E-02	.1110612E-01

Covariance Matrix for Group 2, Versicolor

	X1	X2	X3	X4
X1	.2664327			
X2	.8518367E-01	.9846939E-01		
X3	.1828980	.8265306E-01	.2208163	
X4	.5577959E-01	.4120408E-01	.7310204E-01	.3910612E-01

Covariance Matrix for Group 3, Virginica

	X1	X2	X3	X4
X1	.4043429			
X2	.9376327E-01	.1040041		
X3	.3032898	.7137959E-01	.3045878	
X4	.4909388E-01	.4762857E-01	.4882449E-01	.7543265E-01

Total Covariance Matrix with 149 degrees of freedom

	X1	X2	X3	X4
X1	.6856935			
X2	-.4243400E-01	.1899794		
X3	1.274315	-.3296564	3.116278	
X4	.5162707	-.1216394	1.295609	.5810063

On groups defined by SPECIES Population

Analysis number 1
Direct method: All variables passing the tolerance test are entered.
Minimum Tolerance Level..... .00100

Canonical Discriminant Functions

Maximum number of functions..... 2
Minimum cumulative percent of variance... 100.00
Maximum significance of Wilks' Lambda.... 1.0000

Prior probability for each group is .33333

Canonical Discriminant Functions

Fcn	Eigenvalue	Pct of Variance	Cum Pct	Canonical Corr	After Fcn	Wilks' Lambda	Chisquare	DF	Sig
					: 0	.0234	546.115	8	.0000
1*	32.1919	99.12	99.12	.9848	: 1	.7780	36.530	3	.0000
2*	.2854	.88	100.00	.4712	:				

* marks the 2 canonical discriminant functions remaining in the analysis.

Standardized Canonical Discriminant Function Coefficients

	FUNC 1	FUNC 2
X1	-.42695	.01241
X2	-.52124	.73526
X3	.94726	-.40104
X4	.57516	.58104

Structure Matrix:

Pooled-within-groups correlations between discriminating variables and canonical discriminant functions
(Variables ordered by size of correlation within function)

	FUNC 1	FUNC 2
X3	.70607*	.16770
X2	-.11901	.86368*
X4	.63318	.73724*
X1	.22260	.31081*

Unstandardized Canonical Discriminant Function Coefficients

	FUNC 1	FUNC 2
X1	-.8293776	.2410215E-01
X2	-1.534473	2.164521
X3	2.201212	-.9319212
X4	2.810460	2.839188
(constant)	-2.105106	-6.661473

Canonical Discriminant Functions evaluated at Group Means (Group Centroids)

Group	FUNC 1	FUNC 2
1	-7.60760	.21513
2	1.82505	-.72790
3	5.78255	.51277

Classification Results -

Actual Group		No. of Cases	Predicted Group Membership 1	2	3
Group Setosa	1	50	50 100.0%	0 .0%	0 .0%
Group Versicolor	2	50	0 .0%	48 96.0%	2 4.0%
Group Virginica	3	50	0 .0%	1 2.0%	49 98.0%

Percent of "grouped" cases correctly classified: 98.00%

Classification Processing Summary
150 Cases were processed.
0 Cases were excluded for missing or out-of-range group codes.
0 Cases had at least one missing discriminating variable.
150 Cases were used for printed output.

---------- SPECIES=1 ----------

Principal Component Analysis

50 Observations
4 Variables

Simple Statistics

	PETALLEN	PETALWID	SEPALLEN	SEPALWID
Mean	5.006000000	3.428000000	1.462000000	0.2460000000
Std	0.352489687	0.379064369	0.173663996	0.1053855894

Covariance Matrix

	PETALLEN	PETALWID	SEPALLEN	SEPALWID
PETALLEN	0.1242489796	0.0992163265	0.0163551020	0.0103306122
PETALWID	0.0992163265	0.1436897959	0.0116979592	0.0092979592
SEPALLEN	0.0163551020	0.0116979592	0.0301591837	0.0060693878
SEPALWID	0.0103306122	0.0092979592	0.0060693878	0.0111061224

Total Variance = 0.3092040816

Eigenvalues of the Covariance Matrix

	Eigenvalue	Difference	Proportion	Cumulative
PRIN1	0.236456	0.199537	0.764724	0.76472
PRIN2	0.036919	0.010122	0.119399	0.88412
PRIN3	0.026796	0.017763	0.086662	0.97079
PRIN4	0.009033	.	0.029215	1.00000

Eigenvectors

	PRIN1	PRIN2	PRIN3	PRIN4
PETALLEN	0.669078	0.597884	-.439963	-.036077
PETALWID	0.734148	-.620673	0.274607	-.019550
SEPALLEN	0.096544	0.490056	0.832449	-.239901
SEPALWID	0.063564	0.130938	0.195068	0.969930

---------- SPECIES=2 ----------

Principal Component Analysis

50 Observations
4 Variables

Simple Statistics

	PETALLEN	PETALWID	SEPALLEN	SEPALWID
Mean	5.936000000	2.770000000	4.260000000	1.326000000
Std	0.516171147	0.313798323	0.469910977	0.197752680

Covariance Matrix

	PETALLEN	PETALWID	SEPALLEN	SEPALWID
PETALLEN	0.2664326531	0.0851836735	0.1828979592	0.0557795918
PETALWID	0.0851836735	0.0984693878	0.0826530612	0.0412040816
SEPALLEN	0.1828979592	0.0826530612	0.2208163265	0.0731020408
SEPALWID	0.0557795918	0.0412040816	0.0731020408	0.0391061224

Total Variance = 0.6248244898

Eigenvalues of the Covariance Matrix

	Eigenvalue	Difference	Proportion	Cumulative
PRIN1	0.487874	0.415490	0.780818	0.78082
PRIN2	0.072384	0.017608	0.115847	0.89666
PRIN3	0.054776	0.044986	0.087666	0.98433
PRIN4	0.009790	.	0.015669	1.00000

Eigenvectors

	PRIN1	PRIN2	PRIN3	PRIN4
PETALLEN	0.686724	-.669089	0.265083	0.102280
PETALWID	0.305347	0.567465	0.729618	-.228919
SEPALLEN	0.623663	0.343327	-.627165	-.315967
SEPALWID	0.214984	0.335305	-.063661	0.915041

---------- SPECIES=3 ----------

Principal Component Analysis

50 Observations
4 Variables

Simple Statistics

	PETALLEN	PETALWID	SEPALLEN	SEPALWID
Mean	6.588000000	2.974000000	5.552000000	2.026000000
Std	0.635879593	0.322496638	0.551894696	0.274650056

Covariance Matrix

	PETALLEN	PETALWID	SEPALLEN	SEPALWID
PETALLEN	0.4043428571	0.0937632653	0.3032897959	0.0490938776
PETALWID	0.0937632653	0.1040040816	0.0713795918	0.0476285714
SEPALLEN	0.3032897959	0.0713795918	0.3045877551	0.0488244898
SEPALWID	0.0490938776	0.0476285714	0.0488244898	0.0754326531

Total Variance = 0.8883673469

Eigenvalues of the Covariance Matrix

	Eigenvalue	Difference	Proportion	Cumulative
PRIN1	0.695255	0.588704	0.782621	0.78262
PRIN2	0.106551	0.054256	0.119941	0.90256
PRIN3	0.052295	0.018030	0.058867	0.96143
PRIN4	0.034266	.	0.038572	1.00000

Eigenvectors

	PRIN1	PRIN2	PRIN3	PRIN4
PETALLEN	0.741017	-.165259	-.534450	0.371412
PETALWID	0.203288	0.748643	-.325375	-.540684
SEPALLEN	0.627892	-.169428	0.651524	-.390593
SEPALWID	0.123775	0.619288	0.428965	0.645872

Lampiran 3

Factor Analysis of Detergent Study

Correlations

	V1	V2	V3	V4	V5
V1	1.00000	0.41901	0.51840	0.56641	0.18122
V2	0.41901	1.00000	0.57599	0.49886	0.18666
V3	0.51840	0.57599	1.00000	0.64325	0.29080
V4	0.56641	0.49886	0.64325	1.00000	0.38360
V5	0.18122	0.18666	0.29080	0.38360	1.00000
V6	0.17454	0.24648	0.34428	0.39637	0.57915
V7	0.23034	0.22907	0.41083	0.37699	0.59400
V8	0.30647	0.22526	0.34028	0.40391	0.67623
V9	0.24051	0.21967	0.32854	0.42337	0.69269

	V6	V7	V8	V9
V1	0.17454	0.23034	0.30647	0.24051
V2	0.24648	0.22907	0.22526	0.21967
V3	0.34428	0.41083	0.34028	0.32854
V4	0.39637	0.37699	0.40391	0.42337
V5	0.57915	0.59400	0.67623	0.69269
V6	1.00000	0.57756	0.70103	0.62280
V7	0.57756	1.00000	0.67682	0.68445
V8	0.70103	0.67682	1.00000	0.69813
V9	0.62280	0.68445	0.69813	1.00000

Initial Factor Method: Iterated Principal Factor Analysis

Partial Correlations Controlling all other Variables

	V1	V2	V3	V4	V5
V1	1.00000	0.11310	0.20015	0.33260	-0.09494
V2	0.11310	1.00000	0.34766	0.14728	-0.01313
V3	0.20015	0.34766	1.00000	0.34409	-0.01125
V4	0.33260	0.14728	0.34409	1.00000	0.10853
V5	-0.09494	-0.01313	-0.01125	0.10853	1.00000
V6	-0.17318	0.05943	0.06690	0.10531	0.07087
V7	-0.05410	-0.03595	0.22926	-0.05065	0.08684
V8	0.21683	-0.02211	-0.05690	-0.02930	0.26555
V9	0.01440	0.00078	-0.06517	0.11763	0.32215

	V6	V7	V8	V9
V1	-0.17318	-0.05410	0.21683	0.01440
V2	0.05943	-0.03595	-0.02211	0.00078
V3	0.06690	0.22926	-0.05690	-0.06517
V4	0.10531	-0.05065	-0.02930	0.11763
V5	0.07087	0.08684	0.26555	0.32215
V6	1.00000	0.06369	0.39450	0.14688
V7	0.06369	1.00000	0.26431	0.31475

V8	0.39450	0.26431	1.00000	0.17384
V9	0.14688	0.31475	0.17384	1.00000

Kaiser's Measure of Sampling Adequacy: Over-all MSA = 0.86914242

V1	V2	V3	V4	V5
0.803540	0.861835	0.823952	0.859024	0.902492

V6	V7	V8	V9
0.890552	0.894296	0.859837	0.890377

Prior Communality Estimates: ONE

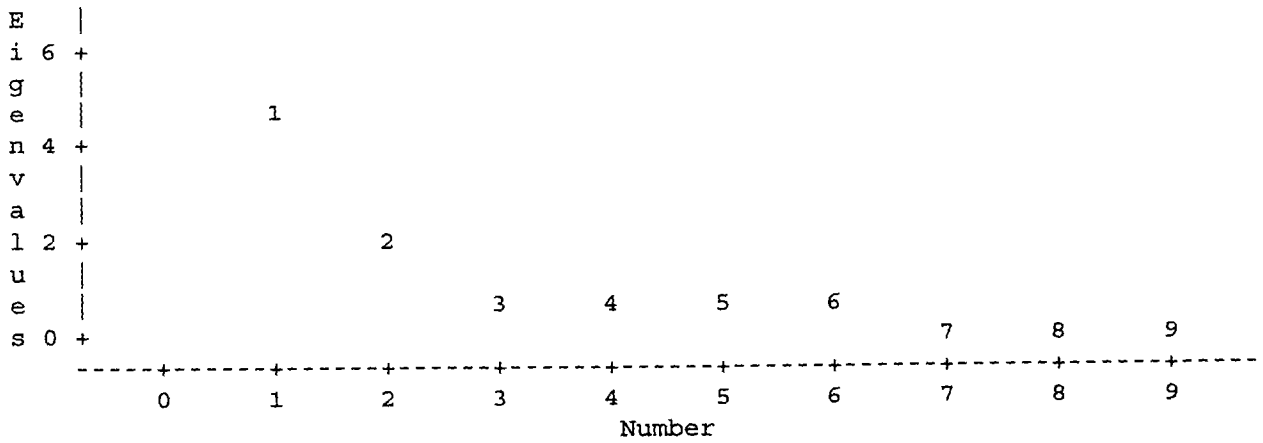
Preliminary Eigenvalues: Total = 9 Average = 1

	1	2	3	4	5
Eigenvalue	4.554639	1.701640	0.617593	0.453211	0.433674
Difference	2.853000	1.084046	0.164382	0.019538	0.003908
Proportion	0.5061	0.1891	0.0686	0.0504	0.0482
Cumulative	0.5061	0.6951	0.7638	0.8141	0.8623

	6	7	8	9
Eigenvalue	0.429765	0.318462	0.264200	0.226815
Difference	0.111303	0.054262	0.037385	
Proportion	0.0478	0.0354	0.0294	0.0252
Cumulative	0.9101	0.9454	0.9748	1.0000

2 factors will be retained by the MINEIGEN criterion.

Scree Plot of Eigenvalues



Iter	Change	Communalities				
1	0.391654	0.60835	0.60955	0.72566	0.70894	0.70498
		0.66774	0.68794	0.77980	0.76332	

2	0.116529	0.49182	0.49351	0.67306	0.65330	0.64097
		0.59718	0.62500	0.74641	0.72253	
3	0.035735	0.45614	0.45777	0.66732	0.64630	0.62583
		0.58187	0.61214	0.74370	0.71634	
4	0.011833	0.44444	0.44594	0.66977	0.64752	0.62174
		0.57832	0.60935	0.74486	0.71580	
5	0.004325	0.44019	0.44161	0.67249	0.64910	0.62046
		0.57741	0.60866	0.74585	0.71594	
6	0.001843	0.43847	0.43987	0.67433	0.65003	0.62000
		0.57714	0.60845	0.74639	0.71606	
7	0.001107	0.43771	0.43911	0.67544	0.65047	0.61982
		0.57705	0.60837	0.74666	0.71612	
8	0.000638	0.43735	0.43876	0.67608	0.65066	0.61975
		0.57702	0.60833	0.74678	0.71615	

Convergence criterion satisfied.

Eigenvalues of the Reduced Correlation Matrix:
Total = 5.47074731 Average = 0.60786081

	1	2	3	4	5
Eigenvalue	4.189535	1.281332	0.120877	0.082178	0.051374
Difference	2.908204	1.160454	0.038700	0.030804	0.064033
Proportion	0.7658	0.2342	0.0221	0.0150	0.0094
Cumulative	0.7658	1.0000	1.0221	1.0371	1.0465

	6	7	8	9
Eigenvalue	-0.012659	-0.022006	-0.086772	-0.133112
Difference	0.009347	0.064765	0.046340	
Proportion	-0.0023	-0.0040	-0.0159	-0.0243
Cumulative	1.0442	1.0402	1.0243	1.0000

Factor Pattern

	FACTOR1	FACTOR2
V1	0.47442	0.46073
V2	0.46575	0.47099
V3	0.64075	0.51529
V4	0.68291	0.42929
V5	0.72117	-0.31569
V6	0.71868	-0.24599
V7	0.74513	-0.23046
V8	0.81052	-0.29973
V9	0.78858	-0.30708

Variance explained by each factor

FACTOR1	FACTOR2

4.189535 1.281332

Final Communality Estimates: Total = 5.470867

V1	V2	V3	V4	V5
0.437347	0.438757	0.676076	0.650656	0.619750

V6	V7	V8	V9
0.577018	0.608332	0.746784	0.716148

Residual Correlations With Uniqueness on the Diagonal

	V1	V2	V3	V4	V5
V1	0.56265	-0.01895	-0.02299	0.04464	-0.01547
V2	-0.01895	0.56124	0.03487	-0.02140	-0.00054
V3	-0.02299	0.03487	0.32392	-0.01553	-0.00862
V4	0.04464	-0.02140	-0.01553	0.34934	0.02662
V5	-0.01547	-0.00054	-0.00862	0.02662	0.38025
V6	-0.05308	0.02761	0.01054	0.01117	-0.01680
V7	-0.01698	-0.00943	0.05215	-0.03293	-0.01612
V8	0.06004	-0.01107	-0.02461	-0.02093	-0.00292
V9	0.00787	-0.00298	-0.01850	0.01667	0.02705

	V6	V7	V8	V9
V1	-0.05308	-0.01698	0.06004	0.00787
V2	0.02761	-0.00943	-0.01107	-0.00298
V3	0.01054	0.05215	-0.02461	-0.01850
V4	0.01117	-0.03293	-0.02093	0.01667
V5	-0.01680	-0.01612	-0.00292	0.02705
V6	0.42298	-0.01464	0.04479	-0.01947
V7	-0.01464	0.39167	0.00380	0.02609
V8	0.04479	0.00380	0.25322	-0.03307
V9	-0.01947	0.02609	-0.03307	0.28385

Root Mean Square Off-diagonal Residuals: Over-all = 0.02632479

V1	V2	V3	V4	V5
0.035165	0.019415	0.027014	0.025801	0.016975

V6	V7	V8	V9
0.028936	0.025862	0.031411	0.021143

Partial Correlations Controlling Factors

	V1	V2	V3	V4	V5
V1	1.00000	-0.03373	-0.05385	0.10068	-0.03345
V2	-0.03373	1.00000	0.08177	-0.04832	-0.00117
V3	-0.05385	0.08177	1.00000	-0.04616	-0.02455
V4	0.10068	-0.04832	-0.04616	1.00000	0.07305

V5	-0.03345	-0.00117	-0.02455	0.07305	1.00000
V6	-0.10882	0.05667	0.02848	0.02906	-0.04189
V7	-0.03618	-0.02011	0.14640	-0.08903	-0.04178
V8	0.15906	-0.02937	-0.08592	-0.07038	-0.00940
V9	0.01970	-0.00747	-0.06102	0.05292	0.08233

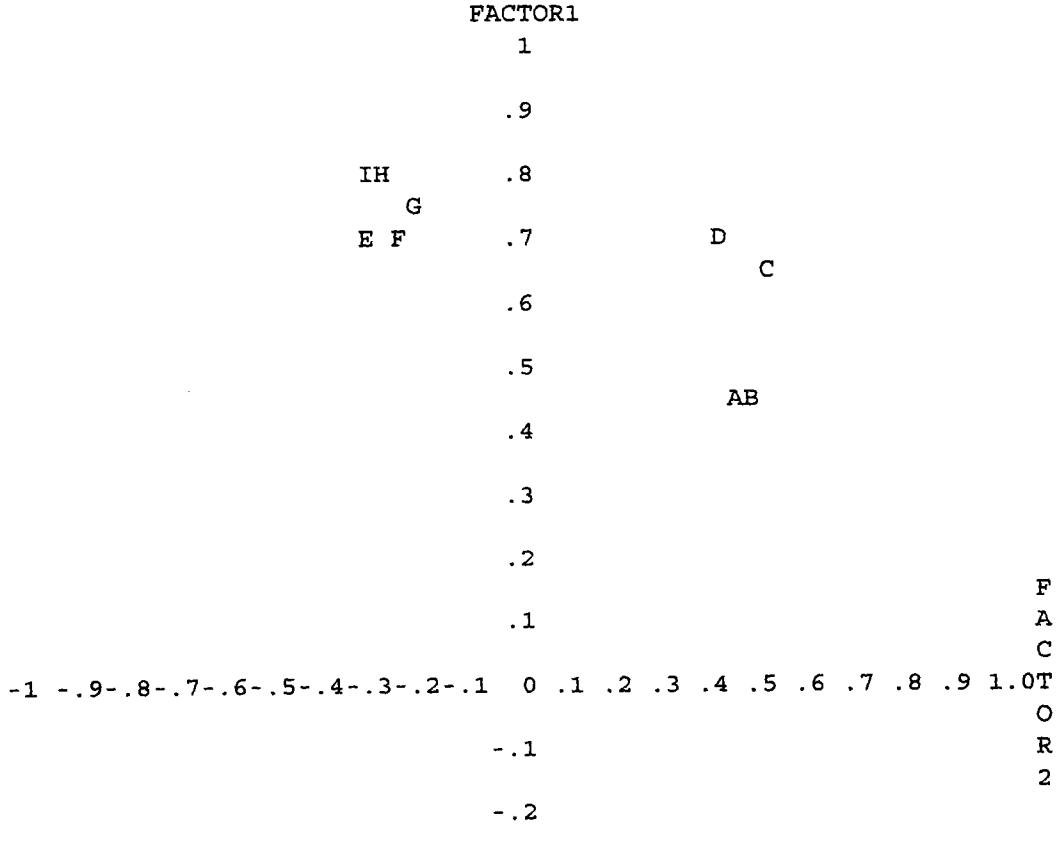
	V6	V7	V8	V9
V1	-0.10882	-0.03618	0.15906	0.01970
V2	0.05667	-0.02011	-0.02937	-0.00747
V3	0.02848	0.14640	-0.08592	-0.06102
V4	0.02906	-0.08903	-0.07038	0.05292
V5	-0.04189	-0.04178	-0.00940	0.08233
V6	1.00000	-0.03598	0.13686	-0.05620
V7	-0.03598	1.00000	0.01206	0.07824
V8	0.13686	0.01206	1.00000	-0.12334
V9	-0.05620	0.07824	-0.12334	1.00000

Root Mean Square Off-diagonal Partials: Over-all = 0.07084171

V1	V2	V3	V4	V5
0.082257	0.042858	0.075558	0.067514	0.046673

V6	V7	V8	V9
0.072151	0.071047	0.095313	0.069139

Plot of Factor Pattern for FACTOR1 and FACTOR2



5

-.3

-.4

-.5

-.6

-.7

-.8

-.9

-1

V1 =A V2 =B V3 =C V4 =D V5 =E
V6 =F V7 =G V8 =H V9 =I

Rotation Method: Varimax

Orthogonal Transformation Matrix

	1	2
1	0.82165	0.57000
2	-0.57000	0.82165

Rotated Factor Pattern

	FACTOR1	FACTOR2
V1	0.12719	0.64897
V2	0.11422	0.65247
V3	0.23275	0.78861
V4	0.31642	0.74198
V5	0.77249	0.15168
V6	0.73072	0.20753
V7	0.74360	0.23536
V8	0.83681	0.21572
V9	0.82296	0.19718

Variance explained by each factor

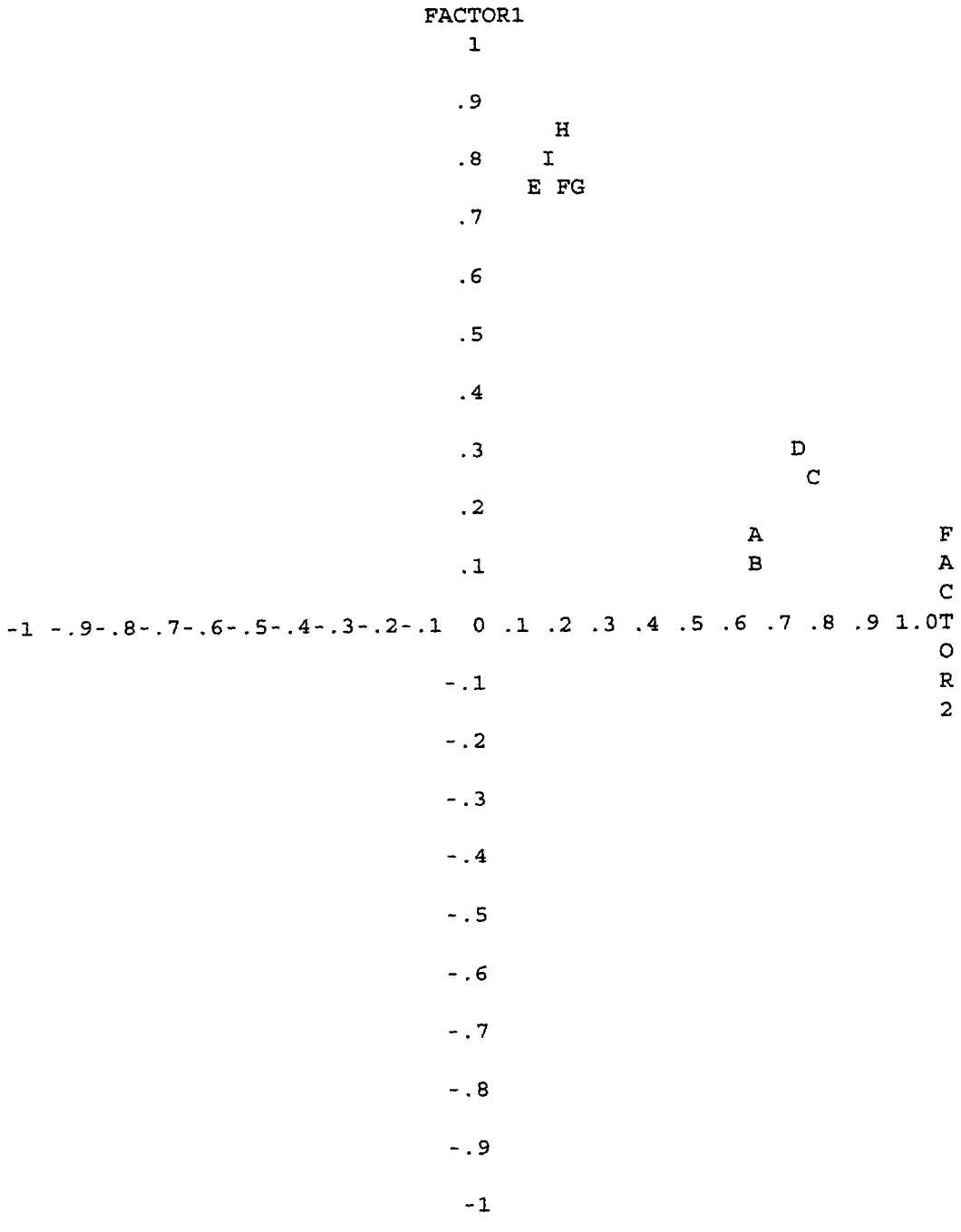
FACTOR1	FACTOR2
3.244671	2.226196

Final Communality Estimates: Total = 5.470867

V1	V2	V3	V4	V5
0.437347	0.438757	0.676076	0.650656	0.619750

V6	V7	V8	V9
0.577018	0.608332	0.746784	0.716148

Plot of Factor Pattern for FACTOR1 and FACTOR2



V1 =A V2 =B V3 =C V4 =D V5 =E
V6 =F V7 =G V8 =H V9 =I

MSG 462 - ANALISIS MULTIVARIAT

LAMPIRAN

Tatatanda adalah seperti di dalam kuliah.

1. Penguraian spektrum bagi suatu matriks simetrik k x k, $\underset{\sim}{A}$ diberikan oleh

$$\underset{\sim}{A} = \lambda_1 \underset{\sim}{e}_1 \underset{\sim}{e}_1' + \lambda_2 \underset{\sim}{e}_2 \underset{\sim}{e}_2' + \ldots + \lambda_k \underset{\sim}{e}_k \underset{\sim}{e}_k'$$

di mana $\lambda_1, \lambda_2, \ldots, \lambda_k$ adalah nilai-nilai eigen $\underset{\sim}{A}$ dan $\underset{\sim}{e}_1, \underset{\sim}{e}_2, \ldots, \underset{\sim}{e}_k$ adalah vektor-vektor eigen terpiawai yang berkaitan.

2. Katakan $\underset{\sim}{X}$ mempunyai $E(\underset{\sim}{X}) = \underset{\sim}{\mu}$ dan $Kov(\underset{\sim}{X}) = \Sigma$. Maka $\underset{\sim}{c}'\underset{\sim}{X}$ mempunyai min. $\underset{\sim}{c}'\underset{\sim}{\mu}$ dan varians. $\underset{\sim}{c}'\Sigma\,\underset{\sim}{c}$.

3. f.k.k. normal bivariat:

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{\sigma_{11}\sigma_{22}(1-\rho_{12}^2)}} \times \exp\left\{ -\frac{1}{2(1-\rho_{12}^2)} \left[\left(\frac{x_1 - \mu_1}{\sqrt{\sigma_{11}}} \right)^2 + \left(\frac{x_2 - \mu_2}{\sqrt{\sigma_{22}}} \right)^2 - 2\rho_{12} \left(\frac{x_1 - \mu_1}{\sqrt{\sigma_{11}}} \right) \left(\frac{x_2 - \mu_2}{\sqrt{\sigma_{22}}} \right) \right] \right\}$$

4. f.k.k. normal multivariat:

$$f(\underset{\sim}{x}) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} e^{-(1/2)(\underset{\sim}{x} - \underset{\sim}{\mu})' \Sigma^{-1} (\underset{\sim}{x} - \underset{\sim}{\mu})}$$

5. Jika $\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)$, maka $A\underset{\sim}{X} \sim N_q(A\underset{\sim}{\mu}, A \Sigma A')$.

...2/-

6. Satu sampel:

(a) $T^2 = n (\bar{X} - \mu)' S^{-1} (\bar{X} - \mu)$

$$\bar{X} = \frac{1}{n} \sum_{j=1}^{n} X_j , \quad S = \frac{1}{n-1} \sum_{j=1}^{n} (X_j - \bar{X})(X_j - \bar{X})'$$

$$T^2 \sim \frac{(n-1)p}{n-p} F_{p,n-p}$$

(b) Lambda Wilks $\Lambda^{2/n} = |\hat{\Sigma}| \Big/ |\hat{\Sigma}_0| = \left(1 + \frac{T^2}{(n-1)}\right)^{-1}$

(c) Selang keyakinan serentak 100(1-α)% bagi $\ell'\mu$:

$$\ell' \bar{X} \pm \sqrt{\frac{p(n-1)}{n(n-p)} F_{p,n-p}(\alpha)\ \ell' S \ell}$$

(d) Selang keyakinan serentak Bonferroni 100(1-α)% bagi

μ_i, $i = 1, \ldots, p$:

$$\bar{X}_i \pm t_{n-1}\left(\frac{\alpha}{2p}\right) \sqrt{\frac{S_{ii}}{n}}$$

7. Dua sampel tak bersandar:

(a) $$T^2 = \left[\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)\right]' \left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right) S_p\right]^{-1} \left[\bar{X}_1 - \bar{X}_2 - \left(\mu_1 - \mu_2\right)\right]$$

$$T^2 \sim \frac{\left(n_1 + n_2 - 2\right)p}{\left(n_1 + n_2 - p - 1\right)} F_{p,\, n_1 + n_2 - p - 1}$$

...3/-

(b) Selang keyakinan serentak $100(1-\alpha)\%$ bagi $\underset{\sim}{\ell}'(\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2)$:

$$\underset{\sim}{\ell}'\left(\bar{\underset{\sim}{X}}_1 - \bar{\underset{\sim}{X}}_2\right) \pm c \sqrt{\underset{\sim}{\ell}'\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\underset{\sim}{S}_p\,\underset{\sim}{\ell}}$$

$$\text{di mana } c^2 = \frac{(n_1 + n_2 - 2)p}{n_1 + n_2 - p - 1} F_{p,\, n_1 + n_2 - p - 1}$$

8. MANOVA satu-hala:

(a) $$\underset{\sim}{B} = \sum_{\ell=1}^{g} n_\ell (\bar{\underset{\sim}{x}}_\ell - \bar{\underset{\sim}{x}})(\bar{\underset{\sim}{x}}_\ell - \bar{\underset{\sim}{x}})'$$

$$\underset{\sim}{W} = \sum_{\ell=1}^{g} \sum_{j=1}^{n_\ell} \left(\underset{\sim}{x}_{\ell j} - \bar{\underset{\sim}{x}}_\ell\right)\left(\underset{\sim}{x}_{\ell j} - \bar{\underset{\sim}{x}}_\ell\right)'$$

$$\Lambda^* = \frac{|\underset{\sim}{W}|}{|\underset{\sim}{B} + \underset{\sim}{W}|}$$

(b) Selang keyakinan serentak $100(1-\alpha)\%$ bagi $\tau_{ki} - \tau_{\ell i}$:

$$\bar{X}_{ki} - \bar{X}_{\ell i} \pm t_{n-g}\left(\frac{\alpha}{pg(g-1)}\right)\sqrt{\frac{W_{ii}}{n-g}\left(\frac{1}{n_k} + \frac{1}{n_\ell}\right)}$$

$i = 1, 2, \ldots, p \quad , \quad \ell < k = 1, 2, \ldots, g$

9. Andaikan $\underset{\sim}{E}$ mempunyai d.k. m_E dan H mempunyai d.k. m_H.

Katakan $\Lambda = \dfrac{|\underset{\sim}{E}|}{|\underset{\sim}{E} + \underset{\sim}{H}|}$.

...4/-

Maka (1) Untuk p = 1,

$$\left[\frac{1-\Lambda}{\Lambda}\right]\frac{m_E}{m_H} \sim F_{m_H,\, m_E} \quad \text{bagi sebarang } m_H.$$

(2) Untuk $m_H = 1$,

$$\left[\frac{1-\Lambda}{\Lambda}\right]\frac{m_E+1-p}{P} \sim F_{P,\, m_E+1-p} \quad \text{bagi sebarang } P.$$

(3) Untuk p = 2,

$$\left[\frac{1-\Lambda^{1/2}}{\Lambda^{1/2}}\right]\left[\frac{m_E-1}{m_H}\right] \sim F_{2m_H,\,2(m_E-1)}$$

untuk $m_H \geq 2$.

(4) Untuk $m_H = 2$,

$$\left[\frac{1-\Lambda^{1/2}}{\Lambda^{1/2}}\right]\left[\frac{m_E+1-p}{p}\right] \sim F_{2p,\,2(m_E+1-p)}$$

untuk $p \geq 2$.

Pembetulan Bartlett: Katakan $n_o = m_E + m_H$

Bagi m_E besar,

$$-f \log \Lambda \sim \chi^2_{pm_H}$$

di mana $f = m_E - \frac{1}{2}\left(p - m_H + 1\right)$

$$= n_o - \frac{1}{2}\left(p + m_H + 1\right)$$

10. MANOVA dua-hala:

$$SSP_{\text{faktor 1}} = \sum_{\ell=1}^{g} bn\left[\bar{x}_{\ell\cdot} - \bar{x}\right]\left[\bar{x}_{\ell\cdot} - \bar{x}\right]'$$

...5/-

$$SSP_{\text{faktor 2}} = \sum_{k=1}^{b} gn\left[\bar{\underset{\sim}{x}}_{.k} - \bar{\underset{\sim}{x}}\right]\left[\bar{\underset{\sim}{x}}_{.k} - \bar{\underset{\sim}{x}}\right]'$$

$$SSP_{\substack{\text{tindakan} \\ \text{bersaling}}} = \sum_{\ell=1}^{g}\sum_{k=1}^{b} n\left[\bar{\underset{\sim}{x}}_{\ell k} - \bar{\underset{\sim}{x}}_{\ell .} - \bar{\underset{\sim}{x}}_{.k} + \bar{\underset{\sim}{x}}\right]$$

$$\left[\bar{\underset{\sim}{x}}_{\ell k} - \bar{\underset{\sim}{x}}_{\ell .} - \bar{\underset{\sim}{x}}_{.k} + \bar{\underset{\sim}{x}}\right]'$$

$$SSP_{\text{residual}} = \sum_{\ell=1}^{g}\sum_{k=1}^{b}\sum_{r=1}^{n}\left[\underset{\sim}{x}_{\ell kr} - \bar{\underset{\sim}{x}}_{\ell k}\right]\left[\underset{\sim}{x}_{\ell kr} - \bar{\underset{\sim}{x}}_{\ell k}\right]'$$

11. Komponen Prinsipal

(a) $\underset{\sim}{Y}_i = \underset{\sim}{e}_i' \underset{\sim}{X}$, $i = 1, 2, \ldots, p.$

$$\rho_{Y_i, X_k} = \frac{e_{ki}\sqrt{\lambda_i}}{\sqrt{\sigma_{kk}}} \quad , \qquad i, k = 1, 2, \ldots, p.$$

(b) $\underset{\sim}{Y}_i = \underset{\sim}{e}_i' \underset{\sim}{Z}$

$$\rho_{Y_i, Z_k} = e_{ki}\sqrt{\lambda_i} \quad , \; i,k = 1, 2, \ldots, p.$$

12. Analisis Faktor

(a) $\underset{\sim}{X} - \underset{\sim}{\mu} = \underset{\sim}{L}\,\underset{\sim}{F} + \underset{\sim}{\epsilon}$

(b) $\underset{\sim}{\text{Kov}}(X) = \underset{\sim}{L}\,\underset{\sim}{L}' + \Psi$

$\underset{\sim}{\text{Kov}}(X, F) = \underset{\sim}{L}$

.. 6. -

(c) $h_i^2 = \ell_{i1}^2 + \ell_{i2}^2 + \ldots + \ell_{im}^2$, $i = 1, 2, \ldots, p.$

$\sigma_{ii} = h_i^2 + \psi_i$, $i = 1, 2, \ldots, p.$

(d) Kriterium varimax: Pilih transformasi ortogon T yang menjadikan

$$V = \frac{1}{p} \sum_{j=1}^{m} \left[\sum_{i=1}^{p} \tilde{\ell}_{ij}^{*4} - \left(\sum_{i=1}^{p} \tilde{\ell}_{ij}^{*2} \right)^2 / p \right]$$

sebesar yang mungkin.

13. Analisis Pembezalayan

(a) $Y = \underset{\sim}{\ell}' \underset{\sim}{X} = (\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2)' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{X}$

$m = \frac{1}{2} \left(\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2 \right)' \underset{\sim}{\Sigma}^{-1} \left(\underset{\sim}{\mu}_1 + \underset{\sim}{\mu}_2 \right)$

(b) $y = \hat{\underset{\sim}{\ell}}' \underset{\sim}{x} = \left(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2 \right)' \underset{\sim}{S}_p^{-1} \underset{\sim}{x}$

$\hat{m} = \frac{1}{2} \left(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2 \right)' \underset{\sim}{S}_p^{-1} \left(\bar{\underset{\sim}{x}}_1 + \bar{\underset{\sim}{x}}_2 \right)$

(c) Petua peruntukan:

Untukkan $\underset{\sim}{x}_o$ kepada $\begin{cases} \pi_1 \text{ jika } y_o \geq \hat{m} \\ \pi_2 \text{ jika } y_o < \hat{m} \end{cases}$

.../7/-

(d) $\underset{\sim}{B}_o = \sum_{i=1}^{g} \left(\underset{\sim}{\mu}_i - \bar{\underset{\sim}{\mu}} \right) \left(\underset{\sim}{\mu}_i - \bar{\underset{\sim}{\mu}} \right)'$

$\lambda_1, \ldots, \lambda_s$ nilai eigen dan

$\underset{\sim}{e}_1, \ldots, \underset{\sim}{e}_s$ vektor eigen $\underset{\sim}{\Sigma}^{-1} \underset{\sim}{B}_o$.

$\underset{\sim}{\ell}_i \underset{\sim}{X} = \underset{\sim}{e}_i \underset{\sim}{X}$ pembezalayan ke-i , $i = 1, 2, \ldots, s$.

(e) $\hat{\underset{\sim}{B}}_o = \sum_{i=1}^{g} \left(\bar{\underset{\sim}{x}}_i - \bar{\underset{\sim}{x}} \right) \left(\bar{\underset{\sim}{x}}_i - \bar{\underset{\sim}{x}} \right)'$

$\underset{\sim}{W} = \sum_{i=1}^{g} \sum_{j=1}^{n_i} \left(\underset{\sim}{x}_{ij} - \bar{\underset{\sim}{x}}_i \right) \left(\underset{\sim}{x}_{ij} - \bar{\underset{\sim}{x}}_i \right)'$

$\hat{\underset{\sim}{\ell}}_i \underset{\sim}{x} = \hat{\underset{\sim}{e}}_i \underset{\sim}{x}$ pembezalayan sampel ke-i, $i = 1, \ldots, s$.

(f) Petua peruntukan:

Untukkan $\underset{\sim}{x}$ kepada π_k jika

$$\sum_{j=1}^{r} \left(\hat{y}_j - \bar{y}_{kj} \right)^2 = \sum_{j=1}^{r} \left[\hat{\underset{\sim}{\ell}}_j' \left(\underset{\sim}{x} - \bar{\underset{\sim}{x}}_k \right) \right]^2$$

$$\leq \sum_{j=1}^{r} \left[\hat{\underset{\sim}{\ell}}_j' \left(\underset{\sim}{x} - \bar{\underset{\sim}{x}}_i \right) \right]^2$$

bagi semua $i \neq k$, $r \leq s$.

- oooOooo -