PENGGUNAAN ANALISIS KOMPONEN UTAMA UNTUK MENGATASI MULTIKOLINIERITAS DALAM ANALISIS REGRESI LINIER BERGANDA

SKRIPSI

Diajukan untuk Memenuhi Persyaratan Penyelesaian Program Sarjana Sains
Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Jember

Oleh :

Vita Desy Andriyani
NIM. 981810101029

JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS JEMBER
2003

MOTTO

" Jadikanlah sabar dan sholat sebagai penolongmu, Dan sesungguhnya yang demikian itu berat, kecuali bagi orang-orang yang beriman "
(Q S. Albaqarah, 153)

"Orang yang gagal meraih sesuatu yang hebat, tidak bisa dikatakan gagal total. Dia selalu yakin dan percaya bahwa paling tidak dia telah memenangkan perang terpenting dalam kehidupan, yaitu mengalahkan rasa takut untuk mencoba"
(Robert H. Schuller)

" Barang siapa yang menghendaki dunia, Maka carilah itu dengan ilmu. Barang siapa menghendaki akhirat, Maka carilah dengan ilmu. Barang siapa yangmenghendaki dua-duanya, Maka carilah dengan ilmu "
(Khutbatul Ali Rodliyallahu' anhu)

"Bukanlah suatu aib jika anda gagal dalam suatu usaha, yang merupakan aib ialah jika anda tidak berusaha bangkit dari kegagalan itu"
(Ali bin Abi Tholib ra.)

"Mencintai kehidupan dengan bekerja adalah menyelami rahasia hidup yang paling dalam"
(Kahlil Gibran)

"Gapailah langit, karena jika melesetpun kau akan tetap berada diantara bintang-bintang"
(Rossa Torcasio)

Kupersembahkan Karya Tulis ini untuk:

- Ayahanda " M. Solikhin " dan Ibunda " Sukartini " tercinta, terima kasih atas semua pengorbanan, kasih sayang, nasehat serta doa yang tiada henti-hentinya.
- Adikku " Andry Wahyudy " yang paling aku sayangi.
- Dosen Pembimbingku " Tanpa Tanda Jasa " yang menjadi pelita harapanku, trimakasih atas bimbingan, nasehat dan ilmu yang telah diberikan menjadi penerang dalam kegelapan, Jasamu Tiada Tara.
- Kekasihku " Edy Satrio Nugroho " yang selalu memberiku semangat dan dorongan.
- Seseorang yang akan membimbing & mendampingiku kelak.
- Almamaterku yang kubanggakan.

DEKLARASI

Karya Tulis Ilmiah (Skripsi) ini berisi hasil kerja penelitian mulai bulan Januari 2003 sampai dengan bulan Juni 2003. Bersama ini saya nyatakan bahwa isi Karya Tulis Ilmiah (Skripsi) ini adalah hasil pekerjaan saya sendiri kecuali jika disebutkan sumbernya dan Karya Tulis Ilmiah (Skripsi) ini belum pernah diajukan pada instansi lain.

Jember, Juni 2003

Penulis,

(Vita Desy Andriyani)

ABSTRAK

Penggunaan Analisis Komponen Utama Untuk Mengatasi Multikolinieritas Dalam Analisis Regresi linier Berganda, Vita Desy Andriyani, 981810101029, Skripsi, Juni 2003, Jurusan Matematika, Fakultas Matematika Dan Ilmu Pengetahuan Alam, Universitas Jember.

Studi ini mengkaji tentang cara mendeteksi multikolinieritas dan penggunaan analisis komponen utama untuk mengatasi multikolinieritas dalam analisis regresi linier berganda. Analisis regresi linier berganda adalah analisis yang mempelajari hubungan kausal antara satu peubah tak bebas dengan dua atau lebih peubah bebas. Salah satu asumsi yang harus dipenuhi dalam analisis regresi linier berganda berhubungan dengan peubah bebas, peubah-peubah bebas tersebut tidak boleh ada korelasi atau terjadi multikolinieritas. Adanya multikolinieritas dapat dideteksi dengan memeriksa matriks korelasi, sedangkan untuk mengukur besarnya multikolinieritas dapat digunakan *Variance Inflation Factor* (VIF). Analisis komponen utama berusaha mereduksi K peubah pengamatan menjadi M peubah baru yang saling ortogonal, masing-masing M peubah baru tersebut merupakan kombinasi linier dari K peubah lama. Data yang dianalisis dalam penelitian ini ada dua jenis yaitu data simulasi dan data riil. Untuk data jenis simulasi didibangkitkan komputer dengan kriteria hubungan diantara dua peubah yang sempurna, kuat dan sedang. Metode analisis terhadap data dilakukan dengan memeriksa matriks korelasi dan menghitung nilai VIF, penerapan metode analisis komponen utama, kemudian membandingkan nilai koefisien regresi yang mengabaikan multikolinieritas dengan nilai koefisien regresi yang menggunakan metode analisis komponen utama. Hasil yang didapat menunjukkan bahwa multikolinieritas terdapat pada data yang memiliki kriteria hubungan diantara dua peubah yang sempurna dan kuat. Metode analisis komponen utama yang diterapkan pada analisis regresi mampu menghilangkan multikolinieritas, tetapi mengakibatkan koefisien regresi mempunyai bias yang lebih besar daripada metode kuadrat terkecil yang diterapkan secara langsung.

Kata kunci: Analisis Regresi Linier berganda, Multikolinieritas, Analisis Komponen Utama, Ortogonal, Variance Inflation Factor (VIF).

PENGESAHAN

Karya Tulis Ilmiah (Skripsi) ini diterima oleh Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember, pada :

Hari : JUM'AT

Tanggal : 27 JUN 2003

Tempat : Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember

Tim Penguji

Ketua (Dosen Pembimbing Utama)	Sekretaris (Dosen Pembimbing Anggota)
Drs. I Made Tirta, Dip.Sc, M.Sc, Ph.D	Rita Ratih T, S.Si, M.Si
NIP : 131 474 500	NIP : 132 243 343
Anggota I	Anggota II
Yuliani Setia Dewi, S.Si, M.Si	M. Fatkurohman, S.Si, M.Si
NIP : 132 258 183	NIP : 132 210 538

Mengesahkan :

Dekan F. MIPA Universitas jember

Ir. Sumadi, MS

NIP : 130 368 784

KATA PENGANTAR

Dengan memanjatkan puji syukur kehadirat Allah SWT, atas segala rahmat, taufiq serta hidayah – Nya sehingga penulis dapat menyelesaikan Tugas Akhir berupa Karya Tulis Ilmiah (Skripsi) ini.

Karya Tulis Ilmiah (Skripsi) ini disusun untuk memenuhi persyaratan Tugas Akhir (MAU 425) pada kurikulum pendidikan di jurusan Matematika, Fakultas MIPA Universitas Jember (UNEJ).

Dengan selesainya Karya Tulis Ilmiah (Skripsi) ini tidak lupa penulis mengucapkan terima kasih kepada :

1. Ibunda dan Ayahanda tercinta atas dorongan semangat dan petuah-petuah yang menuntunku menuju pendewasaan diri,
2. Bapak. Ir. Sumadi MS, selaku Dekan Fakultas MIPA Universitas Jember,
3. Bapak. Drs. Kusno, DEA, Ph.D, selaku Ketua Jurusan Matematika Fakultas MIPA Universitas Jember,
4. Bapak. Drs I Made Tirta M.Sc, Ph.D, selaku Dosen Pembimbing Utama (DPU) dari Jurusan Matematika Fakultas MIPA Universitas Jember,
5. Ibu Rita Ratih T, S.Si, M.Si, selaku Dosen Pembimbing Anggota (DPA) dari Jurusan Matematika Fakultas MIPA Universitas Jember,
6. Ibu Yuliani Setia Dewi, S.Si. M.Si, selaku Dosen Penguji Utama Karya Tulis Ilmiah (Skripsi) dari Jurusan Matematika Fakultas MIPA Universitas Jember,
7. Bapak. Fatkurahman S.Si. M.Si, selaku Dosen Penguji Anggota Karya Tulis Ilmiah (Skripsi) dari jurusan Matematika Fakultas MIPA Universitas Jember,
8. Sahabat-sahabatku sepejuangan " Mifta, Ninip, Lena, Mamad, Anis",
9. Sahabat-sahabatku yang telah membantuku "Indah, Nonik, Rina, Indri, Bagus",
10. Rekan-rekan seperjuangan angkatan '98 Matematika UNEJ,

11. Rekan-rekan seperjuangan angkatan '97 dan '99 Matematika UNEJ,
12. Penghuni kost F mastrip 12 Jember,
13. Saudaraku seperjuangan "Laily, Dewi, Kotery, Anis (sincan)",
14. Sahabat Bermainku "Uud dan Agus, Titin, M' Desy" ,
15. Teman-temanku KKN "Adi, Yuli, Helmi, Soni, Ningsih",
16. Dan seluruh pihak terkait yang tidak dapat saya sebut satu persatu yang telah membantu penulis selama penyusunan Karya Tulis Ilmiah (Skripsi) ini.

Penulis menyadari bahwa Karya Tulis Ilmiah (Skripsi) ini masih jauh dari sempurna oleh karena itu saran dan koreksi atas kekurangan Karya Tulis Ilmiah (Skripsi) ini akan diterima dengan senang hati.

Akhir kata semoga dengan tersusunnya Karya Tulis Ilmiah (Skripsi) ini dapat memberikan manfaat.

Hormat Kami,

Penulis

DAFTAR ISI

DAFTAR TABEL

DAFTAR LAMPIRAN

BAB I
PENDAHULUAN

1.1 Latar Belakang

Pada saat ini model regresi semakin populer dan familiar di kalangan mahasiswa, dosen, praktisi dan terlebih lagi bagi para peneliti. Aplikasi dari model ini sudah tidak didominasi oleh bidang ilmu matematika, statistika ataupun ekonometrika saja, akan tetapi sudah merambah ke berbagai bidang ilmu seperti biologi, pertanian dan seterusnya. Dengan didukung banyaknya paket program komputer para pengguna regresi dengan mudah dapat mengolah data, akan tetapi terkadang penafsiran dan pengambilan keputusan dari model regresi tidak semudah dibandingkan mengolah datanya. Dalam analisis regresi terdapat batasan-batasan (asumsi) yang menjadi dasar apakah model regresi secara tepat dapat digunakan atau tidak. Jika model tersebut bermasalah maka berakibat fatal, misalnya pengambilan kesimpulan (pengujian hipotesis) akan menyesatkan, koefisien regresi menjadi tidak dapat ditafsirkan dan kalaupun dipaksakan hasil yang diperoleh menjadi tidak tepat dan tidak signifikan.

Salah satu asumsi dalam analisis regresi linier berganda berhubungan dengan peubah bebas. Peubah-peubah tersebut tidak boleh ada korelasi. Dengan demikian diasumsikan tidak terjadi multikolinieritas pada peubah-peubah bebas (Gaspersz, 1991). Sebagai ilustrasi situasi adanya multikolinieritas dalam kehidupan sehari-hari, misalnya dalam menganalisis hubungan antara pendapatan dan kekayan terhadap konsumsi. Ahli ekonomi berpendapat selain pendapatan, kekayaan, juga suatu penentu penting dari belanja konsumsi, sehingga modelnya dapat ditulis:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$$

dengan Y konsumsi, X_1 pendapatan, X_2 kekayaan dan β_0 konstanta regresi, β_1 dan β_2 koefisien regresi serta ε galat.

Pada ilustrasi tersebut, dapat dipahami adanya korelasi antara pendapatan dan kekayaan sehigga penggunaan metode kuadrat terkecil (*Ordinary Least Squares Method*) adalah tidak tepat. Ada beberapa metode untuk mengatasi masalah multikolinieritas yang dapat digunakan sebagai petunjuk untuk memperbaiki model regresi diantaranya adalah metode regresi terbaik, dengan menambah data baru, penerapan analisis multivariat seperti analisis komponen utama, dan yang paling mudah adalah dengan membuang peubah-peubah yang diketahui menyebabkan multikolinieritas.

Dalam penelitian ini akan digunakan analisis komponen utama. Analisis komponen utama merupakan cara yang cukup baik untuk mengatasi masalah multikolinieritas (Sembiring, 1995). Cara kerja analisis komponen utama adalah dengan membuat peubah baru yang saling ortogonal dan jumlahnya lebih kecil dari peubah asal. Dari sifat keortogonalan peubah asal ini maka model regresi yang dibentuk tidak lagi mengalami masalah multikolinieritas.

1.2 Perumusan Masalah

1. Bagaimana mendeteksi adanya multikolinieritas pada analisis regresi linier berganda.
2. Bagaimanakah cara menganalisis data yang mengandung multikolinieritas dengan analisis komponen utama.

1.3 Tujuan

1. Mengetahui bagaimana cara mendeteksi adanya multikolinieritas pada analisis regresi linier berganda.
2. Membangun model regresi bila terdapat multikolinieritas pada peubah bebas dengan analisis komponen utama.

1.4 Manfaat

Penulis berharap bagi para pengguna regresi dapat menerapkan hasil penelitian ini sebagai metode perbaikan bila dalam analisis regresi terdapat masalah multikolinieritas pada peubah-peubah bebasnya.

BAB II

TINJAUAN PUSTAKA

2.1 Analisis Regresi Linier Berganda

Dalam penelitian yang menggunakan analisis regresi, sering kali dijumpai lebih dari satu peubah bebas, untuk itu diperlukan model regresi berganda. Jika model regresi berganda yang ditunjukkan oleh koefisiennya adalah linier maka dinamakan model regresi linier berganda (Netty & Linggawati, 1988).

Jika terdapat k peubah bebas $X_1, X_2, X_3, ..., X_k$ dan n pengamatan $Y_1, Y_2, Y_3, ..., Y_n$ maka masing-masing pengamatan Y_i dapat ditulis dalam model:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_k X_{ki} + \varepsilon_i \qquad ...(2.1)$$

dengan

β_0 = konstanta regresi

β_1 = koefisien regresi untuk peubah X_1

β_2 = koefisien regresi untuk peubah X_2

β_k = koefisien regresi untuk peubah X_k

k = banyaknya peubah bebas

ε_i = galat pada pengamatan ke i ; $i = 1, 2, ..., n$.

Persamaan (2.1) dapat ditulis dalam bentuk matriks sebagai berikut:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ . \\ . \\ . \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{21} & . & . & . & X_{n1} \\ 1 & X_{12} & X_{22} & . & . & . & X_{n2} \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 1 & X_{1k} & X_{2k} & . & . & . & X_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ . \\ . \\ . \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_n \end{bmatrix}. \qquad ....(2.2)$$

Persamaan (2.2) dapat ditulis kembali dalam bentuk matriks yang lebih sederhana:

$$\mathbf{Y} = \mathbf{X}\hat{\boldsymbol{\beta}} + \boldsymbol{\varepsilon} \qquad ...(2.3)$$

dengan **Y** matriks berordo $(n \times 1)$, **X** adalah matriks berordo $(n \times (k+1))$, $\boldsymbol{\beta}$ adalah matriks berordo $((k+1) \times 1)$, dan $\boldsymbol{\varepsilon}$ adalah matriks berordo $(n \times 1)$.

2.2 Asumsi-asumsi dalam Model Regresi Linier Berganda

Asumsi-asumsi dasar dalam model regresi linier berganda (Drapper & Smith, 1991):

1. ε_i merupakan suatu peubah acak normal dengan nilai tengah nol dan ragam σ^2 konstan atau $\varepsilon_i \sim NID(0,\sigma^2)$ dan *NID* adalah *Normal Identic Distribution*;
2. ε_i dan ε_j saling bebas untuk $i \neq j$ sehingga peragam $(\varepsilon_i, \varepsilon_j) = 0$; $E(Y_i) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_k X_{ki}$ dan ragam $(Y_i) = \sigma^2$, Y_i independen dengan Y_j untuk $i \neq j$ atau $Y\sim NID(E(Y),\sigma^2)$;
3. tidak terjadi multikolinieritas pada peubah bebas $X_1, X_2, X_3, ..., X_k$.

2.3 Estimasi Parameter dengan Metode Kuadrat Terkecil

Pendugaan terhadap parameter $\hat{\beta}$ dalam model regresi linier berganda dapat dilakukan dengan metode kuadrat terkecil. Fungsi dari metode kuadrat terkecil adalah dengan meminimumkan jumlah dari kuadrat galat.

$$\begin{aligned} \boldsymbol{S} &= \boldsymbol{\varepsilon}^{\mathrm{T}}\boldsymbol{\varepsilon} \quad \text{dengan } \boldsymbol{\varepsilon} = (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \\ &= (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})^{\mathrm{T}}(\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \\ &= \mathbf{Y}^{\mathrm{T}}\mathbf{Y} - 2\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{Y} + \hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{X}\hat{\boldsymbol{\beta}}. \end{aligned} \qquad ...(2.4)$$

Selanjutnya S diturunkan terhadap $\hat{\beta}$ dan disamakan dengan nol

$$\frac{\partial \boldsymbol{S}}{\partial \hat{\boldsymbol{\beta}}} = 0$$

$$\frac{\left(-2\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{Y} + \hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{X}\hat{\boldsymbol{\beta}}\right)}{\partial \hat{\boldsymbol{\beta}}} = \mathbf{0},$$

diperoleh:

$$-2\mathbf{X}^{\mathrm{T}}\mathbf{Y} - 2\mathbf{X}^{\mathrm{T}}\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{0}$$

$$\mathbf{X}^{\mathrm{T}}\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}^{\mathrm{T}}\mathbf{Y}. \qquad ...(2.5)$$

Kedua ruas pada persamaan (2.5) dikalikan dengan $(\mathbf{X}^T\mathbf{X})^{-1}$ dan jika $(\mathbf{X}^T\mathbf{X})^{-1}$ tidak singular maka penduga kuadrat terkecil untuk $\hat{\beta}$ adalah

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}, \qquad \ldots(2.6)$$

dan $var(\hat{\boldsymbol{\beta}})$ adalah

$$var(\hat{\boldsymbol{\beta}}) = var((\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}) . \qquad \ldots(2.7)$$

Ruas kanan pada persamaan (2.7) dikalikan dengan $((\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T)^T$

$$\begin{aligned} var(\hat{\boldsymbol{\beta}}) &= (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T var(\mathbf{Y}) \cdot ((\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T)^T \\ &= (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T(\sigma^2) \cdot \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1} \\ &= \sigma^2(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1} \\ &= \sigma^2(\mathbf{X}^T\mathbf{X})^{-1}. \qquad \ldots(2.8) \end{aligned}$$

2.4 Interval Konfidensi untuk Koefisien Regresi

Menurut Montgomery & Peck (1991), dari asumsi $\varepsilon_i \sim NID\ (0,\sigma^2)$, statistik uji t akan diberikan sebagai berikut:

$$t = \frac{\hat{\beta}_j - \beta_j}{\sqrt{\hat{\sigma}^2 C_{jj}}} \qquad j = 0, 1, 2, \ldots, k.$$

Statistik uji tersebut berdistribusi t dengan derajat kebebasan n-k-1, dengan C_{jj} adalah elemen ke-j dari diagonal matrik $(\mathbf{X}^T\mathbf{X})^{-1}$ dan $\hat{\sigma}^2$ adalah penduga dari ragam galat ($\hat{\sigma}^2 = KTG$ (Kuadrat Tengah Galat)). Interval konfidensi $(100-\alpha)\%$ untuk koefisien regresi $\hat{\beta}_j$ adalah:

$$\hat{\beta}_j - t_{(\alpha/2;n-k-1)}\sqrt{\hat{\sigma}^2 C_{jj}} < \beta_j < \hat{\beta}_j + t_{(\alpha/2;n-k-1)}\sqrt{\hat{\sigma}^2 C_{jj}} .$$

dan biasanya $\sqrt{\hat{\sigma}^2 C_{jj}}$ disebut sebagai standar galat koefisien regresi $\hat{\beta}_j$.

2.5 Pengujian Hipotesis

2.5.1 Pengujian Koefesien Regresi Parsial secara Keseluruhan

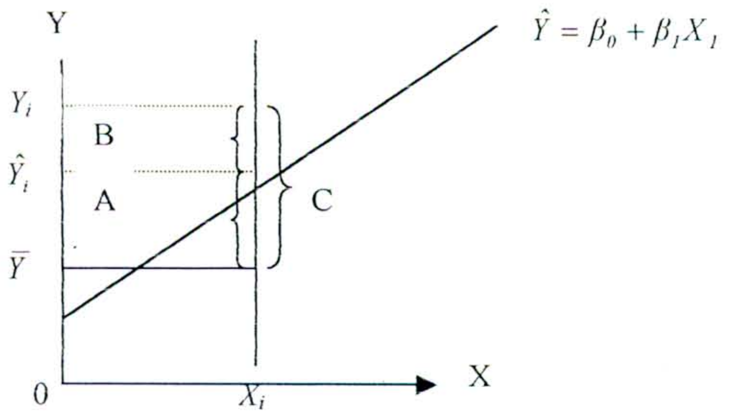
Pengujian hipotesis koefisien regresi parsial secara keseluruhan digunakan untuk menguji kecocokan model regresi yang ditentukan oleh hubungan linier antara peubah tak bebas Y dengan peubah bebas $X_1, X_2, \ldots, X_k$. Pengujian hipotesis ini menggunakan uji statistik F dan pengujian hipotesisnya adalah :

$H_0 : \beta_1 = \beta_2 = \ldots = \beta_k = 0$

H_1 : minimal ada satu $\beta_j \neq 0$.

Dasar dari metode ini adalah pemecahan jumlah kuadrat total (JKT) menjadi jumlah kuadrat regresi (JKR) dan jumlah kuadarat galat (JKG) (Netty & Linggawati, 1988). Perhatikan pemecahannya sebagai berikut:

Pemecahan JKT menjadi JKR dan JKG



Gambar 4.1 Pemecahan JKT menjadi JKR dan JKG
Keterangan gambar

$C = (Y_i - \bar{Y})$

$A = (\hat{Y}_i - \bar{Y})$

$B = (Y_i - \hat{Y}_i)$

$$C = A + B$$

$$\sum (Y_i - \bar{Y})^2 = \sum (\hat{Y}_i - \bar{Y})^2 + \sum (Y_i - \hat{Y})^2.$$

$$JKT = JKR + JKG . \qquad \ldots(2.9)$$

Menurut Gaspersz (1991), persamaan (2.9) dapat diinterpretasikan bahwa jumlah kuadrat total (JKT) merupakan jumlah dari jumlah kuadrat regresi (JKR) dan jumlah kuadrat galat (JKG). Jumlah kuadrat regresi (JKR) merupakan jumlah dari total keragaman yang dijelaskan oleh persamaan regresi. Jumlah kuadrat galat (JKG) merupakan jumlah dari total keragaman yang tidak dapat dijelaskan oleh persamaan regresi.

Prosedur pengujian untuk $H_1 : \hat{\beta}_j \neq 0$.

$$F_0 = \frac{JKR/k}{JKE/n-k-1} = \frac{KTR}{KTG} \; .$$

H_0 ditolak jika $F_0 > F_{(\alpha;k;n-k-1)}$

Selanjutnya, menghitung jumlah kuadrat regresi (JKR) dengan meminimumkan jumlah kuadrat galat.

$$JKG = \sum_{i=1}^{n} (Y_i - \hat{Y})^2 = \varepsilon^{\mathrm{T}} \varepsilon \; .$$

Dari persamaan (2.4) telah diperoleh:

$$JKG = \mathbf{Y}^{\mathrm{T}}\mathbf{Y} - 2\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{Y} + \hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{X}\hat{\boldsymbol{\beta}} \; .$$

Pada persamaan (2.5) telah diketahui bahwa $\mathbf{X}^{\mathrm{T}}\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}^{\mathrm{T}}\mathbf{Y}$ sehingga:

$$JKG = \mathbf{Y}^{\mathrm{T}}\mathbf{Y} - \hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{Y} \; ; \qquad \ldots(2.10)$$

$$\begin{aligned} JKT &= \sum_{i=1}^{n} (Y_i - \overline{Y})^2 \\ &= \sum_{i=1}^{n} (Y_i - 2Y_i\overline{Y} + \overline{Y}^2) \\ &= \sum_{i=1}^{n} \left(Y_i^2 - 2Y_i \frac{\sum_{i=1}^{n} Y_i}{n} + \left(\frac{\sum_{i=1}^{n} Y_i}{n} \right)^2 \right) \end{aligned}$$

$$= \sum_{i=1}^{n} Y_i^2 - 2\frac{\left(\sum_{i=1}^{n} Y_i\right)^2}{n} + n\left(\frac{\sum_{i=1}^{n} Y_i}{n}\right)^2$$

$$= \sum_{i=1}^{n} Y_i^2 - \frac{\left(\sum_{i=1}^{n} Y_i\right)^2}{n} = \mathbf{Y}^{\mathrm{T}}\mathbf{Y} - \frac{\left(\sum_{i=1}^{n} Y_i\right)^2}{n} \quad \text{...(2.11)}$$

Persamaan (2.10) dapat ditulis sebagai berikut:

$$JKG = \mathbf{Y}^{\mathrm{T}}\mathbf{Y} - \frac{\left(\sum_{i=1}^{n} Y_i\right)^2}{n} - \left[\hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{Y} - \frac{\left(\sum_{i=1}^{n} Y_i\right)^2}{n}\right] \quad \text{atau}$$

$JKG = JKT - JKR$ (dari persamaan(2.9)).

Diperoleh $JKR = \hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{Y} - \dfrac{\left(\sum_{i=1}^{n} Y_i\right)^2}{n}$, ...(2.12)

$JKG = \mathbf{Y}^{\mathrm{T}}\mathbf{Y} - \hat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{Y}$ dan

$$JKT = \mathbf{Y}^{\mathrm{T}}\mathbf{Y} - \frac{\left(\sum_{i=1}^{n} Y_i\right)^2}{n}.$$

Tabel 2.1 Analisis Ragam dalam Analisis Regresi Linier Berganda.

Sumber Keragaman	Jumlah Kuadrat (*sum of Square*) (*JK*)	Derajat bebas (*Db*)	Kuadrat Tengah (*Mean Square*) (*KT*)	F_0
Regresi	*JKR*	*k*	$KTR = JKR/k$	$\dfrac{KTR}{KTG}$
Galat	*JKG*	*n-k-1*	$KTG = JKG/n\text{-}k\text{-}1$	
Total	*JKT*	*n-1*		

2.5.2 Pengujian Koefesien Regresi secara Individual

Pengujian hipotesisnya adalah:

$$H_0 : \hat{\beta}_j = 0$$
$$H_1 : \hat{\beta}_j \neq 0$$

H_0 tidak ditolak artinya peubah bebas X_j bisa dihilangkan dari model. Uji statistiknya adalah :

$$t_0 = \frac{\hat{\beta}_j}{\sqrt{\hat{\sigma}^2 C_{jj}}} = \frac{\hat{\beta}_j}{\mathrm{se}(\beta_j)}$$

dengan C_{jj} adalah elemen ke-j dari diagonal matriks $(\mathbf{X}^T\mathbf{X})^{-1}$ yang berkorespondensi dengan $\hat{\beta}_j$ dan $\hat{\sigma}^2 = KTG$.

H_0 ditolak jika $| t_0 | > t_{(\alpha/2 \,;\, n-k-1)}$.

2.6 Koefisien Determinasi Berganda dan Koefisien Korelasi Berganda

Koefisien determinasi berganda (R^2) mengukur proporsi keragaman total peubah tak bebas Y yang dijelaskan oleh peubah bebas X yang ada dalam model persamaan regresi secara bersama, sedangkan koefisien korelasi berganda (R) tidak lain merupakan akar pangkat dua dari koefisien determinasi berganda. Koefisien korelasi berganda (R) mengukur keeratan hubungan linier diantara peubah tak bebas Y dengan semua peubah bebas yang ada dalam model persamaan regresi (Gaspersz, 1991).

Koefisien determinasi dan korelasi berganda ditentukan berdasarkan formula:

$$R^2 = \frac{JKR}{JKT} \quad dan \quad R = \sqrt{\frac{JKR}{JKT}}\,.$$

Menurut Sembiring (1995), makin dekat R^2 dengan 1 makin baik kecocokan data dengan model, dan makin dekat R^2 Dengan 0 makin jelek kecocokan data dengan model.

2.7 Multikolinieritas

Istilah multikolinieritas digunakan untuk hubungan linier pada peubah-peubah bebas dalam model regresi. Multikolinieritas ada yang sempurna dan juga tidak sempurna. Untuk hubungan yang terdiri dari k peubah bebas $X_1, X_2, \dots, X_k$ dan satu peubah tak bebas Y dari persamaan regresi (2.1), maka hubungan linier sempurna terjadi jika:

$$\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k = 0 \,. \qquad \dots(2.13)$$

dengan $\beta_1, \beta_2, \dots, \beta_k$ merupakan konstanta yang tidak seluruhnya nol atau paling sedikit ada satu yang tidak sama dengan nol, yaitu $\beta_j \neq 0$, ($j = 1, 2, \dots, k$), dan untuk hubungan linier tapi tidak sempurna:

$$\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon_i = 0 \,. \qquad \dots(2.14)$$

dengan ε_i = kesalahan pengganggu (galat).

Adanya multikolinieritas yang sempurna dapat menyebabkan koefisien regresi tidak menentu *(Indeterminate)* dan nilai standar galat koefisien regresi menjadi tidak terbatas (*Infinite*). bila multikolinieritas tidak sempurna maka koefisien regresi dapat dicari, tetapi standar galatnya besar, sehingga interval koefisiennya menjadi lebar dan mengakibatkan tingkat ketelitian untuk menguji koefisiennya berkurang. Adanya multikolinieritas yang tinggi tidak memungkinkan untuk memecah pengaruh secara terpisah individu dari peubah bebas X terhadap peubah tak bebas Y (Gujarati, 1992).

Ada baberapa cara mendeteksi multikolinieritas diantaranya adalah dengan memeriksa matriks korelasi, *Variance Inflation Factor* (VIF) dan dengan nilai *Condition Number*. Dalam penelitian ini untuk mendeteksi multikolinieritas akan digunakan matriks korelasi sedangkan untuk mengukur derajat kolinieritas pada peubah bebas akan digunakan *Variance Inflation Factor* (VIF).

Menurut Montgomery & Peack (1991), matriks korelasi dapat ditulis sebagai berikut:

$$\mathbf{R} = \begin{bmatrix} 1 & r_{12} & r_{13} & \dots & r_{1k} \\ r_{12} & 1 & r_{23} & \dots & r_{2k} \\ \dots & \dots & \dots & \dots & \dots \\ r_{1k} & r_{2k} & r_{3k} & \dots & 1 \end{bmatrix} . \qquad \dots(2.15)$$

Dengan $r_{ij} = \dfrac{S_{ij}}{(S_{ii}S_{jj})^{1/2}}$

$$S_{ij} = \sum_{u=1}^{n}(X_{ui} - \overline{X}_{ui})(X_{uj} - \overline{X}_{uj})$$

$$S_{jj} = \sum_{i=1}^{n}(X_{ij} - \overline{X}_{j})^2 .$$

Menurut Sugiarto (1992), besarnya nilai r dapat digunakan untuk menentukan kuat lemahnya hubungan diantara dua peubah. Kuat lemahnya hubungan tersebut dapat ditunjukkan dalam tabel berikut:

Tabel 2.2 Kriteria Hubungan antara Dua Peubah.

Nilai $\|r\|$	Kriteria Hubungan
$r = 0$	Tidak ada korelasi
$0 < r \leq 0{,}5$	Korelasi lemah
$0.5 < r \leq 0{,}8$	Korelasi sedang
$0.8 < r < 1$	Korelasi kuat
$r = 1$	Korelasi sempurna

Menurut Montgomery & Peck (1991), *Variance Inflation Factor* (VIF) merupakan elemen diagonal utama dari invers matriks korelasi dan persamaannya dapat ditulis sebagai berikut:

$$VIF_{ij} = \frac{1}{(1 - R_{ij}^{\ 2})}.$$

dengan R_{ij}^2 adalah koefisien determinasi dari peubah bebas X_{ij} jika diregresikan terhadap peubah bebas yang lain. R_j^2 bernilai kecil jika peubah bebas X_{ij} tidak berkorelasi dengan peubah bebas yang lain sehigga VIF_{ij} bernilai mendekati satu, dan R_{ij}^2 bernilai mendekati satu jika peubah bebas X_{ij} berkorelasi dengan peubah bebas yang lain sehigga VIF_{ij} bernilai besar. Jika dari beberapa nilai VIF_{ij} bernilai lebih dari 10 maka adanya multikolinieritas tidak dapat diabaikan.

2.8 Analisis Komponen Utama

Multikolinieritas dalam analisis regresi dapat diatasi dengan beberapa cara yaitu dengan metode regresi terbaik, penambahan data baru, penerapan analisis multivariat seperti analisis komponen utama, dan yang paling mudah adalah dengan membuang peubah-peubah yang diketahui penyebab multikolinieritas. Dalam penelitian ini akan digunakan analisis komponen utama, karena dengan analisis komponen utama peubah-peubah bebas akan saling ortogonal sehingga tidak ada lagi masalah multikolinieritas.

Analisis komponen utama berusaha mereduksi K peubah pengamatan menjadi M peubah baru yang saling ortogonal, masing-masing M peubah baru tersebut merupakan kombinasi linier dari K peubah lama (Mattjik, dkk, 2002). Pemilihan M peubah baru ini sedemikian rupa sehingga keragaman yang dimiliki oleh K peubah lama sebagian besar dapat diterangkan oleh M peubah baru, dan M peubah baru ini kemudian diregresikan dengan peubah tak bebas Y atau dapat disebut dengan regresi komponen utama.

Tahap-tahap yang dilakukan dalam analisis komponen utama ini dimulai dari prosedur seleksi akar eigen, Jika K peubah bebas tidak menghadapi masalah perbedaan dalam skala pengukuran maka akar eigen diperoleh dari suatu persamaan $|\sum - \lambda\mathbf{I}| = \mathbf{0}$ dengan $\sum$ adalah matriks peragam atau matriks varian-covarian. Matriks peragam dapat dirumuskan sebagai berikut:

$$\sum = \frac{1}{n}\begin{bmatrix} X_{11}-\overline{X}_1 & \cdots & X_{1k}-\overline{X}_1 \\ . & \cdots & . \\ . & \cdots & . \\ . & . & . \\ X_{k1}-\overline{X}_k & \cdots & X_{kn}-\overline{X}_k \end{bmatrix}^{\mathrm{T}} \cdot \begin{bmatrix} X_{11}-\overline{X}_1 & \cdots & X_{1k}-\overline{X}_1 \\ . & \cdots & . \\ . & \cdots & . \\ . & . & . \\ X_{k1}-\overline{X}_k & \cdots & X_{kn}-\overline{X}_k \end{bmatrix}.$$

Jika K peubah bebas menghadapi masalah perbedaan dalam skala pengukuran maka akar eigen diperoleh dari suatu persamaan $|\ \mathbf{R} - \lambda\mathbf{I}\ | = \mathbf{0}$, dengan $\mathbf{R} = \frac{1}{n-1}\mathbf{Z}^{\mathbf{T}}\mathbf{Z}$ dan $\mathbf{Z}$ adalah matriks $\mathbf{X}$ yang dibakukan dengan persamaan.

$$Z_{ji} = \frac{(X_{ji} - \overline{X}_j)}{S_j} \qquad \ldots(2.16)$$

dengan

Z = peubah baku

X = peubah bebas

$\overline{X}$ = nilai rata-rata peubah bebas

S = simpangan baku peubah bebas X

$i = 1, 2, ..., n$

$j = 1, 2, ..., k.$

Jika akar eigen λ diurut dari nilai terbesar sampai nilai terkecil, maka pengaruh komponen utama W_i berpadanan dengan mengurut λ_i. Ini berarti komponen-komponen tersebut menerangkan proporsi keragaman terhadap peubah W_i yang semakin lama semakin kecil. Menurut Gaspersz (1992), dengan menggunakan kriteria Kaiser dipilih komponen utama dengan akar eigen yang lebih besar dari satu. Vektor eigen $\mathbf{V_i}$ diperoleh dari setiap akar eigen λ_i yang memenuhi suatu persamaan:

$$(\sum - \boldsymbol{\lambda}\mathbf{I})\mathbf{V_i} = \mathbf{0} \quad \text{atau} \quad (\mathbf{R} - \boldsymbol{\lambda}_i\mathbf{I})\mathbf{V}_i = \mathbf{0}$$

dengan

$$\mathbf{V_i} = (V_{1i}, V_{2i}, ..., V_{ki}) \text{ dan } \mathbf{V_i}^{\mathbf{T}}\mathbf{V_i} = \mathbf{0}. \qquad ...(2.17)$$

Dari persamaan (2.16) dan persamaan (2.17) komponen utama W_i dapat dirumuskan:

$$W_i = V_{1i}Z_1 + V_{2i}Z_2 + ... + V_{ki}Z_k. \qquad ...(2.18)$$

Dengan demikian $W_1, W_2, W_3, ..., W_m$ sebagai banyaknya komponen utama yang dilibatkan dalam analisis regresi dan komponen-komponen utama ini saling ortogonal sesamanya.

Tahap selanjutnya adalah meregresikan komponen utama $W_1, W_2, W_3, ..., W_m$ dengan peubah tak bebas Y, maka model regresi komponen utama dapat dirumuskan sebagai berikut:

$$Y_i = \hat{a}_0 + \hat{a}_1W_1 + \hat{a}_2W_2 + ... + \hat{a}_mW_m + \varepsilon_i. \qquad ...(2.19)$$

Model persamaan regresi (2.19) identik dengan model persamaan regresi (2.1), dan dapat ditulis dalam bentuk matriks sebagai berikut:

$$\mathbf{Y} = \hat{\mathbf{a}}\mathbf{W} + \varepsilon. \qquad ...(2.20)$$

Persamaan (2.20) identik dengan persamaan (2.3) maka taksiran parameternya adalah:

$$\hat{\mathbf{a}} = (\mathbf{W}^T\mathbf{W})^{-1}\mathbf{W}^T\mathbf{Y} \qquad \text{dan} \qquad var(\hat{\mathbf{a}}) = (\mathbf{W}^T\mathbf{W})^{-1}\sigma^2. \qquad \ldots(2.21)$$

Persamaan regresi komponen utama (2.19) dengan peubah bebas W_i dapat ditransformasikan ke peubah baku Z sebagai berikut:

$$Y = \hat{b}_0 + \hat{b}_1 Z_1 + \hat{b}_2 Z_2 + \ldots + \hat{b}_k Z_k. \qquad \ldots(2.22)$$

Untuk menghitung nilai $\hat{b}$ substitusikan persamaan (2.18) ke persamaan (2.19)

$$\begin{aligned} Y &= \hat{a}_0 + \hat{a}_1(V_{11}Z_1 + V_{21}Z_2 + \ldots + V_{k1}Z_k) + \hat{a}_2(V_{12}Z_1 + V_{22}Z_2 + \ldots + V_{k2}Z_k) \\ &\quad + \ldots + \hat{a}_m(V_{1m}Z_1 + V_{2m}Z_2 + \ldots + V_{km}Z_k) \\ &= \hat{a}_0 + (\hat{a}_1 V_{11} + \hat{a}_2 V_{12} + \ldots + \hat{a}_m V_{1m})Z_1 + (\hat{a}_1 V_{21} + \hat{a}_2 V_{22} + \ldots + \hat{a}_m V_{2m})Z_2 \\ &\quad + \ldots + (\hat{a}_1 V_{k1} + \hat{a}_2 V_{k2} + \ldots + \hat{a}_m V_{km})Z_k, \end{aligned}$$

diperoleh:

$$\hat{b}_0 = \hat{a}_0$$

$$\hat{b}_1 = \hat{a}_1 V_{11} + \hat{a}_2 V_{12} + \ldots + \hat{a}_m V_{1m}$$

$$\hat{b}_2 = \hat{a}_2 V_{21} + \hat{a}_2 V_{22} + \ldots + \hat{a}_m V_{2m}$$

………………………………

$$\hat{b}_k = \hat{a}_1 V_{k1} + \hat{a}_2 V_{k2} + \ldots + \hat{a}_m V_{km}$$

$$\begin{aligned} var(\hat{b}) &= var(\hat{a}_1 V_{i1} + \hat{a}_2 V_{i2} + \ldots + \hat{a}_m V_{km}) \\ &= V_{i1}{}^2 var(\hat{a}_1) + V_{i2}{}^2 var(\hat{a}_2) + \ldots + V_{km}{}^2 var(\hat{a}_m) \qquad \ldots(2.23) \end{aligned}$$

$$i = 1, 2, 3, \ldots, k.$$

Untuk membandingkan hasil taksiran koefisien regresi yang menggunakan analisis komponen utama dengan koefisisen regresi yang menggunakan metode kuadrat terkecil secara langsung, persamaan model regresi dengan peubah baku Z (2.22) ditransformasi ke peubah asal X:

$$Y = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \ldots + \hat{\beta}_k X_k. \qquad \ldots(2.24)$$

Untuk menghitung nilai $\hat{\beta}_i$ substitusikan persamaan (2.16) ke persamaan (2.22)

$$Y = \hat{b}_o + \hat{b}_1\left(\frac{X_1 - \overline{X}_1}{S_1}\right) + \hat{b}\left(\frac{X_2 - \overline{X}_2}{S_2}\right) + \ldots + \hat{b}_k\left(\frac{X_k - \overline{X}_k}{S_k}\right)$$

$$Y = \hat{b}_o - \left(\hat{b}_1 \frac{\overline{X}_1}{S_1} + \hat{b}_2 \frac{\overline{X}_2}{S_2} + ... + \hat{b}_k \frac{\overline{X}_k}{S_k} \right) + \frac{\hat{b}_1}{S_1} X_1 + \frac{\hat{b}_2}{S_2} X_2 + ... + \frac{\hat{b}_k}{S_k} X_k,$$

diperoleh:

$$\hat{\beta}_0 = \hat{b}_0 - \left(\hat{b}_1 \frac{\overline{X}_1}{S_1} + \hat{b}_2 \frac{\overline{X}_2}{S_2} + ... + \hat{b}_k \frac{\overline{X}_k}{S_k} \right)$$

$$\hat{\beta}_1 = \frac{\hat{b}_1}{S_1} \qquad ...(2.25)$$

$$\hat{\beta}_2 = \frac{\hat{b}_2}{S_2}$$

....

$$\hat{\beta}_k = \frac{\hat{b}_k}{S_k}$$

$var(\hat{\beta}_i)$ sama dengan $var(\hat{b}_i)$ pada persamaan (2.23).

Pengujian hipotesis terhadap koefisien regresi parsial komponen utama pada persamaan (2.21) baik secara keseluruhan atau individual digunakan rumus seperti dalam sub bab 2.5.

BAB III
METODOLOGI

3.1 Sumber Data

Sumber data yang digunakan dalam penelitian ini ada dua, yaitu data simulasi dan data riil.

3.1.1 Data simulasi

Data simulasi dibangkitkan melalui komputer mengggunakan program statistika S-plus versi 4.5 dengan peubah tak bebas Y dan peubah bebas X_1, X_2, $X_3, \dots$, X_k. Data simulasi diasumsikan berdistribusi normal dan terjadi multikolinieritas pada peubah bebas. Prosedur simulasi :

1. menentukan nilai parameter $\beta = (\beta_0, \beta_1, \beta_2, \beta_3, \dots, \beta_k)$;
2. membangkitkan peubah bebas X_1, X_2, $X_{3,\dots,}$ X_k yang berkorelasi;

   Untuk membangkitkan data ini $X_3 = \alpha X_2 + \delta$, dalam penelitian ini jenis peubah bebas yang dibangkitkan ada tiga:

   a. membangkitkan peubah bebas yang memiliki hubungan diantara dua peubah yang sempurna. Untuk membangkitkan data ini $\delta = 0$;
   b. membangkitkan peubah bebas yang memiliki hubungan diantara dua peubah yang kuat. Untuk membangkitkan data ini $\delta \sim N(0,0.4)$;
   c. membangkitkan peubah bebas yang memiliki hubungan diantara dua peubah yang sedang. Untuk membangkitkan data ini $\delta \sim N(0,3.3)$.
3. menentukan nilai mean:

$$\mu = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki};$$

4. membangkitkan peubah tak bebas Y:

$$Y = \text{rnorm}\ (n, \mu, \sigma^2).$$

3.1.2 Data Riil

Untuk data riil dalam penelitian ini diambil dari skripsi Tinuk Fitrias yang berjudul "Faktor-Faktor Sosial Ekonomi Terhadap Pendapatan Petani Pisang Kuba" Fakultas Pertanian Jurusan Sosial Ekonomi Pertanian. Penelitian ini dilakukan di kecamatan Ajung Kabupaten Jember.

Adapun peubah yang digunakan dalam analisis adalah:

- peubah tak bebas (Y):

  pendapatan;

- peubah bebas (X):

  X_1 = umur

  X_2 = pendidikan formal

  X_3 = jumlah anggota keluarga

  X_4 = luas lahan

  X_5 = biaya produksi

  X_6 = produksi

  X_7 = harga jual.

3.2 Analisis Data Simulasi

Untuk mengolah data simulasi digunakan program SPSS versi 10.00, sedangkan untuk mencari vektor eigen digunakan program S-plus versi 4.5. Adapun langkah-langkah yang digunakan adalah sebagai berikut:

1. menganalisis semua data dengan metode kuadrat terkecil, pendeteksian multikolinieritas dengan memeriksa matriks korelasi dan menghitung besarnya multikolinieritas dengan *Variance Inflation Factor* (VIF).
2. penerapan metode analisis komponen utama.
   - menentukan peubah Z hasil pembakuan dari peubah X;
   - menentukan nilai akar eigen dengan persamaan :

     $$|\mathbf{R} - \lambda\mathbf{I}| = 0;$$

   - menentukan nilai vektor eigen dengan persamaan:

     $$(\mathbf{R} - \lambda\mathbf{I})\mathbf{V_i} = \mathbf{0};$$

- menentukan komponen utama W_i melalui kombinasi linier antara vektor eigen dan peubah baku Z:

  $$W_i = V_{1i}Z_1 + V_{2i}Z_2 + \ldots + V_{ki}Z_k\ ;$$

- meregresikan W_i terpilih berdasarkan keragamannya dengan peubah tak bebas Y sehingga didapat persamaan model regresi komponen utama:

  $$Y_i = \hat{a}_0 + \hat{a}_1W_1 + \hat{a}_2W_2 + \ldots + \hat{a}_mW_m + \varepsilon_i\ ;$$

- menentukan nilai koefisien regresi $(\hat{a})$;
- menentukan JKG, JKT, JKR serta R^2;
- menguji koefisien regresi $(\hat{a})$ baik secara keseluruhan atau individual dengan menggunakan rumus seperti dalam sub bab 2.5 ;
- mentransformasi persamaan regresi dengan peubah bebas W_i ke peubah baku Z;
- mentransformasi persamaan regresi dengan peubah baku Z ke peubah bebas X.

3.3 Perbandingan Nilai Koefisien Regresi pada Data Simulasi

Membandingkan nilai taksiran koefisien regresi ($\hat{\beta}$) yang mengabaikan multikolinieritas pada persamaan (2.1) dengan koefisien regresi ($\hat{\beta}$) yang menggunakan analisis komponen utama pada persamaan (2.25).

3.4 Analisis Data Riil

Untuk mengolah data riil digunakan program SPSS versi 10.00 dan untuk mencari vektor eigen digunakan program S-plus versi 4.5. Adapun langkah-langkah yang digunakan dalam analisa data riil sama serperti dalam sub bab 3.2.

BAB V
KESIMPULAN DAN SARAN

5.1 Kesimpulan

1. Multikolinieritas dapat dideteksi dengan memeriksa matriks korelasi, multikolinieritas terjadi pada data yang memiliki kriteria hubungan sempurna dan kuat ($0.8 < r \leq 1$). Untuk mengukur besarnya multikolinieritas dapat digunakan *Variance Inflation Factor* (VIF), multikolinieritas tidak dapat diabaikan jika nilai VIF lebih besar dari 10.
2. Analisis komponen utama merupakan salah satu metode yang dapat digunakan untuk menghilangkan multikolinieritas. Cara menganalisis data dengan analisis komponen utama adalah dengan membuat peubah-peubah baru yang saling ortogonal.
3. Hasil simulasi menunjukkan bahwa, jika koefisien regresi dikembalikan pada peubah asal maka penerapan metode analisis komponen utama pada analisis regresi menghasilkan taksiran koefisien regresi yang biasnya lebih besar daripada taksiran koefisien regresi yang menggunakan metode kuadrat terkecil secara langsung.

5.2 Saran

Dari hasil penelitian ini, diharapkan ada yang akan melanjutkan untuk meneliti dengan menggunakan metode lain sehingga hasilnya dapat dibandingkan.