

UNIVERSITI SAINS MALAYSIA

Peperiksaan Semester Pertama  
Sidang 1991/92

Oktober/November 1991

MSG 466 Analisis Multivariat

Masa : [3 jam]

Jawab SEMUA soalan; semua soalan mesti dijawab dalam Bahasa Malaysia.

1. (a) Vektor-vektor rawak  $\tilde{\mathbf{X}}' = (X_1, X_2, X_3)$  dan  $\tilde{\mathbf{Y}}' = (Y_1, Y_2, Y_3)$  adalah tak bersandar dan mempunyai taburan normal trivariat dengan parameter-parameter

$$\mu_{\tilde{\mathbf{X}}_1} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} \quad \Sigma_{\tilde{\mathbf{X}}_1} = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 4 & 1 \\ 1 & 1 & 2 \end{pmatrix} \quad \mu_{\tilde{\mathbf{X}}_2} = \begin{pmatrix} 3 \\ 4 \\ 2 \end{pmatrix} \quad \Sigma_{\tilde{\mathbf{X}}_2} = \begin{pmatrix} 4 & 2 & 0 \\ 2 & 4 & 2 \\ 0 & 2 & 4 \end{pmatrix}$$

masing-masing.

Katakan  $\tilde{\mathbf{A}} = \begin{pmatrix} 2 & -1 & -1 \\ 0 & 1 & -1 \end{pmatrix}$   $\tilde{\mathbf{B}} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}$

(i) Carikan taburan-taburan bagi (a)  $\tilde{\mathbf{A}}\tilde{\mathbf{X}}$

(b)  $\tilde{\mathbf{B}}\tilde{\mathbf{X}}$ .

(ii) Carikan taburan tercantum bagi  $(\tilde{\mathbf{X}}'\tilde{\mathbf{A}}' \tilde{\mathbf{Y}}'\tilde{\mathbf{B}}')$ .

- (b) Diberikan matriks data

$$\tilde{\mathbf{X}} = \begin{pmatrix} 3 & -2 & 2 \\ 5 & -1 & 2 \\ 2 & 3 & 4 \end{pmatrix}$$

.../2

dan gabungan-gabungan linear

$$\tilde{\mathbf{b}}' \tilde{\mathbf{x}} = \begin{pmatrix} 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}$$

$$\text{dan } \tilde{\mathbf{c}}' \tilde{\mathbf{x}} = \begin{pmatrix} 1 & 2 & -1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix},$$

nilaikan min, varians dan kovarians sampel bagi  $\tilde{\mathbf{b}}' \tilde{\mathbf{x}}$  dan  $\tilde{\mathbf{c}}' \tilde{\mathbf{x}}$ .

- (c) Cari anggaran kebolehdjian maksimum bagi vektor min,  $\tilde{\mu}$  dan matriks kovarians,  $\tilde{\Sigma}$  berdasarkan sampel rawak

$$\tilde{\mathbf{x}} = \begin{pmatrix} 6 & 7 & 4 & 7 \\ 3 & 4 & 4 & 5 \end{pmatrix}$$

daripada suatu populasi normal bivariat.

- (d) (i) Dengan menggunakan data

$$\tilde{\mathbf{x}} = \begin{pmatrix} 12 & 9 & 9 & 10 \\ 1 & 7 & 5 & 7 \end{pmatrix}$$

nilaikan  $T^2$  Hotelling untuk menguji

$$H_0: \tilde{\mu}' = (6, 11).$$

- (ii) Nyatakan taburan  $T^2$  bagi keadaan dalam (i) atas. Berikan juga anggapan-anggapan yang telah anda menggunakan.
- (iii) Dengan menggunakan (i) dan (ii) di atas, ujikan  $H_0$  pada paras  $\alpha = .05$ . Apakah kesimpulan anda?

(100/100)

.../3

2. (a) Skor-skor GRE (Graduate Record Examination) "verbal" dan kuantitatif bagi 10 pelajar siswazah yang tidak berjaya menghabiskan pengajian mereka dalam matapelajaran Ekonomiks dan 13 pelajar siswazah yang berjaya menghabiskan pengajian mereka dalam mata-pelajaran Ekonomiks ditunjukkan di bawah:

Berjaya (Kumpulan I)		Tidak Berjaya (Kumpulan II)		
Verbal	Kuantitatif	Verbal	Kuantitatif	
750	590	740	680	
360	600	670	600	
720	750	560	530	
540	710	540	520	
570	700	590	540	
520	670	590	700	
590	790	470	600	
670	700	560	540	
620	730	540	630	
690	840	500	600	
610	680			
550	730			
590	750			
Min	598.46	710.77	576.00	596.00

$$(n_1 - 1)S_1^2 = \begin{pmatrix} 56,240 & 17,240 \\ 17,240 & 33,240 \end{pmatrix}$$

$$(n_2 - 1)S_2^2 = \begin{pmatrix} 121,569 & 25,615 \\ 25,615 & 56,492 \end{pmatrix}$$

Andaikan kita ingin menguji  $H_0 : \mu_1 = \mu_2$  berlawanan  $H_1 : \mu_1 \neq \mu_2$  pada  $\alpha = .01$ .

Apakah kesimpulan anda? Berikan juga anggapan-anggapan yang telah anda menggunakan.

- (b) Cerapan-cerapan pada dua balasan,  $X_1$  dan  $X_2$ , dikutip bagi tiga rawatan. Vektor-vektor cerapan  $\tilde{x}' = (x_1 \ x_2)$  adalah:

.../4

Rawatan 1 : (3 7) (4 7) (5 8) (5 9) (6 10)  
 Rawatan 2 : (4 5) (4 6) (5 7) (6 7) (6 8)  
 Rawatan 3 : (5 5) (6 5) (6 6) (7 7) (7 8)

- (i) Binakan suatu jadual MANOVA satu hala bagi data ini.
- (ii) Nilaikan lambda Wilks,  $\Lambda^*$ , dan uji bagi kesan rawatan. Gunakan  $\alpha = .01$ .
- (iii) Ulangkan ujian dengan menggunakan penghampiran khi-kuasa dua dengan pembedahan Bartlett.

Perhatian: Ringkasan statistik diberikan di bawah:

	Rawatan 1	Rawatan 2	Rawatan 3
$\Sigma x_{1j}$	23	25	31
$\Sigma x_{2j}$	41	33	31
$\Sigma x_{1j}^2$	111	129	195
$\Sigma x_{2j}^2$	343	223	199
$\Sigma x_{1j} x_{2j}$	194	169	196

(100/100)

3. Bagi setiap bahagian yang berikut, tuliskan suatu perenggan yang menghuraikan kesimpulan-kesimpulan anda.

- (a) Tiga jenis dadah, A, B, dan C diketahui mempunyai kesan berbisa yang kelihatan dalam beberapa cirian yang dapat disukatkan daripada suatu organisma yang menerima dadah tersebut.

.../5

Untuk kajian kita, andaikan kehilangan berat dalam minggu-minggu pertama dan kedua adalah cirian disukatkan. Dirasa juga kesan dadah-dadah itu mungkin berbeza bagi binatang jantan dan binatang betina. Jadi, jantina akan digunakan sebagai faktor kedua di dalam eksperimen kita.

Dua belas ekor tikus jantan dengan umur dan berat yang pada hakikatnya sama dipilih daripada sejenis turunan tikus makmal biasa dan diumpukkan secara rawak di antara tiga dadah itu dengan setiap dadah menerima empat ekor tikus.

Dua belas ekor tikus betina dengan umur dan berat yang dapat dibandingkan diperolehi daripada seperinduk tikus daripada turunan yang sama dan diumpukkan dengan cara yang sama.

Data yang dikutip ditunjukkan di bawah:

	Dadah		
	A	B	C
Jantan	(5,6)	(7,6)	(21,15)
	(5,4)	(7,7)	(14,11)
	(9,9)	(9,12)	(17,12)
	(7,6)	(6,8)	(12,10)
Betina	(7,10)	(10,13)	(16,12)
	(6,6)	(8,7)	(14,9)
	(9,7)	(7,6)	(14,8)
	(8,10)	(6,9)	(10,5)

Hasil yang diperolehi daripada SPSS melalui prosedur MANOVA adalah seperti diberikan di bawah:

\* \* ANALYSIS OF VARIANCE -- DESIGN 1 \* \*

EFFECT .. SEX BY DRUG

Multivariate Tests of Significance (S = 2, M = -1/2, N = 7 1/2)

Test Name	Value	Approx. F	Hypoth. DF	Error DF	Sig. of F
Pillais	.22695	1.15199	4.00	36.00	.348
Hotellings	.28969	1.15877	4.00	32.00	.347
Wilks	.77436	1.15933	4.00	34.00	.346
Roys	.22102				

.../6

Univariate F-tests with (2,18) D. F.

Variable	Hypoth. SS	Error SS	Hypoth. MS	Error MS	F	Sig. of F
X1	14.33333	94.50000	7.16667	5.25000	1.36508	.281
X2	32.33333	114.00000	16.16667	6.33333	2.55263	.106

\* \* ANALYSIS OF VARIANCE -- DESIGN 1 \* \*

EFFECT .. DRUG

Multivariate Tests of Significance (S = 2, M = -1/2, N = 7 1/2)

Test Name	Value	Approx. F	Hypoth. DF	Error DF	Sig. of F
Pillais	.88038	7.07686	4.00	36.00	.000
Hotellings	4.63954	18.55815	4.00	32.00	.000
Wilks	.16863	12.19913	4.00	34.00	.000
Roys	.82066				

Univariate F-tests with (2,18) D. F.

Variable	Hypoth. SS	Error SS	Hypoth. MS	Error MS	F	Sig. of F
X1	301.00000	94.50000	150.50000	5.25000	28.66667	.000
X2	36.33333	114.00000	18.16667	6.33333	2.86842	.083

\* \* ANALYSIS OF VARIANCE -- DESIGN 1 \* \*

EFFECT .. SEX

Multivariate Tests of Significance (S = 1, M = 0, N = 7 1/2)

Test Name	Value	Approx. F	Hypoth. DF	Error DF	Sig. of F
Pillais	.00746	.06391	2.00	17.00	.938
Hotellings	.00752	.06391	2.00	17.00	.938
Wilks	.99254	.06391	2.00	17.00	.938
Roys	.00746				

Univariate F-tests with (1,18) D. F.

Variable	Hypoth. SS	Error SS	Hypoth. MS	Error MS	F	Sig. of F
X1	.66667	94.50000	.66667	5.25000	.12698	.726
X2	.66667	114.00000	.66667	6.33333	.10526	.749

.../7

\* \* ANALYSIS OF VARIANCE -- DESIGN 1 \* \*

EFFECT .. CONSTANT

Multivariate Tests of Significance (S = 1, M = 0, N = 7 1/2)

Test Name	Value	Approx. F	Hypoth. DF	Error DF	Sig. of F
Pillai's	.96066	207.56013	2.00	17.00	.000
Hotellings	24.41884	207.56013	2.00	17.00	.000
Wilks	.03934	207.56013	2.00	17.00	.000
Roys	.96066				

Univariate F-tests with (1,18) D. F.

Variable	Hypoth. SS	Error SS	Hypoth. MS	Error MS	F	Sig. of F
X1	2281.50000	94.50000	2281.50000	5.25000	434.57143	.000
X2	1802.66667	114.00000	1802.66667	6.33333	284.63158	.000

(b) Fisher (1936) mengkaji data daripada Anderson yang berkaitan dengan sampel-sampel rawak bunga iris daripada 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 dianalisis dengan menggunakan prosedur DISCRIMINANT daripada pakej SPSS.

Output diperolehi ditunjukkan di bawah:

- - - - - 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   jenis bunga iris

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

SPECIES	Number of Cases		Label
	Unweighted	Weighted	
1	50	50.0	bunga iris 1
2	50	50.0	bunga iris 2
3	50	50.0	bunga iris 3
Total	150	150.0	

Group Means

SPECIES	LSEPAL	WSEPAL	LPETAL	WPETAL
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	LSEPAL	WSEPAL	LPETAL	WPETAL
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

	LSEPAL	WSEPAL	LPETAL	WPETAL
LSEPAL	.2650082			
WSEPAL	.9272109E-01	.1153878		
LPETAL	.1675143	.5524354E-01	.1851878	
WPETAL	.3840136E-01	.3271020E-01	.4266531E-01	.4188163E-01

Pooled Within-Groups Correlation Matrix

	LSEPAL	WSEPAL	LPETAL	WPETAL
LSEPAL	1.00000			
WSEPAL	.53024	1.00000		
LPETAL	.75616	.37792	1.00000	
WPETAL	.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
LSEPAL	.38129	119.3	.0000
WSEPAL	.59922	49.16	.0000
LPETAL	.05863	1180.	.0000
WPETAL	.07112	960.0	.0000

.../9

Covariance Matrix for Group 1, bunga iris 1

	LSEPAL	WSEPAL	LPETAL	WPETAL
LSEPAL	.1242490			
WSEPAL	.9921633E-01	.1436898		
LPETAL	.1635510E-01	.1169796E-01	.3015918E-01	
WPETAL	.1033061E-01	.9297959E-02	.6069388E-02	.1110612E-01

Covariance Matrix for Group 2, bunga iris 2

	LSEPAL	WSEPAL	LPETAL	WPETAL
LSEPAL	.2664327			
WSEPAL	.8518367E-01	.9846939E-01		
LPETAL	.1828980	.8265306E-01	.2208163	
WPETAL	.5577959E-01	.4120408E-01	.7310204E-01	.3910612E-01

Covariance Matrix for Group 3, bunga iris 3

	LSEPAL	WSEPAL	LPETAL	WPETAL
LSEPAL	.4043429			
WSEPAL	.9376327E-01	.1040041		
LPETAL	.3032898	.7137959E-01	.3045878	
WPETAL	.4909388E-01	.4762857E-01	.4882449E-01	.7543265E-01

Total Covariance Matrix with 149 degrees of freedom

	LSEPAL	WSEPAL	LPETAL	WPETAL
LSEPAL	.6856935			
WSEPAL	-.4243400E-01	.1899794		
LPETAL	1.274315	-.3296564	3.116278	
WPETAL	.5162707	-.1216394	1.295609	.5810063

DISCRIMINANT ANALYSIS

On groups defined by SPECIES jenis bunga iris

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

Classification Function Coefficients  
(Fisher's Linear Discriminant Functions)

SPECIES =	1	2	3
	bunga iris 1	bunga iris 2	bunga iris 3
LSEPAL	23.54417	15.69821	12.44585
WSEPAL	23.58787	7.072510	3.685280
LPETAL	-16.43064	5.211451	12.76654
WPETAL	-17.39841	6.434229	21.07911
(constant)	-86.30847	-72.85261	104.3683

(c) Pertimbangkan bahagian (b). Komponen-komponen prinsipal bagi matriks kovarians pun telah diperolehi untuk setiap jenis iris melalui prosedur PRINCOMP daripada pakej SAS.

Output yang diperolehi adalah seperti yang diberikan dibawah:

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

Principal Component Analysis

50 Observations  
4 Variables

Simple Statistics

	LSEPAL	WSEPAL	LPETAL	WPETAL
Mean	5.006000000	3.428000000	1.462000000	0.246000000
Std	0.352489687	0.379064369	0.173663996	0.1053855894

Covariance Matrix

	LSEPAL	WSEPAL	LPETAL	WPETAL
LSEPAL	0.1242489796	0.0992163265	0.0163551020	0.0103306122
WSEPAL	0.0992163265	0.1436897959	0.0116979592	0.0092979592
LPETAL	0.0163551020	0.0116979592	0.0301591837	0.0060693878
WPETAL	0.0103306122	0.0092979592	0.0060693878	0.0111061224

Total Variance = 0.3092040816

.../11

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
LSEPAL	0.669078	0.597884	-.439963	-.036077
WSEPAL	0.734148	-.620673	0.274607	-.019550
LPETAL	0.096544	0.490056	0.832449	-.239901
WPETAL	0.063564	0.130938	0.195068	0.969930

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

Principal Component Analysis

50 Observations  
4 Variables

Simple Statistics

	LSEPAL	WSEPAL	LPETAL	WPETAL
Mean	5.936000000	2.770000000	4.260000000	1.326000000
Std	0.516171147	0.313798323	0.469910977	0.197752680

Covariance Matrix

	LSEPAL	WSEPAL	LPETAL	WPETAL
LSEPAL	0.2664326531	0.0851836735	0.1828979592	0.0557795918
WSEPAL	0.0851836735	0.0984693878	0.0826530612	0.0412040816
LPETAL	0.1828979592	0.0826530612	0.2208163265	0.0731020408
WPETAL	0.0557795918	0.0412040816	0.0731020408	0.0391061224

.../12

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
LSEPAL	0.686724	-.669089	0.265083	0.102280
WSEPAL	0.305347	0.567465	0.729618	-.228919
LPETAL	0.623663	0.343327	-.627165	-.315967
WPETAL	0.214984	0.335305	-.063661	0.915041

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

Principal Component Analysis

50 Observations  
4 Variables

Simple Statistics

	LSEPAL	WSEPAL	LPETAL	WPETAL
Mean	6.588000000	2.974000000	5.552000000	2.026000000
Std	0.635879593	0.322496638	0.551894696	0.274650056

Covariance Matrix

	LSEPAL	WSEPAL	LPETAL	WPETAL
LSEPAL	0.4043428571	0.0937632653	0.3032897959	0.0490938776
WSEPAL	0.0937632653	0.1040040816	0.0713795918	0.0476285714
LPETAL	0.3032897959	0.0713795918	0.3045877551	0.0488244898
WPETAL	0.0490938776	0.0476285714	0.0488244898	0.0754326531

Total Variance = 0.8883673469

.../13

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
LSEPAL	0.741017	-.165259	-.534450	0.371412
WSEPAL	0.203288	0.748643	-.325375	-.540684
LPETAL	0.627892	-.169428	0.651524	-.390593
WPETAL	0.123775	0.619288	0.428965	0.645872

(d) Sinha dan Lee (1970) memperolehi matriks korelasi berikut yang terdiri daripada sifat-sifat dan "arthropod infestation counts" (dijelmakan melalui fungsi logarithma) dari 165 sampel biji-bijian:

Pembolehubah	1	2	3	4	5	6	7	8	9
1. Grade	1	.44	.44	.11	.19	.10	.20	.20	-.24
2. Moisture		1	.34	.25	.32	.40	.49	.16	-.22
3. Dockage			1	.04	.06	.08	.07	.05	-.07
4. Acarus				1	.18	.12	.23	.02	-.20
5. Cheyletus					1	.22	.48	.14	-.08
6. Glycyphagus						1	.40	-.11	-.30
7. Tarsonemus							1	.15	-.13
8. Cryptolestes								1	-.10
9. Psocoptera									1

Pembolehubah "dockage" terdiri daripada berat bahan asing (benih rumput rampai, dll) di dalam sampel.

Analisis faktor telah dijalankan dengan pakej SPSS dengan kaedah kebolehjadian maksimum dan kaedah komponen prinsipal melalui prosedur FACTOR.

Output diperolehi ditunjukkan di bawah:

--- FACTOR ANALYSIS ---

Analysis Number 1 Matrix input  
Extraction 1 for Analysis 1, Maximum Likelihood (ML)

Initial Statistics:

Variable	Communality	*	Factor	Eigenvalue	Pct of Var	Cum Pct
GRADE	.34005	*	1	2.74414	30.5	30.5
MOISTURE	.45430	*	2	1.32169	14.7	45.2
DOCKAGE	.23778	*	3	1.10691	12.3	57.5
ACARUS	.10975	*	4	.98969	11.0	68.5
CHEY	.25194	*	5	.87253	9.7	78.2
GLYCY	.31562	*	6	.63946	7.1	85.3
TARSON	.41068	*	7	.51183	5.7	91.0
CRYPTO	.11361	*	8	.42129	4.7	95.6
PSOCOP	.17645	*	9	.39246	4.4	100.0

ML Extracted 3 factors. 7 Iterations required.

Factor Matrix:

	FACTOR 1	FACTOR 2	FACTOR 3
GRADE	.10095	.60078	.47483
MOISTURE	.40125	.61144	.10261
DOCKAGE	.08054	.36954	.43386
ACARUS	-.12064	.26414	-.07686
CHEY	.22101	.45665	-.26130
GLYCY	.99950	-.00179	.00038
TARSON	.40145	.60551	-.43254
CRYPTO	-.10945	.32539	-.00077
PSOCOP	-.30034	-.16638	-.17374

Final Statistics:

Variable	Communality	*	Factor	Eigenvalue	Pct of Var	Cum Pct
GRADE	.59660	*	1	1.50342	16.7	16.7
MOISTURE	.54539	*	2	1.64986	18.3	35.0
DOCKAGE	.33129	*	3	.71570	8.0	43.0
ACARUS	.09023	*				
CHEY	.32565	*				
GLYCY	.99900	*				
TARSON	.71489	*				
CRYPTO	.11786	*				
PSOCOP	.14807	*				

Varimax Rotation 1, Extraction 1, Analysis 1 - Kaiser Normalization.

Varimax converged in 6 iterations.

Rotated Factor Matrix:

	FACTOR 1	FACTOR 2	FACTOR 3
GRADE	.19845	.74620	.01992
MOISTURE	.50709	.47007	.25938
DOCKAGE	.04219	.57199	.04824
ACARUS	.27771	.10500	.04566
CHEY	.56005	.08036	.07447
GLYCY	.25888	.03942	.96459
TARSON	.82226	.04276	.19225
CRYPTO	.21876	.19333	-.18063
PSOCOP	-.10071	-.25105	-.27368

Extraction 2 for Analysis 1, Principal-Components Analysis (PC)

Initial Statistics:

Variable	Communality	*	Factor	Eigenvalue	Pct of Var	Cum Pct
GRADE	1.00000	*	1	2.74414	30.5	30.5
MOISTURE	1.00000	*	2	1.32169	14.7	45.2
DOCKAGE	1.00000	*	3	1.10691	12.3	57.5
ACARUS	1.00000	*	4	.98969	11.0	68.5
CHEY	1.00000	*	5	.87253	9.7	78.2
GLYCY	1.00000	*	6	.63946	7.1	85.3
TARSON	1.00000	*	7	.51183	5.7	91.0
CRYPTO	1.00000	*	8	.42129	4.7	95.6
PSOCOP	1.00000	*	9	.39246	4.4	100.0

PC Extracted 3 factors.

Factor Matrix:

	FACTOR 1	FACTOR 2	FACTOR 3
GRADE	.60107	.57705	-.06101
MOISTURE	.80212	.07652	-.04473
DOCKAGE	.41610	.64975	-.25175
ACARUS	.40674	-.23264	-.05207
CHEY	.57451	-.28737	.42982
GLYCY	.55948	-.42767	-.41454
TARSON	.70587	-.36332	.25821
CRYPTO	.25589	.32980	.67756
PSOCOP	-.44054	.01707	.39083

Final Statistics:

Variable	Communality	*	Factor	Eigenvalue	Pct of Var	Cum Pct
GRADE	.69799	*	1	2.74414	30.5	30.5
MOISTURE	.65124	*	2	1.32169	14.7	45.2
DOCKAGE	.65869	*	3	1.10691	12.3	57.5
ACARUS	.22227	*				
CHEY	.59739	*				
GLYCY	.66776	*				
TARSON	.69693	*				
CRYPTO	.63333	*				
PSOCOP	.34712	*				

Varimax Rotation 1, Extraction 2, Analysis 1 - Kaiser Normalization.

Varimax converged in 8 iterations.

Rotated Factor Matrix:

	FACTOR 1	FACTOR 2	FACTOR 3
GRADE	.17366	.81581	.04784
MOISTURE	.56502	.53996	-.20109
DOCKAGE	-.08651	.80587	-.04214
ACARUS	.39768	.06866	-.24374
CHEY	.76459	.01858	.11155
GLYCY	.44519	.08580	-.67986
TARSON	.82522	.07410	-.10227
CRYPTO	.34537	.25842	.66879
PSOCOP	-.17600	-.33092	.45457

(100/100)

4. (a) Tuliskan nota pendek tentang tajuk-tajuk di bawah:
- (i) Kaedah kelompok "hierarchical"
  - (ii) Kaedah k-min MacQueen.

.../17

(b) Diberikan lima objek mempunyai matriks jarak berikut:

$$\tilde{D} = \{d_{ik}\} = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{pmatrix} 0 & & & & \\ 3 & 0 & & & \\ 11 & 2 & 0 & & \\ 6 & 9 & 8 & 0 & \\ 9 & 7 & 10 & 5 & 0 \end{pmatrix} \end{matrix},$$

gunakan (i) kaedah pautan tunggal,

(ii) kaedah pautan lengkap

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

(100/100)

-ooo00ooo -

LAMPIRAN

Tatatanda adalah seperti di dalam kuliah.

1. Penguraian spektrum bagi suatu matriks simetrik  $k \times k$ ,  $A$  diberikan oleh

$$A = \lambda_1 e_1 e_1' + \lambda_2 e_2 e_2' + \dots + \lambda_k e_k e_k'$$

di mana  $\lambda_1, \lambda_2, \dots, \lambda_k$  adalah nilai-nilai eigen  $A$  dan  $e_1, e_2, \dots, e_k$  adalah vektor-vektor eigen terpiawai yang berkaitan.

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

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:

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

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

...2/-

6. Satu sampel:

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

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

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

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

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

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

(d) Selang keyakinan serentak Bonferroni 100(1- $\alpha$ )% bagi

$$\mu_i, \quad i = 1, \dots, p :$$

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

7. Dua sampel tak bersandar:

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

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

...3/-

(b) Selang keyakinan serentak  $100(1-\alpha)\%$  bagi  $\mu_1 - \mu_2$ :

$$\bar{X}_1 - \bar{X}_2 \pm c \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} S_p$$

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:

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

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

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

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

$$\bar{X}_{ki} - \bar{X}_{\ell i} \pm t_{n-g} \left( \frac{\alpha}{pg(g-1)} \right)^{1/2} \sqrt{\frac{W_{ii}}{n-g} \left( \frac{1}{n_k} + \frac{1}{n_{\ell}} \right)}$$

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

9. Andaikan E mempunyai d.k.  $m_E$  dan H mempunyai d.k.  $m_H$ .

Katakan  $\Lambda = \frac{|E|}{|E + H|}$

Maka (1) Untuk  $p = 1$ ,

$$\left( \frac{1 - \Lambda}{\Lambda} \right) \frac{m_E}{m_H} \sim F_{m_H, m_E} \text{ bagi sebarang } m_H$$

(2) Untuk  $m_H = 1$ ,

$$\left( \frac{1 - \Lambda}{\Lambda} \right) \frac{m_E + 1 - p}{p} \sim F_{p, m_E + 1 - p} \text{ bagi sebarang } p.$$

(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,

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

$$\begin{aligned} \text{di mana } f &= m_E - \frac{1}{2} \left[ p - m_H + 1 \right] \\ &= n_o - \frac{1}{2} \left[ p + m_H + 1 \right] \end{aligned}$$

10. MANOVA dua-hala:

$$SSP_{\text{faktor 1}} = \sum_{\ell=1}^g bn \left[ \begin{array}{c} \bar{x}_{\ell} \\ \sim \\ \bar{x} \end{array} \right] \left[ \begin{array}{c} \bar{x}_{\ell} \\ \sim \\ \bar{x} \end{array} \right]'$$

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

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

$$\left[ \bar{x}_{lk} - \bar{x}_{.l} - \bar{x}_{.k} + \bar{x} \right]'$$

$$SSP_{\text{residual}} = \sum_{l=1}^g \sum_{k=1}^b \sum_{r=1}^n \left[ x_{lkr} - \bar{x}_{lk} \right] \left[ x_{lkr} - \bar{x}_{lk} \right]'$$

11. Komponen Prinsipal

(a)  $\tilde{Y}_i = e_i' X$  ,  $i = 1, 2, \dots, p.$

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

(b)  $\tilde{Y}_i = e_i' Z$

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

12. Analisis Faktor

(a)  $\tilde{X} - \tilde{\mu} = L \tilde{F} + \tilde{\epsilon}$

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

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

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

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

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

sebesar yang mungkin.

13. Analisis Pembezaian

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

$m = \frac{1}{2} (\underline{\mu}_1 - \underline{\mu}_2)' \Sigma^{-1} (\underline{\mu}_1 + \underline{\mu}_2)$

(b)  $y = \hat{\ell}'x = (\bar{x}_1 - \bar{x}_2)' S_p^{-1} x$

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

(c) Petua peruntukan:

Untuk  $x_o$  kepada  $\begin{cases} \pi_1 & \text{jika } y_o \geq \hat{m} \\ \pi_2 & \text{jika } y_o < \hat{m} \end{cases}$

...7/-

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

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

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

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

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

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

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

(f) Petua peruntukan:

Untukkan  $x$  kepada  $\pi_k$  jika

$$\begin{aligned} \sum_{j=1}^r \left[ \hat{y}_j - \bar{y}_{kj} \right]^2 &= \sum_{j=1}^r \left[ \hat{\underset{\sim}{\ell}}_j' \left[ x - \bar{x}_k \right] \right]^2 \\ &\leq \sum_{j=1}^r \left[ \hat{\underset{\sim}{\ell}}_j' \left[ x - \bar{x}_i \right] \right]^2 \end{aligned}$$

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