

**PERBANDINGAN METODE *RIDGE* DENGAN METODE PCA
(*PRINCIPAL COMPONENT ANALYSIS*) DALAM MENGATASI
*MULTIKOLINEARITAS***

SKRIPSI

Untuk memenuhi sebagian persyaratan
mencapai derajat Sarjana S-1

Program Studi Matematika



Diajukan oleh

Ani Rohmah

06610010

Kepada

PROGRAM STUDI MATEMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA

YOGYAKARTA

2013



SURAT PERSETUJUAN SKRIPSI/TUGAS AKHIR

Hal :

Lamp :

Kepada

Yth. Dekan Fakultas Sains dan Teknologi

UIN Sunan Kalijaga Yogyakarta

di Yogyakarta

Assalamu'alaikum wr. wb.

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengadakan perbaikan seperlunya, maka kami selaku pembimbing berpendapat bahwa skripsi Saudara:

Nama : Ani Rohmah

NIM : 06610010

Judul Skripsi : Perbandingan Metode *Ridge* dengan Metode PCA (*Principal Component Analysis*) dalam Mengatasi *Multikolinearitas*

sudah dapat diajukan kembali kepada Program Studi Matematika Fakultas Sains dan Teknologi UIN Sunan Kalijaga`Yogyakarta sebagai salah satu syarat untuk memperoleh gelar Sarjana Strata Satu dalam bidang matematika.

Dengan ini kami berharap agar skripsi/tugas akhir Saudara tersebut di atas dapat segera dimunaqsyahkan. Atas perhatiannya kami ucapkan terima kasih.

Wassalamu'alaikum wr. wb.

Yogyakarta, 04 Desember 2012

Pembimbing

Epha Diana Supandi, S.Si., M.Sc.
NIP. 19750912 200801 2015



SURAT PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini:

Nama : Ani Rohmah
NIM : 06610010
Prodi : MATEMATIKA
Fakultas : Sains dan Teknologi

Menyatakan bahwa tulisan yang saya tulis ini, bukanlah hasil karya orang lain dan sepanjang sepengetahuan saya tidak berisi materi yang telah dipublikasikan atau ditulis orang lain atau telah dipergunakan dan diterima persyaratan studi pada universitas manapun kecuali pada bagian tertentu yang telah dinyatakan dalam teks. Tulisan ini merupakan hasil karya saya sendiri yang saya tulis dengan penuh cinta, kasih sayang penuh dari yang terkasih, umi dan abah, dan penuh keikhlasan hati untuk terus berjuang demi terciptanya tulisan ini.

Yogyakarta, 04 Desember 2012

Yang Menyatakan,

MEI TERAI
TEMPEL
PAJAK MENBANGUN BANGSA
TGL. 20



7392FABF229022906

ENAM RIBU RUPIAH

6000

DJP

Ani Rohmah
NIM.06610010



Universitas Islam Negeri Sunan Kalijaga

FM-UINSK-BM-05-07/R0

PENGESAHAN SKRIPSI/TUGAS AKHIR

Nomor : UIN.02/D.ST/PP.01.1/318/2013

Skripsi/Tugas Akhir dengan judul : Perbandingan Metode *Ridge* dengan Metode PCA (*Principal Component Analysis*) Dalam Mengatasi *Multikolinearitas*

Yang dipersiapkan dan disusun oleh :

Nama : Ani Rohmah

NIM : 06610010

Telah dimunaqasyahkan pada : Jum'at, 11 Januari 2013

Nilai Munaqasyah : A / B

Dan dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga

TIM MUNAQASYAH :

Ketua Sidang

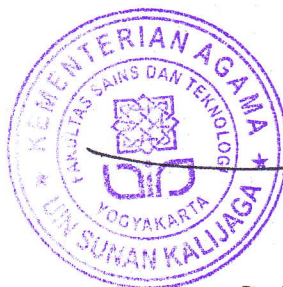
Epha Diana Supandi, S.Si. M.Sc
NIP. 19750912 200001 2 015

Penguji I

M. Farhan Quadratullah, M.Si
NIP.19790922 200801 1 011

Penguji II

Ki Haryadi, S.Si. MPH
NIDN. 0515057601



Yogyakarta, 4 Februari 2013
UIN Sunan Kalijaga
Fakultas Sains dan Teknologi
Dekan

Prof. Drs. H. Akh. Minhaji, M.A, Ph.D
NIP. 19580919 198603 1 002

HALAMAN MOTTO

"Sesungguhnya sesudah kesulitan itu ada kemudahan. Maka apabila kamu telah selesai (dari suatu urusan), kerjakanlah dengan sungguh-sungguh (urusan) yang lain. Dan hanya kepada Tuhan-mulah hendaknya kamu berharap"

(QS. Al Insyirah: 6-8)

"YAKIN"

KATA PENGANTAR

Alhamdulillah Rabbil'alam, segala puji syukur penulis panjatkan kehadirat Allah SWT yang telah melimpahkan rahmat dan hidayah-Nya dengan memberikan kemudahan, kekuatan, kelancaran, dan kesehatan dalam penyusunan karya tulis yang berjudul “Perbandingan Metode *Ridge* dengan Metode PCA (*Principal Component Analysis*) dalam Mengatasi *Multikolinearitas*”. Penulisan skripsi ini bertujuan untuk memenuhi mata kuliah Tugas Akhir sebagai syarat untuk memperoleh derajat kesarjanaan. Pada kesempatan ini, penulis ingin menyampaikan ucapan terima kasih kepada semua pihak yang telah membantu penulis dalam penyusunan skripsi ini, baik secara langsung maupun tidak langsung, terutama penulis tujukan kepada:

1. Bapak Prof. Drs. H. Akh. Minhaji, M.A., Ph.D., selaku Dekan Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
2. Ibu Sri Utami Zuliana, M.Si., selaku Pembimbing Akademik Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
3. Mohammad Abrori M.Si., Ketua Program Studi Matematika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
4. Ibu Epha Diana Supandi, S.Si.,M.Sc. selaku dosen pembimbing yang telah meluangkan waktu memberikan bimbingan, arahan, bantuan dan ilmu dalam menyelesaikan skripsi ini.
5. Bapak dan ibu yang telah membesarkanku hingga sekarang ini dan telah memberi dukungan berupa moril maupun materiil, do'a, perhatian, kasih

sayang dan cinta yang tulus. Serta terima kasih kepada Kakak dan adek yang selalu memberi dukungan.

6. Tidak lupa kepada semua teman-teman yang selalu memberi semangat dan tidak lupa untuk semua pihak yang telah membantu dalam penyelesaian skripsi ini, yang tidak bisa penulis sebutkan satu persatu.

Penulis menyadari masih banyak kekurangan dalam penyusunan skripsi ini, sehingga saran maupun kritik yang membangun sangat penulis harapkan agar proses berkembangnya ilmu bisa terus berlangsung. Semoga skripsi ini dapat memberi manfaat bagi kita semua.

Amin

Yogyakarta, 04 Desember 2012

Penulis

Ani Rohmah
NIM. 06610010

HALAMAN PERSEMBAHAN

Karya ini khusus kupersembahkan kepada:

Bapak dan Ibu tercinta

Yang senantiasa menyayangi, menjaga, dan selalu bermunajjah setiap
malamnya demi belahan jiwanya

DAFTAR ISI

HALAMAN JUDUL	i
SURAT PERSETUJUAN SKRIPSI	ii
HALAMAN PENGESAHAN	iii
PERNYATAAN KEASLIAN	iv
HALAMAN MOTTO	v
KATA PENGANTAR	vi
HALAMAN PERSEMBAHAN	viii
DAFTAR ISI	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN	xiv
DAFTAR SIMBOL.....	xv
ABSTRAKSI	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Batasan Masalah	3
1.3 Rumusan Masalah	4
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	5
1.6 Tinjauan Pustaka.....	5
1.7 Sistematika Penulisan	6
BAB II DASAR TEORI	9
2.1 Matriks	9
2.2 Pengali Lagrang	14
2.3 Distribusi Normal	14
2.4 Distribusi Chi-Kuadrat (χ^2)	16
2.5 Regresi Linear	17

2.3.1 Regresi Linear Sederhana	18
2.3.2 Regresi Linear Berganda	18
2.3.3 Metode Kuadrat Terkecil	19
2.4 Pemusatan dan Penskalan	25
2.5 <i>Multikolinearitas</i>	27
2.5.1 Sebab-sebab <i>Multikolinearitas</i>	27
2.5.2 Akibat-akibat <i>Multikolinearitas</i>	28
2.5.3 Konsekuensi <i>Multikolinearitas</i>	29
2.5.4 Mendeteksi <i>Multikolinearitas</i>	30
2.6 Metode Regresi <i>Ridge</i>	35
2.7 Metode PCA(<i>Principal Component Analysis</i>)	36
2.8 Menentukan Model yang Terbaik	37
2.8.1 Koefisien Determinasi	37
2.8.2 MSE (<i>Mean Square Error</i>)	38
BAB III METODOLOGI PENELITIAN	40
3.1 Jenis Penelitian	40
3.2 Sumber Data	40
3.3 Metode Analisis Data	40
3.4 Metode Pengumpulan Data	42
3.4 Pengujian Hipotesis	42
3.5 Menentukan Model yang Terbaik	43
3.6 Penarikan Kesimpulan	43
BAB IV METODE REGRESI <i>RIDGE</i> DAN METODE PCA	44
(<i>PRINCIPAL COMPONENT ANALYSIS</i>)	
4.1 Regresi <i>Ridge</i>	44
4.1.1 Estimator Regresi <i>Ridge</i>	44
4.1.2 Hubungan Estimator <i>Ridge</i> dengan Penduga Kuadrat	
Terkecil	47
4.1.3 Sifat-sifat Estimator Regresi <i>Ridge</i>	48
4.1.4 Mendeteksi <i>Multikolinearitas</i> dengan Metode <i>Ridge</i>	50
4.2 PCA (<i>Principal Componet analysis</i>).....	53

4.2.1 Menentukan Komponen Utama	53
4.2.2 Komponen Utama Berdasarkan Matriks Kovariansi	58
4.2.3 Komponen Utama Berdasarkan Matriks Korelasi	59
4.2.4 Kriteria Pemilihan Komponen Utama	60
4.2.5 Kontribusi Komponen Utama	60
BAB V STUDI KASUS	62
5.1 Pendeteksian <i>Multikolinearitas</i>	64
5.2 Pemodelan Regresi <i>Ridge</i>	67
5.2.1 Proses Pemusatan dan Penskalaan	67
5.2.2 Proses Pemodelan Regresi <i>Ridge</i>	69
5.2.3 Uji Keberartian Regresi	72
5.3 Pemodelan PCA (<i>Principal Component Analysis</i>)	74
5.3.1 Penanggulangan <i>Multikolinearitas</i> Pada Metode PCA ..	74
5.3.2 Proses Pemodelan PCA (<i>Principal Component Analysis</i>)	77
5.3.3 Uji Keberartian Regresi	78
5.4 Menentukan Model yang Terbaik	80
BAB VI PENUTUP	84
6.1 Kesimpulan	84
6.2 Saran	86
DAFTAR PUSTAKA	87
LAMPIRAN-LAMPIRAN	89

DAFTAR TABEL

Tabel 5.1: Data Pendapatan tahunan dan factor-faktor yang mempengaruhi	63
Tabel 5.2: Estimator Parameter Regresi Kuadrat Terkecil	64
Tabel 5.3: Anava	64
Tabel 5.4: Hasil VIF dan TOL	65
Tabel 5.5: Hasil Proses Pemusatan dan Penskalaan	68
Tabel 5.6: Nilai VIF $\hat{\beta}(k)$ dengan berbagai nilai k	69
Tabel 5.7: Nilai $\hat{\beta}(k)$ dengan berbagai harga k	70
Tabel 5.8: Anava Ridge	73
Tabel 5.9: KMO and Bartlett's Test	74
Tabel 5.10: Anti Image Matrices	75
Tabel 5.11: Total Variance Explained	75
Tabel 5.12: Menentukan Komponen	76
Tabel 5.13: Hasil Regresi dari Nilai Faktor Baru	77
Tabel 5.14: Anava PCA (<i>Principal Component Analysis</i>)	79

DAFTAR GAMBAR

Gambar 5.1: <i>Ridge Trace</i>	71
Gambar 5.2: VIF Plot	71

DAFTAR LAMPIRAN

Lampiran 1: Diskriptif Data	89
Lampiran 2: Uji <i>Multikolinearitas</i>	90
Lampiran 3: Uji Korelasi Parsial	91
Lampiran 4: Anava Ridge	92
Lampiran 5: Anava PCA (<i>Principal Component Analysis</i>)	92
Lampiran 6: Data Faktor Komponen BARu	93
Lampiran 7: Uji Korelasi Parsial Komponen baru PCA	95

DAFTAR SIMBOL DAN SINGKATAN

R	Korelasi populasi
r	Korelasi sampel
R^2	Determinasi populasi
r^2	Determinasi sampel
λ	Nilai <i>eigen</i>
Y	Variabel tak bebas
X	Variabel bebas
β_i	Parameter/koeffisien regresi ke- i
ε	Kesalahan/ <i>error</i>
$\bar{Y}$	Rata-rata dari Y
$\bar{X}$	Rata-rata dari X
S	Standar deviasi
$\hat{\beta}_R$	Estimasi koefisien regresi <i>ridge</i>
z_i	Komponen utama
JKT	Jumlah Kuadrat Total
JKR	Jumlah Kuadrat Regresi
JKE	Jumlah Kuadrat <i>Error</i>
MSE	<i>Means Square Error</i>
$E(X)$	Nilai ekspektasi x
kov	Kovarian
var	variansi
VIF	<i>Variance Inflation Faktors</i>

PERBANDINGAN METODE *RIDGE* DENGAN METODE PCA (*PRINCIPAL COMPONENT ANALYSIS*) DALAM MENGATASI *MULTIKOLINEARITAS*

(Studi Kasus: Pendapatan Tahunan Rumah tangga)

Oleh: Ani Rohmah (06610010)

ABSTRAKSI

Multikolinearitas dalam analisis regresi ganda terjadi bila antara variabel bebas terjadi korelasi. Bila *multikolinearitas* ini terjadi, maka pendugaan dengan metode kuadrat terkecil akan menghasilkan penduga yang masih tetap tak bias dan konsisten, tetapi tidak efisien, sehingga varian dari koefisien regresi menjadi tidak minimum. Dua metode yang dapat digunakan untuk mengatasi *multikolinearitas*, yaitu Metode *ridge* dan metode PCA (*Principal Component Analysis*). Metode *ridge* bertujuan untuk mengurangi *multikolinieritas* dengan menentukan penduga yang bias tetapi mempunyai varians yang lebih kecil dari varians penduga regresi linier berganda. Sedangkan metode PCA (*Principal Component Analysis*) bertujuan untuk membentuk komponen utama sebagai variabel bebas dan dalam kasus *multikolinieritas* regresi komponen utama mempunyai varians penduga yang kecil. Tujuan dari penelitian ini adalah untuk membandingkan antara metode *ridge* dan metode PCA (*Principal Component Analysis*) sebagai model yang terbaik dengan mempertimbangkan nilai koefisien determinasi dan *means square error*. Hasil dari penelitian ini Metode *ridge* lebih efektif dibandingkan dengan metode PCA (*Principal Component Analysis*) dilihat dari nilai *means square error*-nya lebih kecil.

Kata kunci: Koefisien Determinasi, *Means Square Error*, *Multikolinearitas*, PCA (*Principal Component Analysis*), dan Regresi *Ridge*.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Analisis regresi merupakan salah satu metode statistik yang sering digunakan dalam kehidupan sehari-hari. Tujuan analisis regresi yaitu untuk mengetahui sejauh mana hubungan sebuah variabel bebas dengan beberapa variabel tak bebas. Bila dalam analisisnya hanya melibatkan sebuah variabel bebas saja, maka analisis yang digunakan adalah analisis regresi sederhana. Sedangkan, bila analisisnya melibatkan lebih dari satu atau beberapa variabel bebas, maka analisis yang digunakan adalah analisis regresi berganda. Adapun model regresi secara umum adalah sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_i \quad i = 1, 2, \dots, n$$

Dalam sebuah model regresi dikatakan baik, jika dipenuhi asumsi-asumsi sederhana yang sering disebut sebagai asumsi klasik yaitu:¹

1. Nilai rata-rata kesalahan pengganggu nol, yaitu $E(\varepsilon_i) = 0$, untuk $i = 1, 2, \dots, n$
2. Varian $(\varepsilon_i) = E(\varepsilon_i^2) = \sigma^2$, sama untuk semua kesalahan pengganggu (asumsi : *homoscedastic*)
3. Tidak ada *otokorelasi* antara kesalahan pengganggu, berarti kovarian $(\varepsilon_i, \varepsilon_j) = 0, i \neq j$

¹ Supranto J, *Ekonometri Buku Kedua*, (Bogor: Ghalia Indonesia, 2010), hal.10.

4. Variabel Bebas $X_1, X_2, \dots, X_k$, konstan dalam sampling yang terulang dan bebas terhadap kesalahan pengganggu ε_i
5. Tidak ada kolinearitas ganda (*multicollinearity*) di antara variabel bebas X
6. $\varepsilon_i \sim N(0; \sigma^2)$, artinya kesalahan pengganggu mengikuti distribusi normal dengan rata-rata nol dan varian σ^2

Permasalahan pada regresi linear berganda adalah seringkali terjadi korelasi antar variabel-variabel bebas pada model regresi linear berganda yang disebut sebagai multikolinearitas. Jika ada data yang terdapat multikolinearitas berarti salah satu asumsi klasik regresi linear dilanggar. Hal ini berarti bahwa penduga yang dihasilkan dari metode kuadrat terkecil tidak bersifat BLUE. Penduga yang dihasilkan masih tetap tak bias dan konsisten, tetapi tidak efisien sehingga variansi dari koefisien regresi menjadi tidak minimum.

Metode kuadrat terkecil merupakan salah satu cara untuk mendapatkan koefisien regresi pada persamaan regresi linear berganda dengan cara meminimumkan jumlah kuadrat errornya. Jika multikolinearitas yang hampir sempurna terjadi, meskipun metode kuadrat terkecil dapat digunakan tetapi *error* yang dihasilkan akan menjadi besar, variansi dan kovariansi parameter tidak terhingga.

Ada beberapa cara untuk mengatasi masalah *multikolinearitas* yaitu informasi dugaan sebelumnya, mengombinasikan data *cross section* dan *time series*, mengeluarkan suatu variabel dan kesalahan (bias) spesifikasi, pengeluaran atau pengadaan data baru, metode PCA (*prinpicipal analysis*

component), dan metode regresi *ridge*.² Dari berbagai cara untuk mengatasi multikolinearitas di atas, metode *ridge* dan PCA sangat menarik jika dilakukan suatu perbandingan.

Metode regresi *ridge* merupakan modifikasi dari metode kuadrat terkecil dengan cara menambah tetapan bias k yang kecil pada diagonal matrik $X'X$. Tujuan metode ini adalah memperkecil variansi estimator koefisien regresi. Sedangkan, metode PCA (*prinpicipal analysis component*) ini digunakan untuk meminimumkan masalah *multikolinearitas* tanpa harus mengeluarkan variabel bebas yang terlibat hubungan kolinear. Tujuan metode ini adalah untuk menyederhanakan variabel yang diamati dengan cara menyusutkan (mereduksi) dimensinya. Hal ini dilakukan dengan cara menghilangkan korelasi di antara variabel bebas melalui transformasi variabel bebas asal ke variabel baru yang tidak berkorelasi sama sekali atau yang biasa disebut dengan komponen utama.

Berdasarkan penjelasan di atas, peneliti tertarik untuk membandingkan metode regresi *ridge* dan PCA (*principal component analysis*) sebagai penyelesaian masalah multikolinearitas. Oleh karena itu penulis mengangkat judul untuk penelitian ini yaitu “*Perbandingan Metode Ridge Dengan Metode PCA (Principal Component Analysis) Dalam Mengatasi Multikolinearitas*”.

1.2 Batasan Masalah

Batasan masalah merupakan suatu hal yang penting dalam suatu penulisan agar sesuai dengan tujuan yang dimaksud. Agar penyelesaian

² Damodar N Gujarati, *Dasar-dasar Ekonometrika*, (Jakarta: Salemba Empat, 2011), hal.434-440.

masalah tidak menyimpang dari pembahasan maka perlu dibuat suatu pembatasan masalah yaitu membandingkan metode *ridge* dengan metode *PCA* (*principal component analysis*) sebagai penyelesaian masalah *multikolinearitas*.

1.3 Rumusan Masalah

Berdasarkan penjelasan latar belakang di atas, maka dapat dirumuskan masalah-masalah yang terjadi, adalah:

1. Bagaimana mendeteksi ada tidaknya salah satu pelanggaran asumsi regresi linear klasik yaitu *multikolinearitas*.
2. Bagaimana akibat mengabaikan masalah *multikolinearitas*.
3. Bagaimana penanggulangan masalah *multikolinearitas* dengan metode *ridge* dan *PCA* (*principal component analysis*).
4. Membandingkan metode *ridge* dan metode *PCA* (*principal component analysis*) sebagai penyelesaian *multikolinearitas*.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah di atas, maka tujuan penulisan skripsi ini sebagai berikut:

1. Mendeteksi ada tidaknya pelanggaran asumsi regresi linear klasik yaitu *multikolinearitas*.
2. Menerapkan atau menanggulangi masalah *multikolinearitas* dengan metode *ridge* dan *PCA* (*principal component analysis*).

3. Mengatahui akibat yang diperoleh jika mengabaikan masalah *multikolinearitas*
4. Membandingkan metode *ridge* dan *PCA (principal component analysis)* sebagai penyelesaian *multikolinearitas* yang baik.

1.5 Manfaat Penelitian

Hasil dari penelitian ini diharapkan dapat memberikan manfaat, antara lain sebagai berikut:

1. memberikan pengetahuan dasar tentang regresi, metode *ridge* dan *PCA (principal component analysis)*.
2. Memberikan penjelasan tentang akibat mengabaikan masalah *multikolinearitas*.
3. Memberikan penjelasan tentang mengaplikasikan metode *ridge* dan *PCA (principal component analysis)* dalam menyelesaikan masalah *multikolinearitas*.
4. Menambah pengetahuan dan wawasan bagi pembaca dan peneliti yang lain tentang penerapan dan perbandingan metode *ridge* dan *PCA (principal component analysis)*

1.6 Tinjauan Pustaka

Menurut Istikomah pada skripsinya yang berjudul “*Estimasi Principal Component Regresi Untuk Penanganan Kasus Multikolinearitas*” Program Studi Statistika Jurusan Matematika dan Ilmu Pengetahuan Alam Universitas Gadjah Mada Yogyakarta 2009. Dalam skripsi ini juga membahas tentang

PCR dan mengaplikasikannya pada Pemakaian Bensin untuk 25 Jenis Mobil. Dalam penelitian ini dihasilkan kesimpulan bahwa metode PCR merupakan estimasi bias yang menghasilkan nilai estimasi untuk β bias tetapi akan memperkecil nilai variansi yang jauh lebih kecil dibandingkan dengan metode estimasi tak bias, selain itu kesalahan standarnya akan tetap konsisten terhadap adanya perubahan dalam data dan standar errornya jauh lebih kecil dibandingkan dengan metode *least square*.

Menurut Siti Fatimah pada skripsinya yang berjudul “*Mengatasi Multikolinearitas dengan Regresi Ridge*” Program Studi Statistika Jurusan Matematika dan Ilmu Pengetahuan Alam Universitas Gadjah Mada Yogyakarta 2010 yaitu membahas tentang regresi *ridge* dengan mengaplikasikannya pada Pengaruh Banyaknya Usaha, Nilai Produksi Bruto, Biaya Antara dan Upah Gaji Terhadap Jumlah Tenaga Kerja Yang Dibayar Menurut Propinsi, Tahun 2000. Dalam penelitian ini diambil kesimpulan bahwa metode regresi *ridge* merupakan metode yang akurat untuk mengatasi masalah *multikolinearitas* dalam analisis regresi ganda.

Skripsi ini membahas tentang perbandingan metode *ridge* dan *PCA* (*principal component analysis*) pada masalah *multikolinearitas*. Perbedaan penelitian ini dengan penelitian sebelumnya adalah penelitian sebelumnya hanya menggunakan satu metode saja dalam menyelesaikan suatu masalah *multikolinearitas*. Sedangkan, pada penelitian ini dengan menggunakan dua metode yang bersifat membandingkan. Kedua metode tersebut yaitu metode *ridge* dan *PCA* (*principal component analysis*).

1.7 Sistematika Penulisan

Sistematika penulisan ini disusun untuk mempermudah pembahasan pada hasil penelitian. Sistematika penulisan pada skripsi ini yaitu terdiri dari beberapa bab, antara lain:

BAB I : Pendahuluan

Bab ini berisikan tentang latar belakang, batasan masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, tinjauan pustaka, dan sistematika penulisan.

BAB II : Dasar Teori

Bab ini membahas dasar teori yang berisikan tentang teori-teori yang berkaitan tentang metode *ridge* dan *PCA* (*principal component analysis*).

BAB III : Metode Penelitian

Bab ini membahas metodologi yang digunakan dalam mencapai tujuan penelitian

BAB IV : Metode *Ridge* dan Metode *PCA* (*Principal Component Analysis*)

Bab ini merupakan pembahasan dari penelitian yang dilakukan, yaitu tentang penyelesaian masalah *multikolinearitas* dengan membandingkan antara metode *ridge* dengan *PCA* (*principal component analysis*)

BAB V : Studi Kasus

Bab ini membahas tentang penerapan metode regresi *ridge* dengan *PCA* dalam studi kasus pendapatan tahunan rumah tangga.

BAB VI : Penutup

Bab ini berisi tentang kesimpulan dari pembahasan di atas dan saran-saran yang berkaitan tentang penelitian ini.

BAB VI

PENUTUP

6.1 Kesimpulan

Berdasarkan hasil studi yang dilakukan penulis tentang perbandingan metode *ridge* dan metode PCA (*principal component analysis*) dalam mengatasi *multikolinearitas* pada data pendapatan tahunan rumah tangga, maka dapat diambil kesimpulan sebagai berikut:

1. Setelah melakukan pengolahan data pendapatan tahunan rumah tangga, hasil yang diperoleh adalah data pendapatan tahunan rumah tangga terdapat masalah *multikolinearitas*. cara untuk mengetahui *multikolinearitas* pada data tersebut dapat dilihat dari besarnya nilai VIF (Nilainya lebih dari 10) dan nilai korelasi ($-1 \leq r \leq 1$) pada variabel bebas.
2. Akibat jika *multikolinearitas* diabaikan adalah menghasilkan varian yang tinggi dan bias.
3. Metode *ridge* adalah metode untuk menambahkan tetapan bias k pada diagonal matriks $X'X$ yang bertujuan memperkecil variansinya. Nilai $k = 0,02$ yang diperoleh dapat menjadikan koefisien $\hat{\beta}$ lebih stabil. Jadi diperoleh persamaan *ridge* yaitu:

$$\hat{Y}^* = -0,2277X_1^* + 0,5443X_2^* - 0,0813X_3^* + 0,0886X_4^* + 0,3188X_5^*$$

Sehingga terbentuk estimasi yang diperoleh dengan menggunakan regresi *ridge* adalah:

$$\hat{Y} = 10352,24 - 0,9156X_1 + 303,3842X_2 - 0,00727X_3 + 5,168733X_4 + 70,05513X_5$$

4. Metode PCA (*principal component analysis*) bertujuan mendapatkan faktor-faktor komponen baru disebut komponen utama dengan cara mereduksi dimensinya. Faktor-faktor yang diperoleh adalah:

$$F_1 = 0,828X_1 + 0,914X_2 + 0,860X_3 - 0,037X_4 + 0,924X_5 \text{ dan}$$

$$F_2 = -0,030X_1 + 0,060X_2 + 0,472X_3 + 0,993X_4 - 0,0315X_5.$$

Sehingga terbentuk estimasi persamaan:

$$\hat{Y} = 1081,943 + 136,950F_1 - 3,068F_2$$

5. Tabel model terbaik

Menentukan Model yang Terbaik		
Metode	Ridge	PCA
R^2	0,462	0,285
MSE	0,019	50035,100

Menentukan persamaan terbaik dapat dilihat dari nilai R^2 dan MSE . Berdasarkan nilai yang dihasilkan dari R^2 baik dari regresi *ridge* maupun PCA bahwa nilai R^2 kurang dari 0,5 ($R^2_{ridge} = 0,462$ dan $R^2_{PCA} = 0,285$) dan nilai yang dihasilkan dari nilai MSE baik dari regresi *ridge* maupun PCA bahwa nilai MSE *ridge* lebih kecil dari pada nilai MSE PCA. ($MSE_{ridge} = 0,019$ dan $MSE_{PCA} = 50035,100$). Jadi dapat disimpulkan bahwa persamaan yang dapat digunakan sebagai model data di atas adalah persamaan yang dihasilkan dari metode regresi *ridge*.

6.2 Saran

Multikolinearitas merupakan salah satu masalah pada analisis regresi yang menimbulkan model kurang baik untuk peramalan, oleh sebab itu disarankan kepada pembaca untuk terlebih dahulu menghilangkan *multikolonearitas* tersebut. Terdapat beberapa cara untuk menghilangkan *multikolinearitas*, diantaranya yaitu dengan menggunakan metode *ridge*, PCA (*principal component analysis*), menggabungkan data cross section dan time series, mengeluarkan variabel yang terdapat *multikolinearitas*, mengukur ukuran sampel dan lain sebagainya, masih banyak metode-metode yang dapat digunakan untuk menghilangkan masalah *multikolinearitas*.

Demikian saran dari peneliti semoga dapat menjadi inspirasi para peneliti lain dalam bidang statistik khususnya yang mempunyai masalah *multikolinearitas* untuk melanjutkan dan mengembangkan penelitian ini.

DAFTAR PUSTAKA

- Adiningsih, Sri. 2001. *Statistik*. Yogyakarta: BPFE-Yogyakarta.
- Algifari. 2000. *Analisis Regresi*. Yogyakarta: BPFE-Yogyakarta.
- Anton, Howard. 2000. *Dasar-Dasar Aljabar Linear*. Batam Center: Interaksara.
- Drapper. N.R. and Smith. 1981. *Applien Regression Analysis*. Second Edition. New York: Juhn Wiley and Son Inc.
- Diana Supandi, Epha. 2011. *Modul Praktikum Analisis Multivariat Terapan*. Yogyakarta.
- Frank Ayres JR. PhD,. 1984. *MATRIKS*. Jakarta: Erlangga.
- Firdaus, Muhammad. 2004. *Ekonometrika Suatu Pendekatan Aplikasi*. Jakarta: PT Bumi Aksara.
- Gujarati. Damorda N, Dawn C. Porter. 2010. *Dasar-Dasar Ekonometrika*. Edisi 5. Jakarta: Salemba Empat.
- Hoerl, A.E. & R.W. Kennard. 1970. *Ridge Regression: Biased Estimation for Nonorthogonal Problems*. <http://statgen.ucr.edu/file/STAT288/hoerl70a.pdf>. Akses 11 November 2011.
- Johnson, R. and Wichern, D 2007. *Applied Multivariate Statistical Analysis*. Sixth Edition. Pearson Education.
- Jolliffe and Rowlings 1986. *Principal Component Analysis*. New York: Springer Verlag.
- K. A. Stroud. 2003. *Matematika Teknik*. Edisi 5. Jakarta : Erlangga.
- Kutner, M.H., *et al.* 2005. *Applied Linear Statistical Models, Fifth Edition*. McGraw-Hill. New York.
- Lipschutz. Seymour, Ph.D. 2004. *Aljabar Linear*. Edisi 3. Jakarta : Erlangga.
- Montgomery D. C and Peck E. A. 1982. *Introduction to Linear Regression Analysis*, John Wiley & Sons Inc, New York.
- Nachrowi Djalal Nachrowi, M.Sc. 2005. *Penggunaan Teknik Ekonomi*. Jakarta : PT Raja Grafindo.

- Sarwoko, 2005. *Dasar-Dasar Ekonometrika*. Yogyakarta : ANDI.
- Siegel. 1985. *Statistik Nonparametrik*. Jakarta: Gramedia Pustaka.
- Soelistyo. 2001. *Dasar-Dasar Ekonometrika*. Edisi 1. Yogyakarta : BPFE-Yogyakarta.
- Soenarto dan Lincoln Arsyad. 2003. *Metodologi Penelitian Untuk Ekonomi dan Bisnis*. Yogyakarta: UUP AMP YKPN.
- Sumodiningrat, Gunawan M.Ec., Ph.D. 2010. *Ekonometrika Pengantar*. Edisi 2. Yogyakarta : BPFE-Yogyakarta.
- Supranto J. 2010. *Ekonometri Buku Kedua*. Bogor: Ghalia Indonesia.
- .

DAFTAR LAMPIRAN

Lampiran 1

Diskriptif data

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Y	35	342.00	1805.00	1.0819E3	256.62526
X1	35	1985.00	2267.00	2.1363E3	63.81995
X2	35	1.423	3.636	2.74329	.460409
X3	35	1370.00	12710.00	6.0767E3	2868.65252
X4	35	22.40	57.70	39.2429	4.39895
X5	35	6.60000	11.70000	9.9171429 E0	1.16782496
Valid N (listwise)	35				

Lampiran 2

Uji Multikolinearitas

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.679 ^a	.462	.369	203.89804

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1033463.935	5	206692.787	4.972	.002 ^a
	Residual	1205657.950	29	41574.412		
	Total	2239121.886	34			

a. Predictors: (Constant), X5, X4, X1, X3, X2

b. Dependent Variable: Y

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	1126.811	2754.705		.409	.686		
	X1	-.866	1.146	-.215	-.756	.456	.229	4.375
	X2	327.779	284.242	.588	1.153	.258	.071	14.006
	X3	-.013	.041	-.143	-.315	.755	.089	11.208
	X4	6.891	20.133	.118	.342	.735	.156	6.415
	X5	71.931	134.312	.327	.536	.596	.050	20.120

a. Dependent Variable: Y

Lampiran 3

Uji Korelasi Parsial

Control Variables			Correlations				
			X1	X2	X3	X4	X5
Y	X1	Correlation	1.000	.553	.704	-.078	.690
		Significance (2-tailed)	.	.001	.000	.661	.000
		df	0	32	32	32	32
	X2	Correlation	.553	1.000	.736	.065	.818
		Significance (2-tailed)	.001	.	.000	.717	.000
		df	32	0	32	32	32
	X3	Correlation	.704	.736	1.000	.462	.527
		Significance (2-tailed)	.000	.000	.	.006	.001
		df	32	32	0	32	32
	X4	Correlation	-.078	.065	.462	1.000	-.398
		Significance (2-tailed)	.661	.717	.006	.	.020
		df	32	32	32	0	32
	X5	Correlation	.690	.818	.527	-.398	1.000
		Significance (2-tailed)	.000	.000	.001	.020	.
		df	32	32	32	32	0

Lampiran 4

Anava Ridge

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.679 ^a	.462	.369	.13626182

a. Predictors: (Constant), X5, X4, X1, X3, X2

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1033463.935	5	206692.787	4.972	.002 ^a
	Residual	1205657.950	29	41574.412		
	Total	2239121.886	34			

a. Predictors: (Constant), X5, X4, X1, X3, X2

b. Dependent Variable: Y

Lampiran 5

Anava PCA (Principal component Analysis)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.534 ^a	.285	.240	223.68527

a. Predictors: (Constant), F2, F1

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	637998.672	2	318999.336	6.376	.005 ^a
	Residual	1601123.214	32	50035.100		
	Total	2239121.886	34			

a. Predictors: (Constant), F2, F1

b. Dependent Variable: Y

Coefficients^a

		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	1081.943	37.810		28.615	.000		
	F1	136.950	38.362	.534	3.570	.001	1.000	1.000
	F2	-3.068	38.362	-.012	-.080	.937	1.000	1.000

a. Dependent Variable: Y

Lampiran 6

Data Faktor Komponen Baru

F1	F2
-1.606940	-0.445220
-1.650710	-0.480310
-2.740560	0.536623
1.126638	-0.151570
0.687410	-0.390140
-1.018190	0.386749
0.687004	-0.177540
0.179413	4.226496
-0.632670	0.199060
0.486907	0.019036
-0.182190	-0.293390
-0.204200	0.024086
-1.545500	0.366492
1.684355	0.209506
0.532759	-0.225790
0.753047	-0.117310
0.586312	0.026855
0.456507	-0.129460
0.767700	-0.455150
0.762884	0.016730
0.631524	0.075869
0.494492	0.076979
-1.158460	-0.064290
1.251182	0.021360
0.504466	0.139906
-0.987410	-0.016830
-0.102760	-3.784780
0.129114	0.008410
-0.398610	-0.167150
-1.112050	0.058568
0.700443	0.106216
0.002005	0.035693
-0.581500	0.105365
-0.036710	0.030114
1.534293	0.228823

Lampiran 7

Uji Korelasi Parsial Komponen baru PCA

Correlations

Control Variables			F1	F2
Y	F1	Correlation	1.000	.008
		Significance (2-tailed)	.	.966
		df	0	32
	F2	Correlation	.008	1.000
		Significance (2-tailed)	.966	.
		df	32	0

Kepada yang terhormat:

Bapak Ki Haryadi sebagai penguji munaqasah

Assalamu'alaikum Wr.Wb

Setelah mengadakan munaqasah atas skripsi tersebut, maka saya menyerahkan perbaikan skripsi sebagaimana yang telah ditentukan bapak sebagai berikut:

No	Topik	BAB	Uraian Perbaikan
1	Metode Analisis Data	3	Ditambahkan: - Mengolah bagaimana? - Melakukan interpretasi dari analisis data/hasil
2	Kasus	5	- Penjelasan data dan mengapa diambil rata-rata kelompok? - Alasan di pakai rata-rata kelompok untuk setiap variable
3	Penyajian Tabel	5	Diurutkan per nilai Y dari yang terkecil ke yang terbesar atau sebaliknya.
4	Penulisan	Di skripsi lipatan bawah	- Angka desimal dan ribuan. - Kata depan “ di atas”

Berdasarkan tabel di atas, terdapat beberapa perbaikan yang belum saya mengerti yaitu:

1. Metode analisis data

Hal ini saya kurang menguasai metode penelitian, sehingga agak susah untuk saya uraikan

2. Kasus

Data pada penelitian ini berjenis data sekunder. Data yang saya peroleh dari buku “*Dasar-dasar Ekonometrika*” dari Gujarati. Jadi kurang tahu alasan dan penjelasan data tersebut.

Setelah menguraikan maksud saya di atas, saya mohon kepada bapak berkenan untuk membantu dan memberikan perbaikan selanjutnya kepada saya. Dikarenakan tanggal 11 februari adalah batas terakhir perbaikan skripsi saya yang telah ditentukan fakultas.

Mahasiswa

Ani Rohmah

NB: Maaf bapak sebelumnya, mendadak nyerehin perbaikan skripsinya, saya minta solusinya bapak. Makasih pak.... :)