UNIVERSITI SAINS MALAYSIA

Peperiksaan Semester Pertama
Sidang Akademik 2005/2006

November 2005

MSG 366 – Analisis Multivariat

Masa : 3 jam

Sila pastikan bahawa kertas peperiksaan ini mengandungi **DUA PULUH DUA** muka surat yang bercetak sebelum anda memulakan peperiksaan ini.

Jawab **semua empat** soalan.

...2/-

1. (a) Katakan $X' = (X_1, X_2, X_3)$ tertabur $N_3\ (\mu, \Sigma)$ di mana

$$\mu' = [3,1,4] \quad \text{dan} \quad \Sigma = \begin{bmatrix} 6 & 1 & -2 \\ 1 & 13 & 4 \\ -2 & 4 & 4 \end{bmatrix}.$$

(i) Cari taburan bagi $2X_1 - X_2 + 3X_3$.

(ii) Cari taburan tercantum bagi $X_1 + X_2 + X_3$ dan $X_1 - X_2 + 2X_3$.

(b) Biarkan $\mathbf{X}$ tertabur $N_3\ (\mu, \Sigma)$ dengan $\mu' = [2, -3, 4]$ dan

$$\Sigma = \begin{bmatrix} 4 & -3 & 0 \\ -3 & 6 & 0 \\ 0 & 0 & 5 \end{bmatrix}.$$

Bagi setiap pembolehubah rawak berikut, tentukan sama ada ianya bersandar atau tak bersandar? Huraikan.

(i) X_1 dan X_2

(ii) X_1 dan X_3

(iii) X_2 dan X_3

(iv) (X_1, X_2) dan X_3

(v) X_3 dan $X_1 + 2X_2 - 3X_3$

(c) Merujuk kepada bahagian (b), dapatkan taburan berikut:

(i) Taburan bersyarat X_1 diberikan $X_2 = x_2$

(ii) Taburan bersyarat X_1 diberikan $X_2 = x_2$ dan $X_3 = x_3$

(d) Biarkan matriks

$$\mathbf{A} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}.$$

(i) Adakah $\mathbf{A}$ tentu positif?

(ii) Dapatkan pasangan nilai eigen – vektor eigen $\mathbf{A}$.

(iii) Cari matriks punca ganda dua $\mathbf{A}^{1/2}$.

(iv) Periksa bahagian (iii) dengan menunjukkan bahawa $(\mathbf{A}^{1/2})^2 = \mathbf{A}$.

[100 markah]

...3/-

2. (a) (i) Dengan menggunakan data

$$\mathbf{X} = \begin{bmatrix} 3 & 10 \\ 6 & 12 \\ 5 & 14 \\ 10 & 9 \end{bmatrix}$$

nilaikan T^2 Hotelling untuk menguji $H_0: \boldsymbol{\mu}' = (6, 11)$.

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

(iii) Dengan menggunakan (i) dan (ii) di atas, uji H_0 pada paras α=0.01. Apakah kesimpulan anda?

(b) Katakan kita ingin tahu keberkesanan latihan ke atas kepuasan dan prestasi kerja. Pertimbangkan latihan sebagai pembolehubah tak bersandar dan kepuasan dan prestasi kerja sebagai pembolehubah-pembolehubah bersandar. Latihan dibahagi kepada tiga kumpulan : 1 = kawalan (*control*), 2 = latihan secara berhadapan (*face-to-face training*), 3 = latihan online (*online training*). Bilangan cerapan ialah 18. Keputusan-keputusan analisis data dipaparkan dalam Jadual 2(b).

Jadual 2(b). Keputusan-keputusan analisis data dari perisian SAS.

The GLM Procedure

Dependent Variable: kepuasan

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	520.777778	260.388889	1.48	0.2588
Error	15	2637.000000	175.800000		
Corrected Total	17	3157.777778			

R-Square	Coeff Var	Root MSE	kepuasan Mean
0.164919	14.87913	13.25896	89.11111

...4/-

Dependent Variable: prestasi

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	1126.777778	563.388889	4.05	0.0392
Error	15	2086.333333	139.088889		
Corrected Total	17	3213.111111			

R-Square	Coeff Var	Root MSE	prestasi Mean
0.350681	13.52132	11.79360	87.22222

Multivariate Analysis of Variance

E = Error SSCP Matrix

	kepuasan	prestasi
kepuasan	2637	1803.8333333
prestasi	1803.8333333	2086.3333333

Partial Correlation Coefficients from the Error SSCP Matrix / Prob > |r|

DF = 15	kepuasan	prestasi
kepuasan	1.000000	0.769041 0.0005
prestasi	0.769041 0.0005	1.000000

H = Type III SSCP Matrix for latihan

	kepuasan	prestasi
kepuasan	520.77777778	761.72222222
prestasi	761.72222222	1126.7777778

Characteristic Roots and Vectors of: E Inverse * H, where
H = Type III SSCP Matrix for latihan
E = Error SSCP Matrix

Characteristic Root	Percent	Characteristic Vector V'EV=1 kepuasan	prestasi
0.57761841	99.13	-0.00780126	0.02790813
0.00506785	0.87	0.02944975	-0.01985596

...5/-

```
MANOVA Test Criteria and F Approximations for the Hypothesis of No Overall
latihan Effect

                  H = Type III SSCP Matrix for latihan
                           E = Error SSCP Matrix

                         S=2    M=-0.5    N=6
```

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.63067070	1.81	4	28	0.1541
Pillai's Trace	0.37117546	1.71	4	30	0.1740
Hotelling-Lawley Trace	0.58268626	2.00	4	15.818	0.1435
Roy's Greatest Root	0.57761841	4.33	2	15	0.0327

```
NOTE: F Statistic for Roy's Greatest Root is an upper bound.
NOTE: F Statistic for Wilks' Lambda is exact.
```

Huraikan model yang telah digunakan dalam analisis data. Tafsirkan keputusan-keputusan yang diperoleh. Apakah kesimpulan anda?

[100 markah]

3. (a) Tuliskan nota pendek tentang tajuk-tajuk di bawah:

   (i) Analisis sukatan berulang
   (ii) Kontras multivariat
   (iii) Analisis komponen prinsipal
   (iv) Analisis kelompok

(b) Pada tahun 1997, Penyelidikan Pemasaran Cina Selatan Sdn. Bhd. melaksanakan sebuah soalselidik tentang sikap penduduk terhadap pekerjaan dan corak kehidupan dalam tiga buah kota yang major di negeri Cina – Beijing, Shanghai dan GuangZhou. 500 sambutan dari setiap kota diperoleh. Calon-calon diminta menyusun tiga criteria yang paling penting dalam pemilihan pekerjaan daripada 13 kriteria :
1. reputasi syarikat yang baik; 2. syarikat yang besar; 3. peluang kenaikan pangkat yang lebih; 4. peluang latihan yang lebih; 5. persekitaran kerja yang selesa; 6. pendapatan yang tinggi; 7. jam kerja yang stabil; 8. faedah (*fringe benefits*); 9. bakat pekerja sesuai dengan jawatannya; 10. jarak pendek antara tempat kerja dan rumah; 11. mencabar; 12. struktur korporat syarikat; 13. tekanan kerja yang rendah. Analisis faktor dikendalikan dan model tiga faktor dipilih. Pemutaran varimax digunakan untuk memperoleh muatan faktor seperti yang ditunjukkan dalam Jadual 3(b) (*Yu, Lam & Ho, 2005*).

...6/-

Jadual 3(b). Anggaran-anggaran parameter bagi model tiga faktor.

Kriteria	Faktor 1	Faktor 2	Faktor 3	b	σ^2
1	-0.58	-0.41	0.25	-0.79	0.73
2	-0.02	-0.21	0.79	-0.46	0.29
3	0.29	-0.65	-0.04	0.30	0.40
4	0.44	-0.38	0.24	-0.74	1.10
5	-0.10	0.09	-0.14	0.75	0.47
6	-0.05	0.07	-0.73	1.31	0.35
7	-0.09	0.69	-0.17	0.42	0.27
8	-0.22	0.36	0.34	0.94	0.31
9	0.58	0.03	0.21	-0.26	0.53
10	0.14	0.70	-0.15	-0.05	0.19
11	0.39	-0.07	0.09	-0.46	0.88
12	-0.50	0.21	0.08	-1.10	0.82
13	-0.07	0.52	-0.23	-0.17	0.47
Kadaran longgok yang diterangkan	0.15	0.30	0.41		

Huraikan model faktor yang digunakan dalam analisis. Kemudian tafsirkan keputusan-keputusan dalam Jadual 3(b). Apakah kesimpulan anda?

[100 markah]

4. Bagi setiap bahagian yang berikut, bincangkan hasil analisis dan berikan kesimpulan anda.

(a) Elston dan Grizzle (1962) mengkaji panjang tulang ramus pada empat umur (8 tahun, 8.5 tahun, 9 tahun dan 9.5 tahun) bagi 20 orang budak lelaki. Data yang diperoleh dianalisiskan dengan menggunakan teknik multivariat yang sesuai. Hasil yang dijana oleh pakej SPSS dipamerkan seperti berikut:

Descriptive Statistics

	Mean	Std. Deviation	Analysis N
8 yr	48.655	2.5159	20
8.5 yr	49.625	2.5396	20
9 yr	50.570	2.6302	20
9.5 yr	51.450	2.7322	20

...7/-

Correlation Matrix

		8 yr	8.5 yr	9 yr	9.5 yr
Correlation	8 yr	1.000	.969	.873	.807
	8.5 yr	.969	1.000	.921	.854
	9 yr	.873	.921	1.000	.967
	9.5 yr	.807	.854	.967	1.000
Sig. (1-tailed)	8 yr		.000	.000	.000
	8.5 yr	.000		.000	.000
	9 yr	.000	.000		.000
	9.5 yr	.000	.000	.000	

Covariance Matrix[a]

a. Determinant = 1.068

Inverse of Covariance Matrix

	8 yr	8.5 yr	9 yr	9.5 yr
8 yr	2.675	-2.916	.501	-.140
8.5 yr	-2.916	4.365	-2.221	.771
9 yr	.501	-2.221	4.662	-2.948
9.5 yr	-.140	.771	-2.948	2.370

KMO and Bartlett's Test[a]

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.726
Bartlett's Test of Sphericity	Approx. Chi-Square	127.723
	df	6
	Sig.	.000

a. Based on correlations

Anti-image Matrices

		8 yr	8.5 yr	9 yr	9.5 yr
Anti-image Covariance	8 yr	.059	-.039	.006	-.003
	8.5 yr	-.039	.036	-.016	.011
	9 yr	.006	-.016	.031	-.037
	9.5 yr	-.003	.011	-.037	.057
Anti-image Correlation	8 yr	.758[a]	-.853	.142	-.056
	8.5 yr	-.853	.710[a]	-.492	.240
	9 yr	.142	-.492	.708[a]	-.887
	9.5 yr	-.056	.240	-.887	.732[a]

a. Measures of Sampling Adequacy(MSA)

...8/-

Communalities

	Raw		Rescaled	
	Initial	Extraction	Initial	Extraction
8 yr	6.330	6.251	1.000	.988
8.5 yr	6.449	6.358	1.000	.986
9 yr	6.918	6.819	1.000	.986
9.5 yr	7.465	7.405	1.000	.992

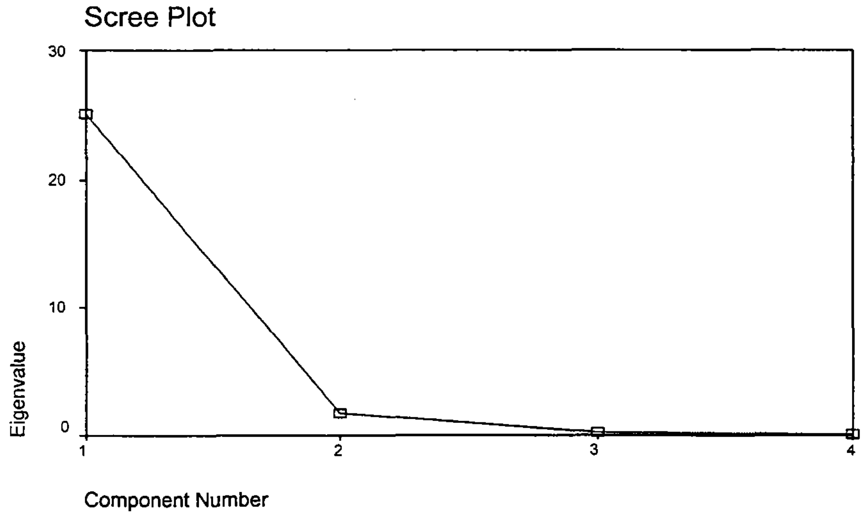
Extraction Method: Principal Component Analysis.

Total Variance Explained

	Component	Initial Eigenvalues[a]			Extraction Sums of Squared Loadings		
		Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
Raw	1	25.089	92.369	92.369	25.089	92.369	92.369
	2	1.744	6.422	98.790	1.744	6.422	98.790
	3	.215	.792	99.582			
	4	.114	.418	100.000			
Rescaled	1	25.089	92.369	92.369	3.695	92.374	92.374
	2	1.744	6.422	98.790	.256	6.404	98.778
	3	.215	.792	99.582			
	4	.114	.418	100.000			

Extraction Method: Principal Component Analysis.

a. When analyzing a covariance matrix, the initial eigenvalues are the same across the raw and rescaled solution.



Component Matrix[a]

	Raw		Rescaled	
	Component		Component	
	1	2	1	2
8 yr	2.374	.783	.944	.311
8.5 yr	2.464	.535	.970	.211
9 yr	2.580	-.401	.981	-.153
9.5 yr	2.593	-.827	.949	-.303

Extraction Method: Principal Component Analysis.

a. 2 components extracted.

...9/-

Reproduced Covariances

		8 yr	8.5 yr	9 yr	9.5 yr
Reproduced Covariance	8 yr	6.251[b]	6.270	5.813	5.508
	8.5 yr	6.270	6.358[b]	6.143	5.946
	9 yr	5.813	6.143	6.819[b]	7.022
	9.5 yr	5.508	5.946	7.022	7.405[b]
Residual[a]	8 yr		-.081	-.036	.040
	8.5 yr	-.081		.010	-.022
	9 yr	-.036	.010		-.075
	9.5 yr	.040	-.022	-.075	

Extraction Method: Principal Component Analysis.

a. Residuals are computed between observed and reproduced covariances.

b. Reproduced communalities

Rotated Component Matrix[a]

	Raw		Rescaled	
	Component		Component	
	1	2	1	2
8 yr	2.223	1.145	.883	.455
8.5 yr	2.109	1.382	.830	.544
9 yr	1.522	2.122	.579	.807
9.5 yr	1.227	2.429	.449	.889

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

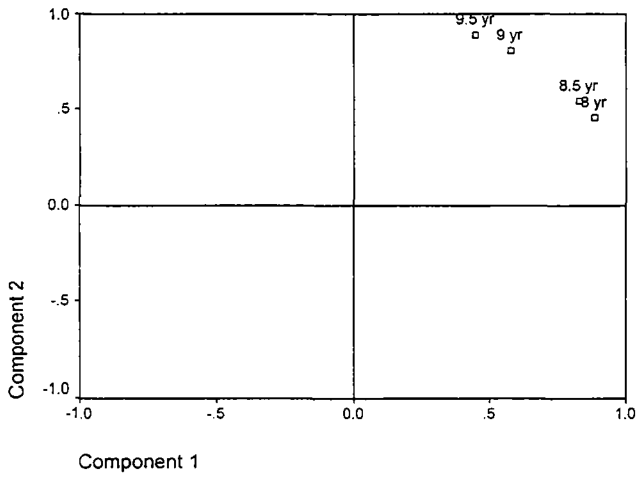
a. Rotation converged in 3 iterations.

Component Transformation Matrix

Component	1	2
1	.701	.713
2	.713	-.701

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

Component Plot in Rotated Space



Component Score Coefficient Matrix [a]

	Component	
	1	2
8 yr	.973	-.622
8.5 yr	.731	-.368
9 yr	-.242	.617
9.5 yr	-.726	1.109

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Coefficients are standardized.

Component Score Covariance Matrix

Component	1	2
1	1.000	.000
2	.000	1.000

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

(b) Gerrild dan Lantz (1969) menganalisis sampel-sampel minyak-mentah (*crude-oil*) dari tiga zon batupasir (*sandstone*):

Z1: Wilhelm
Z2: Sub-Mulinia
Z3: Upper Mulinia

Nilai-nilai bagi unsur-unsur surihan (*trace elements*)

X1 = vanadium (dalam peratus abu)
X2 = besi (dalam peratus abu)
X3 = berilium (dalam peratus abu)

dan dua sukatan hidrokarbon,

X4 = hidrokarbon tepu (dalam peratus luas)
X5 = hidrokarbon aromatik (dalam peratus luas)

diperoleh bagi 56 kes. Analisis data yang sesuai dilaksanakan. Hasil dari pakej SPSS diberikan seperti berikut:

...11/-

Group Statistics

zon batupasir		Mean	Std. Deviation	Valid N (listwise) Unweighted	Weighted
1	besi	43.5714	6.10620	7	7.000
	berilium	.1171	.10012	7	7.000
	hidrokarbon tepu	6.7957	.87195	7	7.000
	hidrokarbon aromatik	11.5400	1.41487	7	7.000
2	besi	33.0909	10.76526	11	11.000
	berilium	.1709	.18393	11	11.000
	hidrokarbon tepu	6.5609	1.26732	11	11.000
	hidrokarbon aromatik	5.4836	3.44068	11	11.000
3	besi	22.2526	8.75883	38	38.000
	berilium	.4321	.33032	38	38.000
	hidrokarbon tepu	4.6582	.99210	38	38.000
	hidrokarbon aromatik	5.7679	2.36888	38	38.000
Total	besi	27.0464	11.60552	56	56.000
	berilium	.3414	.31392	56	56.000
	hidrokarbon tepu	5.2991	1.38724	56	56.000
	hidrokarbon aromatik	6.4336	3.15456	56	56.000

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
besi	.570	20.006	2	53	.000
berilium	.818	5.881	2	53	.005
hidrokarbon tepu	.539	22.673	2	53	.000
hidrokarbon aromatik	.618	16.408	2	53	.000

Pooled Within-Groups Matrices[a]

		besi	berilium	hidrokarbon tepu	hidrokarbon aromatik
Covariance	besi	79.644	.379	1.580	-5.425
	berilium	.379	.084	.163	-.007
	hidrokarbon tepu	1.580	.163	1.076	.627
	hidrokarbon aromatik	-5.425	-.007	.627	6.378
Correlation	besi	1.000	.147	.171	-.241
	berilium	.147	1.000	.543	-.010
	hidrokarbon tepu	.171	.543	1.000	.239
	hidrokarbon aromatik	-.241	-.010	.239	1.000

a. The covariance matrix has 53 degrees of freedom.

...12/-

Covariance Matrices[a]

zon batupasir		besi	berilium	hidrokarbon tepu	hidrokarbon aromatik
1	besi	37.286	-.015	3.348	6.258
	berilium	-.015	.010	.039	.042
	hidrokarbon tepu	3.348	.039	.760	1.159
	hidrokarbon aromatik	6.258	.042	1.159	2.002
2	besi	115.891	.361	4.654	1.555
	berilium	.361	.034	.117	-.031
	hidrokarbon tepu	4.654	.117	1.606	1.198
	hidrokarbon aromatik	1.555	-.031	1.198	11.838
3	besi	76.717	.448	.462	-9.206
	berilium	.448	.109	.196	-.009
	hidrokarbon tepu	.462	.196	.984	.386
	hidrokarbon aromatik	-9.206	-.009	.386	5.612
Total	besi	134.688	-.613	8.318	6.569
	berilium	-.613	.099	.031	-.162
	hidrokarbon tepu	8.318	.031	1.924	1.632
	hidrokarbon aromatik	6.569	-.162	1.632	9.951

a. The total covariance matrix has 55 degrees of freedom.

Box's Test of Equality of Covariance Matrices

Test Results

Box's M		44.692
F	Approx.	1.757
	df1	20
	df2	1193.335
	Sig.	.021

Tests null hypothesis of equal population covariance matrices.

Summary of Canonical Discriminant Functions

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	2.919[a]	82.5	82.5	.863
2	.621[a]	17.5	100.0	.619

a. First 2 canonical discriminant functions were used in the analysis.

...13/-

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 2	.157	95.211	8	.000
2	.617	24.868	3	.000

Standardized Canonical Discriminant Function Coefficients

	Function	
	1	2
besi	.545	.450
berilium	-.786	.474
hidrokarbon tepu	.803	-.806
hidrokarbon aromatik	.259	1.008

Structure Matrix

	Function	
	1	2
hidrokarbon tepu	.531*	-.231
besi	.504*	.139
berilium	-.272*	.092
hidrokarbon aromatik	.328	.702*

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions
Variables ordered by absolute size of correlation within function.

*. Largest absolute correlation between each variable and any discriminant function

Canonical Discriminant Function Coefficients

	Function	
	1	2
besi	.061	.050
berilium	-2.716	1.639
hidrokarbon tepu	.774	-.777
hidrokarbon aromatik	.103	.399
(Constant)	-5.484	-.373

Unstandardized coefficients

Functions at Group Centroids

	Function	
zon batupasir	1	2
1	3.300	1.340
2	1.711	-1.334
3	-1.103	.139

Unstandardized canonical discriminant functions evaluated at group means

...14/-

Classification Statistics

Prior Probabilities for Groups

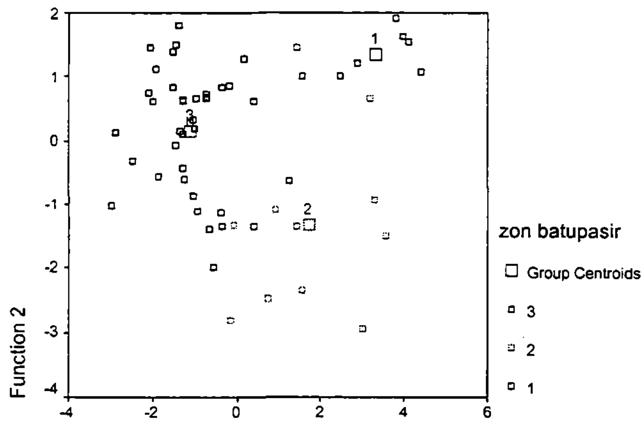
zon batupasir	Prior	Cases Used in Analysis Unweighted	Weighted
1	.333	7	7.000
2	.333	11	11.000
3	.333	38	38.000
Total	1.000	56	56.000

Classification Function Coefficients

	zon batupasir 1	2	3
besi	.595	.363	.266
berilium	-14.007	-14.074	-4.014
hidrokarbon tepu	6.600	7.449	4.127
hidrokarbon aromatik	1.651	.421	.720
(Constant)	-45.201	-31.499	-14.876

Fisher's linear discriminant functions

Canonical Discriminant Functions



(c) Hasil-hasil gandum musim sejuk dalam setiap tahun 1970-1973 di 12 buah lokasi yang berlainan di negeri England (Hand et al., 1994) diperoleh dan dianalisiskan. Keputusannya dipaparkan seperti berikut:

Average Linkage (Between Groups)

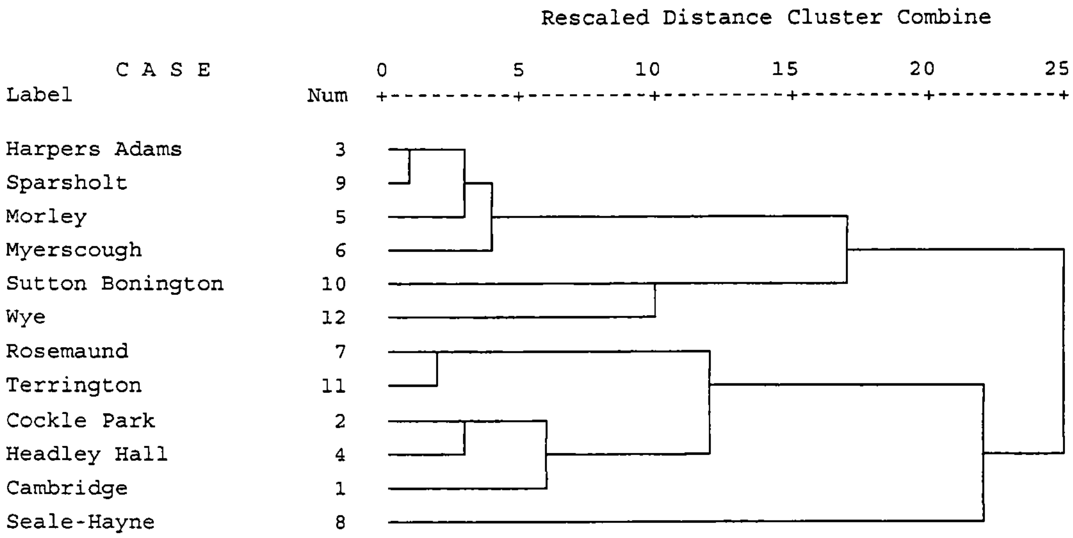
Agglomeration Schedule

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	3	9	34.428	0	0	4
2	7	11	75.825	0	0	8
3	2	4	83.169	0	0	6
4	3	5	93.408	1	0	5
5	3	6	110.151	4	0	9
6	1	2	164.450	0	3	8
7	10	12	246.067	0	0	9
8	1	7	304.429	6	2	10
9	3	10	416.388	5	7	11
10	1	8	516.207	8	0	11
11	1	3	600.522	10	9	0

Dendrogram

* * * * * * H I E R A R C H I C A L   C L U S T E R   A N A L Y S I S * * * * * *

Dendrogram using Average Linkage (Between Groups)



...16/-

Single Linkage

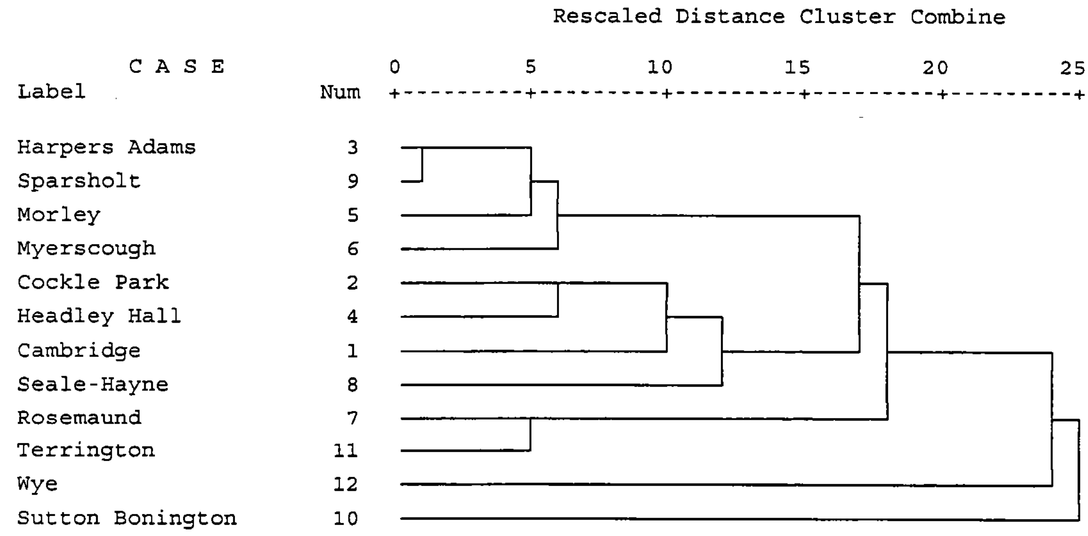
Agglomeration Schedule

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	3	9	34.428	0	0	2
2	3	5	74.899	1	0	5
3	7	11	75.825	0	0	9
4	2	4	83.169	0	0	6
5	3	6	84.605	2	0	8
6	1	2	117.724	0	4	7
7	1	8	133.172	6	0	8
8	1	3	176.705	7	5	9
9	1	7	179.228	8	3	10
10	1	12	232.205	9	0	11
11	1	10	246.067	10	0	0

Dendrogram

* * * * * * H I E R A R C H I C A L   C L U S T E R   A N A L Y S I S * * * * * *

Dendrogram using Single Linkage



...17/-

Complete Linkage

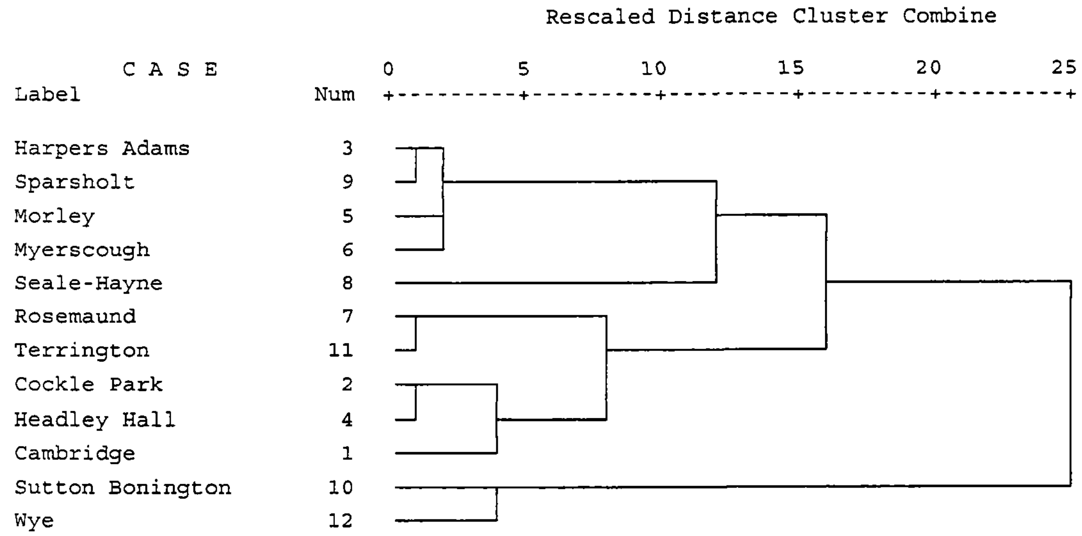
Agglomeration Schedule

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	3	9	34.428	0	0	5
2	7	11	75.825	0	0	8
3	2	4	83.169	0	0	6
4	5	6	102.677	0	0	5
5	3	5	143.173	1	4	9
6	1	2	211.176	0	3	8
7	10	12	246.067	0	0	11
8	1	7	432.647	6	2	10
9	3	8	640.482	5	0	10
10	1	3	853.382	8	9	11
11	1	10	1398.284	10	7	0

Dendrogram

* * * * * * H I E R A R C H I C A L   C L U S T E R   A N A L Y S I S * * * * * *

Dendrogram using Complete Linkage



[100 markah]

...18/-

LAMPIRAN

Tatatanda adalah seperti di dalam kuliah.

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

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

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

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

3. f.k.k. normal bivariat:

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{\sigma_{11}\sigma_{22}(1-\rho_{12}^2)}} \; x \; \exp\left\{-\frac{1}{2(1-\rho_{12}^2)}\right.$$

$$\left.\left[\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)^2 + \left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right) - 2\rho_{12}\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)\left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)\right]\right\}$$

4. f.k.k. normal multivariat:

$$f(x_1, x_2) = \frac{1}{(2\pi)^{p/2}|\sum|^{1/2}} e^{-(1/2)(x-\mu)'\sum^{-1}(x-\mu)}$$

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

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 $\Lambda^{\frac{2}{n}} = \dfrac{|\hat{\sum}|}{|\hat{\sum}_0|} = \left(1 + \dfrac{T^2}{n-1}\right)^{-1}$

...19/-

(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 $\mu_i,\;\; i=1,2,\cdots,p$:

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

7. Dua sampel tak bersandar:

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

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

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

$$l'(\bar{X}_1-\bar{X}_2) \pm c\sqrt{l'\left(\frac{1}{n_1}+\frac{1}{n_2}\right)S_p l}$$

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

8. MANOVA satu-hala:

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

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

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

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

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

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

...20/-

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

Katakan $\wedge = \dfrac{|\boldsymbol{E}|}{|\boldsymbol{E}+\boldsymbol{H}|}$

Maka (1) Untuk p=1,

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

(2) Untuk $m_H = 1$,

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

(3) Untuk $p = 2$,

$$\left(\frac{1-\wedge^{\frac{1}{2}}}{\wedge^{\frac{1}{2}}}\right)\frac{m_E-1}{m_H} \sim F_{2m_H,2(m_E-1)}$$

(4) Untuk $m_H = 2$,

$$\left(\frac{1-\wedge^{\frac{1}{2}}}{\wedge^{\frac{1}{2}}}\right)\left(\frac{m_E+1-p}{p}\right) \sim F_{2p,2(m_E+1-p)}$$

untuk $p \geq 2$.

Pembetulan Bartlett: Katakan $n_0 = m_E + m_H$.
Bagi m_E besar,

$$-f \log \wedge \sim X^2_{pm_H}$$

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

$$= n_0 - \frac{1}{2}(p - m_H + 1)$$

10. MANOVA dua –hala:

$$SSP_{\text{faktor1}} = \sum_{\ell=1}^{g} bn(\bar{x}_{\ell.} - \bar{x})(\bar{x}_{\ell.} - \bar{x})'$$

$$SSP_{\text{faktor 2}} = \sum_{k=1}^{b} gn(\bar{x}_{.k} - \bar{x})(\bar{x}_{.k} - \bar{x})'$$

$$SSP_{\substack{\text{tindakan}\\ \text{bersaling}}} = \sum_{\ell=1}^{g}\sum_{k=1}^{b} n(\bar{x}_{\ell k} - \bar{x}_{\ell.} - \bar{x}_{.k} + \bar{x})(\bar{x}_{\ell k} - \bar{x}_{\ell.} - \bar{x}_{.k} + \bar{x})'$$

$$SSP_{\text{residual}} = \sum_{\ell=1}^{g}\sum_{k=1}^{b}\sum_{r=1}^{n} (x_{\ell kr} - \bar{x}_{\ell k})(x_{\ell kr} - \bar{x}_{\ell k})'$$

...21/-

11. Komponen Prinsipal

(a) $Y_i = e_i' X, \quad i = 1, 2, \ldots, p,$

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

(b) $Y_i = e_i' Z$

$P_{Y_i}, X_k = e_{ki} \sqrt{\lambda_i}, \quad i, \quad k = 1, 2, \ldots, p.$

12. Analisis Faktor

(a) $X - \mu = L F + \in$

(b) $\text{Kov}(X) = L L' + \psi$

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

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

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

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

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

sebesar yang mungkin.

13. Analisis Pembezalayan

(a) $Y = l'X = (\mu_1 - \mu_2)' \sum^{-1} X$

$$m = \frac{1}{2}(\mu_1 - \mu_2)' \sum^{-1} (\mu_1 + \mu_2)$$

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

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

...22/-

(c) Petua peruntukan:

$$\text{Untukkan } \boldsymbol{x}_0 \text{ kepada } \begin{cases} \pi_1 \text{ jika } y_0 \geq \hat{m} \\ \pi_2 \text{ jika } y_0 < \hat{m} \end{cases}$$

(d) $\boldsymbol{B}_0 = \sum_{i=1}^{g} (\mu_i - \bar{\mu})(\mu_i - \bar{\mu})$

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

$\boldsymbol{e}_1, \ \ldots, \ \boldsymbol{e}_s$ vektor eigen $\sum^{-1} \boldsymbol{B}_0$.

$\ell_i \boldsymbol{X} = \boldsymbol{e}_i \boldsymbol{X}$ pembezalayan ke$-i, i = 1, 2, \ldots, s$

(e) $\boldsymbol{B}_0 = \sum_{i=1}^{g} (\bar{\boldsymbol{x}}_i - \bar{\boldsymbol{x}})(\bar{\boldsymbol{x}}_i - \bar{\boldsymbol{x}})'$

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

$\hat{\ell}_i \boldsymbol{x} = \hat{\boldsymbol{e}}_i \boldsymbol{x}$ pembezalayan ke$-i, i = 1, 2, \ldots, s$.

(f) Petua peruntukan :

Untukkan $\boldsymbol{x}$ kepada π_k jika

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

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

- 000 O 000 -