

repository.ub.ac

**ANALISIS KORELASI KANONIK *ROBUST* MENGGUNAKAN
MATRIKS KOVARIAN DENGAN PENDUGA
MINIMUM COVARIANCE DETERMINANT (MCD)
PADA DATA PENCILAN**

SKRIPSI

oleh :

DEWI PRAMITA NINGRUM

0510950011-95



**PROGRAM STUDI STATISTIKA
JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS BRAWIJAYA
MALANG
2009**

UNIVERSITAS BRAWIJAYA



**ANALISIS KORELASI KANONIK *ROBUST* MENGGUNAKAN
MATRIKS KOVARIAN DENGAN PENDUGA
MINIMUM COVARIANCE DETERMINANT (MCD)
PADA DATA PENCILAN**

SKRIPSI

**Sebagai salah satu syarat untuk memperoleh gelar
Sarjana Sains dalam bidang statistika**

**Oleh:
DEWI PRAMITA NINGRUM
0510950011-95**



**PROGRAM STUDI STATISTIKA
JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS BRAWIJAYA
MALANG
2009**

LEMBAR PENGESAHAN SKRIPSI

**ANALISIS KORELASI KANONIK *ROBUST* MENGGUNAKAN
MATRIKS KOVARIAN DENGAN PENDUGA
MINIMUM COVARIANCE DETERMINANT (MCD)
PADA DATA PENCILAN**

Oleh :
DEWI PRAMITA NINGRUM
0510950011-95

Setelah dipertahankan di depan Majelis Penguji
pada tanggal 19 Mei 2009
dan dinyatakan memenuhi syarat untuk memperoleh gelar
Sarjana Sains dalam bidang statistika

Pembimbing I

Pembimbing II

Prof. Dr. Ir. Loekito Adi S., MAg.
NIP. 130 518 961

Suci Astutik, SSi., MSi.
NIP. 132 233 148

Mengetahui,
Ketua Jurusan Matematika
Fakultas MIPA Universitas Brawijaya

Dr. Agus Suryanto, MSc.
NIP. 132 126 049

LEMBAR PERNYATAAN

Saya yang bertanda tangan di bawah ini :

Nama : DEWI PRAMITA NINGRUM
NIM : 0510950011-95
Program Studi : Statistika
Penulis Skripsi Berjudul :

**Analisis Korelasi Kanonik *Robust* Menggunakan Matriks Kovarian
Dengan Penduga *Minimum Covariance Determinant* (MCD)
Pada Data Pencilan**

Dengan ini menyatakan bahwa :

1. Isi dari skripsi yang saya buat adalah benar-benar karya sendiri dan tidak menjiplak karya orang lain, selain nama-nama yang termaktub di isi dan tertulis di daftar pustaka dalam skripsi ini.
2. Apabila di kemudian hari ternyata skripsi yang saya tulis terbukti hasil jiplakan, maka saya akan bersedia menanggung segala resiko yang akan saya terima.

Demikian pernyataan ini dibuat dengan segala kesadaran.

Malang, 19 Mei 2009
Yang menyatakan,

(Dewi Pramita Ningrum)
NIM. 0510950011-95

ANALISIS KORELASI KANONIK *ROBUST* MENGGUNAKAN
MATRIKS KOVARIAN DENGAN PENDUGA
MINIMUM COVARIANCE DETERMINANT (MCD)
PADA DATA PENCILAN

ABSTRAK

Analisis korelasi kanonik adalah analisis statistika multivariat yang memungkinkan identifikasi dan kuantifikasi hubungan antara dua himpunan variabel. Adanya pencilan pada data berpengaruh terhadap matriks kovarian yang dihasilkan, sehingga digunakan pendekatan *robust* multivariat. Salah satu pendekatan *robust* multivariat untuk analisis korelasi kanonik menggunakan matriks kovarian dengan penduga *Minimum Covariance Determinant* (MCD). Berdasarkan hal tersebut pada penelitian ini ingin diketahui seberapa besar peningkatan nilai korelasi dan proporsi keragaman pada analisis korelasi kanonik *robust* menggunakan matriks kovarian dengan penduga MCD pada data yang mengandung pencilan. Dalam penelitian ini digunakan 3 data sekunder. Berdasarkan hasil analisis diperoleh kesimpulan bahwa analisis korelasi kanonik *robust* mampu meningkatkan korelasi di antara dua himpunan variabel pada data yang mengandung pencilan dengan peningkatan sebesar 20%-30% dari nilai korelasi pada analisis korelasi kanonik klasik. Hal ini menunjukkan bahwa analisis korelasi kanonik *robust* lebih baik dalam menjelaskan seberapa besar hubungan di antara kedua himpunan variabel. Pada indikator kedua yaitu proporsi keragaman, hasil analisis menunjukkan bahwa proporsi keragaman variabel kanonik u_r yang dapat dijelaskan oleh variabel kanonik v_r dengan menggunakan analisis korelasi kanonik *robust* mampu meningkatkan proporsi keragaman sebesar 30%-40% daripada analisis korelasi kanonik klasik.

Kata kunci: korelasi kanonik *robust*, *minimum covariance determinant*

repository.ub.ac

ROBUST CANNONICAL CORRELATION ANALYSIS USING COVARIANCE MATRICES WITH MINIMUM COVARIANCE DETERMINANT (MCD) ESTIMATOR IN OUTLIER DATA

ABSTRACT

Analysis of canonical correlation is a multivariate statistical analysis which enables an identifying and quantifying relation between two variable sets. The existence of outliers in a data has influence on resulted covariant matrix. One of multivariate robust approach for canonic correlation analysis is by using covariance matrices with Minimum Covariance Determinant (MCD) estimator. Based on above reasons, this research was meant to find out if the robust canonical correlation analysis with MCD estimator was better used than classical canonical correlation analysis on data with outliers. This research used 3 secondary data. Based on the analysis results, the conclusion is the robust canonical correlation analysis can increase correlation between two set variable in outlier data with increasing 20%-30% than the correlation value of classical canonical correlation analysis which shows that robust canonical correlation analysis is better than classical canonical correlation analysis in terms of clarifying how close the relation between both variable sets is. This things supported by variance proportion which from the analysis result shown that variance proportion of canonic u_k variable can be explained by v_k canonic variable on robust canonical correlation increase 30%-40% than classical canonical correlation analysis.

Keywords: robust canonical correlation, minimum covariance determinant,

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan kehadirat Allah SWT atas segala berkah dan rahmat yang telah diberikan sehingga dapat menyelesaikan penelitian dengan judul Analisis Korelasi Kanonik *robust* Menggunakan Metode Penduga *Minimum Covariance Determinant* (MCD) Pada Data Pencilan.

Dalam penyusunan penelitian ini cukup banyak bantuan yang diberikan berbagai pihak, baik berupa bimbingan maupun saran. Oleh karena itu, penulis dalam kesempatan ini mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Ir. Loekito Adi S., MAgr. selaku dosen pembimbing I dan Ibu Suci Astutik, SSi., MSi. selaku dosen pembimbing II atas motivasi, bimbingan dan waktu yang telah diberikan.
2. Ibu Ir. Heni Kusdarwati, MS, Ibu Dra. Ani Budi Astuti, MSi., dan Bapak Adji Achmad Rinaldo F, SSi., M.Sc selaku dosen penguji yang telah banyak memberikan arahan, saran dan kritik yang membuat skripsi ini menjadi lebih baik
3. Bapak Dr. Agus Suryanto, M.Sc selaku Ketua Jurusan Matematika FMIPA Universitas Brawijaya Malang
4. Bapak, Ibuk, Adek Anis dan saudara-saudaraku atas doa serta dukungannya baik moril maupun materiil
5. Seluruh staf pengajaran dan TU Jurusan Matematika UB
6. Teman-teman statistika 2005 khususnya yang telah banyak membantu, terima kasih untuk kebersamaan kita.
7. Teman-teman statistika 2003, 2004, 2006 atas dukungannya
8. Semua pihak yang telah membantu secara langsung dan tidak langsung yang tidak bisa penulis sebutkan satu per satu.

Penulis menyadari bahwa dalam penulisan skripsi ini masih terdapat banyak kekurangan. Untuk itu saran dan kritik yang membangun sangat diharapkan. Semoga laporan ini dapat bermanfaat bagi semua pihak.

Malang, 19 Mei 2009

Penulis

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
HALAMAN PERNYATAAN	iii
ABSTRAK	iv
ABSTRACT	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR GAMBAR	ix
DAFTAR TABEL	x
DAFTAR LAMPIRAN	xi
BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Permasalahan	2
1.3 Batasan Masalah	2
1.4 Tujuan	2
1.5 Manfaat	2
BAB II TINJAUAN PUSTAKA	
2.1 Distribusi Normal Multivariat	3
2.2 Analisis Korelasi Kanonik	4
2.3 Variabel Kanonik dan Korelasi Kanonik	5
2.4 Variabel Kanonik Sampel	8
2.5 Korelasi Kanonik Sampel	8
2.6 Pendugaan Parameter Korelasi Kanonik dan Variabel Kanonik	9
2.7 Pendekatan <i>Robust</i> untuk Analisis Korelasi Kanonik	9
2.7.1 Deteksi Pencilan Multivariat	10
2.7.2 Matriks Kovarians <i>Robust</i>	11
2.7.3 Matriks Kovarian <i>Robust</i> Menggunakan Penduga <i>Minimum Covariance Determinant (MCD)</i>	11
BAB III METODE PENELITIAN	
3.1 Data	17

3.2 Metode	18
BAB IV HASIL DAN PEMBAHASAN	
4.1 Deskripsi data	21
4.2 Uji Asumsi Normalitas	23
4.3 Deteksi Pencilan Multivariat	23
4.4 Analisis Korelasi Kanonik	25
4.5 Analisis Korelasi Kanonik <i>Robust</i>	26
4.5.1 Analisis Korelasi Kanonik <i>Robust</i> Menggunakan Penduga MCD.....	26
4.6 Perbandingan Analisis Korelasi Kanonik dengan Analisis Korelasi Kanonik <i>Robust</i>	28
4.6.1 Indikator Korelasi Kanonik.....	28
4.6.2 Indikator Persentase Keragaman.....	29
BAB V KESIMPULAN DAN SARAN	
5.1 Kesimpulan	31
5.2 Saran	31
DAFTAR PUSTAKA	33
LAMPIRAN	35



DAFTAR GAMBAR

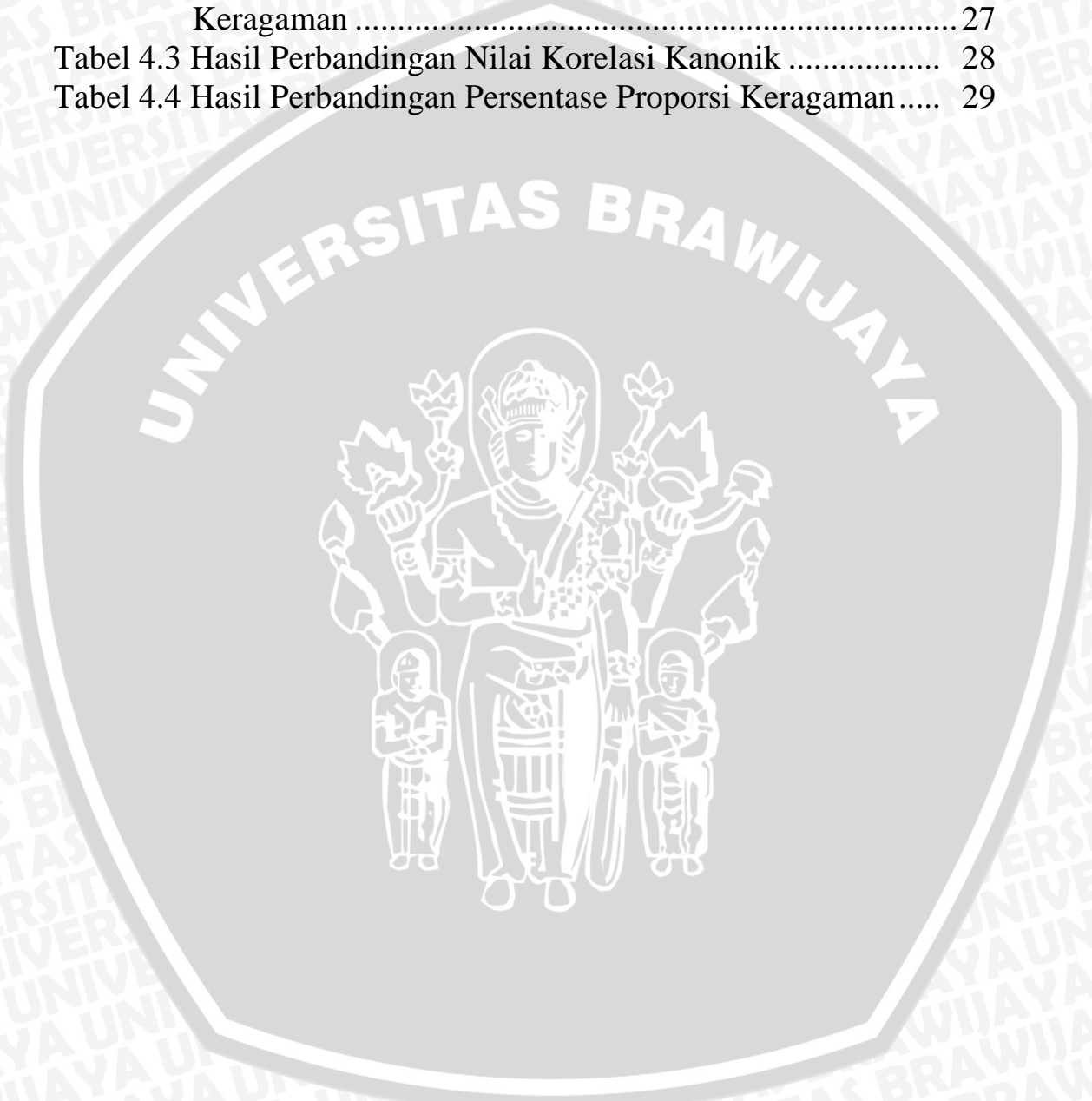
Halaman

Gambar 2.1 Matriks Korelasi	5
Gambar 3.1 Gambar Metode Penelitian	19
Gambar 4.2 Plot Jarak Mahalanobis Data I.....	24
Gambar 4.3 Plot Jarak Mahalanobis Data II.....	24
Gambar 4.2 Plot Jarak Mahalanobis Data III	25



DAFTAR TABEL

	Halaman
Tabel 3.1 Variabel yang Terlibat untuk Setiap Penelitian.....	17
Tabel 4.1 Nilai Korelasi Kanonik dan Proporsi Keragaman	25
Tabel 4.2 Nilai Korelasi Kanonik <i>Robust</i> dan Proporsi Keragaman	27
Tabel 4.3 Hasil Perbandingan Nilai Korelasi Kanonik	28
Tabel 4.4 Hasil Perbandingan Persentase Proporsi Keragaman.....	29



DAFTAR LAMPIRAN

	Halaman
Lampiran 1 Data I.....	35
Lampiran 2 Data II	36
Lampiran 3 Data III.....	37
Lampiran 4 Q-Q Plot Data I, II, III	38
Lampiran 5 Analisis Korelasi Kanonik Data I	40
Lampiran 6 Analisis Korelasi Kanonik Data II.....	42
Lampiran 7 Analisis Korelasi Kanonik Data III.....	44
Lampiran 8 Analisis Korelasi Kanonik <i>Robust</i> Data I	46
Lampiran 9 Analisis Korelasi Kanonik <i>Robust</i> Data II.....	48
Lampiran 10 Analisis Korelasi Kanonik <i>Robust</i> Data III.....	50
Lampiran 11 Makro SPLUS Untuk Analisis Korelasi Kanonik <i>Robust</i> Menggunakan Metode MCD	52
Lampiran 12 Data I yang di Standarisasi	54
Lampiran 13 Data II yang di Standarisasi	55
Lampiran 14 Data III yang di Standarisasi.....	56
Lampiran 15 Macro Minitab Q-Q Plot.....	57





BAB I

PENDAHULUAN

1.1 Latar Belakang

Multivariat merupakan salah satu metode statistika yang digunakan untuk menganalisis data dengan karakteristik lebih dari satu variabel independen dan atau lebih dari satu variabel dependen. Salah satu teknik yang digunakan dalam analisis multivariat untuk mengidentifikasi hubungan antara dua kelompok variabel yaitu analisis korelasi kanonik. Prinsip dari korelasi kanonik yaitu membentuk suatu kombinasi linier dari setiap himpunan variabel sedemikian sehingga korelasi di antara kedua himpunan variabel tersebut menjadi maksimum. Kombinasi linier yang terbentuk pada masing-masing himpunan variabel dinamakan variabel kanonik. Analisis korelasi kanonik tidak dilakukan antar pasangan variabel asal namun antar variabel kanonik di kedua himpunan (Anderson, 1999).

Pendugaan parameter matriks kovarian pada analisis korelasi kanonik umumnya sangat sensitif terhadap adanya pencilan. Keberadaan pencilan pada data dapat berpengaruh terhadap hasil analisis karena dapat memperbesar nilai residual, selain itu adanya pencilan pada data akan menyebabkan dugaan terhadap parameter populasi (ragam) berbias. Oleh karena itu perlu dilakukan identifikasi terhadap pencilan tersebut guna melihat kesalahan sampel observasi. Selanjutnya terdapat pengembangan analisis korelasi kanonik untuk mengatasi adanya pencilan yaitu menggunakan metode *robust* multivariat.

Penelitian ini merujuk pada penelitian Dehon (2000) dengan judul *Robust Methods for Canonical Correlation Analysis*. Hasil yang diperoleh dari penelitian tersebut adalah terdapat beberapa metode pendekatan *robust* multivariat yang dapat digunakan untuk mengatasi pencilan pada analisis korelasi kanonik *robust* yaitu matriks kovarian dengan penduga *Minimum Covariance Determinant* (MCD), *projection pursuit*, *alternating regression* dan *sign test*. Hasil analisis menggunakan keempat metode tersebut menyatakan bahwa matriks kovarian dengan penduga MCD menghasilkan nilai korelasi yang terbesar dengan langkah analisis yang lebih sederhana di banding dengan metode yang lain. Pada penelitian ini akan

dipelajari seberapa besar peningkatan nilai korelasi kanonik dan proporsi keragaman menggunakan pendekatan *robust* multivariat dengan metode matriks kovarian dengan penduga MCD.

1.2 Rumusan Masalah

Seberapa besar analisis korelasi kanonik *robust* mampu meningkatkan keeratan hubungan dan proporsi keragaman di antara dua himpunan variabel pada data yang mengandung pencilan?

1.3 Tujuan

Mengetahui seberapa besar peningkatan keeratan hubungan dan proporsi keragaman di antara dua himpunan variabel pada data yang mengandung pencilan dengan menggunakan analisis korelasi kanonik *robust*.

1.4 Batasan Masalah

Batasan masalah pada penelitian ini yaitu menerapkan analisis korelasi kanonik menggunakan dua himpunan variabel. Data yang digunakan merupakan data yang mengandung pencilan lebih dari satu yaitu sekitar 30% dari jumlah keseluruhan data pengamatan. Penelitian ini mengabaikan sifat pencilan yang terdapat pada data, apakah merupakan pencilan berpengaruh atau tidak berpengaruh.

1.5 Manfaat

Manfaat yang diharapkan dari penelitian ini adalah penduga *minimum covariance determinant* (MCD) dapat digunakan untuk mengatasi data yang mengandung pencilan pada analisis korelasi kanonik *robust*.

BAB II TINJAUAN PUSTAKA

2.1 Distribusi Normal Multivariat

Variabel acak X yang berdistribusi normal univariat dengan rata-rata μ dan ragam σ^2 atau $X \sim N(\mu, \sigma^2)$, memiliki fungsi kepekatan peluang sebagai berikut:

$$f(X) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2\right]; -\infty < X < \infty \quad (2.1)$$

Bila terdapat variabel X_1, X_2, \dots, X_p yang berdistribusi normal multivariat dengan parameter μ dan Σ maka fungsi kepekatan peluang multivariat untuk vektor \mathbf{X} adalah:

$$f(X_1, X_2, \dots, X_p) = \frac{1}{(2\pi)^{\frac{p}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\mathbf{X} - \mu)' \Sigma^{-1} (\mathbf{X} - \mu)\right] \quad (2.2)$$

dengan μ sebagai vektor rata-rata berukuran $p \times 1$ dan Σ adalah matriks ragam-peragam berukuran $p \times p$. Vektor random \mathbf{X} yang berdistribusi normal p variabel dapat ditulis dengan $\mathbf{X} \sim N_p(\mu, \Sigma)$ (Johnson dan Wichern, 2002).

Untuk memeriksa apakah suatu himpunan data mempunyai distribusi normal multivariat adalah dengan menggunakan Q-Q plot yang didasarkan pada jarak kuadrat atau jarak Mahalanobis. Jarak kuadrat dari vektor rata-rata populasi berdistribusi khi-kuadrat dengan derajat bebas p . Langkah-langkah untuk membuat Q-Q plot adalah sebagai berikut (Johnson dan Wichern, 2002) :

1. Hitung nilai d_i^2 sebagai berikut:

$$d_i^2 = (\mathbf{x}_i - \bar{\mathbf{x}})' \mathbf{S}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}) \quad (2.3)$$

Di mana: $i = 1, \dots, n$

n = banyaknya pengamatan

2. Urutkan dari terkecil sampai terbesar sebagai berikut:

$$d_{(1)}^2 \leq d_{(2)}^2 \leq \dots \leq d_{(n)}^2$$

3. Untuk tiap $d_{(i)}^2$ dihitung nilai $\frac{(i-\frac{1}{2})}{n}$, di mana (i) adalah indeks bawah dan merupakan peringkat ke- i .

4. Hitung nilai q_i yaitu nilai khi-kuadrat dari $\frac{(i-\frac{1}{2})}{n}$ dengan derajat bebas p (banyaknya variabel) sebagai berikut:

$$q_i = \chi_p^2, \left(\frac{i-\frac{1}{2}}{n} \right) \quad (2.4)$$

bersesuaian dengan setiap $d_i^2, i=1,2,\dots,n$

5. Plot $d_{(i)}^2$ versus q_i untuk $i=1,2,\dots,n$ dan periksa kelinieran plot. Jika titik-titik tidak membentuk garis lurus, maka kuantil terobservasi tidak mengikuti distribusi khi-kuadrat, sehingga disimpulkan data tidak berdistribusi normal multivariat.

Data dikatakan berdistribusi normal multivariat jika plot cenderung membentuk garis lurus (linier) atau paling tidak 50% nilai d_i^2 lebih kecil dari $\chi_{p,0.05}^2$ ($d_i^2 \leq \chi_{p,0.05}^2$) (Anderson, 1999).

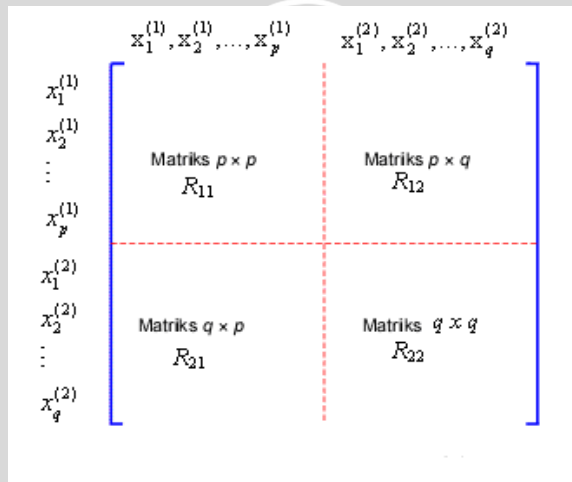
2.2 Analisis Korelasi Kanonik

Analisis korelasi kanonik adalah analisis statistika multivariat yang memungkinkan identifikasi dan kuantifikasi hubungan antara dua himpunan variabel. Fokus analisis korelasi kanonik terletak pada korelasi antara kombinasi linier satu himpunan variabel dengan kombinasi linier himpunan variabel yang lain. Langkah pertama adalah mencari kombinasi linier yang memiliki korelasi terbesar. Selanjutnya, akan dicari pasangan kombinasi linier dengan nilai korelasi terbesar di antara semua pasangan lain yang tidak berkorelasi. Pasangan kombinasi linier disebut sebagai variabel kanonik sedangkan hubungan di antara pasangan tersebut disebut korelasi kanonik. Analisis korelasi kanonik berusaha memusatkan

hubungan dimensi tinggi antara dua himpunan variabel ke dalam sedikit pasangan variabel kanonik (Johnson dan Wichern, 2002).

2.3 Variabel Kanonik dan Korelasi Kanonik

Siregar (2008) menjelaskan bahwa analisis korelasi kanonik dimulai dengan matriks korelasi antara variabel $X_1^{(1)}, X_2^{(1)}, \dots, X_p^{(1)}$ dan variabel $X_1^{(2)}, X_2^{(2)}, \dots, X_q^{(2)}$ dengan $X_p^{(1)}$ dan $X_q^{(2)}$ masing-masing merupakan himpunan variabel pertama dan kedua. Indeks atas (1) dan (2) menyatakan himpunan variabel pertama dan kedua. Dimensi matriks korelasi tersebut adalah $(p+q) \times (p+q)$. Matriks korelasi dapat dipecah menjadi empat partisi yaitu matriks R_{11} R_{12} R_{21} dan R_{22} , seperti disajikan dalam gambar berikut:



Gambar 2.1 Matriks Korelasi

Berdasarkan matriks korelasi dapat dihitung suatu matriks berdimensi $q \times q$ hasil perkalian matriks $R_{22}^{-1} R_{21} R_{11}^{-1} R_{12}$, selanjutnya nilai eigen didapat dari persamaan :

$$(R_{11}^{-1} R_{12} R_{22}^{-1} R_{21} - \lambda I) \mathbf{a} = \mathbf{0} \tag{2.5}$$

Nilai eigen $\lambda_1, \lambda_2, \dots, \lambda_r$ merupakan kuadrat korelasi antara pasangan kombinasi linier. Vektor eigen analisis ini berturut-turut a_1, a_2, \dots, a_r menjadi koefisien variabel $X_j^{(1)}$ untuk pasangan kombinasi linier. Koefisien V_j untuk pasangan kombinasi linier ke- j untuk variabel $X_j^{(2)}$ diperoleh dari elemen vektor:

$$\mathbf{B}_j = \mathbf{R}_{11}^{-1} \mathbf{R}_{12} \mathbf{a}_j \quad (2.6)$$

dari persamaan (2.1) dan (2.2) pasangan kombinasi linier ke- j dihitung dengan perkalian sebagai berikut:

$$U_j = \mathbf{a}'_j \mathbf{X}^{(1)} = (a_{j1}, a_{j2}, \dots, a_{jp}) \begin{bmatrix} X_1^{(1)} \\ X_2^{(1)} \\ \vdots \\ X_p^{(1)} \end{bmatrix}$$

Dan

$$V_j = \mathbf{b}'_j \mathbf{X}^{(2)} = (b_{j1}, b_{j2}, \dots, b_{jq}) \begin{bmatrix} X_1^{(2)} \\ X_2^{(2)} \\ \vdots \\ X_q^{(2)} \end{bmatrix}$$

Di mana:

$$j = 1, \dots, r$$

r = banyaknya kombinasi linier yang terbentuk

p, q = banyaknya himpunan variabel pertama dan kedua

(Siregar, 2008)

Kombinasi linier yang terbentuk:

$$U_r = \mathbf{a}'_r \mathbf{X}^{(1)} \text{ dan } V_r = \mathbf{b}'_r \mathbf{X}^{(2)}$$

Fungsi $U_j = \mathbf{a}'_j \mathbf{X}^{(1)}$ dan $V_j = \mathbf{b}'_j \mathbf{X}^{(2)}$ disebut dengan fungsi kanonik (*canonical function*), sedangkan koefisien \mathbf{a} dan \mathbf{b} disebut dengan bobot kanonik (*canonical weights*).

Selanjutnya Johnson dan Wichern (2002) menjelaskan misalkan terdapat hubungan antara dua kelompok variabel. Kelompok pertama terdiri dari p variabel yang dinyatakan oleh vektor random $\mathbf{X}^{(1)}$ ($p \times 1$) dan kelompok kedua q variabel yang dinyatakan oleh vektor random $\mathbf{X}^{(2)}$ ($q \times 1$). Dari vektor random $\mathbf{X}^{(1)}$ dan $\mathbf{X}^{(2)}$ diperoleh:

$$E(\mathbf{X}^{(1)}) = \bar{\mathbf{x}}^{(1)} ; \text{cov}(\mathbf{X}^{(1)}) = \mathbf{S}_{11}$$

$$E(\mathbf{X}^{(2)}) = \bar{\mathbf{x}}^{(2)} ; \text{cov}(\mathbf{X}^{(2)}) = \mathbf{S}_{22}$$

$$\text{Cov}(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}) = \mathbf{S}_{12} = \mathbf{S}_{21}$$

Sampel random dari n pengamatan pada setiap $(p+q)$ variabel $\mathbf{X}^{(1)}$, $\mathbf{X}^{(2)}$ dapat disusun dalam bentuk matriks data berukuran $n \times (p + q)$ yaitu:

$$\mathbf{X} = [\mathbf{X}^{(1)} : \mathbf{X}^{(2)}] = \begin{bmatrix} x_{11}^{(1)} & x_{12}^{(1)} & \cdots & x_{1p}^{(1)} & x_{11}^{(2)} & x_{12}^{(2)} & \cdots & x_{1q}^{(2)} \\ x_{21}^{(1)} & x_{22}^{(1)} & \cdots & x_{2p}^{(1)} & x_{21}^{(2)} & x_{22}^{(2)} & \cdots & x_{2q}^{(2)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{n1}^{(1)} & x_{n2}^{(1)} & \cdots & x_{np}^{(1)} & x_{n1}^{(2)} & x_{n2}^{(2)} & \cdots & x_{nq}^{(2)} \end{bmatrix}$$

Vektor rata-rata sampel adalah:

$$\bar{\mathbf{x}}_{(p+q) \times 1} = \begin{bmatrix} \bar{x}^{(1)} \\ \vdots \\ \bar{x}^{(2)} \end{bmatrix} \quad \bar{x}^{(1)} = \frac{1}{n} \sum_{i=1}^n x_i^{(1)} \quad \bar{x}^{(2)} = \frac{1}{n} \sum_{i=1}^n x_i^{(2)}$$

di mana i berlaku untuk sejumlah variabel dalam himpunan pertama dan kedua.

Matriks kovarian sampel dapat disusun sebagai berikut:

$$\mathbf{S}_{(p+q) \times (p+q)} = \begin{pmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{pmatrix} \begin{matrix} (p \times p) & (p \times q) \\ (q \times p) & (q \times q) \end{matrix} \quad (2.7)$$

dengan

$$S_{kl} = \frac{1}{n-1} \sum_{i=1}^n (\mathbf{x}_i^{(k)} - \bar{\mathbf{x}}^{(k)})(\mathbf{x}_i^{(l)} - \bar{\mathbf{x}}^{(l)})', \quad k, l = 1, 2$$

di mana:

- \mathbf{S}_{11} : matriks kovarian sampel himpunan variabel $\mathbf{X}^{(1)}$, $(p \times p)$
- \mathbf{S}_{12} : matriks kovarian sampel himpunan variabel $\mathbf{X}^{(1)}$ dan $\mathbf{X}^{(2)}$, $(p \times q)$
- \mathbf{S}_{22} : matriks kovarian sampel himpunan variabel $\mathbf{X}^{(2)}$, $(q \times q)$
- \mathbf{S}_{21} : matriks kovarian sampel himpunan variabel $\mathbf{X}^{(2)}$ dan $\mathbf{X}^{(1)}$, $(q \times p)$

Ragam antara sepasang variabel dari himpunan yang berbeda, yaitu variabel dari $X^{(1)}$ dan variabel dari $X^{(2)}$ termasuk ke dalam S_{12} dan S_{21} .

Ragam dari masing-masing koefisien adalah:

$$\text{Var}(U) = \mathbf{a}' \text{Cov}(\mathbf{X}^{(1)}) \mathbf{a} = \mathbf{a}' \mathbf{S}_{11} \mathbf{a}$$

$$\text{Var}(U) = \mathbf{b}' \text{Cov}(\mathbf{X}^{(2)}) \mathbf{b} = \mathbf{b}' \mathbf{S}_{22} \mathbf{b}$$

$$\text{Cov}(U, V) = \mathbf{a}' \text{Cov}(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}) \mathbf{b} = \mathbf{a}' \mathbf{S}_{12} \mathbf{b}$$

Vektor koefisien a dan b diperoleh dari:

$$\mathbf{r}_{\hat{U}, \hat{V}} = \frac{\hat{\mathbf{a}}' \mathbf{S}_{12} \hat{\mathbf{b}}}{\sqrt{\hat{\mathbf{a}}' \mathbf{S}_{11} \hat{\mathbf{a}}} \sqrt{\hat{\mathbf{b}}' \mathbf{S}_{22} \hat{\mathbf{b}}}} \quad (2.8)$$

(Johnson dan Wichern, 2002).

2.4 Variabel Kanonik Sampel

Misalkan kita menggunakan input matriks \mathbf{S} , maka pasangan variabel kanonik ke-k adalah:

$$\hat{U}_r = \mathbf{a}'_r \mathbf{X}^{(1)} = \hat{\mathbf{e}}'_r \mathbf{S}_{11}^{-1/2} \mathbf{X}^{(1)} \quad (2.9)$$

$$\hat{V}_r = \mathbf{b}'_r \mathbf{X}^{(2)} = \hat{\mathbf{f}}'_r \mathbf{S}_{22}^{-1/2} \mathbf{X}^{(2)} \quad (2.10)$$

di mana:

\mathbf{e}_r = vektor eigen dari $\mathbf{S}_{11}^{-1/2} \mathbf{S}_{12} \mathbf{S}_{22}^{-1} \mathbf{S}_{21} \mathbf{S}_{11}^{-1/2}$

\mathbf{f}_r = vektor eigen dari $\mathbf{S}_{22}^{-1/2} \mathbf{S}_{21} \mathbf{S}_{11}^{-1} \mathbf{S}_{12} \mathbf{S}_{22}^{-1/2}$

r = minimal banyak variabel pada setiap pasangan kombinasi linier.

\mathbf{a}_r = bobot kanonik himpunan pertama

\mathbf{b}_r = bobot kanonik himpunan kedua

(Johnson dan Wichern, 2002).

2.5 Korelasi Kanonik Sampel

Astutik (2007) menjelaskan bahwa korelasi kanonik ke-r merupakan korelasi antara variabel kanonik U_r dan V_r yang dirumuskan sebagai berikut:

$$\text{Corr}(U_r, V_r) = \rho_r^* \quad r=1,2,\dots,p \quad (2.11)$$

di mana $\rho_1^{*2} \geq \rho_2^{*2} \geq \dots \geq \rho_p^{*2}$ adalah nilai Eigen dari matriks kovariansampel $\mathbf{S}_{11}^{-1/2} \mathbf{S}_{12} \mathbf{S}_{22}^{-1} \mathbf{S}_{21} \mathbf{S}_{11}^{-1/2}$ (atau ekuivalen dengan nilai eigen terbesar dari $\mathbf{S}_{22}^{-1/2} \mathbf{S}_{21} \mathbf{S}_{11}^{-1} \mathbf{S}_{12} \mathbf{S}_{22}^{-1/2}$).

Selanjutnya, Johnson dan Wichern (2002) menjelaskan bahwa nilai koefisien korelasi kanonik menyatakan keeratan hubungan antara variabel kanonik yang satu dengan yang lainnya, berada pada kisaran $-1 \leq \rho_r \leq +1$ dan kuadrat korelasi kanonik merupakan proporsi keragaman variabel kanonik U_r yang dapat dijelaskan terhadap variabel kanonik V_r .

2.6 Pendugaan Parameter Korelasi Kanonik dan Variabel Kanonik

Jika terdapat X_1, X_2, \dots, X_n sebanyak N pengamatan dari $N(\mu, \Sigma)$. \mathbf{x} dipartisi ke dalam dua subvektor komponen p dan q , yaitu:

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}^{(1)} \\ \mathbf{X}^{(2)} \end{pmatrix} \quad (2.12)$$

Penduga *maximum likelihood* bagi Σ adalah:

$$\begin{aligned} \hat{\Sigma} &= \begin{pmatrix} \hat{\Sigma}_{11} & \hat{\Sigma}_{12} \\ \hat{\Sigma}_{21} & \hat{\Sigma}_{22} \end{pmatrix} = \frac{1}{N} \sum_{\alpha=1}^N (\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})' \\ &= \frac{1}{N} \begin{pmatrix} \Sigma(\mathbf{x}^{(1)} - \bar{\mathbf{x}}^{(1)})(\mathbf{x}^{(1)} - \bar{\mathbf{x}}^{(1)})' & \Sigma(\mathbf{x}^{(1)} - \bar{\mathbf{x}}^{(1)})(\mathbf{x}^{(2)} - \bar{\mathbf{x}}^{(2)})' \\ \Sigma(\mathbf{x}^{(2)} - \bar{\mathbf{x}}^{(2)})(\mathbf{x}^{(1)} - \bar{\mathbf{x}}^{(1)})' & \Sigma(\mathbf{x}^{(2)} - \bar{\mathbf{x}}^{(2)})(\mathbf{x}^{(2)} - \bar{\mathbf{x}}^{(2)})' \end{pmatrix} \end{aligned}$$

2.7 Pendekatan *Robust* untuk Analisis Korelasi Kanonik

Menurut Huber (1981), metode pendugaan *robust* mendapat perhatian yang cukup menarik karena mempunyai keterkaitan besar dengan masalah dan studi pendugaan lokasi yang sensitif terhadap pencilan.

Vektor rata-rata sampel dan matriks kovarian sampel merupakan landasan dasar dalam analisis multivariat. Keduanya akan mendapatkan hasil yang optimal jika data yang tersedia berdistribusi normal. Distribusi normal mempunyai peran yang penting dalam

analisis multivariat. Dalam hal ini vektor rata-rata sampel dan matriks kovarian sampel akan menjadi kurang efisien ketika terdapat pencilan pada data yang akan dianalisis. Jarak mahalanobis belum mampu untuk mengatasi pencilan dengan jumlah lebih dari satu. Selanjutnya terdapat beberapa pengembangan teori untuk mengatasi pencilan pada kasus multivariat, salah satunya menggunakan pendekatan *robust* multivariat. Salah satu pendekatan *robust* pada analisis korelasi kanonik menggunakan matriks kovarian dengan penduga *Minimum Covariance Determinant* (MCD). Penduga *robust* MCD merupakan rata-rata dan kovarian dari sebagian pengamatan yang meminimumkan determinan matriks kovarian.

2.7.1 Deteksi Pencilan Multivariat

Pencilan adalah suatu pengamatan yang menyimpang dari pengamatan yang lain sehingga menimbulkan kecurigaan bahwa pengamatan tersebut berasal dari distribusi yang berbeda (Marazzi, 1993).

Sebuah pengamatan x_i diidentifikasi sebagai pencilan jika jarak mahalanobis:

$$d_{MD}^2 = (\mathbf{x}_i - \bar{\mathbf{x}})' \mathbf{S}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}) > \chi_{p, (1-\alpha)}$$

di mana :

$\bar{\mathbf{x}}$ = vektor rata-rata

\mathbf{S} = matriks kovarian

p = banyaknya variabel pengamatan

Jarak mahalanobis dapat digunakan untuk mendeteksi adanya pencilan tunggal dalam data. Suatu hal yang sulit untuk melakukan pendeteksian adanya pencilan pada p variabel data di mana $p > 2$. Jika terdapat pencilan tunggal kita dapat menggunakan jarak mahalanobis, namun pendekatan ini kurang efektif untuk pengamatan dengan pencilan lebih dari satu. Untuk itu sebaiknya menggunakan perhitungan jarak dengan penduga *robust* bagi sebaran multivariat. Identifikasi pencilan menjadi kurang optimal karena adanya pengaruh *masking* dan *swamping*. *Masking* terjadi pada saat pengamatan pencilan tidak terdeteksi karena adanya pengamatan pencilan lain yang berdekatan, sedangkan *swamping* terjadi pada saat pengamatan baik teridentifikasi sebagai pengamatan pencilan. Selanjutnya berkembang beberapa penduga *robust* untuk

mengatasi hal tersebut, salah satu penduga *robust* yang mempunyai kemampuan mengukur jarak dan mendeteksi titik *leverage* (pencilan yang disebabkan oleh variabel independen) adalah *Minimum Covariance Determinant* (MCD). Sebuah pengamatan x_i diidentifikasi sebagai pencilan jika jarak mahalanobis *robust* :

$$d_{RD}^2 = (\mathbf{x}_i - \bar{\mathbf{x}}_{MCD})' \mathbf{S}_{MCD}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}_{MCD}) > \chi_{p,(1-\alpha)}^2 \quad (2.13)$$

di mana $\bar{\mathbf{x}}_{MCD}$ dan \mathbf{S}_{MCD} menyatakan vektor rata-rata dan matriks kovarian dari sebagian data X yang mempunyai determinan matriks kovarian terkecil (Hubert, *et al.*, 2007).

2.7.2 Matriks Kovarian Robust

Langkah melakukan analisis korelasi kanonik *robust* adalah menduga \mathbf{S} dengan menghitung vektor eigen dan nilai eigen dari pendugaan pada persamaan 2.5. Beberapa hasil dari pendugaan dengan metode tersebut dikemukakan oleh Croux dan Dehon dalam Dehon, *et al.* (2000). Salah satu metode untuk menduga matriks kovarian *robust* menggunakan penduga M , tetapi penduga ini ternyata kurang efektif pada data dimensi tinggi. Selanjutnya berkembang suatu pendekatan untuk menduga \mathbf{S} dengan *Minimum Covariance Determinant* (MCD) yang dikemukakan oleh Rousseeuw dalam Branco, *et al.*, 2005. Penduga MCD diperoleh dari subhimpunan berukuran h pada data yang mengandung nilai determinan terkecil dari perhitungan matriks kovarian. Hasil pendugaan berupa matriks kovarian dengan subhimpunan paling optimal. Algoritma yang efisien untuk menghitung penduga MCD dikemukakan oleh Rousseeuw dan Van Driessen (1999).

2.7.3 Matriks Kovarian Robust Menggunakan Penduga Minimum Covariance Determinant (MCD)

Rousseeuw dan Van Driessen (1999) menjelaskan bahwa *Minimum Determinant Covariance* (MCD) merupakan salah satu penduga *robust* bagi sebaran multivariat. Penduga MCD cukup obyektif dalam penentuan h pengamatan matriks kovarian yang memiliki determinan terkecil dengan mencari himpunan bagian dari x sejumlah h pengamatan. Jika himpunan bagian tersebut dimisalkan

dengan x_h , maka terdapat C_h^n kombinasi yang harus dicari untuk mendapatkan penduga MCD. Untuk n kecil penduga MCD mudah ditemukan, tetapi jika n besar maka akan banyak kombinasi sub sub sampel yang harus dicari. Berdasarkan hal tersebut, berkembang suatu algoritma yang disebut dengan FAST-MCD. FAST-MCD mampu mendeteksi model yang layak untuk h atau lebih pengamatan jika n berukuran besar. Sehingga algoritma FAST-MCD tersebut sesuai untuk analisis multivariat. MCD memiliki sifat statistik yang baik karena memenuhi sifat *affine equivariant* dan menghasilkan *breakdown point* tingkat tinggi (Rousseeuw dan Leroy, 1987).

Definisi 1. Penduga Minimum Covariance Determinant (MCD)

Diketahui $\mathbf{X} = [\mathbf{X}_1, \dots, \mathbf{X}_n]$ merupakan himpunan data sejumlah n pengamatan terdiri dari p variabel dimana $n \geq p + 1$. Penduga MCD merupakan pasangan matriks definit positif simetri berdimensi pxp dari suatu sub sampel berukuran h pengamatan di mana $[(n + p + 1)/2] < h \leq n$ dengan

$$\mathbf{T}_1 = (1/h) \sum_{i \in H_1} x_i \tag{2.14}$$

$$\mathbf{S}_1 = (1/h) \sum_{i \in H_1} (x_i - T_1)(x_i - T_1)' \tag{2.15}$$

Penduga MCD mencari himpunan bagian dari \mathbf{x} sejumlah h elemen di mana h integer terkecil dari $(n + p + 1) / 2$ (Rousseeuw dan Van Driessen, 1999).

Teorema C-Steps (Rousseeuw dan Van Driessen, 1999)

Terdapat pasangan data $\mathbf{X}_n = \{ \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n \}$ pada pengamatan p variabel. $H_1 \subset \{1, \dots, n\}$ dengan $|H_1| = h$, $\mathbf{T}_1 := (1/h) \sum_{i \in H_1} x_i$ dan $\mathbf{S}_1 = (1/h) \sum_{i \in H_1} (x_i - T_1)(x_i - T_1)'$. Jika $\det(\mathbf{S}_1) \neq 0$, maka jarak relatif:

$$d_1(i) = \sqrt{(x_i - T_1)' \mathbf{S}_1^{-1} (x_i - T_1)} \quad \text{untuk } i = 1, \dots, n \tag{2.16}$$

Misalkan $H_2 = \{ d_1(i); i \in H_2 \} = \{ (d_1)_{1:n}, \dots, (d_1)_{h:n} \}$ di mana $(d_1)_{1:n} \leq (d_1)_{2:n} \leq \dots \leq (d_1)_{n:n}$ adalah orde jarak \mathbf{T}_2 dan \mathbf{S}_2 dihitung berdasarkan \mathbf{H}_2 . Selanjutnya,

$$\det(\mathbf{S}_2) \leq \det(\mathbf{S}_1) \quad (2.17)$$

Jika diasumsikan $\det(\mathbf{S}_2) > 0$ dan jarak relatif $d_2(i) = d_{(\mathbf{T}_2, \mathbf{S}_2)}(i)$ untuk semua $i = 1, \dots, n$. Dengan menggunakan $\#|H_2| = h$ dan definisi $(\mathbf{T}_2, \mathbf{S}_2)$ diperoleh:

$$\begin{aligned} \frac{1}{hp} \sum_{i \in H_2} d_2^2(i) &= \frac{1}{hp} \text{tr} \sum_{i \in H_2} (\mathbf{x}_i - \mathbf{T}_2)^T \mathbf{S}_2^{-1} (\mathbf{x}_i - \mathbf{T}_2) \\ &= \frac{1}{hp} \text{tr} \sum_{i \in H_2} \mathbf{S}_2^{-1} (\mathbf{x}_i - \mathbf{T}_2) (\mathbf{x}_i - \mathbf{T}_2)^T \\ &= \frac{1}{p} \text{tr} \mathbf{S}_2^{-1} \mathbf{S}_2 \\ &= \frac{1}{p} \text{tr}(\mathbf{I}) = 1 \end{aligned} \quad (2.18)$$

selanjutnya

$$\lambda = \frac{1}{hp} \sum_{i \in H_2} d_2^2(i) = \frac{1}{hp} \sum_{i=1}^h (d_1^2)_{i:n} \leq \frac{1}{hp} \sum_{j \in H_1} d_1^2(j) = 1 \quad (2.19)$$

di mana $\lambda > 0$.

Dengan menggabungkan persamaan (2.18) dan (2.19) dihasilkan:

$$\begin{aligned} \sum_{i \in H_2} d_{(\mathbf{T}_1, \lambda \mathbf{S}_1)}^2(i) &= \frac{1}{hp} \sum_{i \in H_2} (\mathbf{x}_i - \mathbf{T}_1)^T \frac{1}{\lambda} \mathbf{S}_1^{-1} (\mathbf{x}_i - \mathbf{T}_1) \\ &= \frac{1}{\lambda hp} \sum_{i \in H_2} d_1^2(i) = \frac{\lambda}{\lambda} = 1 \end{aligned}$$

Sebagai akibatnya $\det(\lambda \mathbf{S}_1) \geq \det(\mathbf{S}_1)$. Dari (2.19) diperoleh pertidaksamaan $\det(\mathbf{S}_2) \leq \det(\lambda \mathbf{S}_2)$ sehingga

$$\det(\mathbf{S}_2) \leq \det(\lambda \mathbf{S}_1) \leq \det(\mathbf{S}_1) \quad (2.20)$$

$\det(\mathbf{S}_2) = \det(\mathbf{S}_1)$ jika dan hanya jika (2.20) menjadi bentuk persamaan.

1. $\det(\mathbf{S}_2) = \det(\lambda \mathbf{S}_1)$ jika dan hanya jika $(\mathbf{T}_2, \mathbf{S}_2) = (\mathbf{T}_1, \lambda \mathbf{S}_1)$
2. $\det(\lambda \mathbf{S}_1) = \det(\mathbf{S}_1)$ jika dan hanya jika $\lambda = 1$
akibatnya $\mathbf{T}_2 = \mathbf{T}_1$ dan $\mathbf{S}_2 = \mathbf{S}_1$.

Teorema tersebut berlaku jika $\det(\mathbf{S}_1) \neq 0$. Jika $\det(\mathbf{S}_1) > 0$ maka \mathbf{S}_2 akan menghasilkan dengan $\det(\mathbf{S}_2) \leq \det(\mathbf{S}_1)$.

Proses iterasi C-step akan berhenti jika $\det(\mathbf{S}_2)=0$ atau $\det(\mathbf{S}_2) = \det(\mathbf{S}_1)$ (Rousseeuw dan Van Driessen, 1999).

Bentuk lain dari penduga MCD adalah dengan menggunakan pembobot. Bobot nol diberikan kepada pengamatan yang tidak disertakan dalam perhitungan penduga rata-rata dan kovarian MCD, sedangkan yang lainnya diberi bobot satu.

Penduga terboboti diperoleh dari:

$$\mathbf{T}_1 = \frac{\left(\sum_{i=1}^n w_i x_i \right)}{\left(\sum_{i=1}^n w_i \right)} \quad (2.21)$$

$$\mathbf{S}_1 = \frac{\left(\sum_{i=1}^n w_i (x_i - \mathbf{T}_1)(x_i - \mathbf{T}_1)' \right)}{\left(\sum_{i=1}^n w_i - 1 \right)} \quad (2.22)$$

di mana:

$$w_i = 1 \quad \text{jika } d_{(T_{MCD}, S_{MCD})}(i) \leq \sqrt{\chi_{p, 975}^2}$$

$$= 0 \quad \text{selainnya}$$

(Rousseeuw dan Van Driessen, 1999).

Definisi 2. Affine Equivariant

Penduga $\mathbf{T}(\mathbf{X})$ dikatakan *affine equivariant* jika $\mathbf{T}(\mathbf{XA} + \mathbf{kv}^T) = \mathbf{T}(\mathbf{X})\mathbf{A} + \mathbf{v}$ untuk semua \mathbf{A} matriks nonsingular berdimensi $p \times p$ di mana \mathbf{v} vektor $p \times 1$ dan $\mathbf{k} = [1, 1, \dots, 1]^T$ vektor $n \times 1$. Penduga kovarian MCD juga bersifat *affine equivariant* jika $\mathbf{S}(\mathbf{XA} + \mathbf{kv}^T) = \mathbf{A}^T \mathbf{S}(\mathbf{X}) \mathbf{A}$ (Lopuhaa dan Rousseeuw, 1991).

Definisi 3. *Breakdown Point*

Selanjutnya Lopuhaa dan Rousseeuw (1991) menjelaskan bahwa *breakdown point* merupakan jumlah pengamatan minimal yang dapat menggantikan sejumlah pengamatan mula-mula yang berakibat pada nilai taksiran yang dihasilkan sangat berbeda dari taksiran sebenarnya. *Breakdown point* dari penaksir parameter lokasi adalah $T(\mathbf{X})$ adalah proporsi pencilan terkecil m/n yang mengakibatkan nilai taksiran menjadi tak terhingga.

Breakdown point untuk penduga kovarian $\mathbf{S}(\mathbf{X})$ didefinisikan sebagai proporsi pencilan terkecil m/n yang mengakibatkan nilai eigen terbesar $\lambda_q(\mathbf{S})$ mencapai tak berhingga atau nilai eigen terkecil $\lambda_p(\mathbf{S})$ mendekati nol.





BAB III METODOLOGI PENELITIAN

3.1 Data

Data yang digunakan pada penelitian ini adalah data sekunder yang merupakan hasil penelitian mahasiswa Fakultas Matematika dan Ilmu Pengetahuan Alam dan mahasiswa Fakultas Pertanian Universitas Brawijaya. Variabel yang terlibat dalam penelitian ini tercantum pada Tabel 3.1:

Tabel 3.1 Variabel yang terlibat untuk setiap penelitian

Data	Sumber Data	Himpunan	Variabel	Judul Skripsi
I	Skripsi Wahyuniarini (2002)	Ciri fisik tanaman kedelai	1.% perkecambahan	Pengaruh Peubah yang Mempunyai Kontribusi Terkecil pada Analisis Korelasi Kanonik
			2.Tinggi tanaman (cm)	
			3.Hari berbunga (hari)	
		Ciri fisik biji kedelai	1.Berat total (kg/plot)	
			2.Berat biji (kg/plot)	
			3.Rasio biji yang baik (%)	
II	Skripsi Dewi, N. (2005)	Biaya produksi	1. Bahan Baku Tempe (Rp)	Faktor-Faktor yang Mempengaruhi Keputusan Pengusaha pada Pemberian Label Produk Kripik Tempe
			2. Bahan Baku Tepung Beras (Rp)	
			3. Bahan Baku Tepung Kanji (Rp)	
		Total penerimaan	1. Jumlah Produksi (kg)	
			2. Harga Jual (Rp)	
III	Skripsi Susanti, D. (2004)	Biaya perawatan	1. harga pupuk urea (Rp)	Analisis Pendapatan dan Efisiensi Penggunaan Faktor-Faktor Produksi pada Usahatani Jagung
			2. Harga pupuk SP36 (Rp)	
		Biaya tetap	1. Pajak (Rp)	
			2. Penyusutan Alat (Rp)	

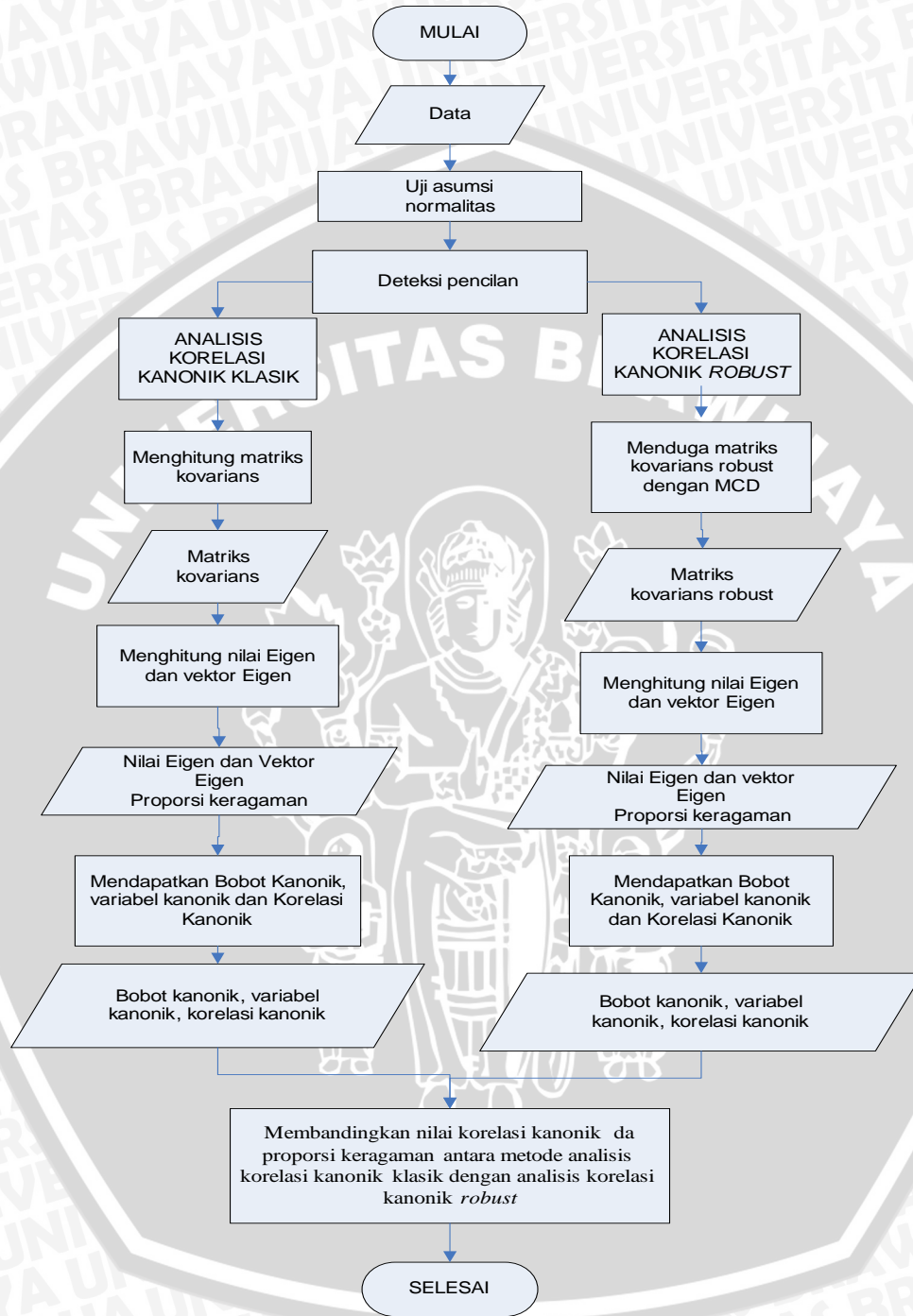
3.2 Metode Penelitian

Tahapan-tahapan yang dilakukan dalam analisis ini:

1. Uji Asumsi normalitas multivariat menggunakan persamaan 2.3
2. Deteksi pencilan menggunakan persamaan 2.13
3. Analisis Korelasi Kanonik:
 - a. Menghitung matriks kovarian
 - b. Menghitung nilai eigen dan vektor eigen
 - c. Mendapatkan bobot kanonik dan korelasi kanonik dari nilai eigen yang diperoleh pada langkah 2.
 - d. Menghitung nilai korelasi kanonik dengan persamaan 2.8
 - e. Menghitung proporsi keragaman yang merupakan akar dari nilai eigen pada langkah 3.b
4. Analisis korelasi kanonik *robust* dengan penduga *minimum covariance determinant*(MCD):
 - a. Menduga matriks kovarian menggunakan MCD dengan:
 - a.1. Mengambil sejumlah h pengamatan yang berbeda sehingga terdapat nC_h sub sampel, h nilai integer terkecil dari $[n + p + 1]/2$
 - a.2. Untuk setiap sub sampel dihitung dengan persamaan (2.14) dan (2.15)
 - a.3. Menghitung nilai jarak relatif menggunakan persamaan (2.16)
 - a.4. Mengurutkan nilai pada langkah a.3 dari yang terkecil hingga terbesar
 - a.5. Iterasi berhenti jika determinan matriks kovarian konvergen.
 - b. Mendapatkan nilai eigen dan vektor eigen berdasarkan matriks kovarian.
 - c. Mendapatkan bobot kanonik dan korelasi kanonik berdasarkan nilai eigen yang diperoleh pada langkah c.
 - d. Menghitung proporsi keragaman berdasarkan nilai eigen pada langkah b.
5. Membandingkan nilai korelasi kanonik dan proporsi keragaman dengan metode analisis korelasi kanonik klasik dan analisis korelasi kanonik *robust*.

Analisis dilakukan menggunakan *software* S-PLUS6 dan MINITAB14.

Tahapan-tahapan yang dilakukan dalam analisis ini disajikan dalam Gambar 3.1:



Gambar 3.1 Langkah-langkah metode penelitian



BAB IV HASIL DAN PEMBAHASAN

4.1 Deskripsi Data

4.1.1 Data I

Data I merupakan data sekunder dari percobaan uji kedelai pada musim kemarau tahun 2000. Percobaan ini bertujuan untuk mengadakan seleksi beberapa galur kedelai yang mempunyai potensi produksi tinggi dan tahan terhadap hama penyakit, khususnya untuk musim kemarau.

Bahan yang diambil untuk analisis ini adalah 6 variabel pada 4 varietas yaitu: bromo; 92-307; 89-210; 88-207. Tanaman kedelai ini ditanam di 17 lokasi di Indonesia, diantaranya: Malang, Pasuruan, Subang, Lombok, dan Jember. Percobaan dilakukan dengan rancangan acak kelompok. Luas plot percobaan 2,4 m x 4,5 m dengan jarak tanam 40 cm x 15 cm dan setiap lubang diisi 4 biji. Kemudian dilakukan penjarangan menjadi 2 tanaman setiap lubangnya setelah daun pertama mulai membuka. Pengambilan contoh tanaman dilakukan secara acak.

Variabel-variabel yang diamati meliputi:

X_1 = Persentase tanaman berkecambah dengan total biji yang ditanam:

$$\frac{\text{tanaman berkecambah}}{\text{Total biji yang ditanam}} \times 100\% = \dots\%$$

X_2 = Tinggi tanaman yang diukur dari pangkal batang sampai tangkai bunga (panjang batang utama).

X_3 = Hari setelah setengah bagian (50%) tanaman dalam plot telah berbunga, dengan perhitungan sebagai berikut:

Y_1 = Berat seluruh tanaman kecuali akar dan daun.

Y_2 = Berat biji tanaman setelah dipanen

Y_3 = Rasio biji tanaman setelah dilakukan seleksi dengan perhitungan sebagai berikut:

$$\frac{\text{Biji yang baik}}{\text{Berat total}} \times 100\% = \dots\%$$

4.1.2 Data II

Kondisi perekonomian dan perdagangan berubah sesuai dengan jumlah permintaan yang semakin meningkat dan beragam. Perubahan tersebut tidak hanya dari segi kuantitas tetapi juga kualitas. Di samping itu, semakin banyak produsen yang menghasilkan produk sejenis. Kondisi ini mengakibatkan terjadinya persaingan yang ketat antar perusahaan. Setiap perusahaan perlu memilih strategi pemasaran yang tepat dan dapat memperkuat keunggulan bersaing. Produk merupakan salah satu elemen penting dalam bauran pemasaran, salah satunya adalah label produk. Bagi produsen label adalah bagian dari atribut produk yang diharapkan mampu mengkomunikasikan nilai dari suatu produk, sedangkan bagi konsumen label akan memberikan informasi mengenai sebuah produk.

Penelitian ini mengangkat permasalahan mengenai faktor-faktor yang mempengaruhi keputusan pengusaha pada pemberian label terhadap peningkatan pendapatan pengusaha kripik tempe yang ada di Sentra Industri Kripik Tempe Sanan Malang. Adapun variabel yang digunakan antara lain lama usaha, tingkat keuntungan, jangkauan pemasaran, dan layanan konsumen. Metode pengambilan sampel ditentukan secara acak sederhana (*simple random sampling*). Dan metode analisis yang digunakan adalah regresi logistik.

4.1.3 Data III

Pertumbuhan penduduk mengakibatkan permintaan akan pangan semakin tinggi. Meningkatnya produksi memberikan kebutuhan masyarakat terpenuhi. Jagung merupakan salah satu komoditi yang mempunyai nilai ekonomis yang tinggi. Di Indonesia, jagung merupakan makanan pokok kedua setelah padi. Sedangkan berdasarkan urutan bahan makanan pokok di dunia, menduduki urutan ketiga setelah padi dan gandum.

Dalam prakteknya pengalokasian faktor-faktor produksi yang tersedia dan dimiliki oleh petani, terdapat kecenderungan bahwa pengetahuan dan informasi petani dalam hal tersebut kurang. Akibatnya produksi tidak dapat dimaksimalkan dan biaya variabel tidak dapat ditekan. Diharapkan dengan diketahuinya efisiensi dari

usahatani jagung ini dapat dijadikan pertimbangan dalam menjalankan usahatani.

Penelitian ini mengangkat permasalahan mengenai usahatani jagung yang ada di desa Ngetal, Kecamatan Pogalan, Kabupaten Trenggalek. Permasalahan yang dihadapi adalah 1) Seberapa besar pendapatan yang diperoleh petani dalam usahatani jagung. 2) Menganalisis efisiensi penggunaan faktor-faktor produksi pada usahatani jagung. Tujuan yang ingin dicapai adalah: 1) Menganalisis tingkat pendapatan, biaya, dan penerimaan pada usahatani jagung 2) Menganalisis efisiensi penggunaan faktor-faktor produksi pada usahatani jagung.

Metode penentuan daerah penelitian dilakukan secara sengaja dengan pertimbangan daerah penelitian merupakan salah satu sentra produksi penghasil jagung di Trenggalek. Metode penentuan sampel untuk petani menggunakan metode sensus pada anggota Kelompok Tani Sumber Makmur, yaitu dari 64 diambil keseluruhan untuk dijadikan sampel.

4.2 Uji Asumsi Normalitas pada Analisis Korelasi Kanonik

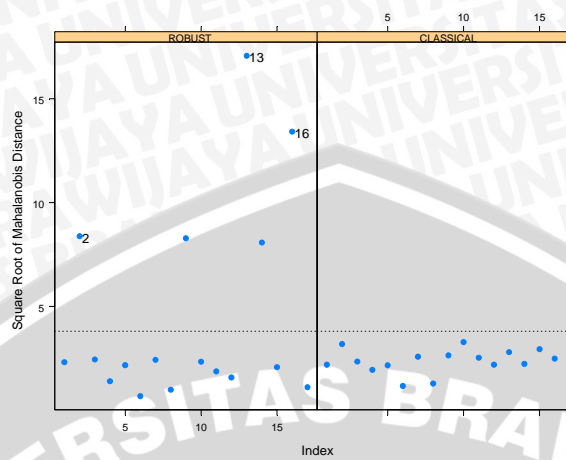
Distribusi normal multivariat mempunyai peranan penting dalam metode statistika multivariat. Oleh karena itu perlu dilakukan pemeriksaan distribusi dari data yang akan dianalisis. Salah satu cara untuk memeriksa sebaran normal multivariat menggunakan Q-Q plot yang didasarkan pada jarak mahalobis. Data dapat dikatakan normal jika plot cenderung membentuk garis lurus (linier) atau 50% dari $d_i^2 \leq \chi_{p;0,05}^2$, p merupakan banyaknya variabel (Anderson, 1999).

Hasil analisis menggunakan Q-Q plot pada setiap himpunan tercantum pada lampiran 4. Dari keenam Q-Q plot tersebut dapat dilihat bahwa plot cenderung membentuk garis lurus (linier) sehingga dapat disimpulkan bahwa keenam data yang digunakan pada penelitian ini memenuhi asumsi distribusi normal multivariat.

4.3 Deteksi pencilan multivariat

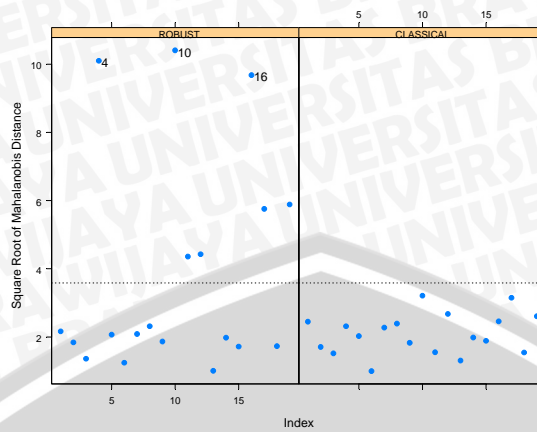
Deteksi pencilan pada analisis multivariat dapat dilihat dari nilai jarak mahalobis yang di bandingkan dengan nilai $\chi_{p,(1-\alpha)}$ dan

dapat juga di gambarkan pada sebuah grafik di mana titik-titik yang berada di luar batas diidentifikasi sebagai pencilan.



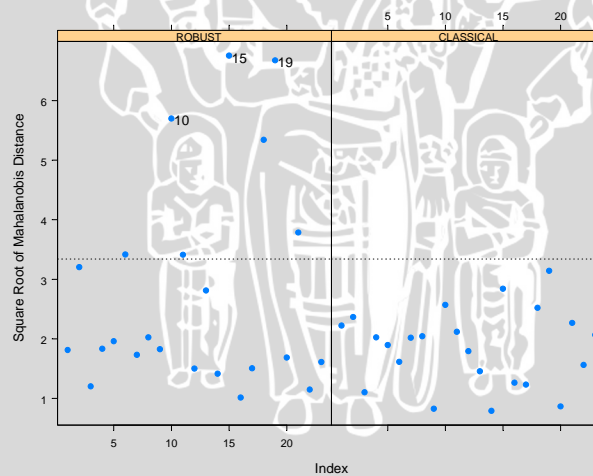
Gambar 4.7 Plot Jarak Mahalanobis Data I

Pada Gambar 4.6 hasil analisis deteksi pencilan *robust* data I menggunakan metode MCD dengan bantuan *software* SPLUS terlihat bahwa 5 titik berada di luar batas. Hal ini mengindikasikan bahwa terdapat 5 pencilan dari 17 data pengamatan. Selain itu, jarak mahalanobis MCD untuk setiap pengamatan juga dapat dilihat pada lampiran 12 di mana pengamatan yang memiliki nilai lebih besar dari $\chi^2_{6,0,975} = 14,45$ dideteksi sebagai pencilan. Sedangkan pada data II terlihat bahwa 7 data pengamatan berada di luar batas, sehingga dapat di tarik kesimpulan bahwa 7 dari 19 data pengamatan dideteksi sebagai pencilan. Sedangkan jarak mahalanobis MCD pada Lampiran 13 juga mengidikasikan hal yang sama dengan ketentuan bahwa pengamatan yang memiliki nilai lebih besar dari $\chi^2_{5,0,975} = 12,83$ dideteksi sebagai pencilan. Gambar 4.8 merupakan hasil deteksi pencilan menggunakan jarak mahalanobis MCD untuk data II.



Gambar 4.8 Plot Jarak Mahalanobis Data II

Pada data III tampak bahwa terdapat 7 data yang berada di luar batas yang juga mengindikasikan bahwa 7 dari 23 data pengamatan terdeteksi sebagai pencilan. Hal yang sama terlihat pada Lampiran 14 di mana data akan dideteksi sebagai pencilan jika jarak mahalanobis MCD memiliki nilai lebih besar dari $\chi^2_{4,0.975} = 11,14$. Gambar 4.9 merupakan hasil deteksi pencilan menggunakan jarak mahalanobis MCD untuk data III.



Gambar 4.9 Plot Jarak Mahalanobis Data III

4.4 Analisis Korelasi Kanonik Klasik

Dari hasil analisis diperoleh beberapa pasangan variabel kanonik yang terbentuk. Ringkasan hasil analisis korelasi kanonik klasik dapat dilihat pada Tabel 4.1, sedangkan hasil lengkapnya dapat dilihat pada Lampiran 5, 6 dan 7.

Tabel 4.1 Nilai korelasi kanonik klasik dan proporsi keragaman

Data	Pasangan variabel kanonik	Korelasi Kanonik Klasik	Proporsi keragaman
I	1	0,7242	0,5244
II	1	0,4307	0,1844
III	1	0,4102	0,1682

Berdasarkan Tabel 4.1 dapat diketahui bahwa besarnya korelasi kanonik pada data I adalah 0,7242 untuk pasangan variabel kanonik pertama. Berdasarkan besarnya korelasi kanonik tersebut, dapat disimpulkan bahwa terdapat korelasi atau hubungan yang kuat antara himpunan variabel ciri fisik tanaman kedelai dan himpunan variabel ciri fisik biji kedelai.

Pada data II besarnya korelasi kanonik adalah 0,4307 untuk pasangan variabel kanonik pertama. Berdasarkan besarnya korelasi kanonik tersebut, dapat disimpulkan bahwa terdapat korelasi atau hubungan yang lemah antara himpunan variabel biaya bahan produksi tempe kripiik dan himpunan variabel total penerimaan penjualan produk tempe kripiik.

Pada data III besarnya korelasi kanonik adalah 0,4102 untuk pasangan variabel kanonik pertama. Berdasarkan besarnya korelasi kanonik tersebut, dapat disimpulkan bahwa terdapat korelasi atau hubungan yang lemah antara himpunan variabel biaya perawatan dan himpunan variabel biaya tetap pada usahatani jagung.

Berdasarkan Tabel 4.1 juga dapat diketahui bahwa besarnya proporsi keragaman untuk data I dengan menggunakan satu pasangan variabel kanonik adalah sebesar 0,5244, artinya dengan menggunakan satu pasangan variabel kanonik, keragaman yang dapat dijelaskan sebesar 52,44%. Besarnya proporsi keragaman untuk data II dengan menggunakan satu pasangan variabel kanonik sebesar 0,1855, dapat di artikan bahwa dengan menggunakan satu pasangan variabel kanonik keragaman yang dapat dijelaskan sebesar 18,55%. Sedangkan pada data III besarnya proporsi keragaman dengan menggunakan satu pasangan variabel kanonik adalah sebesar 0,1682, artinya dengan menggunakan satu pasangan variabel kanonik, keragaman yang dapat dijelaskan sebesar 16,82%.

4.5 Analisis Korelasi Kanonik *Robust* menggunakan matriks kovarian MCD

Salah satu metode *robust* untuk analisis korelasi kanonik yaitu menggunakan matriks kovarian dengan penduga *Minimum Covariance Determinant* (MCD). Hasil analisis korelasi kanonik *robust* dengan penduga MCD dapat diringkas dalam Tabel 4.2:

Tabel 4.2 Nilai korelasi kanonik *robust* dan proporsi keragaman

Data	Pasangan variabel kanonik	Korelasi kanonik	Proporsi keragaman
I	1	0,9406	0,8848
II	1	0,7009	0,4913
III	1	0,7154	0,5118

Berdasarkan Tabel 4.2 dapat diketahui bahwa besarnya proporsi keragaman untuk data I dengan menggunakan satu pasangan variabel kanonik adalah sebesar 0,8848 artinya dengan menggunakan satu pasangan variabel kanonik, keragaman yang dapat dijelaskan sebesar 88,48%. Besarnya korelasi kanonik adalah 0,9406 untuk pasangan variabel kanonik pertama. Berdasarkan besarnya korelasi kanonik tersebut, dapat disimpulkan bahwa terdapat korelasi atau hubungan yang kuat antara himpunan variabel ciri fisik tanaman kedelai dan himpunan variabel ciri fisik biji kedelai.

Pada data II dapat diketahui besarnya korelasi kanonik adalah 0,7009 untuk pasangan variabel kanonik pertama. Berdasarkan besarnya korelasi kanonik tersebut, dapat disimpulkan bahwa terdapat korelasi atau hubungan yang kuat antara himpunan variabel biaya bahan produksi kripik tempe dan himpunan variabel total penerimaan. Dari Tabel 4.2 untuk data II dapat diketahui pula besarnya proporsi keragaman menggunakan satu pasangan variabel kanonik adalah sebesar 0,4913 artinya dengan menggunakan satu pasangan variabel kanonik, keragaman yang dapat dijelaskan sebesar 49,13%.

Selanjutnya, untuk data III besarnya korelasi kanonik yang dapat dijelaskan oleh pasangan variabel kanonik pertama adalah 0,7154. Sedangkan proporsi keragaman yang dapat dijelaskan oleh pasangan variabel kanonik pertama sebesar 0,5118 artinya dengan menggunakan satu pasangan variabel kanonik, keragaman yang dapat dijelaskan sebesar 51,18%.

4.6 Perbandingan Analisis Korelasi Kanonik klasik dengan Korelasi Kanonik *Robust*

4.6.1 Indikator Korelasi Kanonik

Berdasarkan hasil analisis menggunakan korelasi kanonik klasik dan korelasi kanonik *robust* dapat dilihat nilai korelasi kanonik pada setiap pasangan variabel kanonik yang terbentuk. Nilai korelasi kanonik menggunakan analisis korelasi kanonik klasik dan analisis korelasi kanonik *robust* dapat diringkas dalam Tabel 4.3:

Tabel 4.3 Perbandingan Nilai Korelasi Kanonik

Data	Pasangan Variabel Kanonik	Korelasi Kanonik Klasik	Korelasi Kanonik <i>Robust</i>
I	1	0,72417	0,94062
II	1	0,4307	0,7009
III	1	0,4102	0,7154

Berdasarkan Tabel 4.3 dapat dilihat bahwa nilai korelasi kanonik pada pasangan variabel kanonik pertama selalu memiliki nilai yang maksimal dan akan mengalami penurunan pada pasangan variabel kanonik kedua dan seterusnya. Untuk itu korelasi kanonik pasangan pertama mampu menjelaskan seberapa kuat hubungan yang terjadi di antara kedua himpunan variabel. Informasi lain dari Tabel 4.3 berupa nilai korelasi kanonik ketiga data penelitian pada setiap pasangan variabel kanonik menggunakan analisis korelasi kanonik *robust* lebih besar dibanding dengan analisis korelasi kanonik klasik. Hal ini membuktikan bahwa analisis korelasi kanonik *robust* lebih baik dalam mengukur korelasi atau keeratan hubungan di antara dua himpunan variabel. Semakin besar nilai korelasi kanonik semakin

baik pula himpunan variabel kanonik dalam menjelaskan himpunan variabel asalnya (Dehon, *et al.*, 2000).

4.6.2 Indikator Proporsi keragaman

Selain indikator nilai korelasi kanonik, terdapat indikator lain yang dapat digunakan sebagai pembanding yaitu proporsi keragaman. Tabel 4.4 merupakan ringkasan proporsi keragaman menggunakan analisis korelasi kanonik klasik dan analisis korelasi kanonik *robust*.

Tabel 4.4 Perbandingan Nilai Proporsi keragaman

Data	Pasangan Variabel Kanonik	Proporsi keragaman	
		Korelasi Kanonik Klasik	Korelasi Kanonik <i>Robust</i>
I	1	0,5244	0,8848
II	1	0,1844	0,4913
III	1	0,1682	0,5118

Tabel 4.4 menjelaskan perbandingan antara analisis korelasi kanonik dan analisis korelasi kanonik *robust* berdasarkan indikator proporsi keragaman. Analisis korelasi kanonik pada data I dengan menggunakan satu pasangan variabel kanonik mampu menjelaskan keragaman variabel kanonik u_1 terhadap variabel kanonik v_1 sebesar 52,44%. Sedangkan jika digunakan analisis korelasi kanonik *robust*, dengan menggunakan satu pasangan variabel kanonik mampu menjelaskan keragaman data asal sebesar 88,48%. Pada data II untuk analisis korelasi kanonik dengan menggunakan satu pasangan variabel kanonik hanya mampu menjelaskan keragaman variabel kanonik u_1 terhadap variabel kanonik v_1 sebesar 11,44%, sedangkan pada analisis korelasi kanonik *robust* satu pasangan variabel kanonik dapat menjelaskan keragaman sebesar 49,13%. Proporsi keragaman pada data III untuk analisis korelasi kanonik sebesar 16,82% dijelaskan dengan satu pasangan variabel kanonik dan pada analisis korelasi kanonik *robust* 51,18% keragaman variabel kanonik U_1 terhadap variabel kanonik V_1 dapat dijelaskan oleh satu pasangan variabel kanonik.



BAB V KESIMPULAN DAN SARAN

5.1 Kesimpulan

Analisis korelasi kanonik *robust* mampu meningkatkan korelasi di antara dua himpunan variabel pada data yang mengandung pencilan. Hal ini terbukti dari nilai korelasi kanonik untuk analisis korelasi kanonik *robust* meningkat sebesar 20%-30% dari nilai korelasi pada analisis korelasi kanonik klasik. Hal ini menunjukkan bahwa analisis korelasi kanonik *robust* lebih baik dalam menjelaskan seberapa besar hubungan di antara kedua himpunan variabel. Pada indikator kedua yaitu proporsi keragaman, hasil analisis menunjukkan bahwa proporsi keragaman variabel kanonik U_r yang dapat dijelaskan oleh variabel kanonik V_r dengan menggunakan analisis korelasi kanonik *robust* mampu meningkatkan proporsi keragaman sebesar 30%-40% daripada proporsi keragaman yang dihasilkan analisis korelasi kanonik klasik.

5.2 Saran

Penelitian selanjutnya dapat digunakan penduga *minimum covariance determinant* (MCD) untuk pendekatan *robust* multivariat yang lain antara lain analisis faktor.



DAFTAR PUSTAKA

- Anderson, T.W., 1999. **An Introduction to Multivariate Statistical Analysis**. John Wiley & Sons, Inc. New York
- Anonymous. 2007. [Support.spss.com/Student/Documentation/Algorithms/14.0/overalls.pdf](http://support.spss.com/Student/Documentation/Algorithms/14.0/overalls.pdf). Tanggal akses : 18 Januari 2009
- Astutik, S. 2007. **Diktat Analisis Multivariat**. Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Brawijaya, Malang. Tidak dipublikasikan.
- Branco, J.A; C. Croux; P. Filzmoser dan M.R. Oliveira. 2005. **Robust Canonical Correlation: A Comparative Study**. www.proquest.com. Tanggal akses: 30 Oktober 2008
- Dehon, C; P. Filzmoser dan C. Croux. 2000. **Robust Methods for Canonical Correlation Analysis**. <http://www.statistik.tuwien.ac.at/public/filz/papers/namur00.pdf>. Tanggal akses: 29 Oktober 2008
- Dewi, Novita. 2005. **Faktor-faktor yang Mempengaruhi Keputusan Pengusaha pada Pemberian Label Produk Kripik Tempe (Studi Kasus di Sentra Produksi Tempe Sanan Malang)**. Jurusan Sosial Ekonomi Pertanian, Fakultas Pertanian Universitas Brawijaya Malang. Tidak dipublikasikan
- Hair, J.E. Jr.; R. E., Anderson; R. L. Tatham dan W. C. Black, 1998. **Multivariate Data Analysis**, Prentice-Hall International, Inc., New Jersey.
- Huber, P.J. 1981. **Robust Statistics**. Wiley. New York.
- Hubert, M.; J.R. Peter dan V.A. Stevan. 2007. **High-Breakdown Robust Multivariate Methods**. Belgia: K.U. Leuven. Tanggal akses: 4 Januari 2009.

- Johnson, R.A. dan DW. Wichern, 2002. **Applied Multivariate Statistical Analysis**. Fifth edition. Prentice-Hall, Inc., New Jersey.
- Lopuhaa, H.P. dan Rousseeuw, P.J., 1991. **Breakdown Points of Affine Equivariant Estimators of Multivariate Location and Covariance Matrices**. The Annals of Statistics.
- Marazzi, A. 1993. **Algorithms, Routines and S functions for Robust Statistics**. California: Wadworth, Inc.
- Maronna, R. A., R. D. Martin, and V. J. Yohai. 2006. **Robust Statistics**. John Wiley & Sons Ltd: England.
- Rousseeuw, P.J dan Leroy, A.M. 1987. **Robust Regression and Outlier Detection**. Wiley. New York.
- Rousseeuw, P.J., and K. Van Driessen. 1999: **A Fast Algorithm for the Minimum Covariance Determinant Estimator**. , 212-223. *Technometrics*, 41. www.springerlink.com. Tanggal akses 29 Oktober 2008.
- Siregar, L.S. 2008. **Korelasi Kanonik: Komputasi Menggunakan SPSS dan Interpretasi Hasil Analisis**. <http://ssiregar.staff.gunadarma.ac.id/Downloads/files/3024/canonical.pdf>. Tanggal akses: 20 Desember 2008.
- Susanti, Dwi. 2004. **Analisis Pendapatan dan Efisiensi Penggunaan Faktor-Faktor Produksi Pada Usahatani Jagung (*zea mays l.*)**. Jurusan Sosial Ekonomi Pertanian, Fakultas Pertanian Universitas Brawijaya Malang. Tidak dipublikasikan.
- Wahyuniarini, T. 2002 **Pengaruh Peubah yang Mempunyai Kontribusi Terkecil pada Analisis Korelasi Kanonik**. Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Brawijaya. Malang. Tidak dipublikasikan.

Lampiran 1 Data I

Lokasi penelitian	X1	X2	X3	Y1	Y2	Y3
1	75	63.73	37.33	5193	2997	99.7
2	44.8	62.37	35	5637	2283	94.9
3	73.8	81.63	37.33	4410	2270	95.2
4	82.9	60.27	33.67	3230	1870	87.1
5	94.5	63.17	37	5140	2923	94.6
6	82.8	67.93	35	4353	2203	93.5
7	58.4	79.57	36	4530	2047	82.1
8	88.1	65.33	36	4803	2613	95.5
9	64.7	65.9	34	6237	3133	88.8
10	81.6	84.97	38.33	6637	2503	89.7
11	98.8	47.1	36.33	3857	1110	48
12	95	48.23	36	3690	1277	53.5
13	86.7	60.4	42	5187	1403	88.3
14	95	58.33	41	4477	1487	65.3
15	95	54	32	3333	1420	83.9
16	96.3	44.44	39	4167	1640	90.8
17	83.6	63.68	35.33	4797	2420	94.5

(sumber: skripsi Wahyuniarini,2002)

Keterangan:

- X_1 = Persen perkecambahan
 X_2 = Tinggi tanaman (cm)
 X_3 = Hari berbunga (hari)
 Y_1 = Berat total (kg/plot)
 Y_2 = Berat biji (kg/plot)
 Y_3 = Rasio biji yang baik (%)

Lampiran 2 Data II

No.	X1	X2	X3	Y1	Y2
1	12000	5000	2000	15	16000
2	13000	4750	2600	30	17000
3	12000	4800	2300	60	16000
4	12000	4800	1400	20	16000
5	12000	4500	2700	10	17000
6	12000	4800	2500	20	16500
7	10000	2000	1400	10	18000
8	15000	4700	2800	27	15000
9	9000	2300	1400	3	18000
10	12000	4800	1400	12	17500
11	11000	4500	2800	40	17500
12	8000	4800	2600	40	17500
13	12000	2500	1400	30	18000
14	12000	2500	1300	15	17500
15	12000	2400	1200	20	17000
16	12000	2200	2500	12	17000
17	6000	5000	2600	15	17500
18	12000	2300	1300	15	17500
19	10000	4600	1400	50	16500

(sumber: skripsi Novita, 2005)

Keterangan:

 X_1 = Harga Bahan Baku Tempe (Rp) X_2 = Harga Bahan Baku Tepung Beras (Rp) X_3 = Harga Bahan Baku Tepung kanji (Rp) Y_1 = Produksi (kg) Y_2 = Harga jual (Rp/kg)

Lampiran 3 Data III

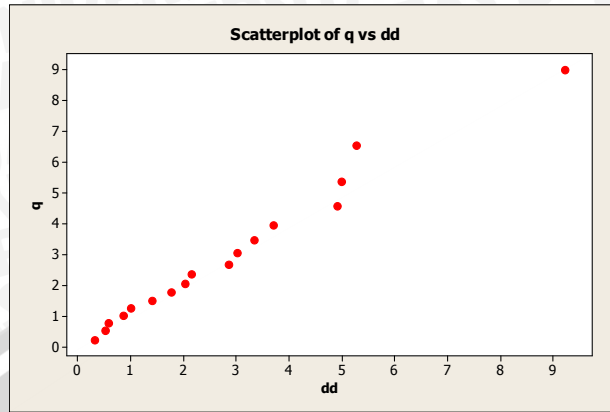
No.	X1	X2	Y1	Y2
1	117500	65000	10333.33	47733.33
2	117500	65000	10333.33	47616.67
3	117500	65000	10333.33	47733.33
4	58750	32500	4583.333	47200
5	47000	32500	4583.333	47200
6	117500	56875	10333.33	54400
7	47000	32500	4583.333	47616.67
8	105750	65000	9166.667	54600
9	117500	65000	10333.33	54566.67
10	117500	56785	10333.33	54566.67
11	70500	32500	5166.667	47616.67
12	70500	40625	5500	47200
13	70500	40625	6200	47900
14	117500	65000	10333.33	47766.67
15	105750	56875	9166.667	47766.67
16	105750	65000	10333.33	47733.33
17	105750	65000	9166.667	54433.33
18	105750	56875	10333.33	47733.33
19	117500	56875	10333.33	47733.33
20	117500	65000	10333.33	54400
21	82250	48750	7233.333	47766.67
22	141000	78000	12400	55100
23	141000	81250	12400	48600

(sumber: skripsi Susanti, 2004)

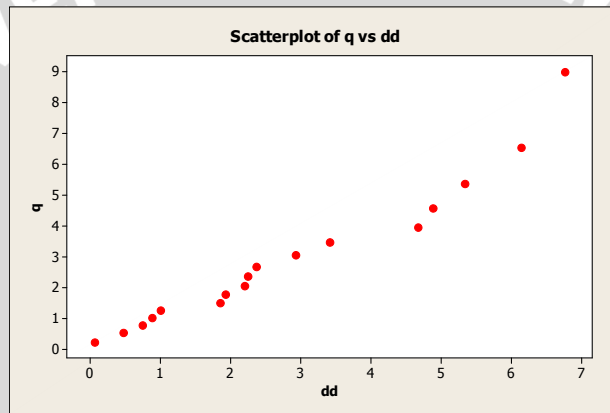
Keterangan:

 X_1 = Harga Pupuk Urea (Rp) X_2 = Harga Pupuk SP36 (Rp) Y_1 = Pajak (Rp) Y_2 = Penyusutan Alat (Rp)

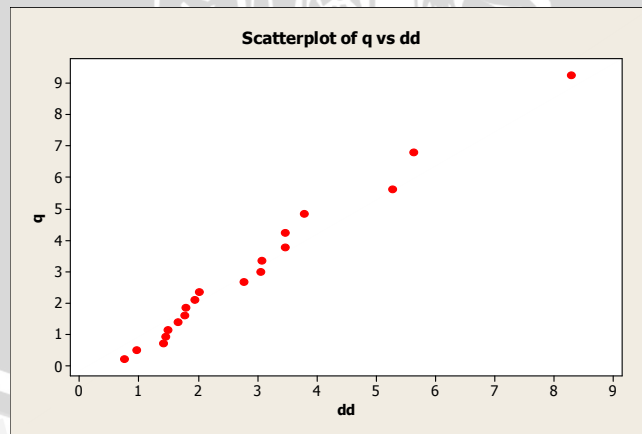
Lampiran 4



Q-Q plot data I Himpunan variabel ciri fisik tanaman kedelai

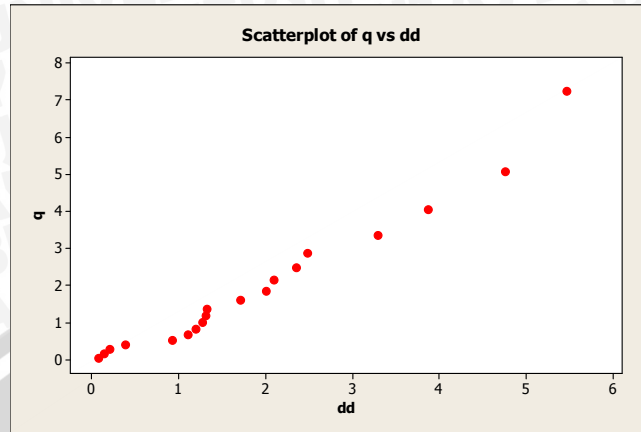


Q-Q plot data I Himpunan variabel ciri fisik biji kedelai

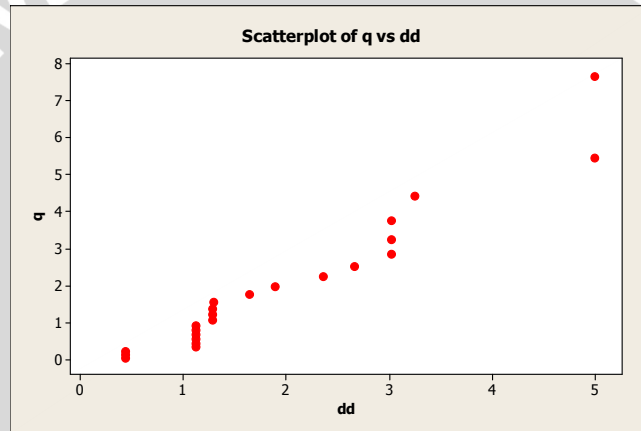


Q-Q plot data II himpunan variabel biaya produksi

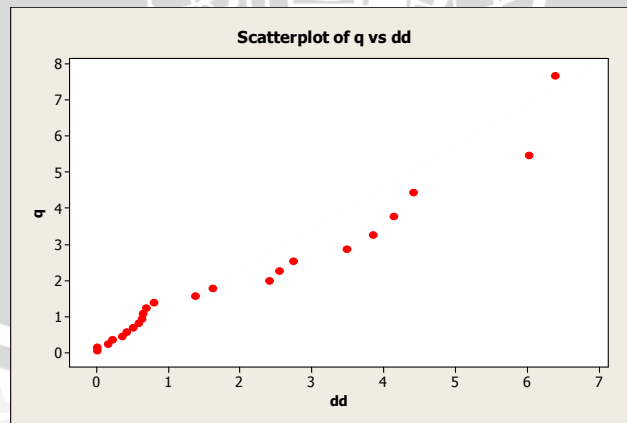
Lampiran 4 (lanjutan)



Q-Q plot data II himpunan variabel total penerimaan



Q-Q plot data III himpunan variabel biaya perawatan



Q-Q plot data III himpunan variabel biaya tetap

Lampiran 5 Analisis Korelasi Kanonik Data I

```
#####
FILE: cc.ssc #
# #
# DESCRIPTION: Canonical correlation, based on a covariance #
# matrix #
# #
# VARIABLES: p<-dimension of the first group #
# q<-dimension of the second group #
# SDF1<-matrix of the SDF1 #
# ind<-indicator of the type of estimator used to #
# obtain the inicial estimation of the cova- #
# riance matrix #
# #
# OBSERVATIONS: ind==1<-Classical estimator #
# ind==2<-MCD estimator #
# ind==3<-M-estimator #
# #
# LAST MODIFICATION: 30 January 2002 #
#####
cc <- function(p, q, data3, ind = 1)
{
  data3 <- as.matrix(data3)
  n <- nrow(data3)
  alpha <- matrix(NA, p, p)
  beta <- matrix(NA, q, p)
  sigma <- matrix(NA, p + q, p + q)
  s11 <- matrix(NA, p, p)
  s22 <- matrix(NA, q, q)
  s12 <- matrix(NA, p, q)
  s <- matrix(NA, p, p)
  a <- matrix(NA, p, p)
  rho <- NULL
  aux <- NULL
  lambda <- NULL
  #sigma<-cor(data3)
  ##### Estimation of the covariance matrix #####
  if(ind == 1) sigma <- var(data3) else if(ind == 2) {
    sigma <- cov.mcd(data3, print = F, cor = T, quan = 0.75
* n)$cov
  }
  else if(ind == 3) {
    sigma <- mesthub(data3, 0.05)$cov
  }
  ##### Estimation of the parameters #####
  s11 <- sigma[1:p, 1:p]
  s12 <- sigma[1:p, (p + 1):(p + q)]
  s22 <- sigma[(p + 1):(p + q), (p + 1):(p + q)]
  s <- solve(s11) %*% s12 %*% solve(s22) %*% t(s12)
  eig <- eigen(s)
  lambda <- eig$values
  rho <- sqrt(lambda)
  a <- eig$vectors
}
```

Lampiran 5 (lanjutan)

```

s <- t(a) %*% s11 %*% a
aux <- sqrt(diag(s))
s <- diag(1/aux)
alpha <- a %*% s
s <- diag(1/rho)
beta <- solve(s22) %*% t(s12) %*% alpha %*% s
list(s = s, au = aux, e = eig, a = alpha, b = beta, r = rho)
}
> cc(3, 3, data3, ind = 1)
$s:
      [,1]      [,2]      [,3]
[1,] 1.380893 0.000000  0.0000
[2,] 0.000000 1.702863  0.0000
[3,] 0.000000 0.000000 157.4804

$au:
[1] 19.058655 31.977714  9.059002

$e:
$$values:
[1] 0.52442105756 0.34485830491 0.00004032244

$$vectors:
      v2      v3      v1
V1  0.7861435 -0.1393806 -0.6584075
V2 -0.5717476 -1.9149860 -0.7495255
V3 -5.1591488  9.9689472 -0.2569593
$a:
      [,1]      [,2]      [,3]
V1  0.04124863 -0.00435868 -0.07267992
V2 -0.02999937 -0.05988502 -0.08273821
V3 -0.27069847  0.31174672 -0.02836508

$b:
      [,1]      [,2]      [,3]
V4 -0.0012138607  0.000728842 -0.0002841178
V5  0.0009936325 -0.002107931 -0.0015273174
V6 -0.0322587588  0.001362461  0.0907847371

$r:
[1] 0.724169219 0.587246375 0.006349995

```

Lampiran 6 Analisis Korelasi Kanonik Data II

```
#####
FILE: cc.ssc #
# #
# DESCRIPTION: Canonical correlation, based on a covariance #
# matrix #
# #
# VARIABLES: p<-dimension of the first group #
# q<-dimension of the second group #
# SDF1<-matrix of the SDF1 #
# ind<-indicator of the type of estimator used to #
# obtain the inicial estimation of the cova- #
# riance matrix #
# #
# OBSERVATIONS: ind==1<-Classical estimator #
# ind==2<-MCD estimator #
# ind==3<-M-estimator #
# #
# LAST MODIFICATION: 30 January 2002 #
#####
cc <- function(p, q, SDF8, ind = 1)
{
  SDF8 <- as.matrix(SDF8)
  n <- nrow(SDF8)
  alpha <- matrix(NA, p, p)
  beta <- matrix(NA, q, p)
  sigma <- matrix(NA, p + q, p + q)
  s11 <- matrix(NA, p, p)
  s22 <- matrix(NA, q, q)
  s12 <- matrix(NA, p, q)
  s <- matrix(NA, p, p)
  a <- matrix(NA, p, p)
  rho <- NULL
  aux <- NULL
  lambda <- NULL
  #sigma<-cor(SDF8)
  ##### Estimation of the covariance matrix #####
  if(ind == 1) sigma <- var(SDF8) else if(ind == 2) {
    sigma <- cov.mcd(SDF8, print = F, cor = T, quan = 0.75 *
n)$cov
  }
  else if(ind == 3) {
    sigma <- mesthub(SDF8, 0.05)$cov
  }
  ##### Estimation of the parameters #####
  s11 <- sigma[1:p, 1:p]
  s12 <- sigma[1:p, (p + 1):(p + q)]
  s22 <- sigma[(p + 1):(p + q), (p + 1):(p + q)]
  s <- solve(s11) %*% s12 %*% solve(s22) %*% t(s12)
  eig <- eigen(s)
  lambda <- eig$values
  rho <- sqrt(lambda)
  a <- eig$vectors
}
```


Lampiran 6 (lanjutan)

```

s <- t(a) %*% s11 %*% a
aux <- sqrt(diag(s))
s <- diag(1/aux)
alpha <- a %*% s
s <- diag(1/rho)
beta <- solve(s22) %*% t(s12) %*% alpha %*% s
list(s = s, au = aux, e = eig, a = alpha, b = beta, r = rho)
}
> cc(2, 3, SDF8, ind = 1)
$s:
      [,1]      [,2]
[1,] 2.321659  0.000000
[2,] 0.000000 10.75748

$au:
[1] 99.34547 144.65888

$e:
$e$values:
[1] 0.185525219 0.008641302

$e$vectors:
      V1      V2
V1 15.97433347 -4.0317840
V2  0.05661928  0.9873165

$a:
      [,1]      [,2]
V1 0.1607957844 -0.027870975
V2 0.0005699231  0.006825136

$b:
      [,1]      [,2]
V3 -5.328638e-006 0.0003709531
V4  7.362603e-004 0.0004828729
V5 -1.859543e-003 0.0002631774

$r:
[1] 0.4307264 0.0929586

```

Lampiran 7 Analisis Korelasi Kanonik Data III

```
#####
FILE: cc.ssc #
# #
# DESCRIPTION: Canonical correlation, based on a covariance #
# matrix #
# #
# VARIABLES: p<-dimension of the first group #
# q<-dimension of the second group #
# SDF1<-matrix of the SDF1 #
# ind<-indicator of the type of estimator used to #
# obtain the inicial estimation of the cova- #
# riance matrix #
# #
# OBSERVATIONS: ind==1<-Classical estimator #
# ind==2<-MCD estimator #
# ind==3<-M-estimator #
# #
# LAST MODIFICATION: 30 January 2002 #
#####
cc <- function(p, q, SDF9, ind = 1)
{
  SDF9 <- as.matrix(SDF9)
  n <- nrow(SDF9)
  alpha <- matrix(NA, p, p)
  beta <- matrix(NA, q, p)
  sigma <- matrix(NA, p + q, p + q)
  s11 <- matrix(NA, p, p)
  s22 <- matrix(NA, q, q)
  s12 <- matrix(NA, p, q)
  s <- matrix(NA, p, p)
  a <- matrix(NA, p, p)
  rho <- NULL
  aux <- NULL
  lambda <- NULL
  #sigma<-cor(SDF9)
  ##### Estimation of the covariance matrix #####
  if(ind == 1) sigma <- var(SDF9) else if(ind == 2) {
    sigma <- cov.mcd(SDF9, print = F, cor = T, quan = 0.75 *
n)$cov
  }
  else if(ind == 3) {
    sigma <- mesthub(SDF9, 0.05)$cov
  }
  ##### Estimation of the parameters #####
  s11 <- sigma[1:p, 1:p]
  s12 <- sigma[1:p, (p + 1):(p + q)]
  s22 <- sigma[(p + 1):(p + q), (p + 1):(p + q)]
  s <- solve(s11) %*% s12 %*% solve(s22) %*% t(s12)
  eig <- eigen(s)
  lambda <- eig$values
  rho <- sqrt(lambda)
  a <- eig$vectors
}
```

Lampiran 7 (lanjutan)

```

s <- t(a) %*% s11 %*% a
aux <- sqrt(diag(s))
s <- diag(1/aux)
alpha <- a %*% s
s <- diag(1/rho)
beta <- solve(s22) %*% t(s12) %*% alpha %*% s
list(s = s, au = aux, e = eig, a = alpha, b = beta, r = rho)
}
> cc(2, 2, SDF9, ind = 1)
$S:
      [,1] [,2]
[1,] 2.437986 0.000000
[2,] 0.000000 9.29405

$au:
[1] 28293.574 4104.497

$e:
$e$values:
[1] 0.16824321 0.01157684

$e$vectors:
      v1      v2
V1  0.8719100 -0.3717169
V2 -0.4896661  1.3556642

$a:
      [,1] [,2]
V1  0.00003081654 -0.00009056334
V2 -0.00001730662  0.00033028755

$b:
      [,1] [,2]
V3 -0.0006109075  0.0007515705
V4  0.0002545695  0.0005072981

$r:
[1] 0.4101746 0.1075957

```


Lampiran 8 Analisis Korelasi Kanonik *Robust* Data I

```
#####
FILE: cc.ssc #
# DESCRIPTION: Canonical correlation, based on a covariance #
# matrix #
# # #
# VARIABLES: p<-dimension of the first group #
# q<-dimension of the second group #
# SDF1<-matrix of the SDF1 #
# ind<-indicator of the type of estimator used to #
# obtain the inicial estimation of the cova- #
# riance matrix #
# # #
# OBSERVATIONS: ind==1<-Classical estimator #
# ind==2<-MCD estimator #
# ind==3<-M-estimator #
# # #
# LAST MODIFICATION: 30 January 2002 #
#####
cc <- function(p, q, data3, ind = 1)
{
  data3 <- as.matrix(data3)
  n <- nrow(data3)
  alpha <- matrix(NA, p, p)
  beta <- matrix(NA, q, p)
  sigma <- matrix(NA, p + q, p + q)
  s11 <- matrix(NA, p, p)
  s22 <- matrix(NA, q, q)
  s12 <- matrix(NA, p, q)
  s <- matrix(NA, p, p)
  a <- matrix(NA, p, p)
  rho <- NULL
  aux <- NULL
  lambda <- NULL
  #sigma<-cor(data3)
  ##### Estimation of the covariance matrix #####
  if(ind == 1) sigma <- var(data3) else if(ind == 2) {
    sigma <- cov.mcd(data3, print = F, cor = T, quan = 0.75
* n)$cov
  }
  else if(ind == 3) {
    sigma <- mesthub(data3, 0.05)$cov
  }
  ##### Estimation of the parameters #####
  s11 <- sigma[1:p, 1:p]
  s12 <- sigma[1:p, (p + 1):(p + q)]
  s22 <- sigma[(p + 1):(p + q), (p + 1):(p + q)]
  s <- solve(s11) %*% s12 %*% solve(s22) %*% t(s12)
  eig <- eigen(s)
  lambda <- eig$values
  rho <- sqrt(lambda)
  a <- eig$vectors
  s <- t(a) %*% s11 %*% a
}
```

Lampiran 8 (lanjutan)

```

aux <- sqrt(diag(s))
s <- diag(1/aux)
alpha <- a %*% s
s <- diag(1/rho)
beta <- solve(s22) %*% t(s12) %*% alpha %*% s
list(s = s, au = aux, e = eig, a = alpha, b = beta, r = rho)
}
> cc(3, 3, data3, ind = 2)
$s:
      [,1] [,2] [,3]
[1,] 1.063127 0.0000 0.000000
[2,] 0.000000 1.1561 0.000000
[3,] 0.000000 0.0000 3.876393

$au:
[1] 19.724659 15.526014 7.076896

$e:
$$values:
[1] 0.88476834 0.74818554 0.06654943

$$vectors:
      V3      V2      V1
V1  0.4428244 -0.7388796 0.8910813
V2  1.6664542 -1.4348275 0.4536927
V3 -12.6725456 -2.8395367 -0.6705613

$a:
      [,1] [,2] [,3]
V1  0.02245029 -0.04758978 0.12591413
V2  0.08448583 -0.09241442 0.06410899
V3 -0.64247222 -0.18288897 -0.09475358

$b:
      [,1] [,2] [,3]
V4  0.0002533699 -0.001250778 0.001167236
V5 -0.0031842264 0.001073537 -0.002974061
V6  0.1214436544 -0.035231274 0.022710770

$r:
[1] 0.9406213 0.8649772 0.2579718

```

Lampiran 9 Analisis Korelasi Kanonik *Robust* Data II

```

#####
FILE: cc.ssc #
# #
# DESCRIPTION: Canonical correlation, based on a covariance #
# matrix #
# #
# VARIABLES: p<-dimension of the first group #
# q<-dimension of the second group #
# SDF1<-matrix of the SDF1 #
# ind<-indicator of the type of estimator used to #
# obtain the inicial estimation of the cova- #
# riance matrix #
# #
# OBSERVATIONS: ind==1<-Classical estimator #
# ind==2<-MCD estimator #
# ind==3<-M-estimator #
# #
# LAST MODIFICATION: 30 January 2002 #
#####
cc <- function(p, q, SDF8, ind = 1)
{
  SDF8 <- as.matrix(SDF8)
  n <- nrow(SDF8)
  alpha <- matrix(NA, p, p)
  beta <- matrix(NA, q, p)
  sigma <- matrix(NA, p + q, p + q)
  s11 <- matrix(NA, p, p)
  s22 <- matrix(NA, q, q)
  s12 <- matrix(NA, p, q)
  s <- matrix(NA, p, p)
  a <- matrix(NA, p, p)
  rho <- NULL
  aux <- NULL
  lambda <- NULL
  #sigma<-cor(SDF8)
  ##### Estimation of the covariance matrix #####
  if(ind == 1) sigma <- var(SDF8) else if(ind == 2) {
    sigma <- cov.mcd(SDF8, print = F, cor = T, quan = 0.75 *
n)$cov
  }
  else if(ind == 3) {
    sigma <- mesthub(SDF8, 0.05)$cov
  }
  ##### Estimation of the parameters #####
  s11 <- sigma[1:p, 1:p]
  s12 <- sigma[1:p, (p + 1):(p + q)]
  s22 <- sigma[(p + 1):(p + q), (p + 1):(p + q)]
  s <- solve(s11) %*% s12 %*% solve(s22) %*% t(s12)
  eig <- eigen(s)
  lambda <- eig$values
  rho <- sqrt(lambda)
  a <- eig$vectors
}

```


Lampiran 9 (lanjutan)

```

s <- t(a) %*% s11 %*% a
aux <- sqrt(diag(s))
s <- diag(1/aux)
alpha <- a %*% s
s <- diag(1/rho)
beta <- solve(s22) %*% t(s12) %*% alpha %*% s
list(s = s, au = aux, e = eig, a = alpha, b = beta, r = rho)
}
> cc(2, 3, SDF8, ind = 2)
$s:
      [,1]      [,2]
[1,] 1.426669 0.000000
[2,] 0.000000 4.165038

$au:
[1] 131.14466  74.97062

$e:
$e$values:
[1] 0.49130737 0.05764504

$e$vectors:
      v2      v1
V1  4.033941 -15.98108406
V2 -1.013454  0.04861167

$a:
      [,1]      [,2]
V1  0.030759473 -0.2131646343
V2 -0.007727758  0.0006484096

$b:
      [,1]      [,2]
V3  0.0001499064  0.0005659679
V4  0.0017140537 -0.0004748210
V5 -0.0035987870  0.0010451878

$r:
[1] 0.7009332 0.2400938

```

Lampiran 10 Analisis Korelasi Kanonik *Robust* Data III

```

#####
FILE: cc.ssc #
# #
# DESCRIPTION: Canonical correlation, based on a covariance #
# matrix #
# #
# VARIABLES: p<-dimension of the first group #
# q<-dimension of the second group #
# SDF1<-matrix of the SDF1 #
# ind<-indicator of the type of estimator used to #
# obtain the inicial estimation of the cova- #
# riance matrix #
# #
# OBSERVATIONS: ind==1<-Classical estimator #
# ind==2<-MCD estimator #
# ind==3<-M-estimator #
# #
# LAST MODIFICATION: 30 January 2002 #
#####
cc <- function(p, q, SDF9, ind = 1)
{
  SDF9 <- as.matrix(SDF9)
  n <- nrow(SDF9)
  alpha <- matrix(NA, p, p)
  beta <- matrix(NA, q, p)
  sigma <- matrix(NA, p + q, p + q)
  s11 <- matrix(NA, p, p)
  s22 <- matrix(NA, q, q)
  s12 <- matrix(NA, p, q)
  s <- matrix(NA, p, p)
  a <- matrix(NA, p, p)
  rho <- NULL
  aux <- NULL
  lambda <- NULL
  #sigma<-cor(SDF9)
  ##### Estimation of the covariance matrix #####
  if(ind == 1) sigma <- var(SDF9) else if(ind == 2) {
    sigma <- cov.mcd(SDF9, print = F, cor = T, quan = 0.75 *
n)$cov
  }
  else if(ind == 3) {
    sigma <- mesthub(SDF9, 0.05)$cov
  }
  ##### Estimation of the parameters #####
  s11 <- sigma[1:p, 1:p]
  s12 <- sigma[1:p, (p + 1):(p + q)]
  s22 <- sigma[(p + 1):(p + q), (p + 1):(p + q)]
  s <- solve(s11) %*% s12 %*% solve(s22) %*% t(s12)
  eig <- eigen(s)
  lambda <- eig$values
  rho <- sqrt(lambda)
  a <- eig$vectors
}

```

Lampiran 10 (lanjutan)

```

s <- t(a) %% s11 %% a
aux <- sqrt(diag(s))
s <- diag(1/aux)
alpha <- a %% s
s <- diag(1/rho)
beta <- solve(s22) %% t(s12) %% alpha %% s
list(s = s, au = aux, e = eig, a = alpha, b = beta, r = rho)
}
> cc(2, 2, SDF9, ind = 2)
$s:
      [,1]      [,2]
[1,] 1.397844 0.000000
[2,] 0.000000 6.787216

$au:
[1] 12820.9766   384.8827

$e:
$e$values:
[1] 0.51177928 0.02170784

$e$vectors:
      V2      V1
V1 0.01301193 -0.06122509
V2 -1.06353061 0.20095050

$a:
      [,1]      [,2]
V1 1.014894e-006 -0.0001590747
V2 -8.295239e-005 0.0005221085

$b:
      [,1]      [,2]
V3 0.0008395973 -0.0008058918
V4 -0.0004372576 -0.0004939500

$r:
[1] 0.7153875 0.1473358

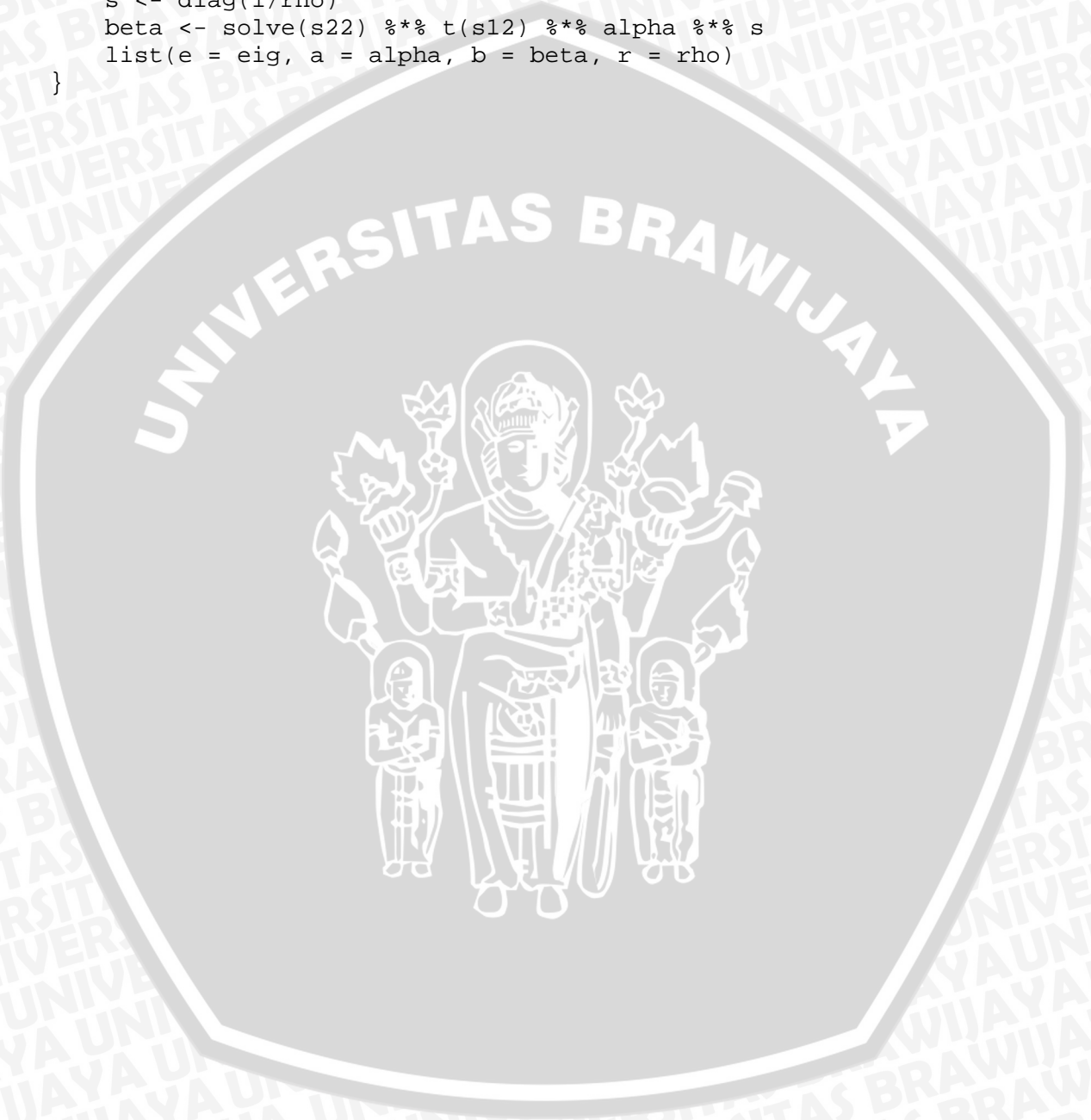
```


Lampiran 11 Macro SPLUS untuk Analisis Korelasi Kanonik *Robust* menggunakan metode MCD

```
#####
FILE: cc.ssc #
# #
# DESCRIPTION: Canonical correlation, based on a covariance #
# matrix #
# #
# VARIABLES: p<-dimension of the first group #
# q<-dimension of the second group #
# SDF1<-matrix of the SDF1 #
# ind<-indicator of the type of estimator used to #
# obtain the inicial estimation of the cova- #
# riance matrix #
# #
# OBSERVATIONS: ind==1<-Classical estimator #
# ind==2<-MCD estimator #
# ind==3<-M-estimator #
# #
# LAST MODIFICATION: 30 January 2002 #
#####
cc <- function(p, q, SDF2, ind = 1)
{
  SDF2 <- as.matrix(SDF2)
  n <- nrow(SDF2)
  alpha <- matrix(NA, p, p)
  beta <- matrix(NA, q, p)
  sigma <- matrix(NA, p + q, p + q)
  s11 <- matrix(NA, p, p)
  s22 <- matrix(NA, q, q)
  s12 <- matrix(NA, p, q)
  s <- matrix(NA, p, p)
  a <- matrix(NA, p, p)
  rho <- NULL
  aux <- NULL
  lambda <- NULL
  #sigma<-cor(SDF2)
  ##### Estimation of the covariance matrix #####
  if(ind == 1) sigma <- var(SDF2) else if(ind == 2) {
    sigma <- cov.mcd(SDF2, print = F, cor = T, quan = 0.75 *
n)$cov
  }
  else if(ind == 3) {
    sigma <- mesthub(SDF2, 0.05)$cov
  }
  ##### Estimation of the parameters #####
  s11 <- sigma[1:p, 1:p]
  s12 <- sigma[1:p, (p + 1):(p + q)]
  s22 <- sigma[(p + 1):(p + q), (p + 1):(p + q)]
  s <- solve(s11) %*% s12 %*% solve(s22) %*% t(s12)
  eig <- eigen(s)
  lambda <- eig$values
  rho <- sqrt(lambda)
}
```

Lampiran 11 (lanjutan)

```
a <- eig$vectors
s <- t(a) %**% s11 %**% a
aux <- sqrt(diag(s))
s <- diag(1/aux)
alpha <- a %**% s
s <- diag(1/rho)
beta <- solve(s22) %**% t(s12) %**% alpha %**% s
list(e = eig, a = alpha, b = beta, r = rho)
}
```



Lampiran 12 Data I yang Distandarisasi

Lokasi	X1	x2	x3	y1	y2	y3
1	-0.48049	0.06382	0.30835	0.54079	1.42789	0.97472
2	-2.50248	-0.05556	-0.61067	1.01526	0.29879	0.65593
3	-0.56083	1.63514	0.30835	-0.29595	0.27823	0.67586
4	0.04844	-0.23991	-1.13526	-1.55693	-0.35432	0.13791
5	0.8251	0.01466	0.17819	0.48415	1.31086	0.63601
6	0.04175	0.43251	-0.61067	-0.35686	0.17228	0.56295
7	-1.59191	1.45431	-0.21624	-0.16771	-0.07442	-0.19416
8	0.3966	0.20428	-0.21624	0.12402	0.82064	0.69578
9	-1.17011	0.25431	-1.0051	1.65644	1.64295	0.25081
10	-0.0386	1.92833	0.70278	2.0839	0.64669	0.31058
11	1.113	-1.39601	-0.08608	-0.8869	-1.55616	-2.45887
12	0.85858	-1.29681	-0.21624	-1.06536	-1.29207	-2.0936
13	0.30286	-0.22849	2.15034	0.53438	-1.09282	0.2176
14	0.85858	-0.4102	1.75591	-0.22435	-0.95999	-1.30991
15	0.85858	-0.7903	-1.79396	-1.44687	-1.06594	-0.07462
16	0.94562	-1.62951	0.96705	-0.55563	-0.71804	0.38364
17	0.09531	0.05943	-0.48051	0.11761	0.51543	0.62937

Deteksi Pencilan Data I

Comparison of Mahalanobis Distances:

	1	2	3	4	5	6	7
Robust	5.3481	70.2228	5.9482	1.9565	4.6884	0.4598	5.9049
Classical	4.8002	10.1479	5.4731	3.7888	4.6817	1.3568	6.6379
	8	9	10	11	12	13	14
Robust	0.9675	68.4945	5.4592	3.4883	2.4815	291.4141	65.2030
Classical	1.6559	6.9662	10.7562	6.3966	4.8141	7.8056	4.9722
	15	16	17				
Robust	4.3098	179.8874	1.2263				
Classical	8.6372	6.1791	0.9303				

Lampiran 13 Data II yang Distandarisasi

No.	X1	X2	X3	Y1	Y2
1	0.37449	0.93825	0.03326	-0.55787	-1.22474
2	0.88273	0.73335	0.98118	0.44208	0
3	0.37449	0.77433	0.50722	2.44198	-1.22474
4	0.37449	0.77433	-0.91466	-0.22455	-1.22474
5	0.37449	0.52844	1.13917	-0.89118	0
6	0.37449	0.77433	0.82319	-0.22455	-0.61237
7	-0.64199	-1.52061	-0.91466	-0.89118	1.22474
8	1.89921	0.69236	1.29716	0.24209	-2.44949
9	-1.15023	-1.27473	-0.91466	-1.35782	1.22474
10	0.37449	0.77433	-0.91466	-0.75786	0.61237
11	-0.13375	0.52844	1.29716	1.10871	0.61237
12	-1.65847	0.77433	0.98118	1.10871	0.61237
13	0.37449	-1.1108	-0.91466	0.44208	1.22474
14	0.37449	-1.1108	-1.07265	-0.55787	0.61237
15	0.37449	-1.19277	-1.23063	-0.22455	0
16	0.37449	-1.35669	0.82319	-0.75786	0
17	-2.67495	0.93825	0.98118	-0.55787	0.61237
18	0.37449	-1.27473	-1.07265	-0.55787	0.61237
19	-0.64199	0.6104	-0.91466	1.77535	-0.61237

Deteksi Pencilan Data II

Comparison of Mahalanobis Distances:

	1	2	3	4	5	6	7
Robust	4.6871	3.3776	1.855	101.957	4.277	1.5434	4.3515
Classical	5.9726	2.9176	2.3076	5.3424	4.0972	0.9953	5.1761
	8	9	10	11	12	13	14
Robust	5.3485	3.4699	108.3116	18.9754	19.5822	1.0164	3.9061
Classical	5.7236	3.334	10.2936	2.4053	7.1553	1.7031	3.9302
	15	16	17	18	19		
Robust	2.9500	93.6653	33.1546	2.9893	34.6249		
Classical	3.5463	6.0387	9.9017	2.3749	6.7840		

Lampiran 14 Data III yang Distandarisasi

No.	X1	X2	Y1	Y2
1	0.3645	0.62129	0.58605	-0.62848
2	0.3645	0.62129	0.58605	-0.66427
3	0.3645	0.62129	0.58605	-0.62848
4	-1.73136	-1.62905	-1.70361	-0.7921
5	-1.73136	-1.62905	-1.70361	-0.7921
6	0.3645	0.0587	0.58605	1.4168
7	-1.73136	-1.62905	-1.70361	-0.66427
8	0.3645	0.62129	0.12148	1.47816
9	0.6639	0.62129	0.58605	1.46793
10	0.6639	0.0587	0.58605	1.46793
11	-1.58165	-1.62905	-1.47133	-0.66427
12	-1.13254	-1.06647	-1.33859	-0.7921
13
14	0.3645	0.0587	0.58605	-0.62848
15	0.6639	0.0587	0.58605	-0.62848
16	0.6639	0.62129	0.58605	1.4168
17	-0.83313	-0.50388	-0.64838	-0.61825
18	1.41242	1.52143	1.409	1.63155
19	1.56213	1.74646	1.409	-0.36259

Deteksi Pencilan Data III

Comparison of Mahalanobis Distances:

	1	2	3	4	5	6	7	8
Robust	3.284	10.2622	1.4447	3.356	3.841	11.6744	2.999	4.0988
Classical	4.932	5.599	1.212	4.1015	3.5924	2.6016	4.072	4.1731
	9	10	11	12	13	14	15	16
Robust	3.332	32.464	11.640	2.248	7.909	2.002	45.633	1.0256
Classical	0.6833	6.599	4.488	3.209	2.115	0.625	8.076	1.5959
	17	18	19	20	21	22	23	
Robust	2.264	28.52	44.55	2.843	14.315	1.316	2.59	2.5925
Classical	1.513	6.350	9.878	0.746	5.135	2.436	4.265	

Lampiran 15

Macro minitab untuk Q-Q plot

```

macro
qq x.1-x.p
mconstant i n p t chis
mcolumn d x.1-x.p dd pi q ss tt
mmatrix s sinv ma mb mc md
let n=count(x.1)
cova x.1-x.p s
invert s sinv
do i=1:p
  let x.i=x.i-rata-rata(x.i)
enddo
do i=1:n
  copy x.1-x.p ma;
  use i.
  transpose ma mb
  multiply ma sinv mc
  multiply mc mb md
  copy md tt
  let t=tt(1)
  let d(i)=t
enddo
set pi
  1:n
end
let pi=(pi-0.5)/n
sort d dd
invcdf pi q;
chis p.
plot q*dd
regress.
invcdf 0.5 chis;
chis p.
let ss=dd<chis
let t=sum(ss)/n
print t
if t>0.5
  note distribusi data multinormal
endif
if t<=0.5
  note distribusi data bukan multinormal
endif
endmacro

```

