INNOVATIVE: Journal Of Social Science Research
Volume 4 Nomor 1 Tahun 2024 Page 10300-10312
E-ISSN 2807-4238 and P-ISSN 2807-4246
Website: https://j-innovative.org/index.php/Innovative

Perbandingan Antara *Latent Root Regression* dan *Ridge Regression* dalam Mengatasi Multikolinearitas

Putu Ariesta Candra W[1✉], I Komang Gde Sukarsa[2], G.K. Gandhiadi[3]
Universitas Udayana, Bukit Jimbaran, Bali
Email: ariestacandraw@gmail.com[1✉]

Abstrak

Analisis regresi adalah metode yang digunakan untuk mengetahui pengaruh variabel independen terhadap variabel dependen. Model regresi dianggap baik ketika asumsi model regresi telah terpenuhi. Salah satu asumsi yang harus dipenuhi dalam analisis regresi linier berganda adalah multikolinearitas. Multikolinearitas terjadi ketika ada hubungan sempurna antara variabel independen. Multikolinearitas dapat dideteksi dengan melihat nilai faktor inflasi varians (VIF) yang lebih besar dari 10. Ada beberapa metode yang dapat digunakan untuk mengatasi masalah multikolinearitas, antara lain *latent root regression* dan *ridge regression*. Regresi akar laten dapat mengatasi multikolinearitas dengan lebih baik daripada regresi ridge dengan membandingkan nilai VIF karena menghasilkan nilai VIF sama dengan 1.
Kata Kunci: *Multikolinearitas, Regresi Akar Laten, Regresi Ridge*

Abstract

Regression analysis is a method used to determine the effect of independent variables on the dependent variable. A regression model is considered good when the assumptions of the regression model have been met. One of the assumptions that must be met in multiple linear regression analysis is multicollinearity. Multicollinearity occurs when there is a perfect relationship between independent variables. Multicollinearity can be detected by looking at the value of the variance inflation factor (VIF) which is greater than 10. There are several methods that can be used to overcome multicollinearity problems, including latent root regression and ridge regression. Latent root regression can better overcome multicollinearity than ridge regression by comparing the VIF value because it produces a VIF value equal to 1.
Keywords: *Multicollinearity, Latent Root Regression, Ridge Regression*

PENDAHULUAN

Analisis Regresi adalah metode yang mampu memodelkan hubungan yang terjadi antara variabel bebas dan variabel terikatnya. Analisis regresi linear yang baik adalah ketika nilai koefisien determinasinya mendekati satu. Koefisien determinasi adalah persentase dari besarnya pengaruh variabel bebas terhadap variabel terikatnya (Kadir, 2019).

Terdapat asumsi-asumsi yang harus dipenuhi sebelum melakukan analisis regresi linear. Satu-satunya asumsi yang harus terpenuhi hanya dalam analisis regresi linear berganda adalah non-multikolinearitas (Suharjo, 2008). Multikolinearitas terjadi apabila terdapat hubungan antar variabel bebas yang bernilai sempurna (Gujarati, 1995). Multikolinearitas akan menyebabkan sensitifnya penafsiran koefisien regresi sehingga berakibat pada kurang pastinya hasil estimasi dan penduga yang berbias. Multikolinearitas dapat dideteksi dengan melihat nilai faktor variansi inflansi (VIF) antara 4 hingga 10 (O'Brien, 2007).

Beberapa cara yang dapat digunakan dalam menyelesaikan masalah multikolinearitas yaitu dengan menghilangkan variabel yang berkorelasi tinggi, menambah jumlah amatan, melakukan transformasikan data ke dalam bentuk lain, atau menggunakan metode regresi lain yang lebih *advance* (Kurniawan & Yuniarto, 2016).

Latent root regression adalah metode perluasan dari metode regresi komponen utama yang menyatukan matriks data dari variabel bebas dan variabel terikatnya yang sudah dibakukan. Komponen utama ini yang akan diregresikan dengan variabel berikat sehingga korelasinya dapat dihilangkan dan masalah multikolinearitas dapat teratasi (Untari & Susanti, 2017). Perbedaan antara *latent root regression* dan *ridge regression* adalah pada nilai akar laten dan vektor laten yang diperoleh dari matriks korelasinya yang merupakan penggabungan dari variabel bebas dan variabel terikatnya yang telah dibakukan. Akar laten yang kecil akan menunjukkan adanya ketergantungan antara variabel bebas dan vektor laten 0 menyatakan bahwa vektor laten tidak mampu meramalkan variabel terikatnya.

Ridge regression adalah metode yang tujuannya adalah untuk mengatasi masalah yang diakibatkan oleh korelasi yang tinggi sehingga menghasilkan penduga yang kurang tepat (Draper & Smith, 1992). Metode ini melakukan penambahan tetapan bias untuk meminimalkan dampak multikolinearitas. Tetapan bias memiliki rata-rata residual yang lebih kecil dari penduga metode kuadrat terkecil. Penentuan tetapan bias dapat meminimumkan varians sehingga diperoleh keadaan yang stabil (Irwan, 2015).

METODE PENELITIAN

Data dalam penelitian ini berupa data simulasi yang terdiri dari sepuluh kelompok data dengan jumlah amatan dan jumlah variabel bebas bervariasi. Model regresi menggunakan sejumlah variabel acak bervariasi dan penambahan satu variabel bebas yang dikorelasikan untuk memunculkan multikolinearitas dengan nilai VIF yang bervariasi dengan menggunakan persamaan berikut (Nisa & Wibowo, 2021).

$$x_i = \sqrt{1-\rho^2}x_{i0} + \rho x_{p0}$$

Data hasil simulasi kemudian diuji untuk mengetahui terdapat atau tidaknya masalah multikolinearitas.

Tahapan analisis data menggunakan *latent root regression* menggunakan langkah-langkah seperti berikut.

a. Membakukan data pada variabel bebas dan variabel terikat. Pembakuan data menjadikan data standar dengan nilai rata-rata 0 dan nilai ragam 1. Pembakuan data dapat dilakukan dengan rumus berikut.

   $Z_Y = \frac{(Y-\bar{y})}{\sqrt{U_y}}$, $U_Y = \sum_{i=1}^{n}(Y_i - \bar{Y})^2$ untuk variabel terikat, dan

   $Z_X = \frac{(X_{ij}-\bar{x}_j)}{\sqrt{U_x}}$, $U_x = \sum_{i=1}^{n}\left(X_{ij} - \bar{x}_j\right)^2$ untuk variabel bebas.

b. Melakukan perhitungan dari matriks korelasi gabungan.

   $$Z^* = [Z_Y, Z_X]$$

   Matriks korelasi digunakan untuk mendeteksi masalah multikolinearitas karena menjelaskan nilai korelasi yang terjadi antar variabelnya (Firdaus, 2015).

c. Melakukan perhitungan akar laten dan vektor latennya dari matriks gabungan.

   Persamaan untuk memperoleh akar laten (Webster et al., 1974):

   $$|Z^{*T}Z^* - \lambda_I| = 0$$

   dan vektor laten diperoleh dari masing-masing akar laten yang memenuhi persamaan:

$$(Z^{*T}Z^{*}-\lambda_j I)|\gamma_j = 0$$

d. Menentukan komponen utama dengan mengeluarkan komponen utama dengan akar laten $\lambda_j \leq 0.05$ dan elemen pertama vektor latennya atau $|\gamma_{0j}| < 0{,}10.$

$$PC_j = Z_y\gamma_{0j} + Z_x\gamma_j^0$$

Nilai akar laten yang kecil menunjukkan kuatnya ketergantungan antar variabel bebasnya dan vektor laten yang bernilai 0 menyatakan ketidakmampuan vektor laten dalam meramalkan variabel terikatnya.

e. Komponen utama terpilih kemudian diregresikan dengan variabel terikatnya.

f. Melakukan pendeteksian terhadap multikolinearitas pada model regresi akar laten yang dihasilkan berdasarkan nilai VIF.

g. Membentuk model regresi ke dalam variabel asal.

Setelah mendapatkan model regresi akar laten, dilakukan pengembalian ke dalam variabel awalnya.

Rumus untuk menghitung koefisien terkecil termodifikasi (Webster et al., 1974):

$$\beta^* = \begin{bmatrix}\beta_1^*\\ \beta_2^*\\ \vdots\\ \beta_r^*\end{bmatrix} = c\sum_{j=0}^{n}\gamma_{oj}\lambda_j^{-1}\begin{bmatrix}\gamma_{1j}\\ \gamma_{2j}\\ \vdots\\ \gamma_{rj}\end{bmatrix}; c = -\left\{\sum_{j=0}^{n}\gamma_{oj}^2\lambda_j^{-1}\right\}^{-1}\left\{\sum_{i=1}^{k}(Y_i-\bar{Y})^2\right\}^{\frac{1}{2}}$$

Untuk penduga koefisien regresi pada variabel awal dilakukan dengan persamaan berikut.

$$\beta_j = \frac{\beta_j^*}{S_j}; S_j = \sqrt{\sum_{j=1}^{n}\left(x_j-\bar{x}_j\right)^2}$$

Dan β_0 diperoleh dari rumus:

$\beta_0 = \bar{y} - \beta_1\bar{x}_1 - \beta_2\bar{x}_2 - \cdots - \beta_n\bar{x}_n.$

Tahapan analisis data menggunakan *ridge regression* menggunakan langkah-langkah seperti berikut.

a. Melakukan transformasikan data dengan pemusatan dan penskalaan

Cara dalam penskalaan adalah dengan melakukan tranformasi variabel bebas dan terikatnya ke dalam persamaan (Maulida, 2022):

$$Z_{ij} = \frac{X_{ij}-\bar{X}_j}{S_{xj}}, Y_i^* = \frac{Y_i-\bar{Y}}{S_Y}, S_{X_j} = \sqrt{\frac{\sum_{i=1}^{n}\left(X_{ij}-\bar{X}_j\right)^2}{n-1}}, S_Y = \sqrt{\frac{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}{n-1}}$$

b. Penentuan nilai tetapan bias (c)

Penentuan nilai tetapan bias (c) dilakukan melalui pendekatan nilai VIF.

c. Melakukan estimasi parameter pada model ridge regression. Estimator ridge regression dapat dinyatakan dalam persamaan:

$$\hat{\beta}_R = (X'X + cI)^{-1}X'Y^*$$

HASIL DAN PEMBAHASAN

Data dalam penelitian ini dihasilkan melalui proses pembangkitan data. Pertama dilakukan pembangkitan data yang memiliki multikolinearitas dengan sejumlah variabel bebas dan variabel terikat sebanyak satu buah. Ada sepuluh kelompok data simulasi hasil pembangkitan data dalam penelitian ini.

Data hasil pembangkitan kemudian diuji kembali mengenai multikolinearitasnya. Uji multikolinearitas pada setiap model regresi didasarkan oleh nilai korelasi antar variabel bebas yang terbentuk dan nilai VIF. Pengujian ini memastikan kembali bahwa data yang digunakan benar mengandung multikolinearitas. Untuk menguji multikolinearitas pada variabel bebas dapat dilihat berdasarkan nilai korelasi dan nilai VIF. Untuk data simulasi 1, nilai korelasi dan nilai VIFnya dapat dilihat dalam Tabel 1.

Tabel 1 Uji Multikolinearitas

X	Korelasi				VIF
	X_1	X_2	X_3	X_4	
$\boldsymbol{X_1}$	1				11,069
$\boldsymbol{X_2}$	0,933	1			11,505
$\boldsymbol{X_3}$	0,939	0,931	1		13,067
$\boldsymbol{X_4}$	0,925	0,939	0, 944	1	12,455

Pada model regresi 1 terlihat bahwa nilai korelasi antarvariabelnya semua berada di atas 0,75 yang menandakan bahwa terjadi ketidakbebasan linear diantara variabel bebas. Hal ini juga diperkuat dengan besarnya nilai VIF dari keempat variabel bebas tersebut yang lebih dari 10, sehingga terbukti bahwa terjadi masalah multikolinearitas dalam model regresi 1. Hal ini juga sama untuk model regresi yang lainnya yang telah diuji yaitu memiliki masalah multikolinearitas jika nilai korelasinya tinggi dan nilai VIF di atas 10.

Latent Root Regression dalam Mengatasi Multikolinearitas

Langkah awal dalam metode *latent root regression* adalah melakukan pembakuan data pada variabel bebas dan variabel terikatnya. Pembakuan data bertujuan menjadikan data memiliki rata-rata 0 dan standar deviasi 1. Setelah pembakuan data, maka dibentuk matriks gabungan dari matriks variabel bebas dan matriks variabel terikat yang telah dibakukan. Matriks untuk data simulasi 1 yang terbentuk adalah sebagai berikut.

$$Z^* = \begin{bmatrix} 0,885 & 0,763 & \dots & 1,066 \\ -0,169 & -0,104 & \dots & 0,010 \\ \vdots & \vdots & \ddots & \vdots \\ 0,415 & 0,424 & \dots & 0,020 \end{bmatrix}$$

Kemudian dibentuk matriks Z^{*T} yang akan dikalikan dengan Z^* untuk memperoleh nilai eigen dan vektor eigen.

$$Z^{*T} = \begin{bmatrix} 0,885 & -0,169 & \dots & 0,415 \\ 0,763 & -0,104 & \dots & 0,424 \\ \vdots & \vdots & \ddots & \vdots \\ 1,066 & 0,010 & \dots & 0,020 \end{bmatrix}$$

Dari matriks korelasi $Z^{*T}Z^*$, dibentuk akar laten λ_j dan vektor laten Γ_j yang bersesuaian, dengan $j = 1,2, \dots, p$. Dengan bantuan Minitab 20, diperoleh akar laten λ_j yaitu:

$\lambda_0 = 4,767$, $\lambda_1 = 0,0797$, $\lambda_2 = 0,07$, $\lambda_3 = 0,0537$, $\lambda_4 = 0,0297$

dan dari akar laten λ_j, diperoleh vektor laten Γ_j seperti berikut:

$$\Gamma_0 = \begin{bmatrix} 0,451 \\ 0,445 \\ 0,446 \\ 0,447 \\ 0,448 \end{bmatrix}, \Gamma_1 = \begin{bmatrix} -0,274 \\ 0,763 \\ -0,037 \\ 0,127 \\ -0,570 \end{bmatrix}, \Gamma_2 = \begin{bmatrix} 0,066 \\ 0,011 \\ -0,806 \\ 0,565 \\ 0,162 \end{bmatrix}, \Gamma_3 = \begin{bmatrix} -0,348 \\ -0,432 \\ 0,386 \\ 0,681 \\ -0,283 \end{bmatrix}, \Gamma_4 = \begin{bmatrix} -0,772 \\ 0,184 \\ 0,031 \\ -0,043 \\ 0,606 \end{bmatrix}$$

Penentuan komponen utama terpilih dilakukan dengan mengeluarkan komponen utama yang memiliki nilai akar laten $\lambda_j \leq 0,05$ dan elemen pertama vektor laten $|\gamma_{0j}| < 0,10$.

Berdasarkan kriteria tersebut maka:

a. Nilai eigen $\lambda_0 = 4,767 \geq 0,05$ dan $|\gamma_{00}| = 0,451 \geq 0,10$ sehingga dipertahankan.
b. Nilai eigen $\lambda_1 = 0,0797 \geq 0,05$ dan $|\gamma_{01}| = 0,274 \geq 0,10$ sehingga dipertahankan.
c. Nilai eigen $\lambda_2 = 0,07 \geq 0,05$ dan $|\gamma_{02}| = 0,066 \leq 0,10$ dipertahankan karena tidak menunjukkan adanya ketidakbebasan linear.
d. Nilai eigen $\lambda_3 = 0,0537 \geq 0,05$dan $|\gamma_{03}| = 0,348 \geq 0,10$ sehingga dipertahankan.
e. Nilai eigen $\lambda_4 = 0,0297 \leq 0,05$ dan $|\gamma_{04}| = 0,772 \geq 0,10$ dipertahankan karena kontribusi koefisien vektor latennya cukup tinggi menunjukkan adanya keteramalan yang tinggi.

Selanjutnya dibentuk komponen utama berdasarkan vektor latennya. Beberapa komponen utama yang terbentuk adalah sebagai berikut:

$$PC_1(C_0) = 0,451\, Z_y + 0,445\, Z_1 + 0,446\, Z_2 + 0,447\, Z_3 + 0,448\, Z_4$$
$$PC_2(C_1) = -0,274\, Z_y + 0,763\, Z_1 - 0,037\, Z_2 + 0,127\, Z_3 - 0,570 Z_4$$
$$PC_3(C_2) = 0,066\, Z_y + 0,011\, Z_1 - 0,806\, Z_2 + 0,565\, Z_3 + 0,162\, Z_4$$

$$PC_4(C_3) = -0{,}348\ Z_y - 0{,}432\ Z_1 + 0{,}386\ Z_2 + 0{,}681\ Z_3 - 0{,}283\ Z_4$$
$$PC_5(C_4) = -0{,}772\ Z_y + 0{,}184\ Z_1 + 0{,}031\ Z_2 - 0{,}043\ Z_3 + 0{,}606\ Z_4$$

Kombinasi yang terbentuk antar variabel bebas yang telah terstandarisasi akan membentuk komponen baru yang tidak berkorelasi atau saling tegak lurus. Dari skor komponen utama yang terbentuk, kemudian diregresikan dengan variabel terikat (Y). Model persamaan regresi yang terbentuk disajikan dalam Tabel 2.

Tabel 2 Model Regresi Akar Laten

Data	Model Regresi Akar Laten
1	$Y = -0{,}825 + 1{,}850\ PC_1 - 1{,}128\ PC_2 + 0{,}270\ PC_3 - 1{,}428\ PC_4 - 3{,}167\ PC_5$
2	$Y = 0{,}416 + 2{,}107\ PC_1 - 4{,}048\ PC_2 - 0{,}722\ PC_3 + 1{,}094\ PC_4$
3	$Y = -0{,}073 + 1{,}860\ PC_1 - 3{,}763\ PC_2 - 0{,}526\ PC_3$
4	$Y = 1{,}269 + 1{,}597\ PC_1 - 3{,}270\ PC_2$
5	$Y = 1{,}057 + 1{,}893\ PC_1 - 3{,}861\ PC_2$
6	$Y = 2{,}626 + 2{,}830\ PC_1 - 0{,}606\ PC_2 + 0{,}535\ PC_3 + 0{,}464\ PC_4 + 0{,}2365\ PC_5 - 2{,}163\ PC_6 + 1{,}643\ PC_7 - 7{,}231\ PC_8 - 0{,}626\ PC_9$
7	$Y = 1{,}696 + 2{,}389\ PC_1 - 0{,}992\ PC_2 - 1{,}688\ PC_3 - 4{,}516\ PC_4 - 1{,}728\ PC_5 - 4{,}162\ PC_6$
8	$Y = -0{,}164 + 2{,}506\ PC_1 - 1{,}788\ PC_2 - 4{,}880\ PC_3 - 2{,}277\ PC_4 - 3{,}591\ PC_5 + 1{,}187\ PC_6 + 1{,}778\ PC_7$
9	$Y = 0{,}700 + 2{,}630\ PC_1 - 7{,}379\ PC_2$
10	$Y = 1{,}752 + 2{,}634\ PC_1 - 7{,}522\ PC_2$

Selanjutnya akan diuji kembali nilai VIF untuk memastikan bahwa multikolinearitas telah teratasi. Untuk model regresi 1 diperoleh seperti berikut.

Tabel 3 Pemeriksaan Multikolinearitas pada Model Regresi Akar Laten

Variabel	VIF	Keputusan
$\boldsymbol{PC_1}$	1	Tidak terdapat multikolinearitas
$\boldsymbol{PC_2}$	1	Tidak terdapat multikolinearitas
$\boldsymbol{PC_3}$	1	Tidak terdapat multikolinearitas
$\boldsymbol{PC_4}$	1	Tidak terdapat multikolinearitas
$\boldsymbol{PC_5}$	1	Tidak terdapat multikolinearitas

Nilai $VIF = 1 < 10$ menunjukkan tidak ada korelasi yang kuat antarvariabel komponen utama. Hal tersebut menunjukkan bahwa masalah multikolinearitas telah teratasi. Sehingga terbukti bahwa *latent root regression* dapat mengatasi masalah multikolinearitas dengan baik.

Selanjutnya untuk memperoleh penduga koefisien regresi akar laten pada Untuk model regresi 1, penduga koefisien pada data yang dibakukan adalah:

$$\beta_1^* = c\left\{\frac{(0{,}451)(0{,}445)}{4{,}767} + \frac{(-0{,}274)(0{,}763)}{0{,}0797} + \cdots + \frac{(-0{,}772)(0{,}184)}{0{,}0297}\right\} = c\{-4{,}554\}$$

dengan

$$c = -\left\{\frac{(0{,}451)^2}{4{,}767} + \frac{(-0{,}274)^2}{0{,}797} + \cdots + \frac{(-0{,}772)^2}{0{,}0297}\right\}^{-1} (488{,}32)^{\frac{1}{2}} = -\{20{,}067\}^{-1}(22{,}098)$$
$$= -0{,}946$$

Jadi, $\beta_1^* = -0{,}946(-4{,}554) = 4{,}306$. Ini adalah koefisien bagi variabel yang telah dibakukan. Selanjutnya koefisien regresi bagi X_1 diperoleh melalui β_1^*/S_1 sehingga

$$\beta_1 = \frac{4{,}306}{5{,}71} = 0{,}754.$$

Dengan cara yang sama diperoleh $\beta_2, \beta_3, \mathrm{dan}\ \beta_4$. Selanjutnya menghitung nilai koefisien regresi β_0 diperoleh melalui:

$$\beta_0 = \bar{y} - \beta_1\bar{x}_1 - \beta_2\bar{x}_2 - \beta_3\bar{x}_3 - \beta_4\bar{x}_4$$
$$= -0{,}825 - (0{,}754)(-0{,}3) - (0{,}705)(-0{,}443) - \cdots - (2{,}06)(-0{,}491) = 0{,}923$$

Berdasarkan nilai-nilai $\beta_0, \beta_1, \beta_2, \beta_3$ dan β_4, maka dapat dibentuk model regresi akar laten pada variabel awal yaitu sebagai berikut:

$$Y = 0{,}923 + 0{,}754\, X_1 + 0{,}705\, X_2 + 0{,}538\, X_3 + 2{,}06\, X_4$$

Model tersebut menginterpretasikan bahwa variabel terikat akan bernilai $0{,}923$ apabila semua variabel bebas konstan. Variabel terikat akan menurun sebesar $0{,}754$ satuan setiap X_1 naik sejumlah satu satuan apabila variabel bebas yang lain konstan. Variabel terikat akan menurun sebesar $0{,}705$ satuan setiap X_2 naik sejumlah satu satuan apabila variabel bebas yang lain konstan. Variabel terikat akan menurun sebesar $0{,}538$ satuan setiap X_3 naik sejumlah satu satuan apabila variabel bebas yang lain konstan. Variabel terikat akan menurun sebesar $2{,}06$ satuan setiap X_4 naik sejumlah satu satuan apabila variabel bebas yang lain konstan. Perhitungan dan interpretasi juga dilakukan dengan langkah yang sama pada model regresi yang lain. Model regresi akar laten secara lengkap disajikan dalam Tabel 4.

Tabel 4 Model Regresi Akar Laten pada Variabel Awal

Data	Model Regresi Akar Laten pada Variabel Awal
1	$Y = 0{,}923 + 0{,}754\, X_1 + 0{,}705\, X_2 + 0{,}538\, X_3 + 2{,}06\, X_4$
2	$Y = 0{,}955 + 0{,}606\, X_1 - 1{,}499\, X_2 + 4{,}179\, X_3 + 0{,}847\, X_4$
3	$Y = 1{,}038 + 2{,}834\, X_1 - 1{,}208\, X_2 + 2{,}604\, X_3 - 0{,}075\, X_4$
4	$Y = 0{,}985 + 1{,}025\, X_1 + 0{,}961\, X_2 + 1{,}177\, X_3 + 0{,}814\, X_4$
5	$Y = 1{,}252 + 1{,}143\, X_1 + 1{,}034\, X_2 + 1{,}051\, X_3 + 0{,}944\, X_4$
6	$Y = 1{,}150 + 1{,}279\, X_1 + 2{,}228\, X_2 + 0{,}923\, X_3 + 1{,}985\, X_4 - 0{,}325\, X_5 + 0{,}997\, X_6 - 0{,}374\, X_7 + 1{,}259\, X_8$
7	$Y = 0{,}902 + 0{,}594\, X_1 - 0{,}172\, X_2 + 2{,}308\, X_3 + 1{,}481\, X_4 + 0{,}984\, X_5 - 1{,}047\, X_6 + 2{,}982\, X_7 + 0{,}869\, X_8$
8	$Y = 1{,}269 - 0{,}907\, X_1 + 3{,}053\, X_2 + 0{,}890\, X_3 + 2{,}492\, X_4 + 2{,}733\, X_5 + 0{,}472\, X_6 + 0{,}317\, X_7 - 1{,}386\, X_8$
9	$Y = 1{,}067 + 1{,}759\, X_1 + 1{,}111\, X_2 + 0{,}690\, X_3 + 0{,}435\, X_4 + 1{,}862\, X_5 + 0{,}303\, X_6 + 1{,}258\, X_7 + 0{,}813\, X_8$
10	$Y = 1{,}047 + 1{,}047\, X_1 + 1{,}203\, X_2 + 0{,}870\, X_3 + 1{,}057\, X_4 + 0{,}850\, X_5 + 0{,}969\, X_6 + 1{,}313\, X_7 + 0{,}827\, X_8$

Ridge Regression dalam Mengatasi Multikolinearitas

Langkah awal dalam metode ridge regression adalah memusatkan dan menskalakan data pada variabel bebas dan variabel terikat. Pemusatan dan penskalaan bertujuan untuk menstandarisasi variabel yang dilakukan dengan mengeluarkan penduga β_0 sehingga terwujud perhitungan model regresi yang lebih sederhana.

Langkah selanjutnya adalah melakukan estimasi regresi ridge dengan melakukan pemilihan tetapan bias (c) yang dilakukan melalui pendekatan nilai VIF. Nilai tetapan bias merupakan bilangan kecil positif yang bernilai antara 0 sampai 1 yang ditambahkan pada diagonal utama $X'X$ agar determinan $X'X$ tidak mendekati 0 akibat multikolinearitas. Nilai VIF dari koefisien $\hat{\beta}_1(c)$ dengan berbagai kemungkinan nilai tetapan bias (c) untuk model regresi 1 disajikan dalam Tabel 5.

Tabel 5 Pendekatan Nilai VIF berdasarkan Nilai Tetapan Bias

$\boldsymbol{Nilai\ c}$	$\boldsymbol{VIF\ X_1}$	$\boldsymbol{VIF\ X_2}$	$\boldsymbol{VIF\ X_3}$	$\boldsymbol{VIF\ X_4}$
0.000000	11.0694	11.5054	13.0672	12.4549
0.001000	10.7427	11.1571	12.6149	12.0388
0.002000	10.4308	10.8249	12.1864	11.6440
0.003000	10.1329	10.5077	11.7800	11.2691
0.004000	9.8482	10.2046	11.3941	10.9127
0.005000	9.5758	9.9148	11.0275	10.5736
0.006000	9.3150	9.6375	10.6787	10.2507
0.007000	9.0651	9.3720	10.3467	9.9430

0.008000	8.8255	9.1176	10.0303	9.6494
0.009000	8.5957	8.8737	9.7287	9.3690
0.010000	8.3750	8.6396	9.4407	9.1012
0.020000	6.5794	6.7414	7.1614	6.9689
0.030000	5.3167	5.4152	5.6303	5.5221
0.040000	4.3925	4.4507	4.5501	4.4922
0.050000	3.6944	3.7265	3.7584	3.7315

Dari berbagai nilai tetapan bias (c) di atas, nilai VIF terlihat mengalami penurunan pada $c = 0{,}005$ yang menunjukkan bahwa koefisien $\hat{\beta}$ lebih stabil pada $c = 0{,}005$, ditunjukkan dengan semua nilai VIF berada di bawah 4. Setelah ditentukan nilai tetapan bias (c), kemudian dilakukan analisis regresi ridge dan diperoleh model regresinya seperti berikut:

Tabel 6 Model *Ridge Regression*

Data	Model Regresi Ridge
1	$Y = 0{,}852 + 0{,}791\,X_1 + 0{,}869\,X_2 + 0{,}763\,X_3 + 1{,}574\,X_4$
2	$Y = 0{,}989 + 0{,}755\,X_1 + 0{,}588\,X_2 + 1{,}992\,X_3 + 0{,}794\,X_4$
3	$Y = 1{,}042 + 1{,}661\,X_1 + 0{,}596\,X_2 + 1{,}080\,X_3 + 0{,}816\,X_4$
4	$Y = 0{,}998 + 0{,}956\,X_1 + 1{,}002\,X_2 + 0{,}626\,X_3 + 1{,}361\,X_4$
5	$Y = 1{,}250 + 0{,}359\,X_1 + 1{,}098\,X_2 + 0{,}977\,X_3 + 1{,}730\,X_4$
6	$Y = 1{,}042 + 1{,}274\,X_1 + 1{,}552\,X_2 + 0{,}903\,X_3 + 1{,}287\,X_4 + 0{,}572\,X_5 + 0{,}836\,X_6 + 0{,}551\,X_7 + 0{,}936\,X_8$
7	$Y = 0{,}963 + 1{,}008\,X_1 + 0{,}644\,X_2 + 1{,}027\,X_3 + 1{,}005\,X_4 + 1{,}106\,X_5 + 0{,}561\,X_6 + 1{,}433\,X_7 + 0{,}910\,X_8$
8	$Y = 1{,}168 + 0{,}651\,X_1 + 1{,}104\,X_2 + 1{,}086\,X_3 + 1{,}224\,X_4 + 1{,}092\,X_5 + 0{,}977\,X_6 + 0{,}721\,X_7 + 0{,}880\,X_8$
9	$Y = 1{,}044 + 0{,}677\,X_1 + 0{,}973\,X_2 + 1{,}188\,X_3 + 1{,}313\,X_4 + 0{,}653\,X_5 + 1{,}355\,X_6 + 0{,}946\,X_7 + 1{,}095\,X_8$
10	$Y = 1{,}053 + 0{,}920\,X_1 + 0{,}446\,X_2 + 1{,}457\,X_3 + 0{,}9033\,X_4 + 1{,}545\,X_5 + 1{,}176\,X_6 + 0{,}110\,X_7 + 1{,}585\,X_8$

Kemudian dilakukan uji multikolinearitas melalui nilai VIF pada model 1 yaitu:

Tabel 7 Pemeriksaan Multikolinearitas pada Model *Ridge Regression*

Variabel	VIF	Keputusan
$\boldsymbol{X_1}$	3,694	Tidak terdapat multikolinearitas
$\boldsymbol{X_2}$	3,726	Tidak terdapat multikolinearitas
$\boldsymbol{X_3}$	3,758	Tidak terdapat multikolinearitas
$\boldsymbol{X_4}$	3,731	Tidak terdapat multikolinearitas

Semua nilai $VIF < 10$ menunjukkan tidak ada korelasi yang kuat antarvariabel bebas. Hal tersebut menunjukkan bahwa masalah multikolinearitas telah teratasi. Sehingga terbukti bahwa *ridge regression* dapat mengatasi masalah multikolinearitas dengan baik.

Perbandingan antara *Latent Root Regression* dan *Ridge Regression*

Latent Root Regression dan *Ridge Regression* adalah metode yang mampu mengatasi masalah multikolinearitas. Kedua metode ini mempunyai cara yang berbeda dalam mengatasi masalah multikolinearitas walaupun ada Langkah-langkah yang hampir mirip. Kedua metode ini sama-sama diawali dengan melakukan standarisasi data pada variabel awalnya yaitu variabel bebas dan variabel terikatnya. Akan tetapi, pada *laten root regression*, hasil pembakuan data kemudian dibentuk matriks korelasi gabungan yang kemudian dicari nilai akar laten dan vektor latennya untuk menentukan komponen utama yang yang tidak berkorelasi atau saling tegak lurus. Setelah mendapatkan komponen utama maka dibentuk skor komponen utama yang kemudian diregresikan dengan variabel terikat untuk membentuk model regresi akar laten. Sedangkan dalam *ridge regression*, setelah pembakuan data maka dilakukan penambahan tetapan bias yang nilainya antara 0 sampai 1 yang bertujuan agar determinan matriks $X'X$ mendekati 0 akibat multikolinearitas.

Pada model regresi yang dibentuk melalui simulasi data, kedua metode yaitu *latent root regression* dan *ridge regression* mampu mengatasi masalah multikolinearitas dengan baik. Hal ini ditandai dengan nilai VIF yang digunakan sebagai uji multikolinearitas telah menunjukkan nilai yang lebih kecil dari 10. Untuk membandingkan kedua metode maka diperlukan adanya pembanding. Kriteria pembanding didasarkan pada nilai VIF berdasarkan nilai VIF awal yang bervariasi tingkat besarnya nilai VIF. Nilai VIF dari variabel X_1 dipilih untuk melihat perbandingan nilai VIF awal dengan nilai VIF setelah data dianalisis dengan kedua metode.

Tabel 8 Perbandingan Nilai VIF

Model	Nilai VIF Awal	Nilai VIF setelah LRR	Nilai VIF setelah RR
1	11,069	1	3,694
2	47,728	1	3,877
3	75,613	1	3,069
4	396,286	1	1,552
5	3190,400	1	3,604
6	19,990	1	3,204
7	43,978	1	3,744
8	87,460	1	2,809

9	730,390	1	1,124
10	4368,000	1	3,674

Keterangan: LRR (*Latent Root Regression*), RR (*Ridge Regression*)

Dari rataan nilai VIF awal yang bervariasi, dapat dilihat bahwa *laten root regression* mampu mengatasi masalah multikolinearitas dengan sempurna karena semua nilai VIF akhirnya bernilai 1 sedangkan *ridge regression* mampu mengatasi multikolinearitas dengan baik tetapi tidak sempurna karena nilai VIF akhirnya tidak lebih kecil atau sama dengan 1. Sehingga dengan kriteria pembanding yaitu nilai VIF, maka dapat disimpulkan bahwa *laten root regression* dapat mengatasi masalah multikolinearitas lebih baik dari *ridge regression* pada nilai VIF awal dengan tingkat rendah (di atas 10) hingga tinggi (hingga ribuan).

SIMPULAN

Berdasarkan pembahasan, dapat disimpulkan bahwa *latent root regression* dan *ridge regression* dapat mengatasi masalah multikolinearitas dengan baik. Latent root regression dapat mengatasi masalah multikolinearitas dengan sempurna yang ditunjukkan dengan nilai VIFnya yang bernilai 1. Sehingga dapat dikatakan bahwa latent root regression lebih baik dari ridge regression dalam penelitian ini berdasarkan kriteria pembandingnya yaitu nilai VIF.

DAFTAR PUSTAKA

Draper, N., & Smith, H. (1992). *Analisis Regresi Terapan*. PT Gramedia Pustaka, Jakarta.

Firdaus. (2015). Bangkitan Dan Tarikan Perjalanan Di Kecamatan Medan Labuhan. *Angewandte Chemie International Edition, 6(11), 951–952.*, *6*, 48.

Gujarati, D. (1995). *Ekonometrika Dasar* (G. Hutauruk (ed.); 4 ed.). Erlangga, Jakarta.

Irwan, M. (2015). Least Square and Ridge Regression Estimation. *Jurnal MSA*, *3*(2), 7–13.

Kadir. (2019). *Satistika Terapan: Konsep, Contoh dan Analisis Data dengan Program SPSS/Lisrel dalam Penelitian* (3 ed.). RajaGrafindo Persada, Jakarta.

Kurniawan, R., & Yuniarto, B. (2016). *Analisis Regresi: Dasar dan Penerapannya dengan R* (1 ed.). Kencana: Jakarta.

Maulida, R. (2022). *Perbandingan Principal Component Regression Dan Regresi Ridge Pada Analisis Faktor-Faktor Indeks Pembangunan Manusia*. Universitas Islam Negeri Malang.

Nisa, K., & Wibowo, R. A. (2021). *Simulasi Data Statistik Menggunakan R: Membuat Aneka Program dan Simulasi Statistik dengan Pemrograman R untuk Berbagai Kajian dalam Statistik* (1 ed.). Teknosain: Yogyakarta.

O'Brien, R. M. (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality and Quantity*, *41*(5), 673–690. https://doi.org/10.1007/s11135-006-9018-6

Suharjo, B. (2008). *Analisis Regresi Terapan dengan SPSS* (1 ed.). Graha Ilmu.

Untari, D. P., & Susanti, M. (2017). Latent root regression dalam mengatasi multikolinearitas. *Pythagoras: Jurnal Pendidikan Matematika*, *12*(1), 23. https://doi.org/10.21831/pg.v12i1.11633

Webster, J. T., Gunst, R. F., & Mason, R. L. (1974). Latent Root Regression Analysis. *Technometrics*, *16*(4), 513–522. https://doi.org/10.1080/00401706.1974.10489232.