

**ANALISIS KOMPONEN UTAMA *ROBUST* DENGAN METODE  
PENDUGAAN *REWEIGHTED MINIMUM COVARIANCE*  
*DETERMINANT***

**SKRIPSI**

oleh:

**NENSI PRADENI**

**0810950053-95**



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

**2013**

**ANALISIS KOMPONEN UTAMA *ROBUST* DENGAN METODE  
PENDUGAAN *REWEIGHTED MINIMUM COVARIANCE*  
*DETERMINANT***

**SKRIPSI**

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

**Oleh:  
NENSI PRADENI  
0810950053-95**



**PROGRAM STUDI STATISTIKA  
JURUSAN MATEMATIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS BRAWIJAYA  
MALANG  
2013**

**LEMBAR PENGESAHAN SKRIPSI**

**ANALISIS KOMPONEN UTAMA *ROBUST* DENGAN METODE  
PENDUGAAN *REWEIGHTED MINIMUM COVARIANCE*  
*DETERMINANT***

oleh:  
**NENSI PRADENI**  
**0810950053-95**

Setelah dipertahankan di depan Majelis Penguji  
pada tanggal 21 Januari 2013  
dan dinyatakan memenuhi syarat untuk memperoleh gelar  
Sarjana Sains dalam bidang Statistika

**Dosen Pembimbing I**

**Dosen Pembimbing II**

**Dr. Ir. M. Bernadetha Mitakda**  
**NIP. 19520521 198103 2 001**

**Prof. Dr. Ir. Loekito Adi S., M.Agr**  
**NIP. 194703271 97412 1 001**

**Mengetahui,**  
**Ketua Jurusan Matematika**  
**Fakultas MIPA**  
**Universitas Brawijaya**

**Dr. Abdul Rouf Alghofari, M.Sc**  
**NIP. 196709071 99203 1 001**

## LEMBAR PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : NENSI PRADENI

NIM : 0810950053-95

Jurusan : Matematika

Program Studi : STATISTIKA

Skripsi berjudul :

**ANALISIS KOMPONEN UTAMA *ROBUST* DENGAN METODE  
PENDUGAAN *REWEIGHTED MINIMUM COVARIANCE*  
*DETERMINANT***

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.

Demikian pernyataan ini saya buat dengan segala kesadaran.

Malang, 21 Januari 2013

Yang menyatakan,

(NENSI PRADENI)

NIM. 0810950053-95

# ANALISIS KOMPONEN UTAMA *ROBUST* DENGAN METODE PENDUGAAN *REWEIGHTED MINIMUM COVARIANCE DETERMINANT* (RMCD)

## ABSTRAK

*Classic Principal Component Analysis* (CPCA) adalah suatu metode pendekatan klasik untuk menjelaskan struktur matriks ragam-peragam melalui sejumlah kecil komponen yang tidak saling berkorelasi, dan merupakan kombinasi linier dari variabel asal sehingga memiliki ragam maksimum. Data sering mengandung beberapa pencilan sehingga penduga klasik kurang diandalkan. Pada penelitian ini penulis menganalisis data multivariat yang mengandung pencilan dengan *Robust Principal Component Analysis* (ROBPCA). Salah satu metode *robust* yang memiliki kemampuan mengukur jarak dan mendeteksi pencilan adalah *Minimum Covariance Determinant* (MCD). Metode ini menduga rata-rata dan ragam-peragam dengan mencari himpunan bagian data yang menghasilkan determinan matriks ragam-peragam terkecil. Untuk meningkatkan efisiensi dari MCD digunakan *Reweighted Minimum Covariance Determinant* (RMCD), yaitu metode MCD yang ditambahi bobot  $w_j$  berdasarkan jarak *robust* pengamatan. Tujuan penelitian ini adalah menentukan metode yang lebih baik (CPCA atau ROBPCA). Penulis menggunakan data Survei Sosial Ekonomi Nasional 2010-2011 Provinsi Jawa Timur, karena Jatim terdiri dari 29 kabupaten dan 9 kota sehingga ditentukan 38 pengamatan. Hasil analisis menunjukkan, untuk data yang mengandung pencilan, metode ROBPCA lebih baik dibanding CPCA karena menghasilkan komponen utama terpilih lebih sedikit menurut kriteria akar ciri  $\sqrt{1}$  dan dengan jumlah komponen terpilih yang sama banyak ROBPCA mampu menjelaskan keragaman kumulatif yang lebih besar dibanding CPCA.

Kata kunci : CPCA, ROBPCA, MCD, RMCD, *robust*.

# **ROBUST PRINCIPAL COMPONENT ANALYSIS USING REWEIGHTED MINIMUM COVARIANCE DETERMINANT (RMCD) ESTIMATOR**

## **ABSTRACT**

Classic Principal Component Analysis (CPCA) is a classical approach to describe the structure of a variety of matrix variance-covariance through a small number of components that are not correlated with each other, and is a linear combination of the original variables that has the maximum range. The data often contain some outliers so that classical estimators are less reliable. In this study the authors analyze multivariate data containing outliers with Robust Principal Component Analysis (ROBPCA). One robust method that has ability to measure distance and detect the outliers are the Minimum Covariance determinant (MCD). This method assumed average and range variance-covariance to find subsets of data that make up the determinant matrix variance-covariance smallest range. To improve the efficiency of the MCD used Reweighted Minimum Covariance determinant (RMCD), the MCD method which added weight ( $w_i$ ) robust distance based observations. The purpose of this study was to determine better methods (CPCA or ROBPCA). The authors used data 2010-2011 National Socio-Economic Survey of East Java, because Java consists of 29 counties and 9 cities so determined 38 observations. The analysis showed, for data containing outliers, the method is better than the CPCA ROBPCA because it produces fewer major components selected according to the criteria of the roots of characteristic 1 and with the same number of selected components can explain the diversity ROBPCA much larger than the cumulative CPCA.

Keywords : CPCA, ROBPCA, MCD, RMCD, *robust*.

## KATA PENGANTAR

Puji syukur atas segala limpahan rahmat dan karunia Allah SWT, hanya dengan ijin-Nya penyusunan tugas akhir ini dapat selesai dengan baik. Sholawat serta salam semoga tetap tercurah pada junjungan kita, Nabi Muhammad SAW serta pengikutnya.

Skripsi berjudul "ANALISIS KOMPONEN UTAMA *ROBUST* DENGAN METODE PENDUGAAN *REWEIGHTED MINIMUM COVARIANCE DETERMINANT* " disusun sebagai salah satu syarat mendapatkan gelar Sarjana Sains dalam bidang Statistika di Fakultas MIPA universitas Brawijaya.

Penyusunan skripsi ini tak lepas dari bantuan yang diberikan oleh beberapa pihak, maka dari itu penulis banyak mengucapkan terima kasih kepada:

1. Ibu Dr. Ir. Maria Bernadetha M. selaku dosen pembimbing I yang telah memberikan motivasi, bimbingan, pengarahan, nasehat, dan kesabaran selama proses pembimbingan sehingga skripsi ini terselesaikan dengan baik.
2. Bapak Prof. Dr. Ir. Loekito Adi S. selaku dosen pembimbing II yang telah memberikan bimbingan dan nasehat hingga skripsi ini terselesaikan dengan baik.
3. Ibu Eni Sumarminingsih, S.Si., MM, selaku dosen penguji yang telah memberikan saran dan masukan hingga skripsi ini terselesaikan dengan baik.
4. Bapak Dr. Abdul Rouf Alghofari, M.Sc. selaku Ketua Jurusan Matematika Fakultas MIPA Universitas Brawijaya Malang.
5. Keluargaku terutama Mami, Ibu, Emak, Pakde, Kakak, Adek, Bapak yang telah memberikan kasih sayang, doa, kepercayaan dan dorongan. Semoga Allah memberikan kesempatan kepada ananda untuk mewujudkan kebahagiaan dan membalas semua kasih yang tercurah.
6. Sahabatku tersayang, teman seperjuangan (Amri, Virgi, Cira, Zee, Dita, Septi, Kartika, Arifa, Acix, Tetta ) yang senantiasa selalu

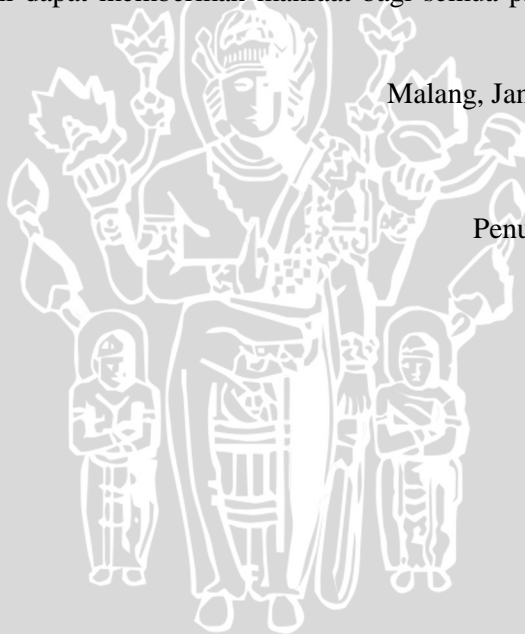
memberi dukungan, motivasi, canda dan tawa, serta indahnya kebersamaan dalam menjalani kuliah di sini.

7. Teman-teman Statistika dari berbagai angkatan, terutama angkatan 2008.
8. Semua pihak yang telah membantu menyelesaikan penyusunan skripsi ini yang tidak dapat disebutkan satu per satu. Terima kasih atas bantuan yang telah diberikan dan semoga Allah memberikan balasan dengan kebaikan.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih banyak kekurangan, oleh karena itu saran dan kritik yang membangun demi kesempurnaan penulisan selanjutnya sangat penulis harapkan. Semoga skripsi ini dapat memberikan manfaat bagi semua pihak yang membutuhkan.

Malang, Januari 2013

Penulis





## DAFTAR ISI

	Halaman
<b>HALAMAN JUDUL</b> .....	i
<b>HALAMAN PENGESAHAN</b> .....	ii
<b>HALAMAN PERNYATAAN</b> .....	iii
<b>ABSTRAK/ ABSTRACT</b> .....	iv
<b>KATA PENGANTAR</b> .....	vi
<b>DAFTAR ISI</b> .....	viii
<b>DAFTAR GAMBAR</b> .....	x
<b>DAFTAR TABEL</b> .....	xi
<b>DAFTAR LAMPIRAN</b> .....	xii
<b>BAB I. PENDAHULUAN</b>	
1.1. Latar Belakang .....	1
1.2. Rumusan Masalah .....	3
1.3. Batasan Masalah.....	3
1.4. Tujuan Penelitian.....	3
1.5. Manfaat Penelitian.....	3
<b>BAB II. TINJAUAN PUSTAKA</b>	
2.1. Deteksi Pencilan Multivariat .....	4
2.2. Analisis Komponen Utama .....	7
2.2.1. Tipe Matriks Masukan.....	7
2.2.1.1. Matriks Ragam-peragam .....	7
2.2.1.2. Matriks Korelasi .....	9
2.2.2. Vektor Ciri Komponen Utama .....	11
2.2.3. Penentuan Banyaknya Komponen Utama .....	12
2.2.4. Skor Komponen Utama .....	12
2.3. Analisis Komponen Utama Robust.....	13
2.4. <i>Reweighted Minimum Covariance Determinant</i> .....	13
2.5. Indikator Pembanding.....	19
<b>BAB III. METODE PENELITIAN</b>	
3.1. Sumber Data.....	20
3.2. Metode Penelitian.....	22

## **BAB IV. HASIL DAN PEMBAHASAN**

4.1.	Identifikasi Pencilan.....	24
4.2.	Analisis Komponen Utama .....	25
4.2.1.	Matriks Masukan.....	25
4.2.2.	Akar Ciri dan Vektor Ciri.....	26
4.2.3.	Penentuan Banyaknya Komponen Utama.....	26
4.2.4.	Skor Komponen Utama.....	28
4.3.	Metode RMCD.....	28
4.4.	Jarak <i>Robust</i> .....	30
4.5.	Analisis Komponen Utama <i>Robust</i> .....	33
4.5.1.	Akar Ciri dan Vektor Ciri.....	33
4.5.2.	Penentuan Komponen Utama ROBPCA.....	34
4.2.1.	Skor Komponen Utama.....	36
4.6.	Penentuan Metode yang Lebih Baik .....	37

## **BAB V. PENUTUP**

5.1.	Kesimpulan .....	40
5.2.	Saran.....	40

<b>DAFTAR PUSTAKA</b> .....	41
-----------------------------	----

<b>LAMPIRAN</b> .....	42
-----------------------	----

## DAFTAR GAMBAR

	Halaman
Gambar 3.1 Diagram Alir Penelitian.....	23
Gambar 4.1 Plot Jarak <i>Mahalanobis</i> .....	25
Gambar 4.2 Plot Komponen Utama terhadap Akar Ciri.....	27
Gambar 4.3 Plot SK1 terhadap SK2 (CPCA).....	28
Gambar 4.4 Plot Jarak <i>Robust</i> .....	32
Gambar 4.5 Plot Komponen Utama terhadap Akar Ciri.....	36
Gambar 4.6 Plot SK1 terhadap SK2 (ROBPCA) .....	37



## DAFTAR TABEL

	Halaman
Tabel 2.1 Tabel $n$ Pengamatan terhadap $p$ Peubah .....	5
Tabel 2.2 Matriks $S$ .....	5
Tabel 2.3 Matriks $S$ .....	6
Tabel 2.4 Tabel $h$ pengamatan Contoh .....	14
Tabel 2.5 Unsur-unsur matriks $S_1$ .....	15
Tabel 2.6 Matriks Ragam-peragam .....	15
Tabel 3.1 Keterangan Tentang Data .....	20
Tabel 4.1 Identifikasi Pencilan .....	24
Tabel 4.2 Banyaknya Pencilan .....	25
Tabel 4.3 Akar Ciri, Proporsi Keragaman dan Proporsi Keragaman Kumulatif CPCA .....	26
Tabel 4.4 Jumlah Komponen Utama Terpilih CPCA .....	26
Tabel 4.5 Persen Keragaman Kumulatif Komponen Utama Terpilih .....	27
Tabel 4.6 Jumlah Komponen Utama Terpilih .....	28
Tabel 4.7 Pengamatan yang Terpilih .....	29
Tabel 4.8 Pengamatan dengan Determinan Matriks .....	30
Ragam-peragam Terkecil .....	29
Tabel 4.9 Jarak <i>Robust</i> .....	30
Tabel 4.10 Pembobot .....	31
Tabel 4.11 Banyaknya Pencilan .....	32
Tabel 4.12 Akar Ciri, Proporsi Keragaman dan Proporsi Keragaman Kumulatif ROBPCA .....	33
Tabel 4.13 Vektor Ciri .....	33
Tabel 4.14 Jumlah Komponen Utama Terpilih ROBPCA .....	35
Tabel 4.15 Persen Keragaman Kumulatif .....	35
Tabel 4.16 Jumlah Komponen Utama Terpilih .....	36
Tabel 4.17 Banyaknya Komponen Utama Terpilih .....	37
Tabel 4.18 Persen Keragaman Kumulatif .....	38
Tabel 4.19 Banyaknya Pencilan Teridentifikasi .....	39
Tabel 4.20 Jumlah Komponen Utama Terpilih .....	39

## DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Data Survei Sosial Ekonomi Nasional 2010-2011 Provinsi Jawa Timur tentang Rata-rata Pengeluaran Makanan/kapita dalam Juta Rupiah Menurut Kabupaten atau Kota (Data I). .....	41
Lampiran 2. Data Survei Sosial Ekonomi Nasional 2010-2011 Provinsi Jawa Timur tentang Rata-rata Pengeluaran Non Makanan/kapita dalam Juta Rupiah Menurut Kabupaten atau Kota (Data II).....	43
Lampiran 3. Data Survei Sosial Ekonomi Nasional 2010-2011 Provinsi Jawa Timur tentang Rata-rata Pendapatan Rumah Tangga yang Menerima Kredit Usaha dalam Juta Rupiah Menurut Kabupaten atau (Data III). .....	45
Lampiran 4. Skor Baku Data I. ....	47
Lampiran 5. Skor Baku Data II. ....	48
Lampiran 6. Skor Baku Data III.....	49
Lampiran 7. Jarak Mahalanobis. ....	50
Lampiran 8. Plot Jarak Mahalanobis dan Jarak Robust. ....	52
Lampiran 9. Akar Ciri, Proporsi Keragaman dan Proporsi Keragaman Kumulatif Metode CPCA dan ROBPCA .....	54
Lampiran 10. Scree Plot.....	55
Lampiran 11. Skor Komponen Utama. ....	57
Lampiran 12. <i>Tolerance Ellipse</i> .....	60
Lampiran 13. Pengamatan dengan Determinan Matriks Ragam-peragam Terkecil yang Tersusun sesuai Jarak Mahalanobis .....	62
Lampiran 14. Vektor Rata-rata dan Matriks Ragam-peragam <i>Robust</i> dari Penduga MCD .....	63
Lampiran 15. Jarak <i>Robust</i> .....	65
Lampiran 16. Vektor Rata-rata dan Matriks Ragam-peragam <i>Robust</i> dari Penduga RMCD .....	67
Lampiran 17. Vektor Ciri dengan Penduga RMCD.....	68

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

*Principal Component Analysis* (PCA) atau analisis komponen utama merupakan salah satu analisis multivariat yang digunakan untuk mereduksi dimensi data dari yang berukuran besar dan saling berkorelasi menjadi dimensi yang lebih kecil dan tidak saling berkorelasi. Namun walaupun dimensi data menjadi lebih kecil, tidak akan banyak informasi yang hilang karena keragaman tetap dipertahankan minimum 80% (Johnson dan Wichern, 2002).

Analisis komponen utama dilakukan untuk memperoleh komponen-komponen utama yang dapat menjelaskan sebagian besar keragaman data. Komponen pertama memiliki persentase terbesar dibandingkan komponen-komponen lain dan akan menjelaskan keragaman terbesar data, komponen utama kedua menjelaskan keragaman terbesar kedua dan seterusnya (Garmezy, 1972).

Perkembangan analisis komponen utama diperkenalkan pertama kali oleh Pearson pada tahun 1901. Sejalan dengan perkembangan teknologi komputer dan kemajuan di bidang matematika, hingga kini analisis komponen utama masih terus mengalami perkembangan. Perkembangan analisis komponen utama selanjutnya dipengaruhi oleh kebutuhan model analisis komponen utama yang *robust* (kekar/tangguh) terhadap data yang mengandung pencilan. Suryana (2007), menyatakan bahwa analisis komponen utama klasik sangat rentan terhadap kehadiran pencilan karena didasarkan pada matriks ragam-peragam yang juga sangat sensitif terhadap keberadaan pencilan.

Pendeteksian pencilan merupakan langkah penting sebelum analisis data, karena pencilan dapat mempengaruhi keseimbangan data. Menurut Sujatmiko (2005), untuk mengatasi pencilan data multivariat digunakan metode penduga yang tangguh terhadap pencilan sehingga analisis komponen utama *robust* merupakan metode komponen utama yang tidak lagi dipengaruhi pencilan.

Sujatmiko (2005), menyatakan bahwa penduga *robust* pada analisis komponen utama adalah dengan menghasilkan penduga *robust*

pada matriks korelasi atau ragam-peragam yang digunakan pada analisis komponen utama. Setelah itu, analisis komponen utama klasik dapat dilakukan.

Salah satu metode *robust* yang memiliki kemampuan mengukur jarak sekaligus mendeteksi pencilan adalah *Minimum Covariance Determinant* (MCD) yang diperkenalkan oleh Rousseeuw pada tahun 1984. MCD adalah suatu metode penduga rata-rata dan ragam-peragam dengan mencari himpunan bagian dari data yang menghasilkan determinan matriks ragam-peragam terkecil (Hubert et al., 2003).

Untuk meningkatkan efisiensi dari MCD digunakan *Reweighted Minimum Covariance Determinant* (RMCD), yaitu metode MCD yang ditambahi bobot  $w_j$  berdasarkan jarak *robust* pengamatan. RMCD mewarisi sifat baik MCD seperti *breakdown point* yang tinggi untuk mengganti matriks ragam-peragam klasik (Hubert et al., 2003).

Pembangunan disegala bidang telah dilaksanakan oleh pemerintah secara terus menerus sebagai upaya meningkatkan kesejahteraan masyarakat. Pelaksanaan program pemerintah tersebut disesuaikan dengan potensi yang ada pada masing-masing daerah. Untuk itulah diperlukan adanya data yang mencerminkan keadaan masing-masing daerah tersebut. Survei Sosial Ekonomi Nasional (SUSENAS) merupakan salah satu survei yang dilaksanakan setiap tahun. Dirancang untuk dapat memenuhi kebutuhan data yang berkaitan dengan kualitas sumber daya manusia, khususnya yang berhubungan dengan karakteristik sosial ekonomi. Survei Sosial Ekonomi Nasional 2010-2011 Provinsi Jawa Timur merupakan data multivariat yang mengandung lebih dari satu peubah. Sebanyak 38 pengamatan untuk data I, II dan III diambil secara acak karena Provinsi Jawa Timur terdiri dari 29 kabupaten dan 9 kota, dengan jumlah contoh sebanyak 29.960 rumah tangga pada masing-masing kabupaten/kota maka data SUSENAS dapat menghasilkan informasi tentang perkembangan kesejahteraan masyarakat (Susenas, 2011).

Analisis Komponen Utama dapat diterapkan dalam studi kasus untuk mendapatkan dan mengetahui peubah yang memberikan sokongan keragaman terbesar dalam data. Karena analisis data

multivariat rentan akan keberadaan pencilan sehingga analisis komponen utama klasik kurang diandalkan. Untuk mengatasi hal ini digunakanlah metode yang tangguh yaitu Analisis Komponen Utama *Robust*. Pemilihan metode mana yang lebih baik antara Analisis Komponen Utama Klasik dengan Analisis Komponen Utama *Robust* didasarkan pada komponen utama terbentuk dan proporsi keragaman yang dihasilkan oleh kedua metode.

### **1.2. Rumusan Masalah**

Rumusan masalah yang akan dikaji adalah bagaimana perubahan banyaknya komponen utama yang terbentuk jika menggunakan metode *Clasic Principal Component Analysis (CPCA)* dan *Robust Principal Component Analysis (ROBPCA)*?

### **1.3. Batasan Masalah**

Penulis hanya mempelajari cara mendapatkan komponen utama dengan metode ROBPCA menggunakan penduga RMCD. Data yang digunakan pada penelitian ini adalah data multivariat yang mengandung pencilan.

### **1.4. Tujuan Penelitian**

Berdasarkan rumusan masalah yang telah diuraikan, penelitian ini bertujuan untuk mengetahui perubahan banyaknya komponen utama yang terbentuk jika menggunakan metode CPCA dan ROBPCA.

### **1.5. Manfaat Penelitian**

1. Dapat memperkaya ilmu pengetahuan dan memahami pengembangan konsep statistika multivariat, khususnya analisis komponen utama.
2. Dapat menerapkan dan mengembangkan metode *robust* pada data yang mengandung pencilan baik pada analisis komponen utama maupun analisis multivariat lain.



## BAB II

### TINJAUAN PUSTAKA

#### 2.1. Deteksi Pencilan Multivariat

Pencilan didefinisikan sebagai satu atau lebih pengamatan yang menyimpang cukup jauh dari pengamatan lain sehingga menimbulkan kecurigaan bahwa pengamatan-pengamatan tersebut berasal dari sebaran data berbeda (Draper and Smith, 1981).

Setyawan (2007), berpendapat bahwa pendeteksian pencilan sangat penting dilakukan dalam prosedur statistika karena pencilan mampu mempengaruhi informasi data. Pencilan yang terdeteksi dari suatu contoh mungkin saja dihilangkan agar dapat dilakukan analisis lebih lanjut. Tindakan tersebut dapat dilakukan apabila hanya terdapat satu pencilan, namun hal ini tidak mungkin dilakukan jika pencilan yang terdeteksi lebih dari satu, karena mungkin pencilan mengandung informasi penting. Pencilan secara univariat adalah kasus yang muncul pada satu peubah. Sedangkan pencilan multivariat muncul pada lebih dari satu peubah.

Pengidentifikasian pencilan pada multivariat tidaklah mudah untuk dilakukan, dikarenakan efek *masking* (mengaburkan data) dan *swamping* (kesalahan pengidentifikasian data *nonoutlier* sebagai *outlier*). *Masking* terjadi pada saat pencilan tidak terdeteksi karena adanya pencilan lain yang berdekatan, sedangkan *swamping* terjadi pada saat pengamatan baik, teridentifikasi sebagai pencilan. Jarak *Mahalanobis* atau jarak *Robust* adalah cara yang dapat digunakan untuk mendeteksi pencilan multivariat (Retno, 2005).

Identifikasi pencilan data multivariat yang didasarkan pada kuadrat jarak *Mahalanobis* didefinisikan berikut:

(Chen dan Jackson, 2004):

$$MD_j^2 = [\mathbf{x}_{ij} - \bar{\mathbf{x}}_i]' \mathbf{S}^{-1} [\mathbf{x}_{ij} - \bar{\mathbf{x}}_i] \quad (2.1)$$

di mana :

$\mathbf{x}_{ij}$  : vektor pengamatan ke- $j$  pada peubah ke- $i$  ( $px1$ )

$\bar{\mathbf{x}}_i$  : vektor rata-rata peubah ke- $i$  ( $px1$ )

$\mathbf{S}$  : matriks ragam-peragam ( $pxp$ )

$i : 1, 2, \dots, p = p$ : peubah prediktor

$j : 1, 2, \dots, n = n$ : pengamatan

Tabel 2.1. Tabel  $n$  Pengamatan Contoh terhadap  $p$  Peubah

$i$	$j$					Rata-rata
	1	2	3	...	n	
1	$X_{11}$	$X_{12}$	$X_{13}$	...	$X_{1n}$	$\bar{X}_1$
2	$X_{21}$	$X_{22}$	$X_{23}$	...	$X_{2n}$	$\bar{X}_2$
3	$X_{31}$	$X_{32}$	$X_{33}$	...	$X_{3n}$	$\bar{X}_3$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	$\vdots$
$\bar{X}_i$	$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_3$	...	$\bar{X}_n$	$\bar{X}_p$
$\bar{X}_p$	$X_{p1}$	$X_{p2}$	$X_{p3}$	...	$X_{pn}$	$\bar{X}_p$

Dalam bentuk matriks  $X = \begin{bmatrix} X_{11} & X_{12} & X_{13} & \dots & X_{1n} \\ X_{21} & X_{22} & X_{23} & \dots & X_{2n} \\ X_{31} & X_{32} & X_{33} & \dots & X_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ X_{p1} & X_{p2} & X_{p3} & \dots & X_{pn} \end{bmatrix}$ ,  $\bar{X} = \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \bar{X}_3 \\ \vdots \\ \bar{X}_p \end{bmatrix}$

Matriks deviasi  $[x_{ij} - \bar{x}_i]$

$$j = 1, \begin{bmatrix} X_{11} \\ X_{21} \\ X_{31} \\ \vdots \\ X_{p1} \end{bmatrix} - \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \bar{X}_3 \\ \vdots \\ \bar{X}_p \end{bmatrix} = \begin{bmatrix} X_{11} - \bar{X}_1 \\ X_{21} - \bar{X}_2 \\ X_{31} - \bar{X}_3 \\ \vdots \\ X_{p1} - \bar{X}_p \end{bmatrix}; j = n, \begin{bmatrix} X_{1n} \\ X_{2n} \\ X_{3n} \\ \vdots \\ X_{pn} \end{bmatrix} - \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \bar{X}_3 \\ \vdots \\ \bar{X}_p \end{bmatrix} = \begin{bmatrix} X_{2n} - \bar{X}_1 \\ X_{2n} - \bar{X}_2 \\ X_{3n} - \bar{X}_3 \\ \vdots \\ X_{pn} - \bar{X}_p \end{bmatrix}$$

(px1) (px1) (px1) (px1) (px1) (px1)

Tabel 2.2. Matriks  $S$

$X_{1j} - \bar{X}_1$	$\frac{\sum (X_{1j} - \bar{X}_1)^2}{(\sum (X_{1j} - \bar{X}_1)^2)}$	$\frac{\sum (X_{2j} - \bar{X}_2)(X_{1j} - \bar{X}_1)}{(\sum (X_{2j} - \bar{X}_2)^2)}$	$\frac{\sum (X_{3j} - \bar{X}_3)(X_{1j} - \bar{X}_1)}{(\sum (X_{3j} - \bar{X}_3)^2)}$	...	$\frac{\sum (X_{ij} - \bar{X}_i)(X_{1j} - \bar{X}_1)}{(\sum (X_{ij} - \bar{X}_i)^2)}$
$X_{2j} - \bar{X}_2$	...	$\frac{\sum (X_{2j} - \bar{X}_2)^2}{(\sum (X_{2j} - \bar{X}_2)^2)}$	$\frac{\sum (X_{3j} - \bar{X}_3)(X_{2j} - \bar{X}_2)}{(\sum (X_{3j} - \bar{X}_3)^2)}$	...	$\frac{\sum (X_{ij} - \bar{X}_i)(X_{2j} - \bar{X}_2)}{(\sum (X_{ij} - \bar{X}_i)^2)}$
$X_{3j} - \bar{X}_3$	...	...	$\frac{\sum (X_{3j} - \bar{X}_3)^2}{(\sum (X_{3j} - \bar{X}_3)^2)}$	...	$\frac{\sum (X_{ij} - \bar{X}_i)(X_{3j} - \bar{X}_3)}{(\sum (X_{ij} - \bar{X}_i)^2)}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$X_{ij} - \bar{X}_i$	...	...	...	...	$\frac{\sum (X_{ij} - \bar{X}_i)^2}{(\sum (X_{ij} - \bar{X}_i)^2)}$

Kemudian dituliskan menjadi:

$$\begin{bmatrix} \Sigma(X_{1j} - \bar{X}_1)^2 & \Sigma(X_{1j} - \bar{X}_1)(X_{2j} - \bar{X}_2) & \Sigma(X_{1j} - \bar{X}_1)(X_{3j} - \bar{X}_3) & \dots & \Sigma(X_{1j} - \bar{X}_1)(X_{ij} - \bar{X}_i) \\ \dots & \Sigma(X_{2j} - \bar{X}_2)^2 & \Sigma(X_{2j} - \bar{X}_2)(X_{3j} - \bar{X}_3) & \dots & \Sigma(X_{2j} - \bar{X}_2)(X_{ij} - \bar{X}_i) \\ \dots & \dots & \Sigma(X_{3j} - \bar{X}_3)^2 & \dots & \Sigma(X_{3j} - \bar{X}_3)(X_{ij} - \bar{X}_i) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \dots & \dots & \dots & \dots & \Sigma(X_{ij} - \bar{X}_i)^2 \end{bmatrix}$$

Sehingga menghasilkan

Tabel 2.3. Matriks  $S$

	$(\frac{X_{11} - \bar{X}_1}{\dots} \dots \frac{X_{1j} - \bar{X}_1}{\dots})$	$(\frac{X_{21} - \bar{X}_2}{\dots} \dots \frac{X_{2j} - \bar{X}_2}{\dots})$	$(\frac{X_{31} - \bar{X}_3}{\dots} \dots \frac{X_{3j} - \bar{X}_3}{\dots})$	...	$(\frac{X_{i1} - \bar{X}_i}{\dots} \dots \frac{X_{ij} - \bar{X}_i}{\dots})$
$(\frac{X_{11} - \bar{X}_1}{\dots} \dots \frac{X_{1j} - \bar{X}_1}{\dots})$	<b>Var</b> $(\frac{X_{1j} - \bar{X}_1}{\dots})$	<b>Cov</b> $(\frac{X_{1j} - \bar{X}_1}{\dots}, \frac{X_{2j} - \bar{X}_2}{\dots})$	<b>Cov</b> $(\frac{X_{1j} - \bar{X}_1}{\dots}, \frac{X_{3j} - \bar{X}_3}{\dots})$	...	<b>Cov</b> $(\frac{X_{1j} - \bar{X}_1}{\dots}, \frac{X_{ij} - \bar{X}_i}{\dots})$
$(\frac{X_{21} - \bar{X}_2}{\dots} \dots \frac{X_{2j} - \bar{X}_2}{\dots})$	...	<b>Var</b> $(\frac{X_{2j} - \bar{X}_2}{\dots})$	<b>Cov</b> $(\frac{X_{2j} - \bar{X}_2}{\dots}, \frac{X_{3j} - \bar{X}_3}{\dots})$	...	<b>Cov</b> $(\frac{X_{2j} - \bar{X}_2}{\dots}, \frac{X_{ij} - \bar{X}_i}{\dots})$
$(\frac{X_{31} - \bar{X}_3}{\dots} \dots \frac{X_{3j} - \bar{X}_3}{\dots})$	...	...	<b>Var</b> $(\frac{X_{3j} - \bar{X}_3}{\dots})$	...	<b>Cov</b> $(\frac{X_{3j} - \bar{X}_3}{\dots}, \frac{X_{ij} - \bar{X}_i}{\dots})$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$(\frac{X_{i1} - \bar{X}_i}{\dots} \dots \frac{X_{ij} - \bar{X}_i}{\dots})$	...	...	...	...	<b>Var</b> $(\frac{X_{ij} - \bar{X}_i}{\dots})$

Kuadrat jarak Mahalanobis untuk  $j=1$

$$MD_j^2 = [x_{ij} - \bar{x}_i]' S^{-1} [x_{ij} - \bar{x}_i]$$

$$[x_{11} - \bar{x}_1 \quad \dots \quad x_{p1} - \bar{x}_p] \begin{bmatrix} \text{Var}(x_{1j}) & \text{Cov}(x_{1j}, x_{2j}) & \text{Cov}(x_{1j}, x_{3j}) & \dots & \text{Cov}(x_{1j}, x_{ij}) \\ \dots & \text{Var}(x_{2j}) & \text{Cov}(x_{2j}, x_{3j}) & \dots & \text{Cov}(x_{2j}, x_{ij}) \\ \dots & \dots & \text{Var}(x_{3j}) & \dots & \dots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \dots & \dots & \dots & \dots & \text{Var}(x_{ij}) \end{bmatrix}^{-1} \begin{bmatrix} x_{11} - \bar{x}_1 \\ x_{21} - \bar{x}_2 \\ x_{31} - \bar{x}_3 \\ \vdots \\ x_{p1} - \bar{x}_p \end{bmatrix}$$

(1xp) (p x p) (p x 1)

Hipotesis yang melandasi pengujian pencilan:

$$H_0 : x_{ij} = \text{pengamatan biasa}$$

$$H_1 : x_{ij} = \text{pencilan}$$

Jika  $H_0$  benar, statistik uji:

$$MD_j^2 \sim \chi_{(p)}^2$$

$H_0$  akan ditolak jika  $MD_j^2 > \chi_{\alpha/2(p)}^2$ , berarti pengamatan  $x_{ij}$  diidentifikasi sebagai pencilan.

Penggunaan jarak Mahalanobis dalam mendeteksi pencilan dapat dikatakan kurang baik, karena efek *masking* dan *swamping*. Efek *masking* mengurangi jarak Mahalanobis sehingga jarak antar titik terpencil saling berdekatan, sedangkan efek *swamping* menambah jarak Mahalanobis sehingga dimungkinkan terjadinya kesalahan di mana

pengamatan baik diidentifikasi sebagai pencilan (Retno, 2005).

## 2.2. Analisis Komponen Utama

Analisis komponen utama merupakan teknik statistika yang digunakan untuk menjelaskan struktur ragam-peragam sekumpulan peubah melalui beberapa peubah baru yang merupakan kombinasi linier dari peubah asal dan bersifat saling bebas. Peubah baru ini dinamakan komponen utama (Johnson dan Wichern, 2002).

Analisis komponen utama  $p$  peubah prediktor, dalam bentuk vektor adalah:

$\mathbf{X} = [X_1, X_2, X_3, \dots, X_p]$  menjadi vektor

$\mathbf{K} = [K_1, K_2, K_3, \dots, K_q]$

dalam bentuk persamaan sebagai berikut:

(Johnson dan Wichern, 2002):

$$K_k = b_{k1}X_1 + b_{k2}X_2 + b_{k3}X_3 + \dots + b_{kp}X_p$$

di mana:

$$b_{k1}^2 + b_{k2}^2 + b_{k3}^2 + \dots + b_{kp}^2 = 1$$

Ragam ( $K_k$ ) =  $\mathbf{b}'_k \mathbf{S} \mathbf{b}_k$

Peubah-peubah  $[K_1, K_2, K_3, \dots, K_q]$  bebas satu dengan yang lain dan vektor peubah baru  $\mathbf{K} = [K_1, K_2, K_3, \dots, K_q]$  menjelaskan sebesar mungkin proporsi keragaman vektor peubah prediktor  $\mathbf{X} = [X_1, X_2, X_3, \dots, X_p]$  (Johnson dan Wichern, 2002).

### 2.2.1. Tipe Matriks Masukan

Terdapat dua tipe matriks masukan pada PCA yaitu matriks ragam-peragam dan matriks korelasi (Johnson dan Wichern, 2002).

#### 2.2.1.1. Matriks Ragam-peragam

Apabila satuan pengukuran semua peubah yang dianalisis mempunyai satuan pengukuran yang sama maka matriks satuan yang digunakan adalah matriks ragam-peragam. Menurut Johnson dan Wichern (2002:437), jika  $\mathbf{B}$  didefinisikan sebagai matriks konstan berukuran  $p \times q$ , maka komponen utama didefinisikan sebagai kombinasi linier dari  $p$  peubah asal yang dinyatakan dalam bentuk persamaan matriks (Johnson dan Wichern, 2002):

$$\mathbf{K}_{(q,1)} = \mathbf{B}'_{(qp)} \mathbf{X}_{(p,1)}$$

$$\begin{bmatrix} K_1 \\ K_2 \\ \vdots \\ K_q \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{q1} & b_{q2} & \dots & b_{qp} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}$$

$(qx1) \qquad \qquad (qxp) \qquad \qquad (px1)$

di mana:

$k_k$  : komponen utama ke- $k$  peubah asal (PC)

$b_{ki}$  : vektor ciri komponen utama ke- $k$  untuk peubah ke- $i$  ( $pxq$ )

Komponen utama ini tergantung pada matriks ragam-peragam atau matriks korelasi dari peubah asal  $[X_1, X_2, X_3, \dots, X_p]$ . Secara umum kombinasi linier yang dimaksud adalah (Johnson dan Wichern, 2002):

$$K_1 = b_{11}X_1 + b_{12}X_2 + \dots + b_{1p}X_p = \mathbf{b}'_1 \mathbf{X}$$

$$K_2 = b_{21}X_1 + b_{22}X_2 + \dots + b_{2p}X_p = \mathbf{b}'_2 \mathbf{X}$$

$$K_3 = b_{31}X_1 + b_{32}X_2 + \dots + b_{3p}X_p = \mathbf{b}'_3 \mathbf{X}$$

$$\vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots$$

$$K_q = b_{q1}X_1 + b_{q2}X_2 + \dots + b_{qp}X_p = \mathbf{b}'_q \mathbf{X}$$

Berdasarkan kombinasi linier tersebut didapatkan ragam dan peragam masing-masing komponen utama yaitu:

$$\text{Ragam } (K_i) = \mathbf{b}'_i \mathbf{S} \mathbf{b}_i$$

$$\text{Peragam } (k_i, k_k) = \mathbf{b}'_i \mathbf{S} \mathbf{b}_k \qquad i, k = 1, 2, \dots, p \text{ di mana } i \neq k$$

Komponen utama pertama berupa kombinasi linier  $K_1 = \mathbf{b}'_1 \mathbf{X}$  yang bertujuan memaksimumkan ragam ( $K_1$ ) dengan batasan  $\mathbf{b}'_1 \mathbf{b}_1 = 1$  dan akan didapat:

$$\text{Ragam } (K_1) = \mathbf{b}'_1 \mathbf{S} \mathbf{b}_1 = S_{k1}^2 = \lambda_1$$

$$= \begin{bmatrix} b_{11} & b_{12} & b_{13} & \dots & b_{1p} \end{bmatrix} \begin{bmatrix} \text{Var}(x_{1j}) & \text{Cov}(x_{1j}, x_{2j}) & \text{Cov}(x_{1j}, x_{3j}) & \dots & \text{Cov}(x_{1j}, x_{ij}) \\ \dots & \text{Var}(x_{2j}) & \text{Cov}(x_{2j}, x_{3j}) & \dots & \text{Cov}(x_{2j}, x_{ij}) \\ \dots & \dots & \text{Var}(x_{3j}) & \dots & \dots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \dots & \dots & \dots & \dots & \text{Var}(x_{ij}) \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{12} \\ b_{13} \\ \vdots \\ b_{1p} \end{bmatrix}$$

$(1xp) \qquad \qquad \qquad (pxp) \qquad \qquad \qquad (px1)$

di mana  $\lambda_1$  merupakan akar ciri terbesar pertama matriks  $\mathbf{S}$ . Selanjutnya komponen utama pertama ditulis:

$$K_1 = b_{11}X_1 + b_{12}X_2 + \dots + b_{1p}X_p = \mathbf{b}'_1 \mathbf{X}$$

Sedangkan komponen utama kedua berupa kombinasi linier  $K_2 = \mathbf{b}'_2 \mathbf{X}$  yang bertujuan memaksimumkan ragam ( $K_2$ ) yang tidak berkorelasi dengan komponen utama pertama tetapi bersifat ortogonal dengan komponen utama pertama (Johnson dan Wichern, 2002), sehingga harus memenuhi batasan

$$\mathbf{b}'_2 \mathbf{b}_2 = 1 \text{ dan Peragam } (K_1, K_2) = 0 \text{ menghasilkan}$$

$$\text{Ragam } (K_2) = \mathbf{b}'_2 \mathbf{S} \mathbf{b}_2 = S_{k_2}^2 = \lambda_2$$

$\lambda_2$  = akar ciri terbesar kedua dari matriks  $\mathbf{S}$  dan komponen utama kedua adalah:

$$K_2 = b_{21}X_1 + b_{22}X_2 + \dots + b_{2p}X_p = \mathbf{b}'_2 \mathbf{X}$$

Jadi komponen utama ke- $k$  merupakan kombinasi linier  $K_k = \mathbf{b}'_k \mathbf{X}$  yang bertujuan memaksimumkan ragam ( $K_k$ ) dan tidak berkorelasi dengan komponen utama yang lain tetapi bersifat ortogonal dengan komponen utama yang lain, dan harus memenuhi :

$$\mathbf{b}'_k \mathbf{b}_k = 1 \text{ dan Peragam } (k_i, k_k) = 0$$

(Johnson dan Wichern, 2002)

### 2.2.1.2. Matriks Korelasi

Matriks ini digunakan jika satuan pengukuran semua peubah berbeda, kemudian peubah perlu dibakukan sehingga komponen utama ditentukan oleh peubah baku. Johnson dan Wichern (2002), menyarankan pembakuan peubah asal  $X$  ke dalam peubah baku  $Z$  yang

dapat dilakukan melalui  $Z_{ij} = \frac{(x_{ij} - \bar{x}_i)}{\sqrt{S_{ii}}}$

dalam bentuk matriks adalah :

$$\mathbf{Z}_{ij} = \mathbf{V}^{-\frac{1}{2}} [\mathbf{x}_{ij} - \bar{\mathbf{x}}_i]$$

Matriks diagonal simpangan baku:

$$\mathbf{V}^{\frac{1}{2}} = \begin{bmatrix} \sqrt{S_{11}} & 0 & \dots & 0 \\ 0 & \sqrt{S_{22}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sqrt{S_{pp}} \end{bmatrix} = \text{diagonal } [S_{ii}^{1/2}]$$

Invers matrik simpangan baku

$$(\mathbf{V}_2)^{-1} = \begin{bmatrix} \frac{1}{\sqrt{s_{11}}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{s_{22}}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sqrt{s_{pp}}} \end{bmatrix} = \text{diagonal} \left[ S_{ii}^{-1/2} \right]$$

$$\text{Cov}(\mathbf{Z}_{ij}) = \left\{ \left( \text{diagonal} \left[ S_{ii}^{\frac{1}{2}} \right] \right) \times (S) \times \left( \text{diagonal} \left[ S_{ii}^{\frac{1}{2}} \right] \right) \right\} = \rho$$

$$\rho = \begin{bmatrix} \frac{1}{\sqrt{s_{11}}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{s_{22}}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sqrt{s_{pp}}} \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} & \dots & S_{1p} \\ S_{21} & S_{22} & \dots & S_{2p} \\ S_{31} & S_{32} & \dots & S_{3p} \\ \vdots & \vdots & \ddots & \vdots \\ S_{p1} & S_{p2} & \dots & S_{pp} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{s_{11}}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{s_{22}}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sqrt{s_{pp}}} \end{bmatrix}$$

$$\rho = \begin{bmatrix} \frac{S_{11}}{\sqrt{S_{11}}\sqrt{S_{11}}} & \frac{S_{12}}{\sqrt{S_{11}}\sqrt{S_{22}}} & \dots & \frac{S_{1p}}{\sqrt{S_{11}}\sqrt{S_{pp}}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{S_{p1}}{\sqrt{S_{11}}\sqrt{S_{pp}}} & \frac{S_{p2}}{\sqrt{S_{22}}\sqrt{S_{pp}}} & \dots & \frac{S_{pp}}{\sqrt{S_{pp}}\sqrt{S_{pp}}} \end{bmatrix} = \begin{bmatrix} 1 & r_{12} & \dots & r_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \dots & 1 \end{bmatrix}$$

$r_{ij} = r_{ji}$

Dengan demikian, komponen-komponen utama dari  $\mathbf{Z}$  dapat ditentukan dari vektor ciri yang diperoleh melalui matriks korelasi  $R$  peubah asal  $X$ . Model komponen utama ke- $k$  (Johnson dan Wichern, 2002) adalah:

$$K_1 = b_{11}X_1 + b_{12}X_2 + \dots + b_{1p}X_p = \mathbf{b}'_1 \mathbf{X}$$

$$K_2 = b_{21}X_1 + b_{22}X_2 + \dots + b_{2p}X_p = \mathbf{b}'_2 \mathbf{X}$$

$$K_3 = b_{31}X_1 + b_{32}X_2 + \dots + b_{3p}X_p = \mathbf{b}'_3 \mathbf{X}$$

$$\vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots$$

$$K_q = b_{q1}X_1 + b_{q2}X_2 + \dots + b_{qp}X_p = \mathbf{b}'_q \mathbf{X}$$

Cara menentukan nilai vektor ciri  $b_k$  dan akar ciri  $\lambda_k$  matriks korelasi  $R$  sama seperti pada penggunaan matriks ragam-peragam.

### 2.2.2 Koefisien Pembobot (Vektor ciri) Komponen Utama

Koefisien pembobot komponen utama adalah hal penting dalam analisis komponen utama. Menentukan komponen utama

$$K_k = b_{k1}X_1 + b_{k2}X_2 + \dots + b_{kp}X_p$$

berarti menghitung vektor ciri  $b_k = [b_{k1}, b_{k2}, \dots, b_{kp}]$  sedemikian sehingga  $b'_k S b_k$  yang merupakan ragam komponen utama, mencapai nilai maksimum jika  $b'_k b_k = 1$ .

$$k = 1, \quad b'_1 = [b_{11}, b_{12}, \dots, b_{1p}] \text{ sehingga,}$$

$$b'_1 b_1 = [b_{11}, b_{12}, \dots, b_{1p}] \begin{bmatrix} b_{11} \\ b_{12} \\ \vdots \\ b_{1p} \end{bmatrix} = 1$$

$$k = q, \quad b'_q = [b_{q1}, b_{q2}, \dots, b_{qp}] \text{ sehingga,}$$

$$b'_q b_q = [b_{q1}, b_{q2}, \dots, b_{qp}] \begin{bmatrix} b_{q1} \\ b_{q2} \\ \vdots \\ b_{qp} \end{bmatrix} = 1$$

Menurut Harris (1975), permasalahan ini dapat diselesaikan dengan menggunakan fungsi *Lagrange*. Maksimum *Lagrange*  $b'_k S b_k$  dengan kendala  $b'_k b_k = 1$  atau  $b'_k b_k - 1 = 0$ . Menggunakan teknik pemaksimalan berkendala *Lagrange* dapat dibentuk fungsi (Harris, 1975):

$$L = b'_k S b_k - \lambda_k (b'_k b_k - 1)$$

Apabila  $L$  diturunkan terhadap vektor  $b_k$  kemudian dinormalisasi akan diperoleh:

$$\begin{aligned} \frac{\partial L}{\partial b} &= 2b_k S - 2\lambda_k b_k = 0 \\ &= 2(S - \lambda_k I)b_k = 0 \\ &= (S - \lambda_k I)b_k = 0 \end{aligned} \tag{2.3}$$

Persamaan (2.3) akan menghasilkan jawab *non trivial* karena menghasilkan nilai jawab bukan 0.

Matriks  $(S - \lambda_k I)$  merupakan matriks *singular* karena

$$|S - \lambda_k I| = 0 \tag{2.4}$$



Penyelesaian persamaan ciri (2.4) akan menghasilkan akar-akar ciri  $\lambda_1, \lambda_2, \dots, \lambda_q$  di mana  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_q = 0$ . Setiap akar ciri  $\lambda_k$  akan menentukan vektor ciri  $b_k$ .

Penentuan akar ciri dari  $p$  buah akar ciri yang dipergunakan dalam komponen utama pertama adalah :

Kembali pada persamaan (2.3)

$$(\mathbf{S} - \lambda_k \mathbf{I})\mathbf{b}_k = 0$$

$$\mathbf{S}\mathbf{b}_k - \lambda_k \mathbf{I}\mathbf{b}_k = 0$$

$$\mathbf{S}\mathbf{b}_k = \lambda_k \mathbf{I}\mathbf{b}_k$$

Kedua sisi persamaan dikalikan  $\mathbf{b}'_k$  menjadi

$$\mathbf{b}'_k \mathbf{S}\mathbf{b}_k = \mathbf{b}'_k \lambda_k \mathbf{I}\mathbf{b}_k$$

Dengan memperhatikan batasan  $\mathbf{b}'_k \mathbf{b}_k = 1$  maka diperoleh  $\mathbf{b}'_k \mathbf{S}\mathbf{b}_k = \lambda_k$ .

Diketahui bahwa ragam setiap komponen utama dipengaruhi oleh nilai setiap akar ciri. Agar ragam komponen utama pertama maksimum, maka harus dipilih akar ciri terbesar matriks  $\mathbf{S}$  untuk digunakan dalam komponen utama pertama sehingga komponen ini akan menjelaskan keragaman terbesar data. Komponen-komponen utama lain menjelaskan proporsi keragaman yang semakin kecil (Harris, 1975).

### 2.2.3. Penentuan banyaknya Komponen Utama

Salah satu tujuan dari analisis komponen utama adalah mereduksi dimensi data asal dari  $p$  peubah prediktor menjadi  $k$  komponen utama (di mana  $k < p$ ), berdasarkan kriteria :

1. Akar ciri lebih besar dari satu ( $\lambda_i \geq 1$ ) (Draper and Smith, 1981).
2. Proporsi kumulatif keragaman data asal yang dijelaskan oleh  $k$  komponen utama minimum 80% (Draper and Smith, 1981).
3. *Scree plot* merupakan plot antara akar ciri dan komponen utama yang terbentuk dilandasi dari patahan siku, sehingga diketahui banyak komponen utama yang terbentuk dari  $p$  peubah asal (Johnson dan Wichern, 2002).

### 2.2.4. Skor Komponen Utama

Menurut Johnson dan Wichern (2002), berdasarkan komponen utama dihitung skor komponen dari setiap  $p$  prediktor yang akan digunakan dalam analisis berikut. Pandang matriks data:

$$(X_{ij})_{p \times n} = \begin{bmatrix} X_{11} & \dots & X_{1n} \\ \vdots & \ddots & \vdots \\ X_{p1} & \dots & X_{pn} \end{bmatrix}$$

Skor komponen pengamatan ke- $j$  pada komponen utama  $k_k$  yang dihasilkan dari matriks ragam-peragam adalah:

$$sk_{jk} = [b_{1k}, b_{2k}, \dots, b_{pk}] \begin{bmatrix} X_{1j} - \bar{X}_1 \\ X_{2j} - \bar{X}_2 \\ \vdots \\ X_{ij} - \bar{X}_i \end{bmatrix} \text{ atau}$$

$$sk_{jk} = b'_k [X_{ij} - \bar{X}_i] \quad (2.5)$$

Jika komponen utama dihasilkan dari matriks korelasi  $\mathbf{R}$ , maka

$$sk_{jk} = b'_k V^{-1/2} [X_{ij} - \bar{X}_i] \quad (2.6)$$

### 2.3. Analisis Komponen Utama Robust

Menurut Hubert et al. (2003), *Robust Principal Component Analysis* (ROBPCA) adalah suatu metode yang bertujuan untuk menghasilkan komponen utama yang tidak dipengaruhi oleh pencilan. *Minimum Covariance Determinant* (MCD) dan *Minimum Volume Ellipsoid* (MVE) dikenalkan oleh Rousseeuw (1999) sebagai suatu penduga dengan breakdown point yang tinggi. Pendekatan lain untuk mendapatkan komponen utama yang tangguh adalah dengan menggunakan *Projection Pursuit* (PP). Penduga MCD hampir mirip dengan penduga MVE, perbedaan terletak pada kriteria pemilihan amatan. Prinsip dasar ROBPCA adalah mengganti matriks ragam-peragam dan vektor rata-rata dengan penduga *robust* bagi matriks ragam-peragam dan vektor rata-rata dengan metode *Reweighted Minimum Covariance Determinant*. Langkah-langkah untuk mendapatkan komponen utama metode ROBPCA hampir sama dengan CPCA, yang membedakan adalah pada ROBPCA matriks ragam-peragam dan vektor rata-rata diduga dengan RMCD.

### 2.4. Reweighted Minimum Covariance Determinant (RMCD)

*Minimum Covariance Determinant* (MCD) adalah suatu metode pencarian sejumlah  $h$  pengamatan bagian dari himpunan matriks populasi data  $X$  di mana  $h = \frac{n+p+1}{2}$  dengan  $n > p$  untuk menduga

vektor rata-rata dan matriks ragam-peragam populasi yang menghasilkan determinan terkecil. Untuk meningkatkan efisiensi MCD digunakan *Reweighted Minimum Covariance Determinant*, yaitu metode MCD yang ditambahi bobot  $w_j$  berdasarkan jarak *Robust* pengamatan. RMCD mewarisi sifat-sifat baik dari MCD seperti *breakdown point* yang tinggi (Hubert et al., 2003).

*Breakdown point* adalah jumlah pengamatan minimal yang dapat menggantikan sejumlah pengamatan semula yang berakibat pada nilai penduga yang dihasilkan sangat berbeda dari penduga sebenarnya (Hubert et al., 2007). *Breakdown point* yang dipakai adalah sebesar 50% (Branden, 2005).

Rousseeuw dan Van Driessen pada tahun 1999 memperkenalkan algoritma berikut:

Diketahui himpunan data sejumlah  $n$  pengamatan. Misal  $H_1 \subset \{1, \dots, n\}$  dengan jumlah elemen  $H_1 = h$

1. Ambil secara acak sejumlah  $h$  pengamatan berbeda,

$$h = \frac{n+p+1}{2} \tag{2.7}$$

Tabel 2.4. Tabel  $h$  Pengamatan Contoh terhadap  $p$  Peubah

$i$	$l$					Rata-rata
	1	2	3	...	$h$	
1	$\begin{matrix} x_{11} \\ \dots \\ x_{1h} \end{matrix}$	$\begin{matrix} x_{12} \\ \dots \\ x_{1h} \end{matrix}$	$\begin{matrix} x_{13} \\ \dots \\ x_{1h} \end{matrix}$	...	$\begin{matrix} x_{1h} \\ \dots \\ x_{1h} \end{matrix}$	$\begin{matrix} \bar{x}_1 \\ \dots \\ \bar{x}_1 \end{matrix}$
2	$\begin{matrix} x_{21} \\ \dots \\ x_{2h} \end{matrix}$	$\begin{matrix} x_{22} \\ \dots \\ x_{2h} \end{matrix}$	$\begin{matrix} x_{23} \\ \dots \\ x_{2h} \end{matrix}$	...	$\begin{matrix} x_{2h} \\ \dots \\ x_{2h} \end{matrix}$	$\begin{matrix} \bar{x}_2 \\ \dots \\ \bar{x}_2 \end{matrix}$
3	$\begin{matrix} x_{31} \\ \dots \\ x_{3h} \end{matrix}$	$\begin{matrix} x_{32} \\ \dots \\ x_{3h} \end{matrix}$	$\begin{matrix} x_{33} \\ \dots \\ x_{3h} \end{matrix}$	...	$\begin{matrix} x_{3h} \\ \dots \\ x_{3h} \end{matrix}$	$\begin{matrix} \bar{x}_3 \\ \dots \\ \bar{x}_3 \end{matrix}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	$\vdots$
$p$	$\begin{matrix} x_{p1} \\ \dots \\ x_{ph} \end{matrix}$	$\begin{matrix} x_{p2} \\ \dots \\ x_{ph} \end{matrix}$	$\begin{matrix} x_{p3} \\ \dots \\ x_{ph} \end{matrix}$	...	$\begin{matrix} x_{ph} \\ \dots \\ x_{ph} \end{matrix}$	$\begin{matrix} \bar{x}_p \\ \dots \\ \bar{x}_p \end{matrix}$

$l = 1, 2, \dots, h$

Dalam bentuk matriks:

$$X = \begin{bmatrix} X_{11} & X_{12} & X_{13} & \dots & X_{1h} \\ X_{21} & X_{22} & X_{23} & \dots & X_{1h} \\ X_{31} & X_{32} & X_{33} & \dots & X_{1h} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ X_{p1} & X_{p2} & X_{p3} & \dots & X_{ph} \end{bmatrix}$$

$(pxh)$

2. Hitung vektor rata-rata dari  $\mathbf{X}$  dengan (Rousseeuw, 1999):

$$\left. \begin{aligned} \bar{x}_1 &= \left(\frac{1}{h}\right) \sum_{l=1}^h x_{1l} \\ \bar{x}_2 &= \left(\frac{1}{h}\right) \sum_{l=1}^h x_{2l} \\ \bar{x}_3 &= \left(\frac{1}{h}\right) \sum_{l=1}^h x_{3l} \\ &\vdots \\ \bar{x}_p &= \left(\frac{1}{h}\right) \sum_{l=1}^h x_{il} \end{aligned} \right\} \bar{\mathbf{X}}_i = h^{-1} \sum_{l=1}^h \mathbf{X}_{il}$$

yang menghasilkan vektor  $\mathbf{t}_1 = \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \bar{X}_3 \\ \vdots \\ \bar{X}_p \end{bmatrix}$  (2.8)

3. Hitung matriks ragam-peragam (Rousseeuw, 1999):

$$\mathbf{S}_1 = \left(\frac{1}{h-1}\right) \sum_{l=1}^h (\mathbf{X}_{il} - \mathbf{t}_1)(\mathbf{X}_{il} - \mathbf{t}_1)' \quad (2.9)$$

Sebelum didapatkan vektor rata-rata dan ragam-peragam *robust* MCD, digunakan inialisasi  $\mathbf{t}_1$  sebagai vektor rata-rata dan  $\mathbf{S}_1$  sebagai matriks ragam-peragam.

Tabel 2.5. Unsur-unsur matriks  $\mathbf{S}_1$

$x_{11} - \bar{x}_1$	$\frac{1}{h-1} \sum_{l=1}^h (x_{1l} - \bar{x}_1)(x_{1l} - \bar{x}_1)$	$\frac{1}{h-1} \sum_{l=1}^h (x_{2l} - \bar{x}_2)(x_{1l} - \bar{x}_1)$	$\frac{1}{h-1} \sum_{l=1}^h (x_{3l} - \bar{x}_3)(x_{1l} - \bar{x}_1)$	...	$\frac{1}{h-1} \sum_{l=1}^h (x_{pl} - \bar{x}_p)(x_{1l} - \bar{x}_1)$
$x_{21} - \bar{x}_2$	...	$\frac{1}{h-1} \sum_{l=1}^h (x_{2l} - \bar{x}_2)(x_{2l} - \bar{x}_2)$	$\frac{1}{h-1} \sum_{l=1}^h (x_{3l} - \bar{x}_3)(x_{2l} - \bar{x}_2)$	...	$\frac{1}{h-1} \sum_{l=1}^h (x_{pl} - \bar{x}_p)(x_{2l} - \bar{x}_2)$
$x_{31} - \bar{x}_3$	...	...	$\frac{1}{h-1} \sum_{l=1}^h (x_{3l} - \bar{x}_3)(x_{3l} - \bar{x}_3)$	...	$\frac{1}{h-1} \sum_{l=1}^h (x_{pl} - \bar{x}_p)(x_{3l} - \bar{x}_3)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$x_{p1} - \bar{x}_p$	...	...	...	...	$\frac{1}{h-1} \sum_{l=1}^h (x_{pl} - \bar{x}_p)(x_{pl} - \bar{x}_p)$

Tabel 2.6. Matriks Ragam-peragam

$\frac{1}{h-1} \sum_{l=1}^h (x_{1l} - \bar{x}_1)(x_{1l} - \bar{x}_1)$	$\text{Var} \left( \frac{1}{h-1} \sum_{l=1}^h (x_{1l} - \bar{x}_1) \right)$	$\text{Cov} \left( \frac{1}{h-1} \sum_{l=1}^h (x_{2l} - \bar{x}_2), \frac{1}{h-1} \sum_{l=1}^h (x_{1l} - \bar{x}_1) \right)$	$\text{Cov} \left( \frac{1}{h-1} \sum_{l=1}^h (x_{3l} - \bar{x}_3), \frac{1}{h-1} \sum_{l=1}^h (x_{1l} - \bar{x}_1) \right)$	...	$\text{Cov} \left( \frac{1}{h-1} \sum_{l=1}^h (x_{pl} - \bar{x}_p), \frac{1}{h-1} \sum_{l=1}^h (x_{1l} - \bar{x}_1) \right)$
$\frac{1}{h-1} \sum_{l=1}^h (x_{2l} - \bar{x}_2)(x_{2l} - \bar{x}_2)$	...	$\text{Var} \left( \frac{1}{h-1} \sum_{l=1}^h (x_{2l} - \bar{x}_2) \right)$	$\text{Cov} \left( \frac{1}{h-1} \sum_{l=1}^h (x_{3l} - \bar{x}_3), \frac{1}{h-1} \sum_{l=1}^h (x_{2l} - \bar{x}_2) \right)$	...	...
$\frac{1}{h-1} \sum_{l=1}^h (x_{3l} - \bar{x}_3)(x_{3l} - \bar{x}_3)$	...	...	$\text{Var} \left( \frac{1}{h-1} \sum_{l=1}^h (x_{3l} - \bar{x}_3) \right)$	...	...
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$\frac{1}{h-1} \sum_{l=1}^h (x_{pl} - \bar{x}_p)(x_{pl} - \bar{x}_p)$	...	...	...	...	$\text{Var} \left( \frac{1}{h-1} \sum_{l=1}^h (x_{pl} - \bar{x}_p) \right)$

$$\mathbf{S}_1 = \left( \frac{1}{h-1} \right) \begin{bmatrix} X_{11} - \bar{X}_1 & X_{12} - \bar{X}_1 & X_{13} - \bar{X}_1 & \cdots & X_{1l} - \bar{X}_1 \\ X_{21} - \bar{X}_2 & X_{22} - \bar{X}_2 & X_{23} - \bar{X}_2 & \cdots & X_{2l} - \bar{X}_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ X_{p1} - \bar{X}_p & X_{p2} - \bar{X}_p & X_{p3} - \bar{X}_p & \cdots & X_{pl} - \bar{X}_p \end{bmatrix} \begin{bmatrix} X_{11} - \bar{X}_1 & X_{21} - \bar{X}_2 & X_{31} - \bar{X}_3 & \cdots & X_{p1} - \bar{X}_p \\ X_{12} - \bar{X}_1 & X_{22} - \bar{X}_2 & X_{32} - \bar{X}_3 & \cdots & X_{p2} - \bar{X}_p \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ X_{1l} - \bar{X}_1 & X_{2l} - \bar{X}_2 & X_{3l} - \bar{X}_3 & \cdots & X_{pl} - \bar{X}_p \end{bmatrix}$$

(pxl) (lpx)

4. Hitung determinan  $\mathbf{S}_1$  (Rousseeuw, 1999):

Jika  $|\mathbf{S}_1| = 0$ , maka berhenti. Jika tidak, hitung jarak relatif

$$d_{1(j)} \approx \sqrt{(\mathbf{x}_{ij} - \mathbf{t}_1)' \mathbf{S}_1^{-1} (\mathbf{x}_{ij} - \mathbf{t}_1)} \quad (2.10)$$

Matriks deviasi  $(\mathbf{x}_{ij} - \mathbf{t}_1)$  untuk  $j=1$

$$\begin{bmatrix} X_{11} \\ X_{12} \\ \vdots \\ X_{1p} \end{bmatrix} - \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \vdots \\ \bar{X}_p \end{bmatrix} = \begin{bmatrix} X_{11} - \bar{X}_1 \\ X_{12} - \bar{X}_2 \\ \vdots \\ X_{1p} - \bar{X}_p \end{bmatrix}$$

(px1)    (px1)    (px1)

5. Selanjutnya ambil  $H_2$  sedemikian sehingga,

$\{d_{1(j)}; j \in H_2\} := \{(d_1)_{1:n}, (d_1)_{2:n}, (d_1)_{3:n}, (d_1)_{4:n}, \dots, (d_1)_{h:n}\}$  di mana  $(d_1)_{1:n} \leq (d_1)_{2:n} \leq (d_1)_{3:n} \leq \dots \leq (d_1)_{h:n}$  merupakan statistik peringkat  $d_j$ . Kemudian hitung  $\mathbf{t}_2$  dan  $\mathbf{S}_2$  berdasarkan himpunan  $H_2$  sehingga  $|\mathbf{S}_2| \leq |\mathbf{S}_1|$ . Dengan demikian didapatkan penduga vektor rata-rata  $\bar{\mathbf{X}}_{MCL}$  dan matriks ragam-peragam  $\mathbf{S}_{MCL}$  berdasarkan anggota himpunan  $H_2$ .

Identifikasi pencilan dengan persamaan (2.1) tidak berjalan maksimal bila data mengandung lebih dari satu pencilan, hal ini karena efek *masking* dan *swamping*. Efek ini diatasi dengan menggunakan penduga *robust* untuk vektor rata-rata dan matriks ragam-peragam sehingga dihasilkan kuadrat jarak *Robust* (Rousseeuw, 1999):

$$RD_j^2 = (\mathbf{x}_{ij} - \bar{\mathbf{X}}_{MCL})' \mathbf{S}_{MCL}^{-1} (\mathbf{x}_{ij} - \bar{\mathbf{X}}_{MCL}) \quad (2.11)$$

Hipotesis yang melandasi pengujian pencilan:

$H_0: x_{ij} =$  pengamatan biasa

$H_1: x_{ij} =$  pencilan

Jika  $H_0$  benar, statistik uji:

$$RD_j^2 \sim \chi_{(p)}^2$$

$H_0$  akan ditolak jika  $RD_j^2 > \chi_{\frac{\alpha}{2}(p)}^2$ , berarti pengamatan  $x_{ij}$  diidentifikasi sebagai pencilan.

Dengan menggunakan pembobot ( $w_j$ ) berdasarkan jarak *Robust* pengamatan, dihitung  $\bar{X}_{RMCD}$  dan  $S_{RMCD}$ . Pengamatan yang tidak disertakan dalam perhitungan diberi bobot nol, lainnya diberi bobot satu (Suryana, 2005).

$$w_j = \begin{cases} 1, & RD_j^2 < \chi_{\frac{\alpha}{2}(p)}^2 \\ 0, & \text{lainnya} \end{cases} \quad (2.12)$$

di mana :

$\bar{X}_{MCD}$  : vektor rata-rata *Robust* MCD

$S_{MCD}$  : matriks ragam-peragam *Robust* MCD

Penduga RMCD dihitung dengan (Rousseeuw, 1999):

$$\bar{X}_{RMCD} = \frac{\sum_{j=1}^n w_j x_{ij}}{\sum_{j=1}^n w_j} \quad (2.13)$$

$$S_{RMCD} = \frac{\sum_{j=1}^n w_j (x_{ij} - \bar{X}_{RMCD}) (x_{ij} - \bar{X}_{RMCD})'}{(\sum_{j=1}^n w_j) - 1} \quad (2.14)$$

$\bar{X}_{RMCD}$  : vektor rata-rata *Robust* RMCD

$S_{RMCD}$  : matriks ragam-peragam *Robust* RMCD

Berikut diagram alur algoritma yang dikembangkan oleh Rousseeuw dan Van Driessen.

mulai

Ambil sejumlah  $h$  pengamatan berbeda, dengan jumlah elemen  $h=(n+p+1)/2$

Pilih anggota himpunan ( $H_1$ )

Hitung berdasarkan anggota  $H_1$

- $t_1 = \left(\frac{1}{h}\right) \sum_{l=1, l \in H_1}^h X_{il}$
- $S_1 = \left(\frac{1}{h-1}\right) \sum_{l=1, l \in H_1}^h (X_{il} - t_1)(X_{il} - t_1)'$
- Det ( $S_1$ ), jika  $|S_1| = 0$ , maka berhenti. Jika tidak, maka hitung jarak relatif  $d_{1(i)}$

Tentukan anggota himpunan  $H_2$  dengan cara:

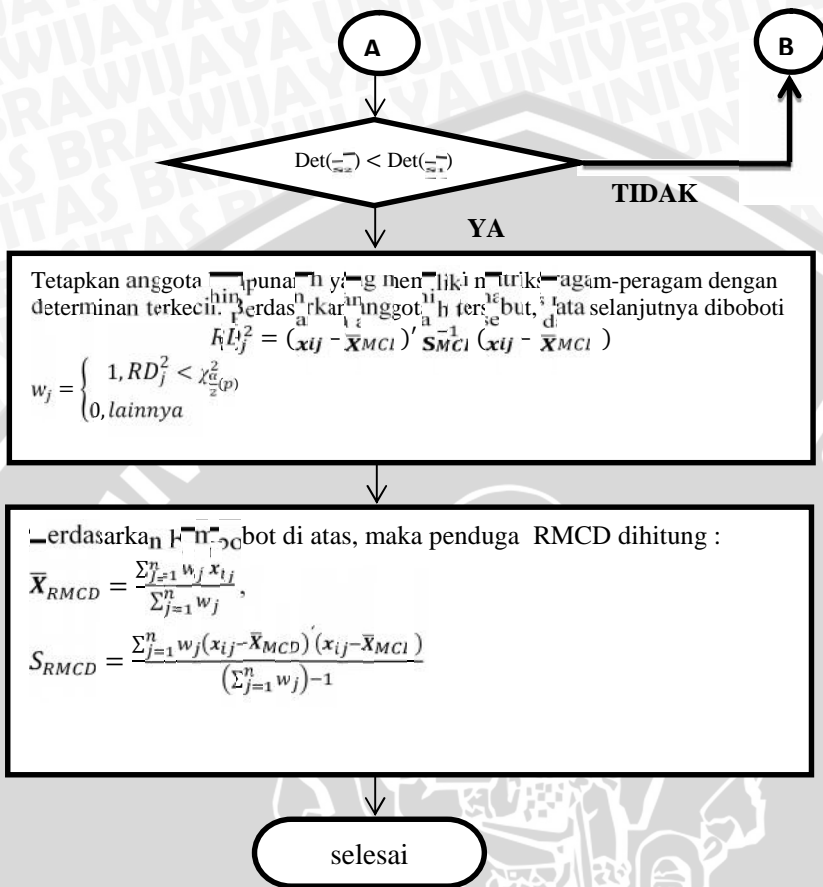
- Hitung  $d_{1(i)} = \sqrt{(X_{ij} - t_1)' S_1^{-1} (X_{ij} - t_1)}$
- Susun statistik peringkat  $(d_1)_{1:n} \leq (d_1)_{2:n} \leq \dots \leq (d_1)_{n:n}$
- Anggota himpunan  $H_2$  adalah  $\{d_{1(i)}; i \in H_2\} := \{(d_1)_{1:n}, (d_1)_{2:n}, (d_1)_{3:n}, \dots, (d_1)_{h:n}\}$

Hitung berdasarkan anggota  $H_2$

- $t_2 = \left(\frac{1}{h}\right) \sum_{l=1, l \in H_2}^h X_{il}$
- $S_2 = \left(\frac{1}{h-1}\right) \sum_{l=1, l \in H_2}^h (X_{il} - t_2)(X_{il} - t_2)'$
- Det ( $S_2$ ),

A

B



Gambar 2.1. Diagram Alur Metode Penduga RMCD

Rousseuw dan Van Driessen (1999), menyatakan cukup dengan menggunakan sepuluh himpunan bagian permulaan untuk memulai pencarian determinan matriks ragam-peragam terkecil.

### 2.5. Indikator Pemanding

Jika metode CPCA dan ROBPCA menghasilkan komponen utama terpilih yang sama banyak, maka digunakan indikator pemanding yaitu besarnya proporsi keragaman yang mampu dijelaskan oleh penduga akar ciri (Johnson dan Wichern, 2002).

$$\frac{\sum_{k=1}^q \lambda_k}{\sum_{j=1}^p \lambda_k} \quad 80\%.$$



## BAB III METODE PENELITIAN

### 3.1. Sumber Data

Data sekunder dalam penelitian merupakan hasil Survei Sosial Ekonomi Nasional 2010-2011 Provinsi Jatim. Data multivariat ini mengandung pencilan. Sebanyak 38 pengamatan ditentukan untuk data I,II dan III karena Jawa Timur terdiri dari 29 kabupaten dan 9 kota. Pengolahan data menggunakan paket program LIBRA-MATLAB 2008.

Tabel 3.1. Keterangan tentang data

Data	Peubah	Sumber	<i>n</i>	<i>p</i>
I	1. Beras (rupiah/minggu)	Survei Sosial Ekonomi Nasional 2010-2011 Provinsi Jatim tentang Rata-Rata Pengeluaran Makanan /kapita Menurut Kabupaten atau Kota dan Jenis Pengeluaran Tahun 2011	3	10
	2. Umbi-umbian (rupiah/minggu)		8	
	3. Daging (rupiah/minggu)			
	4. Telur dan susu (rupiah/bulan)			
	5. Sayur-sayuran (rupiah/bulan)			
	6. Kacang-kacangan (rupiah/bulan)			
	7. Buah-buahan (rupiah/bulan)			
	8. Minyak dan lemak (rupiah/bulan)			
	9. Bahan minuman (rupiah/bulan)			
	10. Bumbu-bumbuan (rupiah/bulan)			

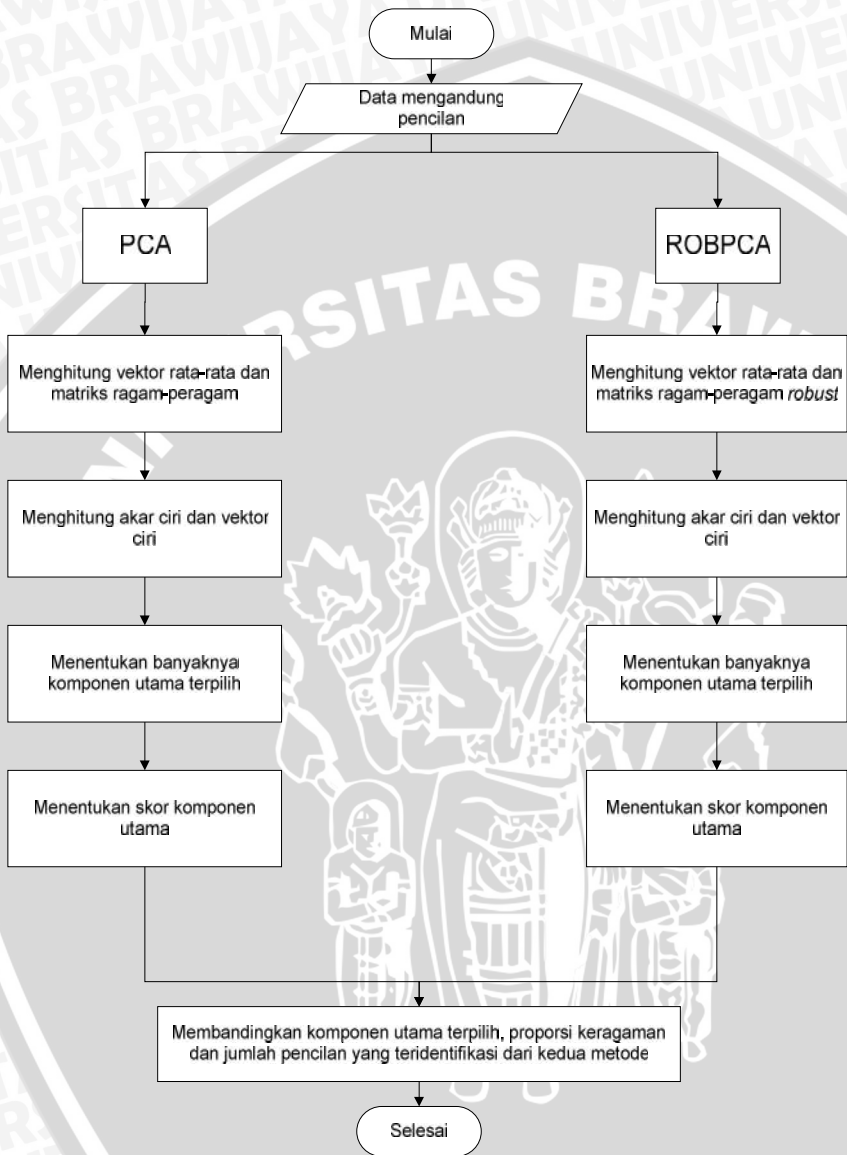
Tabel 3.1. (lanjutan)

Data	Peubah	Sumber	<i>n</i>	<i>p</i>
II	<ol style="list-style-type: none"> <li>1. Perumahan dan fasilitas rumah tangga (rupiah/minggu)</li> <li>2. Aneka barang dan jasa termasuk kesehatan dan pendidikan (rupiah/minggu)</li> <li>3. Pakaian, alas kaki dan tutup kepala (rupiah/bulan)</li> <li>4. Barang tahan lama (rupiah/bulan)</li> <li>5. Pajak dan asuransi (rupiah/bulan)</li> <li>6. Keperluan pesta dan upacara (rupiah/bulan)</li> </ol>	<p>Survei Sosial Ekonomi Nasional 2010-2011 Provinsi Jatim tentang Rata-Rata Pengeluaran Non Makanan/kapita Menurut Kabupaten atau Kota</p>	38	6
III	<ol style="list-style-type: none"> <li>1. Program Nasional Pemberdayaan Masyarakat (PNPM) (rupiah/bulan)</li> <li>2. Program pemerintah lain (rupiah/bulan)</li> <li>3. Kredit Usaha Rakyat (KUR) (rupiah/bulan)</li> <li>4. Program Koperasi (rupiah/bulan)</li> <li>5. Perorangan (rupiah/bulan)</li> <li>6. Kredit Bank (rupiah/minggu)</li> <li>7. Lain-lain (rupiah/bulan)</li> </ol>	<p>Survei Sosial Ekonomi Nasional 2010-2011 Provinsi Jatim tentang Rumah Tangga yang Menerima Kredit Usaha Dalam Setahun Terakhir Menurut Kabupaten/Kota dan Jenis Kredit yang Diterima</p>	38	7

### 3.2. Metode Penelitian

Analisis yang dilakukan dalam penelitian ini adalah:

1. Prosedur deteksi pencilan menggunakan persamaan (2.1).
  2. Menganalisis data dengan metode CPCA.
  3. Menganalisis data dengan metode ROBPCA.
    - 3.1 Mendapatkan matriks ragam-peragam yang *robust* dengan metode RMCD
      - 3.1.1 Menentukan himpunan  $h$  menggunakan persamaan (2.7)
      - 3.1.2 Untuk setiap himpunan bagian, hitung vektor rata-rata dan matriks ragam-peragam  $(\mathbf{t}_1, \mathbf{S}_1)$  menggunakan persamaan (2.8) dan (2.9)
      - 3.1.3 Menghitung jarak relatif  $(d_{1(j)})$  menggunakan persamaan (2.10), mengambil  $h$  pengamatan dengan jarak terkecil kemudian menjadikan  $h$  pengamatan tersebut sebagai unsur himpunan  $H_2$
      - 3.1.4 Menghitung  $(\mathbf{t}_2, \mathbf{S}_2)$ , kemudian kembali ke langkah (3.1.2) sampai diperoleh determinan matriks ragam-peragam terkecil sehingga diperoleh  $(\bar{\mathbf{X}}_{MCL}, \mathbf{S}_{MCL})$ .
      - 3.1.5 Menghitung bobot  $(w_j)$  menggunakan persamaan (2.11).
      - 3.1.6 Menghitung  $\bar{\mathbf{X}}_{RMCL}$  dan  $\mathbf{S}_{RMCL}$  menggunakan persamaan (2.12) dan (2.13).
    - 3.2 Menentukan akar ciri dan vektor ciri komponen utama *robust* menggunakan persamaan (2.4).
    - 3.3 Menentukan komponen utama yang digunakan berdasarkan sub bab 2.2.3
    - 3.4 Menentukan skor komponen utama *robust* menggunakan persamaan (2.5) untuk satuan pengukuran sama dan persamaan (2.6) untuk satuan pengukuran berbeda.
  4. Membandingkan komponen utama terbentuk dan proporsi keragaman yang dihasilkan oleh metode CPCA dan RMCD.
- Prosedur analisis dijelaskan pada Gambar 3.1.



Gambar 3.1. Diagram Alir Penelitian

## BAB IV HASIL DAN PEMBAHASAN

### 4.1. Identifikasi Pencilan

Sebelum analisis dilakukan, akan diperiksa terlebih dahulu apakah komponen utama memuat pencilan. Jarak Mahalanobis (data I) untuk identifikasi pencilan tertera dalam Tabel 4.1.

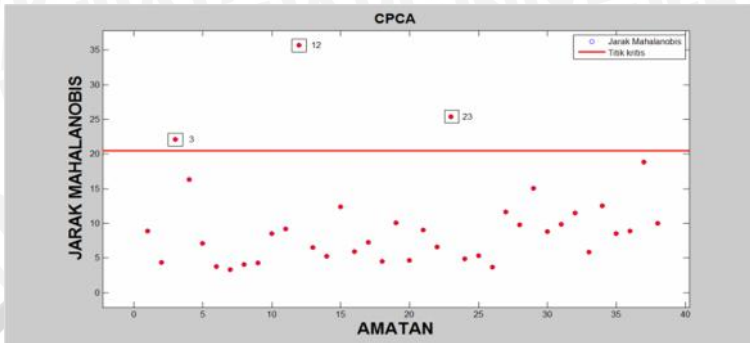
Tabel 4.1. Jarak *Mahalanobis*

Pengamatan	MD <sup>2</sup>	Pengamatan	MD <sup>2</sup>
1	8,877	20	4,657
2	4,332	21	9,003
3	22,09	22	6,548
4	16,28	23	25,35
5	7,082	24	4,902
6	3,766	25	5,337
7	3,297	26	3,667
8	4,071	27	11,63
9	4,310	28	9,803
10	8,499	29	15,05
11	9,177	30	8,835
12	35,66	31	9,872
13	6,480	32	11,46
14	5,243	33	5,855
15	12,38	34	12,51
16	5,880	35	8,502
17	7,232	36	8,909
18	4,514	37	18,87
19	10,04	38	10,02

Keterangan :

MD<sup>2</sup> = kuadrat jarak Mahalanobis

Karena kuadrat jarak *Mahalanobis* amatan ke-3 (22,09), ke-12 (35,66) dan ke-23 (25,35) >  $\chi^2_{0,025}(10) = 20,48$  sehingga terdapat 3 pengamatan berupa pencilan (Chen dan Jackson, 2004). Tabel 4.1 akan menghasilkan plot jarak *Mahalanobis* berikut:



Gambar 4.1. Plot Jarak Mahalanobis

Identifikasi pencilan menggunakan jarak Mahalanobis untuk data II dan III disajikan dalam Tabel 4.2 dan hasil perhitungan pada Lampiran 7. Plot jarak Mahalanobis data II dan III disajikan pada Lampiran 8.

Tabel 4.2. Banyaknya Pencilan

No	Data	n	p	Banyak pencilan	
1	I	38	10	3	20,48
2	II	38	6	4	14,45
3	III	38	7	3	16,01

Pandang Tabel 4.2 dan Lampiran 8, karena jarak Mahalanobis  $>$  , sehingga 4 pengamatan (data II) dan 3 pengamatan (data III) ini berupa pencilan.

## 4.2. Analisis Komponen Utama

### 4.2.1. Matriks Masukan

Data I, II dan III, mempunyai satuan pengukuran peubah berbeda, sehingga analisis komponen utama didasarkan pada data yang telah dibakukan dan tersaji pada Lampiran 4-6. Karena menggunakan peubah baku, maka digunakan matriks ragam-peragam (Johnson dan Wichern, 2002).

#### 4.2.2. Akar Ciri dan Vektor Ciri Komponen Utama

Sebelum membentuk sub ruang yang direntang oleh 38 pengamatan, akan dihitung akar ciri, proporsi keragaman dan proporsi keragaman kumulatif dengan analisis komponen utama klasik.

Tabel 4.3. Akar Ciri, Proporsi Keragaman dan Proporsi Keragaman Kumulatif dengan analisis komponen utama klasik

KU	CPCA		
	Akar ciri	Proporsi keragaman	Proporsi kumulatif
1	4,4855	0,449	0,449
2	1,9167	0,192	0,640
3	1,2765	0,128	0,768
4	0,8232	0,082	0,850
5	0,5563	0,056	0,906
6	0,3028	0,030	0,936
7	0,2738	0,027	0,963
8	0,1749	0,017	0,981
9	0,1250	0,013	0,993
10	0,0653	0,007	1,000

#### 4.2.3. Penentuan banyaknya Komponen Utama Klasik (CPCA)

Penentuan jumlah komponen utama yang dipilih dilandasi kriteria:

❖ Akar ciri.

Diperoleh 3 buah akar ciri yang memiliki nilai akar ciri 1 (Tabel 4.3). Berdasarkan hasil, terpilih jumlah komponen utama sebanyak 3 meskipun hanya mampu menjelaskan 76,8% keragaman data (Draper and Smith, 1981).

Untuk data II dan III, jumlah komponen utama yang terpilih tersaji dalam Tabel 4.4. Lampiran 9 menyajikan hasil perhitungan akar ciri, proporsi keragaman dan proporsi keragaman kumulatif.

Tabel 4.4. Jumlah Komponen Utama Terpilih dengan Metode PCA

Data	Jumlah komponen terpilih
I	3
II	2
III	3

❖ Proporsi Kumulatif Keragaman

Persen keragaman kumulatif data I metode CPCA disajikan dalam Tabel 4.5, untuk data II dan III disajikan pada Lampiran 9.

Tabel 4.5. Persen Keragaman Kumulatif Komponen Utama Terpilih

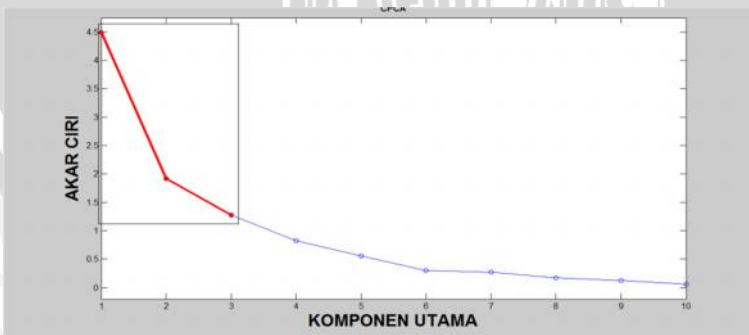
Data	Proporsi Kumulatif	Banyak Komponen
I	0,850	4
II	0,855	2
III	0,807	4

Pandang Tabel 4.3, komponen utama pertama menjelaskan 44,9% total keragaman, diikuti komponen utama kedua secara kumulatif menjelaskan 64% dan komponen ketiga sebesar 76,8% sampai dengan komponen keempat secara bersama-sama menjelaskan 85% total keragaman data.

Data I (Tabel 4.5) membentuk 4 komponen utama terpilih, mampu menjelaskan 85% total keragaman data. Data II membentuk 2 komponen utama terpilih yang mampu menjelaskan 85,5% total keragaman data dan data III membentuk 4 komponen utama terpilih yang mampu menjelaskan 80,7% total keragaman data (Draper and Smith, 1981).

❖ *Scree Plot*

Gambar 4.2 merupakan plot antara akar ciri terhadap komponen utama. Penentuan jumlah komponen dilandasi pada patahan siku dari *scree plot*.



Gambar 4.2. Plot Komponen Utama terhadap Akar Ciri Data I



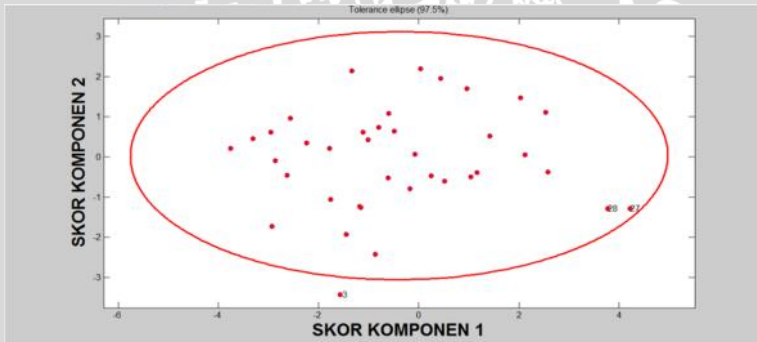
Pada Gambar 4.2 memperlihatkan bahwa setelah komponen ketiga, perubahan nilai akar ciri antar komponen cukup kecil, menurut kriteria ini cukup dibutuhkan 3 komponen utama (Johnson dan Wichern, 2002). *Scree plot* data II dan III tersaji pada Lampiran 10. Jumlah komponen terpilih berdasarkan kriteria *scree plot* disajikan dalam Tabel 4.6.

Tabel 4.6. Jumlah Komponen Utama Terpilih

Data	Banyak Komponen
I	3
II	2
III	3

#### 4.2.4. Skor Komponen Utama

Skor komponen utama merupakan data baru hasil analisis komponen utama yang disajikan pada Lampiran 11 (data I, II dan III).



Gambar 4.3. Plot Skor Komponen 1 terhadap Skor Komponen 2

Gambar 4.3 merupakan plot antara skor komponen 1 terhadap skor komponen 2. *Ellipse* yang tampak melebar dipengaruhi rata-rata yang tidak *robust* terhadap pencilan. *Tolerance ellipse* data II dan III tersaji pada Lampiran 12.

#### 4.3. Metode *Reweighted Minimum Covariance Determinant*

Gambar 4.1 merepresentasikan titik-titik pengamatan yang teridentifikasi sebagai pencilan, sehingga pereduksian dimensi dengan penduga biasa kurang dipercaya. Dengan demikian, untuk mengatasi

masalah pencilan diperlukan suatu metode tangguh yaitu *Reweighted Minimum Covariance Determinant*. Akan ditentukan banyak pengamatan minimal ( $h$ ) yang menggantikan sejumlah pengamatan awal sehingga berakibat pada nilai penduga yang dihasilkan sangat berbeda dari penduga sebenarnya. Hal ini dilakukan untuk proses penghitungan penduga *robust* RMCD,  $h = \frac{38+10+1}{2} \approx 24$  pengamatan (data I), 22 pengamatan untuk data II dan 23 pengamatan untuk data III. Berikut adalah pengamatan yang terpilih (data I):

Tabel 4.7. Pengamatan yang Terpilih

Pengamatan											
26	5	8	2	28	6	24	17	14	16	12	7

Pengamatan											
11	13	9	18	21	20	10	4	29	1	19	27

Pengamatan yang menghasilkan determinan matriks ragam-peragam terkecil, yang tersusun meningkat sesuai jarak relatif disajikan dalam Tabel 4.8.

Tabel 4.8. Pengamatan dengan Determinan Matriks Ragam-peragam Terkecil

Pengamatan											
7	26	6	8	9	2	18	20	24	14	16	13

Pengamatan											
5	17	10	1	21	11	28	19	27	29	4	12

Setelah didapat 24 pengamatan dengan determinan matriks ragam-peragam terkecil, dihitung vektor rata-rata dan matriks ragam-peragam *robust* MCD.

Pengamatan dengan determinan matriks ragam-peragam terkecil data II dan III tersaji pada Lampiran 13.

#### 4.4. Jarak *Robust* ( $w_j$ )

Jarak *Robust* data II dan III (Lampiran 15) dihitung setelah didapatkan vektor rata-rata dan matriks ragam-peragam *robust* MCD (Lampiran 14).

Tabel 4.9. Jarak *Robust*

Pengamatan	RD <sup>2</sup>	Pengamatan	RD <sup>2</sup>
1	13,28	20	12,61
2	7,546	21	10,93
3	91,54	22	12,92
4	11,60	23	430,9
5	6,022	24	7,851
6	7,140	25	61,48
7	7,475	26	4,705
8	6,061	27	39,73
9	6,948	28	8,868
10	10,31	29	12,94
11	8,750	30	87,59
12	401,5	31	147,6
13	13,69	32	67,98
14	8,427	33	84,25
15	80,05	34	214,6
16	8,456	35	163,8
17	8,145	36	193,7
18	9,697	37	406,2
19	12,89	38	12,69

Keterangan :

RD<sup>2</sup> = kuadrat jarak *Robust*

Karena kuadrat jarak *Robust* amatan ke-3 (91,547), ke-12 (11620), ke-15 (80,05), ke-23 (430,9), ke-25 (61,48), ke-27 (39,73), ke-30 (87,59), ke-31 (147,6), ke-32 (67,98), ke-33 (84,25), ke-34 (214,6), ke-35 (163,8), ke-36 (193,7) dan ke-37 (406,02) >  $\chi^2_{0,25(10)} = 20,48$  sehingga 14 pengamatan ini berupa pencilan. Pengamatan yang teridentifikasi sebagai pencilan diberi nilai 0 dan 1 selainnya.

Tabel 4.10. Pembobot

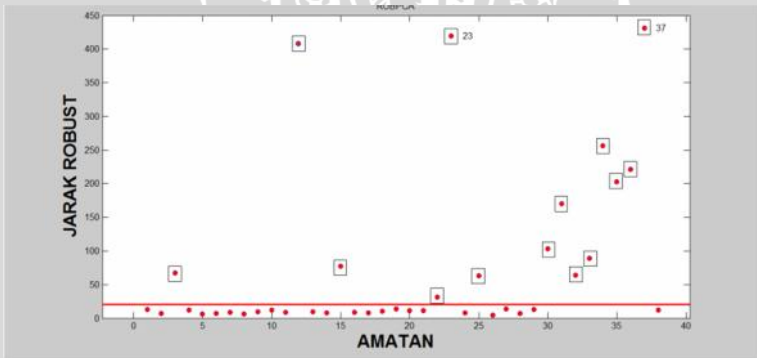
Pengamatan	Nilai pembobot	Pengamatan	Nilai pembobot
1	1	20	1
2	1	21	1
3	0	22	1
4	1	23	0
5	1	24	1
6	1	25	0
7	1	26	1
8	1	27	0
9	1	28	1
10	1	29	1
11	1	30	0
12	0	31	0
13	1	32	0
14	1	33	0
15	0	34	0
16	1	35	0
17	1	36	0
18	1	37	0
19	1	38	1

Kemudian penduga RMCD diperoleh:

$$\bar{X}_{RMCD} = \begin{bmatrix} 0,2310 \\ -0,144 \\ -0,232 \\ -0,424 \\ -0,437 \\ -0,046 \\ -0,017 \\ -0,359 \\ 0,0670 \\ 0,0540 \end{bmatrix}$$

=


Matriks ragam-peragam *robust* RMCD data II dan III (Lampiran 16).  
Tabel 4.9 akan menghasilkan plot jarak *Robust* berikut:



Gambar 4.4. Plot Jarak *Robust*

Identifikasi pencilan menggunakan jarak *Robust* untuk data II dan III disajikan dalam Tabel 4.11 dan hasil perhitungan pada Lampiran 15. Plot jarak *Robust* data II dan III disajikan pada Lampiran 8.

Tabel 4.11. Banyaknya Pencilan

No	Data	n	p	Banyak pencilan	
1	I	38	10	14	20,48
2	II	38	6	14	14,45
3	III	38	7	15	16,01

Berdasarkan Tabel 4.11 dan Lampiran 15, karena jarak *Robust* >  $\chi^2_{0,025(p)}$ , sehingga 14 pengamatan ini sebagai pencilan (data I). Pada data II mengandung 14 pencilan dan data III memuat 7 pencilan.

#### 4.5. Analisis Komponen Utama *Robust*

##### 4.5.1. Akar Ciri dan Vektor Ciri

Setelah didapatkan penduga vektor rata-rata dan matriks ragam-peragam yang tangguh, dilanjutkan dengan perhitungan vektor ciri dan akar ciri menggunakan penduga RMCD.

Tabel 4.12. Akar Ciri, proporsi keragaman, dan proporsi kumulatif

KU	ROBPCA		
	Akar ciri	Proporsi keragaman	Proporsi kumulatif
1	4,186	0,5783	0,5783
2	1,520	0,2100	0,7882
3	0,497	0,0687	0,8569
4	0,484	0,0669	0,9237
5	0,244	0,0337	0,9575
6	0,190	0,0262	0,9837
7	0,078	0,0108	0,9945
8	0,027	0,0037	0,9982
9	0,011	0,0015	0,9997
10	0,002	0,0003	1,0000

Akar ciri dalam analisis komponen utama berkaitan erat dengan vektor ciri yang terbentuk. Berikut vektor ciri komponen utama data I metode ROBPCA dengan penduga RMCD:

Tabel 4.13. Vektor ciri

Peubah	Komponen utama 1	Komponen utama 2	Komponen utama 3	...	Komponen utama 10
Padi-padian	0,326	0,327	-0,102	...	-0,180
Umbi-umbian	-0,256	-0,288	0,393	...	-0,045
Daging	0,075	0,009	-0,067	...	0,943
Telur & susu	-0,131	0,045	-0,183	...	-0,159

Tabel 4.13. (lanjutan)

Peubah	Komponen utama 1	Komponen utama 2	Komponen utama 3	...	Komponen utama 10
Sayur-sayuran	-0,198	0,075	-0,156	...	0,007
Kacang-kacangan	<b>-0,522</b>	-0,177	-0,270	...	0,142
Buah-buahan	-0,268	<b>0,618</b>	-0,353	...	0,052
Minyak & lemak	-0,208	-0,048	-0,485	...	-0,158
Bahan minuman	-0,492	-0,236	0,079	...	-0,047
Bumbu	-0,369	0,576	0,578	...	0,031

Pandang Tabel 4.13, komponen utama pertama data I, peubah kacang-kacangan ( $X_6$ ) memberikan kontribusi terbesar (0,522) terhadap komponen utama pertama, artinya setiap satu juta rupiah pengeluaran konsumsi kacang-kacangan mampu menurunkan 0,522 juta rata-rata pengeluaran makanan perkapita dengan asumsi peubah lain dianggap konstan atau dengan kata lain peningkatan konsumsi kacang-kacang akan menghemat pengeluaran makanan/kapita Provinsi Jawa Timur. Pada komponen utama kedua, peubah buah-buahan memberikan kontribusi terbesar (0,618).

Komponen utama pertama data II, peubah keperluan pesta dan upacara ( $X_6$ ) memberikan sokongan terbesar komponen utama pertama sebesar 0,664 yang artinya setiap satu juta rupiah pengeluaran keperluan pesta dan upacara mampu meningkatkan 0,664 juta rata-rata pengeluaran non-makanan perkapita dengan asumsi peubah lain dianggap konstan.

Komponen utama pertama data III, peubah kredit bank ( $X_6$ ) memberikan dukungan terbesar terhadap komponen utama pertama sebesar 0,661 yang artinya setiap satu juta rupiah peminjaman kredit usaha bank mampu meningkatkan 0,661 juta rata-rata pendapatan rumah tangga dengan asumsi peubah lain dianggap konstan.

#### 4.5.2. Penentuan Komponen Utama *Robust* (ROBPCA)

Penentuan jumlah komponen utama yang dipilih dilandasi pada kriteria:

❖ Akar ciri

Diperoleh 2 buah akar ciri yang memiliki nilai akar ciri 1 (Tabel 4.12). Berdasarkan hasil, terpilih 2 komponen utama meskipun hanya mampu menjelaskan 78,82% keragaman data (Draper and Smith, 1981:313). Jumlah komponen utama terpilih tersaji dalam Tabel 4.14. Lampiran 9 menyajikan hasil perhitungan akar ciri, proporsi keragaman dan proporsi kumulatif.

Tabel 4.14. Jumlah Komponen Utama Terpilih Metode ROBPCA

Data	Jumlah komponen terpilih
I	2
II	1
III	3

❖ Proporsi Kumulatif Keragaman

Persen keragaman kumulatif data I menggunakan metode ROBPCA ditampilkan dalam Tabel 4.15, untuk data II dan III disajikan pada Lampiran 9.

Tabel 4.15. Persen Keragaman Kumulatif Komponen Utama Terpilih

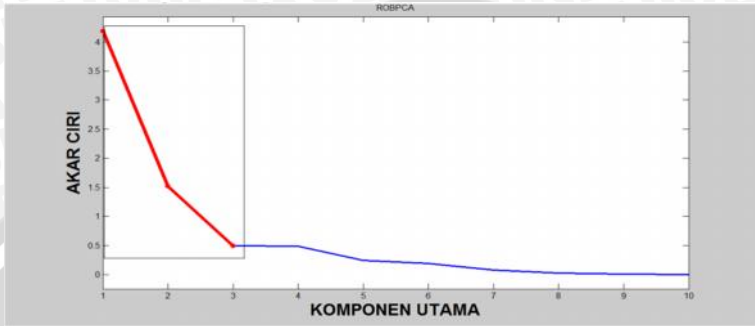
Data	Proporsi Kumulatif	Banyak Komponen
I	0,856	3
II	0,893	2
III	0,870	4

Pandang Tabel 4.12, komponen utama pertama menjelaskan 57,83% total keragaman, diikuti komponen utama kedua secara kumulatif menjelaskan 78,82% total keragaman data.

Data I (Tabel 4.15) membentuk 3 komponen utama terpilih, mampu menjelaskan 85,6% total keragaman data. Data II membentuk 2 komponen utama terpilih yang mampu menjelaskan 89,3% total keragaman data sedangkan data III membentuk 4 komponen utama terpilih yang mampu menjelaskan 87% total keragaman data (Draper and Smith, 1981).



## ❖ Scree Plot



Gambar 4.5. Plot Akar Ciri terhadap Komponen Utama Data I

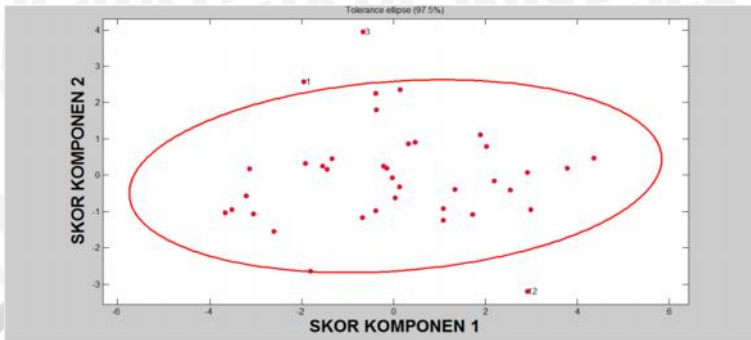
Gambar 4.5 memperlihatkan bahwa setelah komponen ketiga, perubahan nilai akar antar komponen cukup kecil. Menurut kriteria ini cukup dibutuhkan 3 komponen utama. *Scree plot* Data II dan III tersaji pada Lampiran 10. Jumlah komponen terpilih berdasarkan kriteria *scree plot* disajikan berikut (Johnson dan Wichern, 2002):

Tabel 4.16. Jumlah Komponen Utama Terpilih

Data	Banyak Komponen
I	3
II	2
III	3

### 4.5.3 Skor Komponen Utama

Skor komponen utama data I, II dan III disajikan pada Lampiran 11. Gambar 4.8 merupakan plot skor komponen 1 *robust* terhadap skor komponen 2 *robust*.



Gambar 4.6. Plot Skor Komponen 1 terhadap Skor Komponen 2

Ellipse tampak menyempit dibandingkan dengan Gambar 4.3, hal ini karena dipengaruhi rata-rata yang *robust* terhadap pencilan. *Tolerance ellipse* data II dan III tersaji pada Lampiran 12.

#### 4.6. Penentuan Metode yang Lebih Baik

Penentuan metode yang lebih baik antara *Clasic Principal Component Analysis* (CPCA) atau *Robust Principal Component Analysis* (ROBPCA) ditentukan berdasarkan:

- ❖ Komponen utama terpilih menurut kriteria akar ciri

Banyak komponen utama terpilih yang dihasilkan metode CPCA dan ROBPCA menurut kriteria akar ciri 1 disajikan dalam Tabel 4.17.

Tabel 4.17. Banyaknya Komponen Utama Terpilih

No	Data	Banyak komponen dengan akar ciri 1	
		CPCA	ROBPCA
1	I	3	2
2	II	2	1
3	III	3	3

Metode ROBPCA menghasilkan lebih sedikit jumlah komponen dibandingkan CPCA. Data I, metode ROBPCA menghasilkan 2 komponen utama, sedangkan metode CPCA menghasilkan 3. Untuk data II metode ROBPCA menghasilkan 1 komponen utama, sedangkan metode CPCA menghasilkan 2. Berdasarkan kriteria akar ciri 1 ,

ROBPCA lebih baik digunakan untuk data multivariat mengandung pencilan (Hubert et al., 2003).

- ❖ Komponen utama terpilih berdasarkan proporsi kumulatif keragaman

Banyaknya komponen utama terpilih berdasarkan persen keragaman kumulatif (ukuran yang mampu menjelaskan keragaman total data) minimum 80% yang dihasilkan metode CPCA dan ROBPCA disajikan dalam Tabel 4.18.

Tabel 4.18. Persen Keragaman Kumulatif Komponen Utama Terpilih

Data	Proporsi kumulatif (%)		Banyak komponen	
	CPCA	ROBPCA	CPCA	ROBPCA
I	0,850	0,857	4	3
II	0,855	0,893	2	2
III	0,807	0,870	4	4

Dengan jumlah komponen utama ROBPCA lebih sedikit dibanding CPCA, metode ROBPCA mampu menjelaskan total keragaman lebih besar (85,7%) bila dibandingkan dengan CPCA yang hanya mampu menjelaskan total keragaman sebesar 85%. Perbedaan 0,7% keragaman dianggap berarti karena keragaman berhubungan dengan koefisien ciri. Data II dengan jumlah komponen utama yang sama, metode ROBPCA mampu menjelaskan total keragaman lebih besar (89,3%) bila dibandingkan dengan CPCA yang hanya mampu menjelaskan total keragaman sebesar 85,5%. Demikian juga untuk data III, dengan jumlah komponen utama yang sama, metode ROBPCA mampu menjelaskan total keragaman lebih besar (87%) bila dibandingkan dengan CPCA yang hanya mampu menjelaskan total keragaman sebesar 80,7%.

Berdasarkan kriteria proporsi kumulatif keragaman, ROBPCA lebih baik digunakan untuk data multivariat mengandung pencilan (Hubert et al., 2003).

- ❖ Jumlah pencilan yang teridentifikasi.

Menurut kriteria jumlah pencilan yang teridentifikasi, metode CPCA dan ROBPCA yang disajikan berikut:

Tabel 4.19. Banyaknya Pencilan Teridentifikasi

No	Data	Jumlah pencilan yang teridentifikasi	
		CPCA	ROBPCA
1	I	3	14
2	II	4	14
3	III	3	15

Pandang Tabel 4.19, lebih banyak pencilan yang dideteksi oleh ROBPCA dibandingkan CPCA.

❖ *Scree plot*

Berikut jumlah komponen utama terpilih menurut kriteria ini:

Tabel 4.20. Jumlah Komponen Utama Terpilih

No	Data	Banyak Komponen	
		CPCA	ROBPCA
1	I	3	3
2	II	2	2
3	III	3	3

ROBPCA dan CPCA menghasilkan komponen utama yang sama banyak untuk data I, II dan III (Tabel 4.20).

Berdasarkan kriteria-kriteria yang telah disebutkan, maka untuk data yang mengandung pencilan, metode ROBPCA lebih baik digunakan dibanding metode CPCA (Hubert et al., 2003).

## BAB V KESIMPULAN DAN SARAN

### 5.1. Kesimpulan

Berdasarkan hasil analisis, disimpulkan bahwa metode ROBPCA menghasilkan lebih sedikit jumlah komponen dibandingkan CPCA menurut kriteria akar ciri 1 dan proporsi kumulatif keragaman. Data I, metode ROBPCA menghasilkan 2 komponen utama, sedangkan metode CPCA menghasilkan 3. Untuk data II metode ROBPCA menghasilkan 1 komponen utama, sedangkan metode CPCA menghasilkan 2. Dengan jumlah komponen utama ROBPCA lebih sedikit atau sama dibanding jumlah komponen CPCA, metode ROBPCA mampu menjelaskan total keragaman lebih besar dibandingkan dengan CPCA.

Berdasarkan kriteria-kriteria yang telah disebutkan, maka untuk data yang mengandung pencilan, metode ROBPCA lebih baik digunakan dibanding metode CPCA (Hubert et al., 2003).

### 5.2. Saran

Disarankan pada peneliti lain untuk mempelajari metode *robust* lain seperti *Minimum Volume Ellipsoid* (MVE) dan *Projection Pursuit* (PP) kemudian membandingkan dengan metode klasik dengan tidak mengikutsertakan pencilan dalam analisis perhitungan.

## DAFTAR PUSTAKA

- Branden, K.V. 2005. *Robust methods for high-dimensional data and a theoretical study of depth-related estimators*. Departement Wiskunde. Wetenschappen.
- Chen, Y. and D.A. Jackson. 2004. *Robust Principal Component Analysis and Outlier Detection with Ecological Data*. John Wiley and Son, Ltd. New York.
- Draper, N.R. and H. Smith. 1981. *Applied Regression Analysis*. Second Edition. John Wiley and Son, Inc. New York.
- Garzey, N. 1972. *Applied Multivariate Analysis Series in Psychology*. Mc Graw-Hill. New York.
- Harris, R.J. 1975. *A Primer Of Multivariate Statistics*. By Academic Pres, Inc. Universitas of New Mexico. Mexico.
- Hubert, M., A. Struyf and S.V. Aelst. 2003. *Theory and Applications of Recent Robust Methods*. EG&G Idaho, Inc. New York.
- Hubert, M., P.J. Rousseeuw and K.V. Branden. 2003. *ROBPCA: A New Approach to Robust Principal Component*. *Technometrics*. Vol.1 (64-79).  
<http://www.wis.kulueuven.ac.be/stat/ROBPCA.html>. Diakses 9 Juli 2012.
- Hubert, M. and P.J. Rousseeuw. 2007. *High Breakdown Robust Multivariate Methods*. K.U.Leuven. Belgia.  
<http://www.wis.kulueuven.ac.be/stat/robustmultivariat.html>. Diakses 9 Juli 2012.
- Johnson, R.A. and D.W. Wichern. 2002. *Applied Multivariate Statistical Analysis*. 5 th Ed. Prentice-Hall, Inc. New Jersey.
- Retno, A. 2005. *Minimum Covariance Determinant*. Tesis. Fakultas MIPA Universitas Pendidikan Indonesia. Jakarta.  
<http://www.upi.edu.ac.id.html>. Diakses 9 Juli 2012.
- Rousseeuw, P.J. and A. Leroy. 1987. *Robust Regression and Outlier Detection*. Wiley. New York.
- Rousseeuw, P.J. and K.V. Drissen. 1999. *A Fast Algorithm for the Minimum Covariance Determinant Estimator*. *Technometrics*, Vol.4 (212-223).

- Setyawan, B. 2007. *Minimum Volume Ellipsoid*. Tesis. Jakarta: Fakultas Mipa Universitas Pendidikan Indonesia. <http://www.upi.edu.ac.id.html>. Diakses 9 Juli 2012.
- Sujatmiko, I. 2005. *Analisis Komponen Utama dengan Menggunakan Matriks Varian Kovarian yang Robust*. Tesis. Surabaya: Fakultas MIPA Institut Teknologi Sepuluh Nopember.
- Suryana. 2007. *Analisis Data Outlier Pada Data Pengeluaran Rumah Tangga Di Kota Kupang Nusa Tenggara Timur Tahun 2005 Dengan Metode ROBPCA*. Paper. Surabaya: Fakultas MIPA Institut Teknologi Sepuluh Nopember.
- Susenas. 2011. Survei Sosial Ekonomi Nasional 2010-2011 Provinsi Jawa Timur. [www.jatim.bps.go.id/e-pub/susenas2011pub/](http://www.jatim.bps.go.id/e-pub/susenas2011pub/) . Diakses 9 Juli 2012.
- Verboven, S. and M. Hubbert. 2004. *LIBRA: A MATLAB Library for Robust Analysis*. <http://www.wis.kulueuven.ac.be/stat/robust.html>



Lampiran 1. Data Survei Sosial Ekonomi Nasional 2010-2011 Provinsi Jawa Timur tentang Rata-rata Pengeluaran Makanan/kapita dalam juta rupiah menurut Kabupaten atau Kota (Data I).

j	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	X <sub>8</sub>	X <sub>9</sub>	X <sub>10</sub>
1	36,564	3,421	69,75	91,18	13,420	23,002	99,88	86,02	123,30	12,010
2	40,097	2,682	54,83	68,82	12,512	20,794	94,93	76,12	110,15	11,346
3	39,164	4,578	90,23	59,88	12,122	20,573	77,50	69,66	121,47	9,347
4	42,280	1,438	86,50	46,52	11,874	22,658	78,28	99,45	125,48	8,961
5	42,012	1,891	77,75	53,25	12,722	23,397	77,65	97,85	121,73	10,724
6	41,027	1,606	55,22	55,42	11,668	19,900	106,56	86,86	104,09	10,092
7	43,496	1,227	86,83	60,14	13,858	21,232	113,96	91,45	107,94	12,380
8	52,731	0,706	142,58	63,79	11,548	17,708	123,39	63,32	981,1	11,960
9	50,135	0,709	104,77	50,58	9,806	16,890	105,14	55,85	101,76	10,655
10	48,924	1,701	126,26	60,50	12,318	23,536	114,22	95,69	108,69	9,831
11	50,920	1,105	124,75	47,81	11,493	17,645	126,15	65,55	102,37	14,082
12	50,407	0,538	1761,12	51,37	9,117	16,143	114,69	76,70	92,74	11,669
13	52,743	1,002	136,93	40,41	10,399	15,473	109,43	56,40	82,63	10,540
14	49,029	1,188	144,90	65,45	11,617	16,176	139,66	76,86	98,56	12,412
15	43,335	1,819	260,52	186,64	22,319	21,154	134,82	158,92	121,47	12,396
16	43,699	1,713	142,37	106,81	14,694	17,831	142,65	106,98	99,91	11,522
17	40,365	1,203	104,34	90,72	14,026	17,532	121,98	98,16	102,27	10,381
18	44,261	1,630	87,49	67,67	12,026	20,605	131,08	91,26	102,41	11,919
19	44,070	1,874	72,60	63,57	16,143	22,016	137,75	84,55	122,86	13,549
20	38,920	1,713	65,88	95,84	16,713	22,105	132,32	94,78	112,97	12,370
21	47,145	1,731	89,30	70,80	13,603	21,352	103,67	93,54	107,09	13,484
22	52,303	1,142	105,39	48,05	8,125	20,385	101,61	66,51	106,33	9,121
23	47,003	0,639	215,70	59,17	8,409	23,624	96,25	8,04	106,25	9,865
24	48,527	0,877	220,34	83,19	12,427	18,402	93,68	115,05	97,38	9,461
25	47,373	1,174	221,62	131,32	18,262	19,551	92,12	128,41	99,51	9,858



### Lampiran 1. (lanjutan)

26	52,600	0,925	216,35	67,25	9,749	16,539	94,64	76,61	94,75	9,528
27	51,242	1,712	206,67	25,45	6,390	13,192	71,42	55,86	77,32	7,274
28	47,825	0,997	216,53	60,65	7,539	12,431	67,40	56,22	86,94	7,677
29	48,011	1,074	208,21	71,44	8,130	13,221	62,05	49,88	93,91	11
30	37,164	1,959	100,34	115,51	21,406	21,597	88,68	154,55	103,63	11,613
31	38,382	1,596	98,65	104,12	23,089	21,405	87,45	156,97	101,06	10,393
32	37,392	2,060	173,21	193,55	22,691	21,500	138,46	171,43	113,34	12,554
33	45,560	2,014	218,15	106,60	20,160	18,062	130,82	138,44	96,11	10,914
34	44,767	1,228	184,27	138,33	27,699	18,356	153,43	143,60	96,76	12,459
35	36,708	1,697	169,69	152,43	26,958	19,410	133,41	177,86	97,56	10,806
36	33,865	1,741	115,19	172	27,502	21,045	110,33	150,75	108,07	11,851
37	42,750	2,215	236,62	215,67	31,005	22,130	108,97	151,43	112,71	12,088
38	41,817	2,082	128,40	124,15	18,819	26	133,56	158,63	120,51	12,883

Sumber : [Jatim.bps.go.id/e-pub/susenas2011pub](http://Jatim.bps.go.id/e-pub/susenas2011pub)

#### Keterangan

$X_1$  : Padi-padian (rupiah/minggu)

$X_2$  : Umbi-umbian (rupiah/minggu)

$X_3$  : Daging (rupiah/minggu)

$X_4$  : Telur dan susu (rupiah/bulan)

$X_5$  : Sayur-sayuran (rupiah/bulan)

$X_6$  : Kacang-kacangan (rupiah/bulan)

$X_7$  : Buah-buahan (rupiah/bulan)

$X_8$  : Minyak dan lemak (rupiah/bulan)

$X_9$  : Bahan minuman (rupiah/bulan)

$X_{10}$  : Bumbu-bumbuan (rupiah/bulan)

Lampiran 2. Data Survei Sosial Ekonomi Nasional 2010-2011 Provinsi Jawa Timur tentang Rata-rata Pengeluaran Non Makanan/kapita dalam juta rupiah menurut Kabupaten atau Kota (Data II).

j	$\frac{1}{x_1}$	$\frac{1}{x_2}$	$\frac{1}{x_3}$	$\frac{1}{x_4}$	$\frac{1}{x_5}$	$\frac{1}{x_6}$
1	70,998	669,23	13,328	26,673	56,67	23,03
2	63,932	612	12,470	17,558	62,69	34,78
3	60,886	601,77	12,049	34,825	55,33	180,10
4	78,947	696,61	15,163	38,256	72,14	85,55
5	81,393	681,88	13,775	33,306	59,11	40,27
6	72,676	754,10	13,185	25,421	63,46	90,80
7	74,030	944,65	15,774	26,566	56,37	69,17
8	63,850	569,03	13,053	19,257	60,56	6,27
9	59,540	645,79	14,863	23,170	48,85	79,19
10	78,352	763,74	15,256	43,420	84,25	28,35
11	61,287	507,80	13,169	13,267	41,08	25,68
12	60,879	593,27	16,107	12,569	45,86	34,17
13	59,060	557,35	11,787	20,869	42,75	42,88
14	68,149	631,86	11,545	12,065	47,57	56,01
15	130,809	1530,26	20,531	29,948	160,97	50,29
16	80,265	883,99	15,747	28,668	90,40	30,30
17	66,433	945,30	14,672	6,631	72,56	217,55
18	73,062	753,13	11,681	20,903	63,70	59,45
19	73,215	777,79	10,338	24,327	75,76	43,51
20	68,464	898,65	13,049	31,309	87,75	94,79
21	69,188	526,78	9,263	20,143	70,34	59,37
22	69,317	574,32	12,270	13,361	57,05	111,71
23	64,118	608,10	10,277	18,796	53,81	111,20
24	77,828	763,74	17,833	23,981	55,59	96,24

Lampiran 2. (lanjutan)

25	88,056	1048,77	15,171	33,247	99,03	182,56
26	59,047	461,81	11,518	8,323	35,05	24,09
27	49,337	486,87	11,375	8,249	23,80	48,64
28	58,214	397,98	12,832	14,426	27,15	18,32
29	59,986	381,03	12,819	8,611	21,92	5,29
30	146,171	1113,22	17,003	59,136	131,40	16,71
31	130,934	1325,25	24,574	46,982	121,11	24,26
32	188,025	1816,67	29,106	42,598	153,86	108,97
33	105,725	1222,65	18,095	41,474	97,49	99,70
34	139,288	1661,00	33,701	56,177	150,73	145,59
35	135,365	1621,17	19,217	75,094	173,08	63,42
36	170,076	1215,95	22,223	50,058	128,29	90,50
37	278,742	1966,42	25,004	27,404	271,95	36,04
38	98,158	1047,47	21,761	52,388	83,25	199,13

Sumber : [Jatim.bps.go.id/e-pub/konsumsi2010-2011](http://Jatim.bps.go.id/e-pub/konsumsi2010-2011)

Keterangan

$X_1$  : Perumahan dan fasilitas rumah tangga (rupiah/minggu)

$X_2$  : Aneka barang dan jasa termasuk kesehatan dan pendidikan (rupiah/minggu)

$X_3$  : Pakaian, alas kaki dan tutup kepala (rupiah/bulan)

$X_4$  : Barang tahan lama (rupiah/bulan)

$X_5$  : Pajak dan asuransi (rupiah/bulan)

$X_6$  : Keperluan pesta dan upacara (rupiah/bulan)

Lampiran 3. Data Survei Sosial Ekonomi Nasional 2010-2011 Provinsi Jawa Timur tentang rata-rata pendapatan Rumah Tangga yang menerima Kredit Usaha dalam juta rupiah Menurut Kabupaten atau Kota (Data III).

j	$\frac{1}{x_1}$	$\frac{1}{x_2}$	$\frac{1}{x_3}$	$\frac{1}{x_4}$	$\frac{1}{x_5}$	$\frac{1}{x_6}$	$\frac{1}{x_7}$
1	20,50	7,31	9,68	14,81	18,89	33,32	19,94
2	19,79	8,38	9,43	27,62	40,54	15,83	9,2
3	19,04	13,73	10,37	25,67	43,38	16,59	11,22
4	22,35	5,01	6,73	25,68	40,41	16,01	4,34
5	21,02	2,29	14,9	17,62	21,40	27,93	9,22
6	20,71	2,14	7,46	34,74	9,62	36,02	9,84
7	21,14	5,83	9,15	37,21	19,41	21,63	3,28
8	41,74	2,78	4,91	23,62	8,76	29,84	5,87
9	9,99	3,17	10,48	24,03	7,99	44,65	9,98
10	15,9	5,41	5,97	27,32	20,27	19,79	8,68
11	44,93	7,75	9,99	23,37	11,73	13,97	2,39
12	29,17	4,28	2,07	24,03	24,59	25,10	4,85
13	26,59	3,14	10,93	11,33	33,16	29,52	8,19
14	47,04	4,99	6,34	20,73	30,04	12,13	5,86
15	32,62	5,57	5,38	29,90	6,47	28,43	0
16	32,53	15,03	9,74	13,27	21,56	14,31	14,38
17	36,38	5,86	6,63	16,67	29,61	16,61	12,25
18	23,32	3,49	10,3	25,32	11,36	27,23	16,25
19	22,21	13,46	4,02	24,41	25,40	16,28	11,01
20	22,73	11,64	14,25	12,28	18,03	34,52	8,65
21	18,58	1,43	3,08	30,62	35,26	28,28	9,75
22	19,58	7,13	3,67	27,87	24,22	17,37	16,59
23	18,95	7,19	3,82	17,96	20,83	31,12	11,32
24	24,23	7,41	9,28	22,74	35,23	17,38	9,68

Lampiran 3. (lanjutan)

25	34,17	5,16	17,18	15,72	14,97	23,68	5,14
26	8,95	5,46	27,2	14,12	47,93	0	8,21
27	25,92	20,52	43,5	19,70	33,87	0	0
28	25,09	13,64	20,71	13,64	39,92	28,63	26,57
29	40,71	8,94	10,27	27,71	12,57	14,55	11,32
30	10,32	1,69	8,29	40,25	25,14	20,54	6,18
31	2,93	4,08	21,64	24,81	11,26	33,63	13,74
32	15,93	16,07	5,4	23,29	16,87	27,39	7,27
33	20,66	3,75	30,1	26,44	13,96	34,22	9,68
34	26,31	5,80	9	18,19	4,8	40,85	2,31
35	12,80	15,37	13,62	36,80	15,56	11,84	12,83
36	15,44	10,01	3,38	41,43	27,55	14,13	28,73
37	9,72	6,59	2,69	24,34	17,83	32,12	19,80
38	15,08	3,85	8,98	29,16	9,95	35,75	10,27

Sumber : [Jatim.bps.go.id/e-pub/susenas2011pub](http://Jatim.bps.go.id/e-pub/susenas2011pub)

Keterangan

$X_1$  : PNPM (rupiah/bulan)

$X_2$  : Program pemerintah lain (rupiah/bulan)

$X_3$  : KUR (rupiah/bulan)

$X_4$  : Program koperasi (rupiah/bulan)

$X_5$  : Perorangan (rupiah/bulan)

$X_6$  : Kredit Bank (rupiah/minggu)

$X_7$  : Lain-lain (rupiah/bulan)

### Lampiran 4. Skor baku Data I

j	$\frac{1}{n_1}$	$\frac{2}{n_2}$	$\frac{3}{n_3}$	$\frac{4}{n_4}$	$\frac{5}{n_5}$	$\frac{6}{n_6}$	$\frac{7}{n_7}$	$\frac{8}{n_8}$	$\frac{9}{n_9}$	$\frac{10}{n_{10}}$
1	-1,523	2,386	-0,417	0,056	-0,260	1,089	-0,398	-0,337	1,626	0,599
2	-0,853	1,421	-0,472	-0,427	-0,404	0,383	-0,610	-0,583	0,474	0,172
3	-1,030	3,898	-0,341	-0,620	-0,466	0,313	-1,354	-0,743	1,466	-1,113
4	-0,439	-0,205	-0,355	-0,908	-0,505	0,979	-1,321	-0,005	1,817	-1,361
5	-0,490	0,387	-0,387	-0,763	-0,371	1,215	-1,348	-0,044	1,489	-0,228
6	-0,677	0,014	-0,471	-0,716	-0,537	0,098	-0,113	-0,317	-0,057	-0,634
7	-0,208	-0,481	-0,353	-0,614	-0,191	0,523	0,203	-0,203	0,280	0,837
8	1,543	-1,162	-0,147	-0,535	-0,556	-0,603	0,606	-0,900	-0,581	0,567
9	1,051	-1,158	-0,287	-0,820	-0,832	-0,864	-0,174	-1,085	-0,261	-0,272
10	0,821	0,139	-0,207	-0,606	-0,435	1,259	0,214	-0,098	0,346	-0,802
11	1,199	-0,640	-0,213	-0,880	-0,565	-0,623	0,723	-0,845	-0,208	1,931
12	1,102	-1,381	5,857	-0,803	-0,941	-1,103	0,234	-0,568	-1,052	0,380
13	1,545	-0,775	-0,168	-1,040	-0,738	-1,317	0,009	-1,071	-1,938	-0,346
14	0,841	-0,532	-0,138	-0,499	-0,546	-1,092	1,300	-0,564	-0,542	0,857
15	-0,239	0,293	0,291	2,117	1,149	0,498	1,094	1,469	1,466	0,847
16	-0,170	0,154	-0,147	0,394	-0,058	-0,563	1,428	0,182	-0,423	0,285
17	-0,802	-0,512	-0,288	0,046	-0,164	-0,659	0,545	-0,037	-0,216	-0,448
18	-0,063	0,046	-0,351	-0,451	-0,481	0,323	0,934	-0,208	-0,204	0,540
19	-0,100	0,365	-0,406	-0,540	0,171	0,774	1,219	-0,374	1,588	1,588
20	-1,076	0,154	-0,431	0,157	0,261	0,802	0,987	-0,120	0,721	0,830
21	0,484	0,178	-0,344	-0,384	-0,231	0,562	-0,237	-0,151	0,206	1,547
22	1,462	-0,592	-0,285	-0,875	-1,099	0,253	-0,325	-0,821	0,139	-1,258
23	0,457	-1,249	0,125	-0,635	-1,054	1,288	-0,554	-2,269	0,132	-0,780
24	0,746	-0,938	0,142	-0,116	-0,417	-0,381	-0,663	0,382	-0,645	-1,040
25	0,527	-0,550	0,147	0,923	0,507	-0,014	-0,730	0,713	-0,458	-0,784
...	...	...	...	...	...	...	...	...	...	...
38	-0,527	0,636	-0,199	0,768	0,595	2,047	1,040	1,462	1,382	1,160

Lampiran 5. Skor baku Data II

j	$\frac{y_j - \bar{y}}{s_y}$	$\frac{y_j - \bar{y}}{s_y}$	$\frac{y_j - \bar{y}}{s_y}$	$\frac{y_j - \bar{y}}{s_y}$	$\frac{y_j - \bar{y}}{s_y}$	$\frac{y_j - \bar{y}}{s_y}$
1	-0,421	-0,505	-0,465	-0,125	-0,499	-0,898
2	-0,574	-0,645	-0,625	-0,694	-0,379	-0,682
3	-0,641	-0,670	-0,703	0,385	-0,526	1,985
4	-0,248	-0,438	-0,124	0,599	-0,190	0,250
5	-0,195	-0,474	-0,382	0,290	-0,450	-0,581
6	-0,384	-0,297	-0,492	-0,203	-0,363	0,346
7	-0,355	0,170	-0,011	-0,131	-0,505	-0,051
8	-0,576	-0,750	-0,516	-0,588	-0,421	-1,205
9	-0,670	-0,562	-0,180	-0,344	-0,655	0,133
10	-0,261	-0,273	-0,107	0,922	0,052	-0,800
11	-0,632	-0,900	-0,495	-0,963	-0,810	-0,849
12	-0,641	-0,691	0,051	-1,006	-0,715	-0,693
13	-0,680	-0,779	-0,752	-0,487	-0,777	-0,533
14	-0,483	-0,596	-0,797	-1,038	-0,681	-0,292
15	0,879	1,605	0,874	0,080	1,584	-0,397
16	-0,219	0,022	-0,016	0,000	0,175	-0,764
17	-0,520	0,172	-0,215	-1,377	-0,182	2,672
18	-0,376	-0,299	-0,771	-0,485	-0,358	-0,229
19	-0,373	-0,239	-1,021	-0,271	-0,118	-0,522
20	-0,476	0,057	-0,517	0,165	0,122	0,419
21	-0,460	-0,854	-1,221	-0,533	-0,226	-0,231
22	-0,457	-0,737	-0,662	-0,957	-0,491	0,730
23	-0,570	-0,654	-1,032	-0,617	-0,556	0,720
24	-0,272	-0,273	0,372	-0,293	-0,520	0,446
25	-0,050	0,425	-0,123	0,286	0,347	2,030
...	...	...	...	...	...	...
38	0,169	0,422	1,102	1,482	0,032	2,334

Lampiran 6. Skor baku Data III

j	$\frac{1}{n_1}$	$\frac{1}{n_2}$	$\frac{1}{n_3}$	$\frac{1}{n_4}$	$\frac{1}{n_5}$	$\frac{1}{n_6}$	$\frac{1}{n_7}$
1	-0,249	0,014	-0,134	-1,211	-0,306	0,961	1,559
2	-0,319	0,246	-0,164	0,465	1,595	-0,742	-0,147
3	-0,393	1,404	-0,052	0,210	1,845	-0,668	0,174
4	-0,067	-0,484	-0,485	0,211	1,584	-0,725	-0,919
5	-0,198	-1,073	0,488	-0,843	-0,086	0,436	-0,144
6	-0,228	-1,106	-0,398	1,397	-1,120	1,224	-0,045
7	-0,186	-0,307	-0,197	1,720	-0,261	-0,177	-1,087
8	1,842	-0,967	-0,702	-0,058	-1,196	0,622	-0,676
9	-1,284	-0,883	-0,039	-0,004	-1,264	2,065	-0,023
10	-0,702	-0,398	-0,575	0,426	-0,185	-0,357	-0,230
11	2,156	0,109	-0,097	-0,091	-0,935	-0,924	-1,229
12	0,605	-0,642	-1,040	-0,004	0,194	0,160	-0,838
13	0,351	-0,889	0,015	-1,666	0,943	0,591	-0,308
14	2,364	-0,488	-0,531	-0,436	0,673	-1,103	-0,678
15	0,944	-0,363	-0,646	0,764	-1,397	0,485	-1,608
16	0,935	1,685	-0,127	-1,412	-0,072	-0,890	0,676
17	1,315	-0,300	-0,497	-0,967	0,635	-0,666	0,337
18	0,029	-0,813	-0,060	0,164	-0,968	0,368	0,973
19	-0,081	1,345	-0,808	0,045	0,266	-0,699	0,140
20	-0,029	0,951	0,410	-1,542	-0,382	1,078	-0,234
21	-0,438	-1,259	-0,919	0,858	1,132	0,470	-0,060
22	-0,339	-0,025	-0,849	0,498	0,162	-0,592	1,027
23	-0,402	-0,012	-0,831	-0,799	-0,136	0,747	0,190
24	0,118	0,035	-0,181	-0,173	1,129	-0,591	-0,071
25	1,097	-0,452	0,759	-1,092	-0,651	0,022	-0,792
...	...	...	...	...	...	...	...
38	-0,783	-0,735	-0,217	0,667	-1,091	1,198	0,023



Lampiran 7. Jarak Mahalanobis

• Data II

j	MD <sup>2</sup>	j	MD <sup>2</sup>
1	1,234	20	3,205
2	1,801	21	4,055
3	3,01	22	2,982
4	1,958	23	2,474
5	1,958	24	1,869
6	0,652	25	5,445
7	6,238	26	2,214
8	3,184	27	2,817
9	0,891	28	3,026
10	4,668	29	4,449
11	1,881	30	9,308
12	3,951	31	5,404
13	0,956	32	14,64
14	2,163	33	4,834
15	10,092	34	19,94
16	2,762	35	15,492
17	12,938	36	10,312
18	1,826	37	28,83
19	3,075	38	9,422

Titik kritis  $\chi^2_{0,025(6)} = 14,45$

Lampiran 7. (lanjutan)

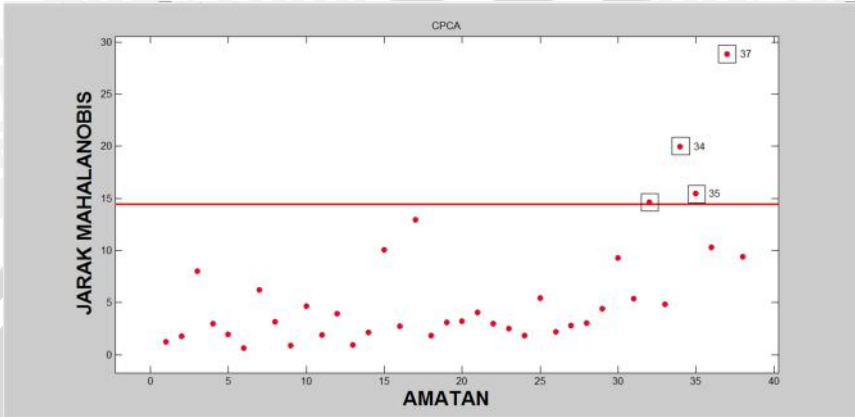
• Data III

j	MD <sup>2</sup>	j	MD <sup>2</sup>
1	4,777	20	5,647
2	3,817	21	6,676
3	7,374	22	3,127
4	5,237	23	3,009
5	2,905	24	1,539
6	4,496	25	3,732
7	4,511	26	20,19
8	5,216	27	21,24
9	5,425	28	18,32
10	3,176	29	6,529
11	6,814	30	7,185
12	2,598	31	8,177
13	5,567	32	9,969
14	7,359	33	12,08
15	5,065	34	6,021
16	7,411	35	9,176
17	4,090	36	15,03
18	4,088	37	4,971
19	4,061	38	2,378

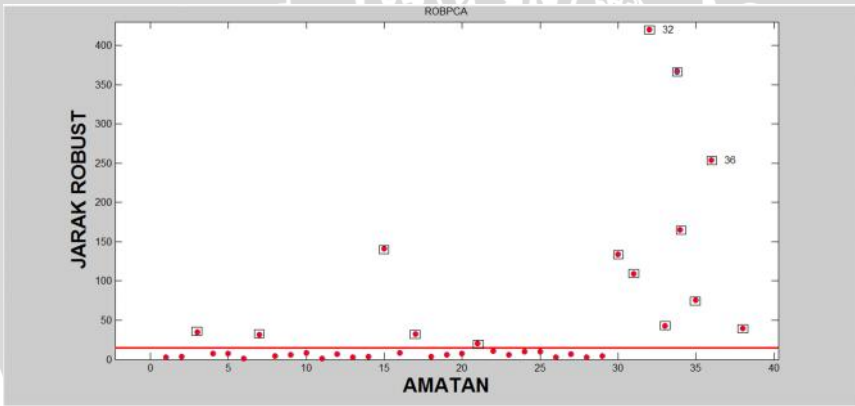
Titik kritis  $\chi^2_{0,025(7)} = 16,01$

Lampiran 8. Plot Jarak Mahalanobis dan Jarak *Robust*

- Data II



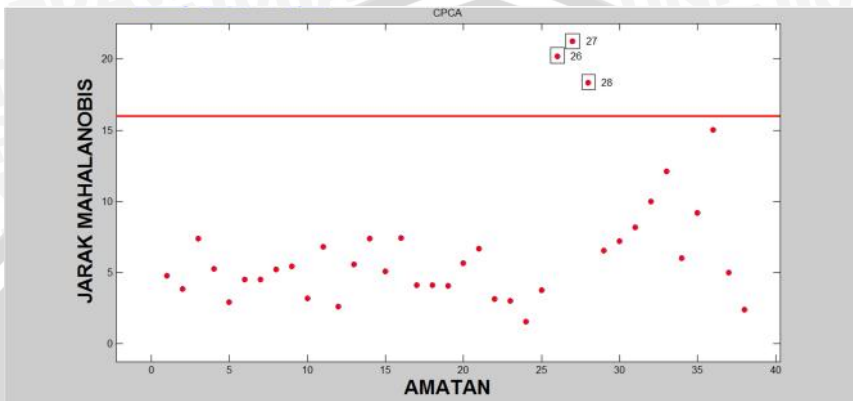
Gambar 4.9. Plot Jarak *Mahalanobis*



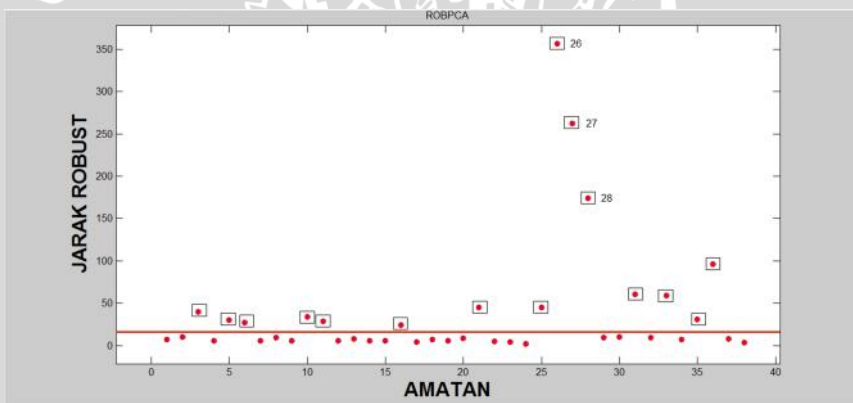
Gambar 4.10. Plot Jarak *Robust*

Lampiran 8. (lanjutan)

- Data III



Gambar 4.11. Plot Jarak Mahalanobis



Gambar 4.12. Plot Jarak Robust

Lampiran 9. Akar ciri, Proporsi Keragaman dan Proporsi Keragaman Kumulatif Metode CPCA dan ROBPCA.

- Data II

KU	CPCA			ROBPCA		
	Akar ciri	Proporsi keragaman	Proporsi kumulatif	Akar ciri	Proporsi keragaman	Proporsi kumulatif
1	4,1107	0,685	0,685	1,050	0,663	0,663
2	1,0193	0,170	0,855	0,363	0,229	0,893
3	0,4997	0,083	0,938	0,197	0,061	0,954
4	0,2707	0,045	0,983	0,055	0,035	0,989
5	0,0764	0,013	0,996	0,010	0,006	0,995
6	0,0234	0,004	1,000	0,008	0,005	1,000

- Data III

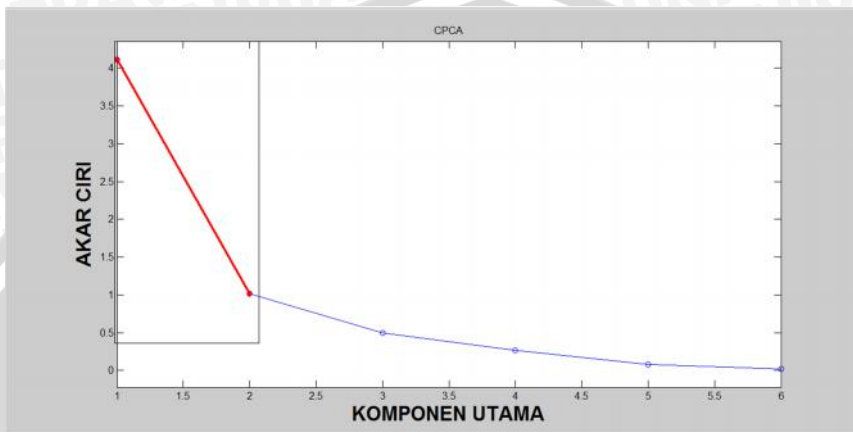
KU	CPCA			ROBPCA		
	Akar ciri	Proporsi keragaman	Proporsi kumulatif	Akar ciri	Proporsi keragaman	Proporsi kumulatif
1	2,1048	0,301	0,301	1,532	0,324	0,324
2	1,4718	0,210	0,511	1,101	0,233	0,557
3	1,1217	0,160	0,671	1,053	0,202	0,759
4	0,9496	0,136	0,807	0,724	0,111	0,870
5	0,7864	0,112	0,919	0,482	0,102	0,971
6	0,4080	0,058	0,977	0,124	0,026	0,998
7	0,1577	0,023	1,000	0,011	0,002	1,000

Keterangan:

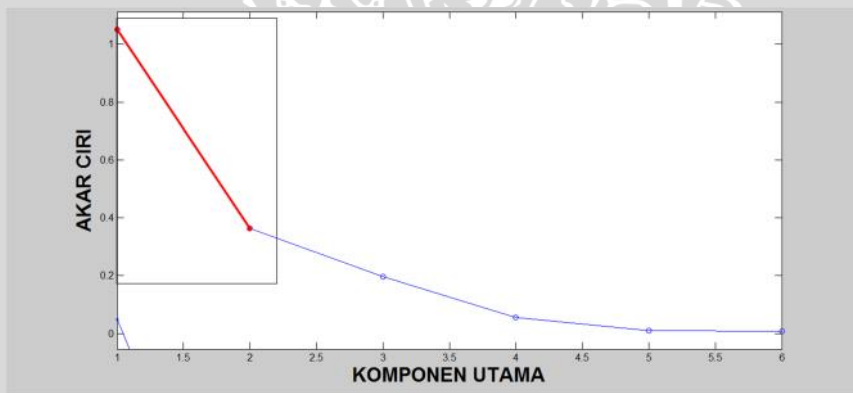
KU : Komponen Utama

Lampiran 10. *Scree Plot*

- Data II



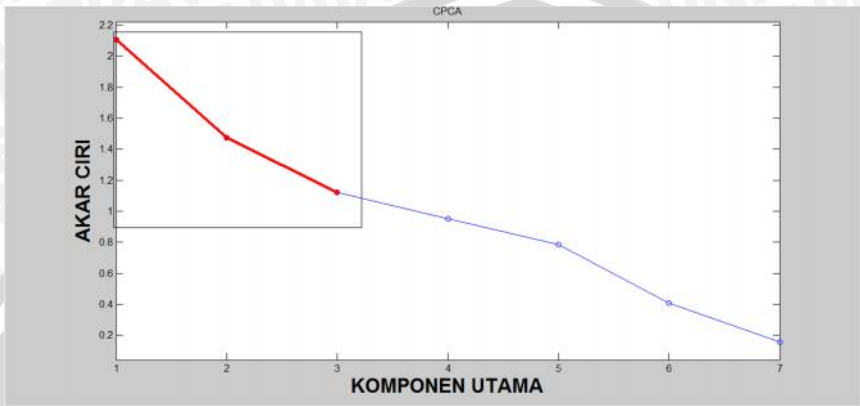
Gambar 4.13. *Scree Plot* Metode CPCA



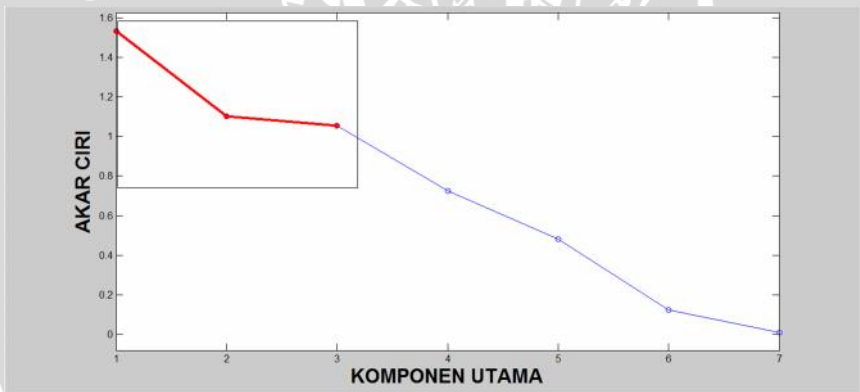
Gambar 4.14. *Scree Plot* Metode ROBPCA

Lampiran 10. (lanjutan)

- Data III



Gambar 4.15. Scree Plot Metode CPCA



Gambar 4.16. Scree Plot Metode ROBPCA

Lampiran 11. Skor Komponen Utama (SK)

• Data I

j	CPCA		ROBPCA	
	SK 1	SK 2	SK 1	SK 2
1	-1,943	2,580	-2,927	-1,727
2	-0,368	1,807	-1,173	-1,232
3	-0,663	3,950	-1,561	-3,428
4	0,149	2,362	-0,865	-2,428
5	-0,384	2,255	-1,446	-1,926
6	0,471	0,914	-0,172	-0,786
7	-0,139	0,191	-1,005	0,424
8	1,727	-1,074	0,966	1,699
9	2,196	-0,160	1,422	0,517
10	0,325	0,860	-0,605	-0,523
11	1,080	-0,913	0,043	2,190
12	2,913	-3,194	2,033	1,478
13	2,991	-0,942	2,542	1,111
14	1,082	-1,236	0,443	1,948
15	-3,046	-1,063	-2,949	0,621
16	-0,386	-0,972	-0,590	1,089
17	0,140	-0,313	-0,069	0,060
18	-0,025	-0,062	-0,799	0,735
19	-1,337	0,451	-2,560	0,961
20	-1,541	0,259	-2,235	0,348
21	-0,212	0,259	-1,108	0,612
22	2,029	0,802	1,042	-0,499
23	1,898	1,118	0,521	-0,598
24	1,344	-0,392	1,171	-0,389
25	0,038	-0,619	0,252	-0,475
26	2,536	-0,396	2,130	0,053
27	4,364	0,466	4,222	-1,282
28	3,785	0,197	3,777	-1,290
29	2,921	0,081	2,583	-0,381
30	-1,917	0,321	-1,760	-1,059
⋮	...	...	...	...
38	-3,125	0,184	-3,747	0,214



Lampiran 11. (lanjutan)

- Data II

j	CPCA		ROBPCA	
	Sk1	Sk2	Sk1	Sk2
1	0,245	-0,646	-1,017	0,722
2	0,437	-0,114	-1,357	0,580
3	-1,772	1,234	-0,805	-2,116
4	-1,088	-0,423	-0,215	-0,388
5	-0,248	-0,722	-0,649	0,425
6	-0,671	0,235	-0,741	-0,419
7	-0,642	-0,225	-0,374	-0,100
8	0,771	-0,551	-1,398	1,050
9	-0,255	0,240	-1,073	-0,331
10	-0,698	-1,452	-0,016	0,604
11	0,921	0,043	-1,769	0,681
12	0,645	0,014	-1,387	0,516
13	0,470	0,026	-1,617	0,334
14	0,448	0,462	-1,594	0,235
15	-2,154	-1,636	2,279	0,763
16	-0,448	-0,977	-0,099	0,704
17	-1,902	2,449	-0,554	-2,507
18	-0,113	0,045	-1,027	0,182
19	-0,095	-0,306	-0,948	0,489
20	-1,187	-0,078	-0,252	-0,476
21	0,152	0,218	-1,487	0,223
22	-0,311	1,089	-1,360	-0,717
23	-0,402	0,979	-1,438	-0,762
24	-0,796	0,199	-0,393	-0,557
25	-2,667	0,771	0,620	-1,912
26	1,231	0,318	-2,155	0,735
27	1,034	0,722	-2,289	0,225
28	1,197	0,028	-2,073	0,746
29	1,580	0,096	-2,287	1,009
30	-1,996	-2,772	1,990	1,157
⋮	...	...	...	...
38	-3,558	0,078	1,592	-2,423

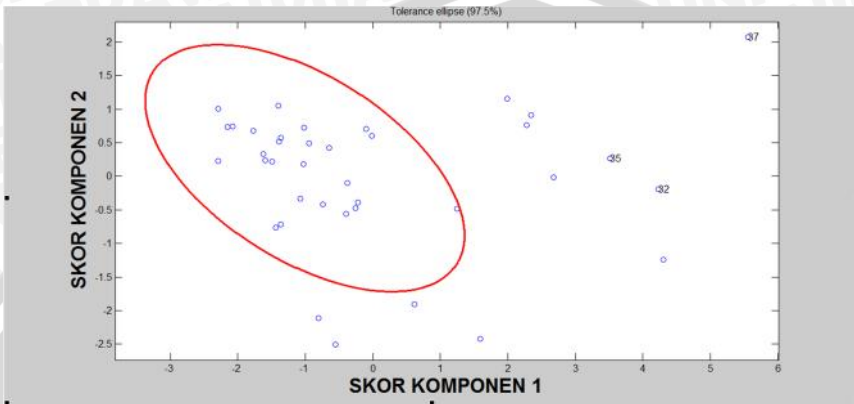
Lampiran 11. (lanjutan)

• Data III

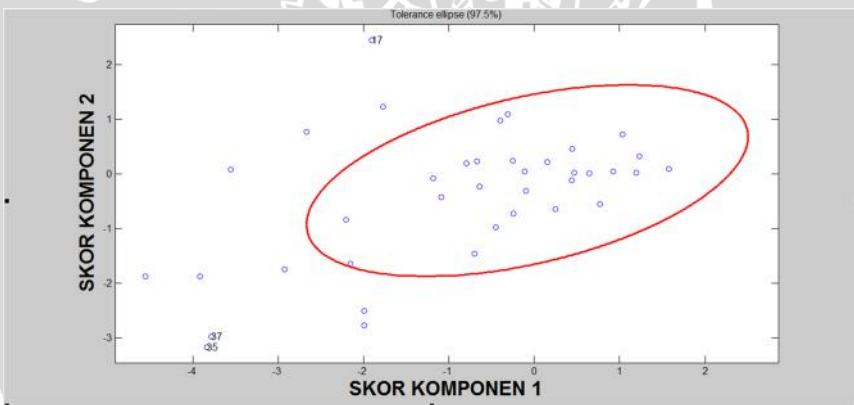
j	CPCA		ROBPCA	
	Sk1	Sk2	Sk1	Sk2
1	-0,468	0,586	1,125	0,008
2	1,114	0,713	-1,402	1,427
3	1,850	1,047	-1,439	1,659
4	0,710	-0,075	-1,593	0,843
5	-0,432	-0,452	0,423	-0,298
6	-2,275	0,254	1,352	0,175
7	-0,680	0,016	-0,283	0,700
8	-1,480	-1,987	-0,028	-1,997
9	-2,336	0,404	2,592	0,108
10	-0,475	0,435	0,009	0,836
11	0,300	-2,374	-1,454	-1,870
12	-0,643	-0,860	-0,565	-0,439
13	0,145	-0,918	-0,308	-0,691
14	0,840	-1,967	-2,431	-1,390
15	-1,472	-1,636	0,190	-1,251
16	1,648	-0,520	-1,009	-0,517
17	0,705	-0,856	-1,429	-0,663
18	-1,134	0,304	0,826	0,014
19	0,838	0,442	-0,778	0,771
20	0,192	-0,706	0,849	-0,789
21	-0,910	0,741	-0,264	1,136
22	-0,083	1,127	-0,344	1,203
23	-0,658	0,104	0,681	0,041
24	0,886	0,105	-1,194	0,623
25	0,109	-1,856	-0,269	-1,601
26	3,122	0,619	-2,035	1,829
27	4,820	-1,041	-2,484	0,342
28	1,942	1,370	-0,224	0,825
29	0,264	-0,996	-1,076	-0,876
30	-1,041	1,158	-0,030	1,902
⋮	...	...	...	...
38	-1,877	0,430	1,632	0,329

Lampiran 12. *Tolerance ellipse*

- Data II



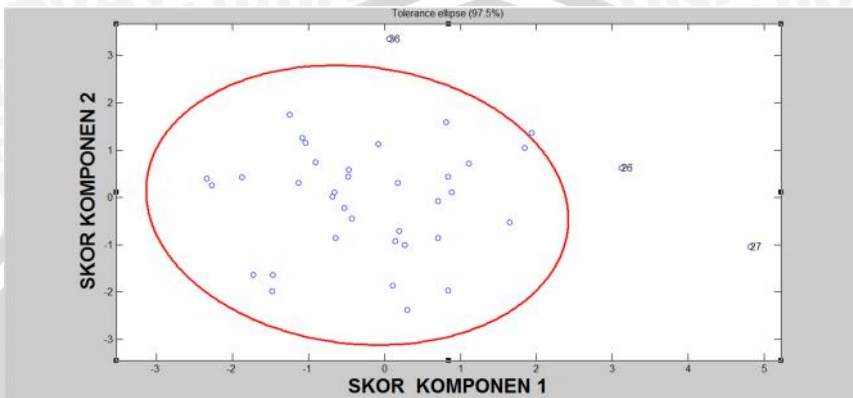
Gambar 4.17. *Tolerance ellipse* CPCA



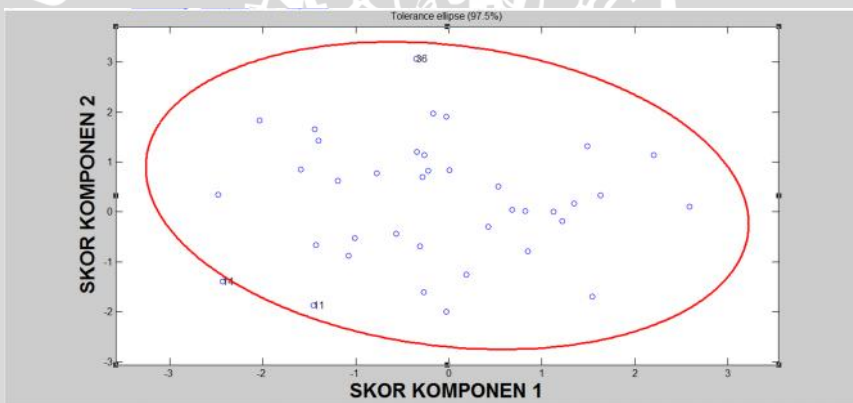
Gambar 4.18. *Tolerance ellipse* ROBPCA

Lampiran 12. (lanjutan)

- Data III



Gambar 4.19. *Tolerance ellipse CPCA*



Gambar 4.20. *Tolerance ellipse ROBPCA*

Lampiran 13. Pengamatan dengan Determinan Matriks Ragam-peragam Terkecil yang tersusun sesuai Jarak Mahalanobis.

- Data II

Pengamatan											
6	9	13	1	2	18	24	11	5	14	26	

Pengamatan											
23	16	27	4	28	19	8	20	12	29	10	

- Data III

Pengamatan											
24	38	12	5	23	22	10	25	2	17	19	18

Pengamatan											
6	7	1	37	15	8	4	9	13	34	14	



Lampiran 14. Vektor Rata-rata dan Matriks Ragam-peragam *Robust* dari Penduga MCD

- Data I

Vektor Rata-rata

$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	$\bar{x}_4$	$\bar{x}_5$	$\bar{x}_6$	$\bar{x}_7$	$\bar{x}_8$	$\bar{x}_9$	$\bar{x}_{10}$
0,29	-0,197	0,036	-0,51	-0,513	-0,273	-0,104	-0,455	-0,14	-0,028

Matriks Ragam-peragam

	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	$\bar{x}_4$	$\bar{x}_5$	$\bar{x}_6$	$\bar{x}_7$	$\bar{x}_8$	$\bar{x}_9$	$\bar{x}_{10}$
$\bar{x}_1$	0,834									
$\bar{x}_2$	-0,556	0,710								
$\bar{x}_3$	0,34	-0,402	1,578							
$\bar{x}_4$	-0,192	0,101	-0,086	0,161						
$\bar{x}_5$	-0,229	0,139	-0,166	0,101	0,16					
$\bar{x}_6$	-0,614	0,526	-0,385	0,108	0,327	1,155				
$\bar{x}_7$	-0,070	-0,017	0,008	0,152	0,255	0,296	1,000			
$\bar{x}_8$	-0,226	0,110	-0,062	0,099	0,13	0,337	0,144	0,205		
$\bar{x}_9$	-0,703	0,521	-0,401	0,131	0,312	1,042	0,185	0,279	1,189	
$\bar{x}_{10}$	-0,165	0,148	-0,01	0,133	0,269	0,462	0,742	0,069	0,498	1,224

- Data II

Vektor Rata-rata

$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	$\bar{x}_4$	$\bar{x}_5$	$\bar{x}_6$
-0,492	-0,544	-0,479	-0,408	-0,491	-0,382

Matriks Ragam-peragam

	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	$\bar{x}_4$	$\bar{x}_5$	$\bar{x}_6$
$\bar{x}_1$	0,036	0,049	0,028	0,087	0,055	0,025
$\bar{x}_2$		0,121	0,039	0,147	0,118	0,094
$\bar{x}_3$			0,129	0,087	0,036	0,015
$\bar{x}_4$				0,360	0,180	0,099
$\bar{x}_5$					0,150	0,064
$\bar{x}_6$						0,324

Lampiran 14. (lanjutan)

- Data III

Vektor Rata-rata

$\frac{\sum x_1}{N}$	$\frac{\sum x_2}{N}$	$\frac{\sum x_3}{N}$	$\frac{\sum x_4}{N}$	$\frac{\sum x_5}{N}$	$\frac{\sum x_6}{N}$	$\frac{\sum x_7}{N}$
0,124	-0,38	-0,36	-0,071	-0,149	0,251	-0,131

Matriks Ragam-peragam

	$\frac{\sum x_1^2}{N}$	$\frac{\sum x_2^2}{N}$	$\frac{\sum x_3^2}{N}$	$\frac{\sum x_4^2}{N}$	$\frac{\sum x_5^2}{N}$	$\frac{\sum x_6^2}{N}$	$\frac{\sum x_7^2}{N}$
$\frac{\sum x_1 x_2}{N}$	0,823	-0,059	0,002	-0,199	0,068	-0,285	-0,336
$\frac{\sum x_1 x_3}{N}$	-0,059	0,287	-0,075	-0,001	0,170	-0,197	0,091
$\frac{\sum x_1 x_4}{N}$	0,002	-0,075	0,204	-0,110	-0,028	0,062	-0,042
$\frac{\sum x_1 x_5}{N}$	-0,199	-0,001	-0,110	0,685	-0,143	-0,050	-0,111
$\frac{\sum x_1 x_6}{N}$	0,068	0,170	-0,028	-0,143	0,875	-0,581	0,048
$\frac{\sum x_1 x_7}{N}$	-0,285	-0,197	0,062	-0,050	-0,581	0,737	0,051
$\frac{\sum x_2 x_3}{N}$	-0,336	0,091	-0,042	-0,111	0,048	0,051	0,689

Lampiran 15. Jarak *Robust*

• Data II

j	RD <sup>2</sup>	j	RD <sup>2</sup>
1	2,579	20	7,880
2	3,420	21	20,489
3	35,03	22	10,402
4	7,653	23	6,184
5	7,906	24	9,727
6	1,384	25	10,329
7	31,69	26	2,427
8	4,727	27	7,129
9	6,307	28	3,167
10	8,459	29	4,468
11	1,361	30	133,667
12	6,883	31	109,087
13	3,162	32	420,284
14	3,890	33	42,931
15	140,7	34	165,069
16	8,471	35	75,566
17	31,993	36	254,1
18	3,763	37	407
19	6,322	38	39,48

Titik kritis  $\chi_{0.025(6)}^2 = 14,4$



Lampiran 15. (lanjutan)

• Data III

j	RD <sup>2</sup>	j	RD <sup>2</sup>
1	6,943	20	8,794
2	10,039	21	45,354
3	39,764	22	5,291
4	5,379	23	3,941
5	30,401	24	2,344
6	27,066	25	45,339
7	5,549	26	356,355
8	9,246	27	262,008
9	5,798	28	173,824
10	33,574	29	9,730
11	28,571	30	9,796
12	5,422	31	60,472
13	7,826	32	9,613
14	6,015	33	58,942
15	5,835	34	6,959
16	24,613	35	31,262
17	4,544	36	96,298
18	7,388	37	8,282
19	6,017	38	3,248

Titik kritis  $\chi^2_{0.025(7)} = 16,01$

Lampiran 16. Vektor Rata-Rata Dan Matriks Ragam Peragam *Robust* dari Penduga RMCD

- Data II

Vektor Rata-rata

$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\mu}_3$	$\hat{\mu}_4$	$\hat{\mu}_5$	$\hat{\mu}_6$
-0,479	-0,540	-0,475	-0,438	-0,486	-0,270

Matriks Ragam-peragam

	$\hat{\sigma}_{11}$	$\hat{\sigma}_{12}$	$\hat{\sigma}_{13}$	$\hat{\sigma}_{14}$	$\hat{\sigma}_{15}$	$\hat{\sigma}_{16}$
$\hat{\mu}_1$	0,042	0,065	0,034	0,097	0,069	0,071
$\hat{\mu}_2$	0,065	0,153	0,054	0,176	0,149	0,182
$\hat{\mu}_3$	0,034	0,054	0,125	0,099	0,048	0,042
$\hat{\mu}_4$	0,097	0,176	0,099	0,364	0,199	0,146
$\hat{\mu}_5$	0,069	0,149	0,048	0,199	0,169	0,149
$\hat{\mu}_6$	0,071	0,182	0,042	0,146	0,149	0,591

- Data III

Vektor Rata-rata

$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\mu}_3$	$\hat{\mu}_4$	$\hat{\mu}_5$	$\hat{\mu}_6$	$\hat{\mu}_7$
0,115	-0,161	-0,397	-0,025	-0,124	0,207	-0,128

Matriks Ragam-peragam

	$\hat{\sigma}_{11}$	$\hat{\sigma}_{12}$	$\hat{\sigma}_{13}$	$\hat{\sigma}_{14}$	$\hat{\sigma}_{15}$	$\hat{\sigma}_{16}$	$\hat{\sigma}_{17}$
$\hat{\mu}_1$	0,975	-0,067	-0,021	-0,217	0,015	-0,352	-0,257
$\hat{\mu}_2$	-0,067	0,565	-0,016	-0,146	0,053	-0,113	0,084
$\hat{\mu}_3$	-0,021	-0,016	0,137	-0,062	-0,020	0,074	-0,007
$\hat{\mu}_4$	-0,217	-0,146	-0,062	0,829	-0,069	-0,242	-0,185
$\hat{\mu}_5$	0,015	0,053	-0,020	-0,069	0,857	-0,554	0,026
$\hat{\mu}_6$	-0,352	-0,113	0,074	-0,242	-0,554	0,773	0,031
$\hat{\mu}_7$	-0,257	0,084	-0,007	-0,185	0,026	0,031	0,690

Lampiran 17. Vektor Ciri dengan Penduga RMCD

- Data I

Peubah	Komponen Utama 1	Komponen Utama 2	Komponen Utama 3	...	Komponen Utama 10
Padi-padian	0,326	0,327	-0,102	...	-0,180
Umbi-umbian	-0,256	-0,288	0,393	...	-0,045
Daging	0,075	0,009	-0,067	...	0,943
Telur dan susu	-0,131	0,045	-0,183	...	-0,159
Sayur-sayuran	-0,198	0,075	-0,156	...	0,007
Kacang-kacangan	-0,522	-0,177	-0,270	...	0,142
Buah-buahan	-0,268	0,618	-0,353	...	0,052
Minyak dan lemak	-0,208	-0,048	-0,485	...	-0,158
Bahan minuman	-0,492	-0,236	0,079	...	-0,047
Bumbu-bumbuan	-0,369	0,576	0,578	...	0,031

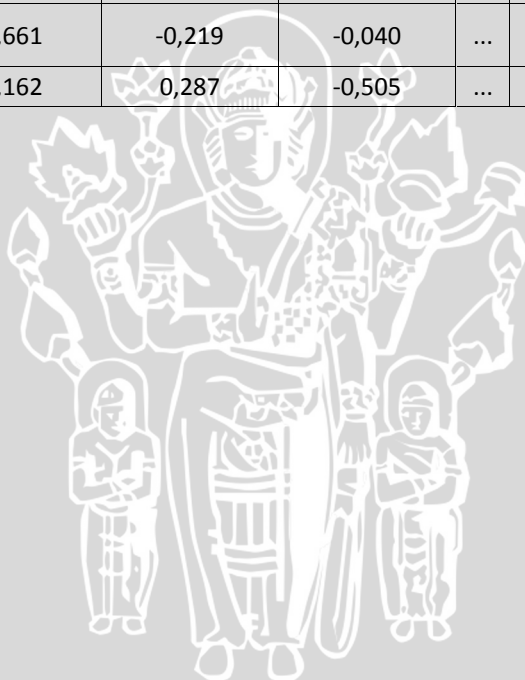
- Data II

Peubah	Komponen Utama 1	Komponen Utama 2	Komponen Utama 3	...	Komponen Utama 6
Perumahan dan fasilitas rumah tangga	-0,171	-0,119	0,004	...	-0,049
Aneka barang dan jasa	-0,370	-0,133	-0,144	...	-0,716
Pakaian, alas kaki dan tutup kepala	-0,152	-0,212	0,935	...	0,101
Barang tahan lama	-0,487	-0,588	-0,073	...	-0,070
Pajak dan asuransi	-0,365	-0,245	-0,305	...	0,683
Keperluan pesta dan upacara	0,664	0,719	0,086	...	0,065

Lampiran 17. (lanjutan)

• Data III

Peubah	Komponen Utama 1	Komponen Utama 2	Komponen Utama 3	...	Komponen Utama 7
PNPM	-0,479	-0,721	0,045	...	0,386
Program pemerintah lain	-0,051	0,104	-0,335	...	0,218
KUR	0,052	-0,041	-0,037	...	-0,007
Program koperasi	-0,063	0,431	0,735	...	0,410
Perorangan	-0,546	0,389	-0,295	...	0,430
Kredit Bank	0,661	-0,219	-0,040	...	0,646
Lain-lain	0,162	0,287	-0,505	...	0,185



Lampiran 18. Nilai VIF

• Data I

Peubah	VIF
X1	2,93
X2	1,98
X3	1,18
X4	5,79
X5	9,76
X6	4,09
X7	2,19
X8	5,30
X9	4,89
X10	2,38

• Data II

Peubah	VIF
X1	11,43
X2	20,43
X3	5,72
X4	1,98
X5	21,94
X6	1,29

• Data III

Peubah	VIF
X1	11,79
X2	1,53
X3	1,47
X4	1,77
X5	2,25
X6	13,05
X7	1,28

Nilai VIF  $\geq 5$  menunjukkan bahwa terjadi multikolinieritas antar peubah prediktor.

Lampiran 19. Nilai VIF setelah dilakukan pereduksian

• Data I

Peubah	VIF
X1	1,33
X2	1,15
X3	1,74
X4	2,26
X5	1,94
X6	3,36
X7	1,28
X8	1,25
X9	3,72
X10	1,52

• Data II

Peubah	VIF
X1	3,65
X2	2,12
X3	1,72
X4	1,98
X5	1,94
X6	1,29

• Data III

Peubah	VIF
X1	1,79
X2	1,53
X3	1,47
X4	1,77
X5	2,25
X6	3,05
X7	1,28

Nilai  $VIF \leq 5$  menunjukkan bahwa tidak terjadi multikolinieritas antar peubah prediktor.