PERBANDINGAN *REGRESI RIDGE* DAN *PRINCIPAL COMPONENT ANALYSIS* DALAM MENGATASI MASALAH MULTIKOLINEARITAS

[1)]Irwan dan Hasriani
[1)]Dosen Pada Jurusan Matematika Fakultas Sains dan Teknologi
Universitas Islam Negeri (UIN) Alauddin Makassar
E-mail: irwan.msi@uin-alauddin.ac.id

Abstract*: Multiple linear regression said to be good if it statistic the assumptions such as: normality assumption, heteroskedastisity, an error does not undergo autocorrelation and not occour multicolinearity. On the assumption that problem often arise in the multiple linear regression assumptions are not fulfilled multicolinearity. Multicollinearity is a condition in which the data of the observations of the independent variables occuror have a relationship that is likely to be high. This study aimed to compare the appropriate method to over come multicollinearity between ridge regression and principal component analysis. Comparison criteria used both methods, the mean square error (MSE) and the coefficient of determination (R^2), from the data is the simulation with Microsoft Excel then the analysis was performed, in order to obtain the data first using ridge regression has a value of MSE of 0.02405 and R^2 of 82.4%, while the principal component analysis MSE value of 14.14 and $R^2$of 37.5% while the data second using ridge regression MSE has a value of 0.00216 and R^2 of 96.9%, while the principal component analysis MSE values of 5.15 and R^2 of 69.5%. From these results it can be concluded that ridge regression method is better used.*

Keywords*: Multiple linear Regression, Multicollinearity, Ridge Regression, Principal Component Analysis*

PENDAHULUAN

Analisis regresi sebagai alat yang digunakan untuk mengetahui hubungan yang terjadi antar beberapa variabel bebas terhadap variabel terikat. Dalam analisis regresi dibutuhkan asumsi-asumsi klasik, asumsi-asumsi tersebut berupa normalitas, heteroskedastisitas, galat tidak mengalami autokorelasi, dan tidak terjadi multikolinearitas. Uji normalitas adalah untuk melihat apakah nilai galat berdistribusi normal atau tidak, sedangkan heteroskedastisitas digunakan untuk melihat apakah terdapat ketidaksamaan varians dari galat satu pengamatan ke pengamatan yang lain, sedangkan untuk uji autokorelasi digunakan untuk melihat apakah terjadi korelasi antara suatu periode ***t*** dengan periode sebelumnya (t-1). Multikolinearitas merupakan suatu kondisi dimana data-data hasil pengamatan dari variabel-variabel bebas terjadi atau memiliki hubungan yang cenderung tinggi. Multikolinearitas dapat menyebabkan suatu model regresi tidak dapat memberikan hasil yang baik jika dijadikan sebagai penduga karena model yang digunakan akan bias. Multikolinearitas dapat dideteksi dengan menggunakan uji *Variance Inflasi Factors* (VIF) dan koefisien korelasi.

Adapun ayat yang menjelaskan tentang tidak ada masalah yang tidak dapat diatasi yaitu terdapat pada (Q.S Al-Insyirah/94: 5-6)

فَإِنَّ مَعَ الْعُسْرِ يُسْراً ﴿٥﴾ إِنَّ مَعَ الْعُسْرِ يُسْراً ﴿٦﴾

Terjemahnya:

> “Maka sesungguhnya bersama kesulitan ada kemudahan, sesungguhmya bersama kesulitan ada kemudahan”

Beberapa metode telah dikemukakan, misalnya penggunaan informasi apriori, mengkombinasikan data *cross section* dan *times series*, pengeluaran dan pengadaan data baru, mengeluarkan suatu variabel yang saling berkolinear, metode *regresi ridge* dan *principal component analysis*. Metode *regresi ridge* merupakan modifikasi dari metode kuadrat terkecil dengan cara menambah tetapan bias **c** yang kecil pada diagonal matriks $\mathbf{Z^TZ}$. Tujuan metode ini adalah memperkecil variabel estimasi koefisien regresi, Sedangkan metode *principal component analysis* ini digunakan untuk meminimumkan masalah multikolinearitas tanpa harus mengeluarkan variabel bebas yang terlibat hubungan kolinear. Tujuan metode *principal component analysis* ini yaitu untuk menyederhanakan variabel yang diamati dengan cara menyusutkan/ mereduksi dimensinya. Hal ini dilakukan dengan cara menghilangkan korelasi diantara variabel bebas melalui transformasi variabel bebas asal ke variabel baru yang tidak berkorelasi sama sekali. Tujuan dari skripsi ini yaitu untuk mendeskripsikan perbandingan hasil *regresi ridge* dan *principal component analysis* dalam mengatasi masalah multikolinearitas.

A. Regresi Linier Berganda

Analisis regresi linear berganda digunakan untuk melihat pengaruh dua variabel predictor atau lebih terhadap satu variabel respon atau untuk membuktikan ada tidaknya hubungan fungsional antara dua variabel bebas (X) atau lebih dengan sebuah variabel terikat (Y)

$$\boldsymbol{Y_i}=\beta_0+\beta_1\boldsymbol{X}_{1i}+\beta_2\boldsymbol{X}_{2i}+\ldots+\beta_k\boldsymbol{X}_{ki}+\varepsilon_i$$

B. Multikolinearitas

Uji multikolinearitas digunakan untuk melihat ada tidaknya korelasi yang tinggi antara variabel-variabel bebas dalam suatu model regresi linier ganda. Jika ada korelasi yang tinggi diantara variabel-variabel bebasnya, maka hubungan antara variabel bebas terhadap variabel terikatnya menjadi terganggu. Penyebab multikolinearitas yaitu Cara pengambilan data dan kecilnya ukuran sampel, Pembatasan pada model atau populasi yang disampel, Spesifikasi model, dan Model yang *overdetermined.* Model yang dimaksud memiliki lebih banyak variabel dibandingkan dengan jumlah sampel.

Pendeteksian multikolinearitas dilakukan dengan *Variansi inflasi factor* (VIF) adalah faktor inflasi penyimpangan baku kuadrat. Jika VIF lebih kecil 0,1 dan lebih besar dari 10 maka hal tersebut menunjukkan adanya kolinearitas antar variabel bebas, rumus VIF: VIF = $\frac{1}{1-R^2}$ dan dapat juga dideteksi dengan melihat nilai Koefisien korelasi dengan rumus:

$$r_{xy} = \frac{n\sum x_i y_i - (\sum x_i)(\sum y_i)}{\sqrt{[n\sum x_i^2 (\sum x_i)^2][n\sum y_i^2 - ((\sum y_i)^2]}}.$$

C. *Regresi Ridge* (*Regresi Gulud*)

Regresi gulud pertama kali diperkenalkan oleh Hoer dan R.W. Kennard yang merupakan salah satu metode untuk mengatasi multikolinearitas dengan cara memodifikasi metode kuadrat terkecil. Metode yang digunakan untuk mendeteksi multikolinearitas adalah *ridge trace* atau jejak gulud. Salah satu kesulitan utama dalam menggunakan regresi gulud adalah dalam menentukan nilai ***c*** yang tepat. *Centering* dan *rescalling* data merupakan bagian dari membakukan variabel. Penduga regresi gulud diperoleh dengan memasukkan suatu konstanta pembiasan kedalam persamaan normal kuadrat terkecil yaitu:

$$\beta^R(c) = (\boldsymbol{Z}^T\boldsymbol{Z} + c\boldsymbol{I})^{-1}\boldsymbol{Z}^T\boldsymbol{Y}^*$$

1. Metode Pemusatan dan Penskalaan(*centering and rescalling*)

$$Y^* = \frac{Q}{(S_{yy})^{1/2}}$$

$$Z_j = \frac{D_j}{(S_{x_j x_j})^{1/2}}$$

Dimana:

S_{yy} : $\sum(Y_i - \bar{Y})^2$

S_{xjxj} : $\sum(X_i - \bar{X})^2$, j = 1, 2,..,p

2. Pemilihan Konstanta c

Pemilihan nilai c merupakan hal yang perlu diperhatikan. Untuk komponen bias didalam kuadrat galat rata-rata (*mean square error*) penduga regresi gulud β^R akan naik jika c bertambah besar (dengan semua β^R cenderung menuju nol) dan keadaan yang sama variansi menjadi lebih kecil. Lebih lanjut juga, bahwa selalu ada nilai ***c*** yang membuat penduga regresi gulud memiliki kuadrat galat rata-rata relative lebih kecil dibandingkan penduga metode kuadrat terkecil. Kesulitannya adalah nilai ***c*** yang optimum itu bervariasi dan penerapan satu dan kepenerapan lainnya tidak.

VIF memiliki nilai mendekati 1 jika variabel bebas **X** tidak saling berkorelasi dengan variabel-variabel bebas lainnya. Oleh karena itu, perlu diperiksa jejak gulud dan nilai-nilai VIF dan kemudian memilik nilai c yang menjadikan koefisien regresi stabil.

D. Principal Component Analysis

Analisis komponen utama merupakan suatu tehnik statistik untuk mengubah dari sebagian besar variable asli yang digunakan yang saling berkorelasi satu dengan yang lainnya menjadi satu set variabel baru yang lebih kecil dan saling bebas (tidak berkorelasi lagi). Jadi analisis komponen utama berguna untuk mereduksi data, sehingga lebih mudah untuk menginterpretasikan data-data tersebut. Bentuk umum persamaan regresi komponen utama adalah:

$$Y = \beta_0 + \beta_1 W_1 + \beta_2 W_2 + \ldots + \beta_k W$$

Vektor eigen adalah kata dari bahasa Jerman dan Inggris. Dalam bahasa Jerman *eigen* diterjemahkan sebagai sebenarnya atau karakteristik, oleh karena itu nilai eigen dapat dikatakan nilai sebenarnya atau karakteristik. Jika A adalah matriks n x n maka vektor tak nol X di dalam R^n dinamakan eigen vektor dari A jika Ax adalah kelipatan scalar dari X_1 yaitu: $A\mathbf{x} = \boldsymbol{\lambda}\mathbf{x}$. Untuk suatu skalar λ. Skalar dinamakan nilai eigen dari A dan x di katakan vektor eigen yang bersesuaian dengan λ untuk menentukan nilai eigen digunakan persamaan: $|\boldsymbol{A} - \boldsymbol{\lambda}\mathbf{I}| = 0$

METODE PENELITIAN

Jenis penelitian yang digunakan yaitu terapan. Terapan yang dimaksud adalah menggunakan metode yang telah ada kemudian akan diterapkan atau diaplikasikan pada penelitian ini.Data yang digunakan dalam penelitian ini yaitu data hasil simulasi yang diperoleh dari program *Microsof Excel* dan data kasus yang diperoleh dari skripsi Nanang Pradipta http://repository.usu.ac.id/ bitstream/ 123456789/ 14037/1 /09E01589.pdf

PROSEDUR PENELITIAN

a. Membangkitkan data
   Melakukan simulasi data pada *Microsof Excel* dengan cara =RAND()

b. Mengatasi multikolinearitas
   1. Mengidentifikasi data yang telah dibangkitkan dengan *Microsof Excel*
   2. Mendeteksi adanya multikolinearitas dengan menggunakan *Varians inflasi factor* (VIF) dan menggunakan uji korelasi antar variabel bebas
   3. Memutuskan hasil pendeteksian adanya multikolinearitas.

c. Langkah mengatasi multikolinearitas dengan menggunakan metode *regresi ridge* (regresi gulud) yaitu:
   1. Berdasarkan data simulasi pada dilakukan transformasi terhadap matriks **X** menjadi **Z** dalam Vektor **y** menjadi Y^* melalui metode *centering* dan *rescalling*.
   2. Bentuk matriks setelah ditransformasi
   3. Mencari nilai c dimana (0 < c < 1)
   4. Tentukan nilai c dengan mempertimbangkan nilai VIF dan β^R
   5. Menguji keberartian regresi
   6. Membuat persamaan model *regresi ridge* (regresi gulud)

d. Langkah mengatasi multikolinearitas dengan menggunakan metode *principal component analysis* atau analisis komponen utama yaitu:
   1. Melakukan standarisasi berdasarkan data simulasi yang terdapat pada Lampiran I
   2. Mencari nilai komponen utama dengan melihat nilai *eigenvalue*
   3. Meregresikan hasil dari komponen utama
   4. Mencari nilai *mean* dan *variance*
   5. Mencari nilai b_0, b_1, b_2, b_3, dan b_4.
   6. Membuat persamaan regresinya

e. Membandingkan hasil *Regresi Ridge* (Regresi Gulud) dan *Principal Component Analysis* (Analisis Komponen Utama) dengan menggunakan nilai *mean square error* (MSE) dan nilai koefisien determinasi (R^2).

HASIL DAN PEMBAHASAN

Data simulasi yang dibangkitkan dengan program *Microsoft Excel*. Pada simulasi data digunakan empat variabel bebas, yaitu (X_1, X_2, X_3, dan X_4) dan variabel terikat Y dengan banyaknya pengamatan yaitu 30. Adapun datanya diberikan pada Tabel berikut.

Tabel 2. Data Simulasi

Y	**X_1**	**X_2**	**X_3**	**X_4**
8	-53	-53	-64	-102
150	156	78	153	73

146	167	116	160	114
142	171	95	166	68
138	137	145	135	170
134	121	127	116	137
130	118	146	113	133
126	90	98	83	94
122	165	167	158	160
118	-38	-74	-43	-83
114	-38	-74	-43	-81
110	116	96	109	90
106	139	112	130	161
102	121	171	133	169
98	107	116	99	113
94	95	95	116	94
6	47	40	38	33
9	46	30	42	29
90	-41	-81	41	-72
86	84	85	77	86
82	83	85	72	80
78	75	60	67	75
9	93	95	82	92
7	-44	-66	-40	-71
74	70	79	63	72
70	65	65	58	67
66	67	70	53	61
62	39	64	48	55
5	-47	-85	-54	-89
3	-53	-34	-54	-89

A. Pendeteksian Multikolinearitas

Hasil pendeteksian VIF

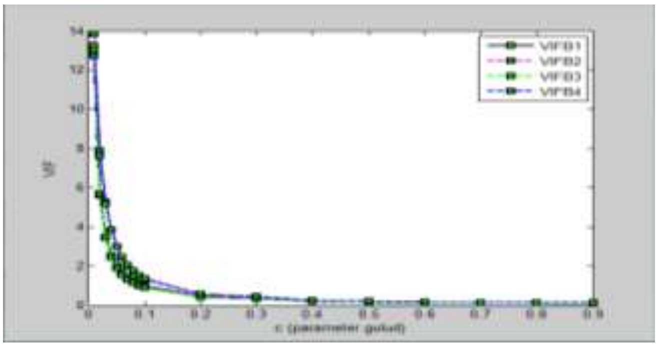
```
Predictor     Coef   SE Coef      T      P    VIF
Constant     53.15     11.11   4.79  0.000
x1         -0.2034    0.9612  -0.21  0.834   99.9
x2         -0.5707    0.4984  -1.15  0.263   32.9
x3          1.0245    0.9888   1.04  0.310  105.8
x4          0.1921    0.4477   0.43  0.672   30.3
```

Tabel 3 Nilai Koefisien Korelasi

Variabel	X_1	X_2	X_3	X_4
X_1	1	0,95	0,99	0,94
X_2	0,95	1	0,96	0,99
X_3	0,99	0,96	1	0,94
X_4	0,94	0,99	0,94	1

B. Penyelesaian Multikolinearitas dengan *Regresi Ridge*

1. Transformasi data dengan metode (*Centered and rescalling*)
2. PenentuanNilai c



Gambar 1. Plot Nilai VIF vs c

Berdasarkan Gambar diatas skala yang digunakan untuk c adalah 0,1 sedangkan untuk VIF dipakai skala 1. Adapun titik-titiknya, ada empat warna dimana untuk besarnya VIF $\beta^R{}_1(c)$ berwarna hitam, untuk warna merah mewakili nilai-nilai VIF $\beta^R{}_2(c)$, sedangkan warna hijau mewakili nilai-nilai VIF $\beta^R{}_3(c)$, dan untuk nilai VIF $\beta^R{}_4(c)$ berwarna biru. Sedangkan metode regresi gulud mengatasi masalah multikolinearitas pada data simulasi ini yaitu ketika c = 0,02, dimana nilai VIF dari masing-masing koefisien regresi gulud lebih kecil dari 10.

3. Menentukan Penduga Koefisien estimator *regresi ridge*

Menentukan nilai penduga koefisien estimator *regresi ridge* dengan rumus

$$\beta^R(c) = (\boldsymbol{Z}^T\boldsymbol{Z} + c\boldsymbol{I})^{-1}\boldsymbol{Z}^T\boldsymbol{Y}^*$$

Sehingga diperoleh nilai c = 0,02 sudah mulai ada penurunan, sedangkan c yang memberikan nilai VIF relatif dekat dengan 1, yaitu pada c = 0.80 ini menunjukkan bahwa pada c = 0,80 koefisien $\hat{\beta}$ lebih stabil. Dengan demikian persamaan *regresi ridge* yang diperoleh jika c yang diambil sebesar 0,80 yaitu:

$$\hat{Y}* = 0{,}1853\ Z_1 + 0{,}0598Z_2 + 0{,}1913Z_3 + 0{,}0800Z_4$$

4. Uji Keberartian Regresi

Table 4. ANOVA Regresi Gulud

Sumber Variasi	SS	Db	MS	F hitung	F Table
Regresi	0,48865	4	0,12216	5,97	2,76
Error	0,51136	25	0,02045		
Total	1,00001	29			

Berdasarkan Tabel 4 dapat diperhatikan nilai MSE regresi gulud (MSE = 0,02045) lebih kecil dari nilai MSE untuk *Principal Component Analysis* (14,14) Artinya bahwa regresi gulud selain dapat mengatasi masalah multikolinearitas dapat juga memperkecil nilai MSE. Selanjutnya, nilai $F_{hitung} > F_{tabel}$ yaitu 5,19 > 2,76. Artinya bahwa hipotesis H_0 ditolak dan disimpulkan bahwa terdapat pengaruh variabel bebas terhadap variabel terikat. Untuk melihat seberapa besar pengaruh variabel bebas terhadap variabel terikat dapat dilihat dari nilai R^2, dimana nilai R^2 untuk regresi gulud yaitu 82,4%.

C. Penyelesaian Multikolinearitas *Principal Component Analysis*

1. MelakukanStandarisasi
2. Mencari nilai komponen utama dengan melihat nilai *eigenvalue*

   Hasil komponen yang terpilih pada data simulasi

```
Eigenvalue  3.8509  0.1270  0.0172  0.0049
Proportion   0.963   0.032   0.004   0.001
Cumulative   0.963   0.994   0.999   1.000

Variable      PC1     PC2     PC3     PC4
z1         -0.500   0.513   0.070   0.694
z2         -0.500  -0.478  -0.720   0.065
z3         -0.501   0.484  -0.038  -0.716
z4         -0.499  -0.523   0.690  -0.042
```

Dari output hasil analisis komponen utama terlihat bahwa hanya ada satu komponen yang terbentuk ini terbukti dengan melihat nilai *eigenvalue* yang lebih besar dari 1, yaitu 3,8509 artinya komponen yang dapat terbentuk dapat menjelaskan 96,3% dari nilai *comulative* dan *proportion*.

3. Meregresikan hasil dari komponen utama

```
The regression equation is
y = 82.8 - 15.2 w1

Predictor      Coef  SE Coef      T      P
Constant     82.833    7.173  11.55  0.000
w1          -15.230    3.718  -4.10  0.000
S = 39.2890   R-Sq = 37.5%   R-Sq(adj) = 35.2%
```

Tabel 5 ANOVA Hasil PCA

Sumber Variasi	SS	Db	MS	F Hitung	F Table
Regresi	33777	4	84,44	5,97	2,76
Error	35347	25			
Total		29	14,14		

sehingga persamaannya:

$Y = 82{,}8 - 15{,}2\ W_1$

Dimana:

$W_1 = -0{,}500\ Z_1 - 0{,}500\ Z_2 - 0{,}501\ Z_3 - 0{,}499\ Z_4$

Artinya bahwa jika variabel terikat (Y) tidak dipengaruhi oleh variabel bebas (X) maka variabel Y akan bernilai 82,8. Selanjutnya untuk setiap kenaikan satu satuan pada variabel X akan mengakibatkan meningkatnya variabel Y sebesar 15,2. Pada output hasil analisis regresi terlihat bahwa signifikansi bernilai 0.000 artinya bahwa nilai sig0,000 < 0,05. Maka dapat disimpulkan bahwa variabel bebas memiliki pengaruh terhadap variabel terikat.

Untuk mengetahui berapa besar pengaruh variabel bebas terhadap variabel terikat dapat dilihat melalui nilai R^2. Berdasarkan pada Persamaan (2.6) diperoleh nilai R^2 = 0,375 artinya bahwa besarnya pengaruh variabel bebas terhadap variabel terikat sebesar 37,5 %.

PEMBAHASAN

1. Pendeteksian multikolinearitas

Pendeteksian multikolinearitas dengan menggunakan VIF semua variabel memiliki nilai yang lebih besar dari 10 yaitu X_1 = 243,1 X_2 = 12,3 X_3 = 338,5 X_4 = 67,8. Dapat disimpulkan bahwa data di atas terdapat multikolinearitas.

Pendeteksian multikolinearitas yang kedua menggunakan koefisien korelasi dimana nilai koefisien korelasi yang terdapat pada TabelVI.2 semua nilainya mendekati 1 sehingga dapat pula disimpulkan bahwa hal ini terdapat multikolinearitas.

2. **Penyelesaian Multikolinearitas dengan *Metode Regresi Ridge***

Nilai dugaan regresi gulud yang diperoleh dengan memilih nilai c = 0,80. Sehingga model yang diperoleh berdasarkan nilai koefisien regresi yaitu:

$$Y^* = 0{,}1853\ Z_1 + 0{,}0598\ Z_2 + 0{,}1913\ Z_3 + 0{,}0800\ Z_4$$

Model di atas menunjukkan bahwa terjadi hubungan antar variabel bebas terhadap variabel terikat. Artinya bahwa ketika nilai variabel bebas dinaikkan satu satuan maka variabel terikat juga akan meningkat satu satuan. Setelah model didapatkan, hasil dari transformasi ulang dari regresi gulud yaitu:

Sehingga persamaan regresi diperoleh sebagai berikut:

$$\hat{Y} = 82{,}8 + 0{,}6724\ X_1 + 0{,}6079\ X_2 + 0{,}6720\ X_3 + 0{,}5689\ X_4$$

Artinya bahwa jika nilai X_1 dinaikkan satu satuan sedangkan variabel lainnya tetap maka taksiran Y akan meningkat sebesar 0,6724, begitupun untuk X_2, X_3, dan X_4. Sehingga jika semua variabel bebas dinaikkan satu satuan maka nilai taksiran Y akan meningkat sama dengan masing-masing koefisien variabel bebas.

3. **Penyelesaian Multikolinearitas dengan *principal Component Analysis***

Data yang mengandung multikolinearitas diatasi dengan menggunakan *principal component analysis* (PCA), data distandarisasi untuk mengurangi korelasi, nilai komponen yang diperoleh hanya 1 yang memiliki nilai *Eigevalue* yang lebih besar 1 yaitu 3,8509. Sehingga diperoleh persamaan regresinya yaitu:

$$Y = 82.8 - 15{,}2\ W_1$$

Dimana:

$$W_1 = -0{,}500\ Z_1 - 0{,}500\ Z_2 - 0{,}501\ Z_3 - 0{,}499\ Z_4$$

Artinya bahwa jika variabel terikat (Y) tidak dipengaruhi oleh variabel bebas (X) maka variabel Y akan bernilai 82,8. Selanjutnya untuk setiap kenaikan satu satuan pada variabel X akan mengakibatkan meningkatnya variabel Y sebesar 15,2. Pada output hasil analisis regresi terlihat bahwa signifikansi bernilai 0.000 artinya bahwa nilai sig 0,000 < 0,05. Maka dapat disimpulkan bahwa variabel bebas memiliki pengaruh terhadap variabel terikat.

selanjutnya persamaan baru yang dibentuk dengan mencari nilai $b_0, \ldots, b_4$ yaitu:

$$\hat{Y} = 82{,}8 - 0.0065077\ X_1 - 0.0061130\ X_2 - 0.0044607\ X_3 - 0.0037023\ X_4$$

KESIMPULAN

Pada data simulasi, nilai MSE *regresi ridge* (0,02405) < dari nilai MSE *principal component analysis* (14,14), sedangkan untuk nilai R^2*regresi ridge* (82,4%) > *principal component analysis* (37,5%). Dari hasil MSE dan R^2 dapat disimpulkan bahwa metode *regresi ridge* lebih baik dari *principal component analysis*, karena selain nilai MSE untuk *regresi ridge* minimum metode ini juga memberikan nilai R^2 yang besar.

Data kasus, Nilai MSE *regresi ridge* (0,00216) < dari nilai MSE *principal component analysis* (5,15), sedangkan untuk nilai R^2*regresi ridge* (96,9%) > *principal component analysis* (69,5%). Dari hasil MSE dan R^2 dapat disimpulkan bahwa metode *regresi ridge* lebih baik dari *principal component analysis*. karena selain nilai MSE untuk *regresi ridge* minimum metode ini juga memberikan nilai R^2 yang besar. Setelah dilakukan analisis baik data simulasi yang dibangkitkan dengan *Microsoft Excel* maupun data kasus yang telah diteliti sebelumnya semua menunjukkan metode *regresi ridge* yang baik dalam mengatasi multikolinearitas.

DAFTAR PUSTAKA

Baroroh, Ali. *Analisis Multivariat dan Time Series dengan SPSS 21*. Jakarta: PT Elex Media Komputindo: 2013

Boediono. *Statistik dan Probabilitas*. Bandung: PT Remaja Rosdakarya, 2001

Departemen Agama RI. *Al Qur'an danTerjemahannya*. Semarang: PT. Karya Toha Putra Semarang, 2002

Ewalpole, Ronald dan Raymond H Myers.*Ilmu Peluang dan Statistika untuk Insinyur dan Ilmuan Edisi ke-4*.Bandung: ITB, 1990

Ryan, Thomas P. *Modern Regression Method*.Canada: Published Simultaneously, 1997

Norman, Draper dan Smith Harry. *Analisis Regresi Terapan Edisi Kedua*, (Jakarta: Gramedia Pustaka Utama. 1992

Sembiring. *Analisis Regresi Edisi Kedua.* Bandung: ITB, 1995

Tiro, Muhammad Arif. *Analisis korelasi dan regresi edisi* kedua. Makaassar: Makassar university press, 2002

Zain, Sumarno dan Damodar Gurajati. *Ekonometrika Dasar*. Jakarta: Erlangga 1978