LEAST SQUARE AND RIDGE REGRESSION ESTIMATION

ABSTRAK

Metode kuadrat terkecil atau *Ordinary Least Square* (OLS) merupakan suatu metode penaksiran koefisien regresi yang paling sederhana. Jika diantara variabel bebas terjadi multikolinearitas sempurna (koefisien korelasi antar variabel bebas sama dengan 1), maka metode OLS tidak dapat digunakan. Sedangkan jika terdapat multikolinearitas hampir sempurna, meskipun OLS dapat digunakan tapi galat yang dihasilkan akan menjadi besar, serta variansi inflasi faktor akan besar pula, padahal nilai estimasi yang diinginkan haruslah memiliki galat dan variansi yang minimum, sehingga digunakan metode regresi gulud.

Metode regresi gulud merupakan salah satu alternatif yang baik untuk mengatasi multikolinearitas diantara variabel-variabel bebas, karena memberikan tetapan bias yang relatif kecil dan memberikan variansi yang minimum. Adapun estimator regresi gulud yaitu:

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

dimana c bilangan positif, pada umumnya terdapat pada selang [0,1] atau $0 \leq c \leq 1$. Jika c = 0 maka koefisien regresi gulud sama dengan koefisien regresi linear dengan menggunakan metode kuadrat terkecil.

Kata kunci: regresi linear ganda, multikolinearitas, regresi gulud.

A. LATAR BELAKANG

Analisis regresi merupakan salah satu metode dari statistik inferensial yang banyak digunakan oleh peneliti untuk menganalisi data. Analisis regresi bertujuan untuk mengetahui sejauh mana ketergantungan atau hubungan tepat satu variabel terikat dengan satu atau lebih variabel bebas. Jika dalam analisis hanya melibatkan satu variabel bebas, maka analisis yang digunakan adalah analisis regresi linear sederhana, sedangkan jika dalam analisis melibatkan dua atau lebih variabel bebas, maka analisis yang digunakan adalah analisis regresi linear berganda.

Salah satu asumsi yang harus dipenuhi dalam analisis regresi linear ganda adalah tidak terjadi multikolinearitas antar variabel bebas yang termasuk didalam model. Jika asumsi ini tidak terpenuhi, maka sulit bagi peneliti untuk mengetahui varibel bebas yang memiliki pengaruh besar di dalam model regresi. Sehingga diperlukan metode untuk mengatasi hal tersebut. Salah satunya adalah dengan menggunakan regresi gulud.

B. TINJAUAN PUSTAKA

1. Regresi Linear Ganda

Regresi linear merupakan salah satu metode statistika yang digunakan untuk mengetahui hubungan antara variabel terikat dengan satu atau lebih variabel bebas. Apabila banyaknya variabel bebas hanya ada satu maka disebut sebagai regresi linear sederhana. Sedangkan apabila terdapat lebih dari satu variabel bebas, disebut sebagai regresi linear berganda.

Secara umum persamaan regresi linear dengan p variabel bebas dinyatakan dengan:

$$\boldsymbol{Y}_i = \beta_0 + \beta_1\boldsymbol{X}_{1i} + \beta_2\boldsymbol{X}_{2i} + \cdots + \beta_p\boldsymbol{X}_{pi} + \boldsymbol{\varepsilon}_i \qquad (2.1)$$

Analisis regresi memiliki 3 kegunaan, yaitu untuk tujuan deskripsi dari fenomena data atau kasus yang sedang diteliti, untuk tujuan kontrol, serta untuk tujuan prediksi. Selain itu, model regresi juga dapat dimanfaatkan untuk melakukan prediksi pada variabel terikat. Koefisien regresi dapat dibedakan menjadi 2 macam, yaitu:

a. Intersep (β_0)
b. Kemiringan

Beberapa asumsi yang harus diperhatikan dalam model persamaan regresi yaitu:

a. Distribusi normal ($\varepsilon \sim N(\mu, \sigma)$)
b. Heteroskedastisitas.
c. Galat tidak mengalami autokorelasi
d. Uji Multikolinearitas

2. *Ordinary Least Square*

Salah satu metode yang digunakan untuk menaksir koefisien regresi adalah metode kuadrat terkecil atau *Ordinary Least Square (OLS*). Tujuan utama dari metode ini adalah mengestimasi koefisien regresi untuk meminimumkan jumlah kuadrat galat[1].

Penaksir OLS diperoleh dengan meminimumkan jumlah kuadrat galat yaitu:

$$\sum \varepsilon_i^2 = \sum\left(Y_i - \beta_0 - \beta_1 X_1 - \beta_2 X_2 - \cdots - \beta_p X_p\right)^2 \tag{2.2}$$

dalam notasi matriks, sama dengan meminimumkan $\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}$ karena

$$\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon} = [\varepsilon_1 \quad \varepsilon_2 \quad \cdots \quad \varepsilon_n]\begin{vmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{vmatrix} = \varepsilon_1^2 + \varepsilon_2^2 + \cdots + \varepsilon_n^2 = \sum_{i=1}^{n} \varepsilon_i^2$$

sekarang dari Persamaan (2.2) diperoleh

$$\boldsymbol{\varepsilon} = \boldsymbol{Y} - \mathbf{X}\boldsymbol{\beta} \tag{2.2}$$

oleh karena itu,

$$\begin{aligned} \boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon} &= (\boldsymbol{Y} - \mathbf{X}\boldsymbol{\beta})^T(\boldsymbol{Y} - \mathbf{X}\boldsymbol{\beta}) \\ &= \boldsymbol{Y}^T\boldsymbol{Y} - 2\boldsymbol{\beta}^T\mathbf{X}^T\boldsymbol{Y} + \boldsymbol{\beta}^T\mathbf{X}^T\mathbf{X}\boldsymbol{\beta} \end{aligned} \tag{2.3}$$

dengan menggunakan sifat-sifat matriks bahwa

$$\boldsymbol{\beta}^T\mathbf{X}^T\boldsymbol{Y} = (\boldsymbol{\beta}^T\mathbf{X}^T\boldsymbol{Y})^T = \boldsymbol{Y}^T\mathbf{X}\boldsymbol{\beta}$$

selanjutnya menurunkan $\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}$ terhadap $\boldsymbol{\beta}$

$$\begin{aligned} \frac{\partial(\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon})}{\partial\boldsymbol{\beta}} &= \frac{\partial(\boldsymbol{Y}^T\boldsymbol{Y} - 2\boldsymbol{\beta}^T\mathbf{X}^T\boldsymbol{Y} + \boldsymbol{\beta}^T\mathbf{X}^T\boldsymbol{X}\boldsymbol{\beta})}{\partial\boldsymbol{\beta}} \\ &= -2\mathbf{X}^T\boldsymbol{Y} + 2\mathbf{X}^T\mathbf{X}\hat{\boldsymbol{\beta}} \end{aligned}$$

setelah diturunkan untuk mendapatkan persamaan normal, selanjutnya disamadenganlan dengan nol, yaitu:

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

$$\mathbf{X}^T\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}^T\boldsymbol{Y}$$

[1]Sanford, Weisberg, *Applied Linear Regression* (Canada : Published Simultaneously, 1947), h. 7.

dari hasil penyederhanaan persamaan di atas secara aljabar, menghasilkan penaksir untuk β yaitu

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-\mathbf{1}}\mathbf{X}^T\boldsymbol{Y} \tag{2.4}$$

dengan demikian, jika diasumsikan bahwa $E(\boldsymbol{\varepsilon}) = 0$, dan $E(\boldsymbol{\varepsilon\varepsilon^T}) = \sigma^{\mathbf{2}}\mathbf{I}$ dan kolom $\boldsymbol{X}$ adalah independen linear maka,

$$\mathrm{E}(\widehat{\boldsymbol{\beta}}) = E((\mathbf{X}^T\mathbf{X})^{-\mathbf{1}}\mathbf{X}^T\boldsymbol{Y})$$
$$= \boldsymbol{\beta}$$

sehingga $\widehat{\boldsymbol{\beta}}$ adalah sebuah estimasi tak bias untuk $\boldsymbol{\beta}$. Asumsi bahwa $\boldsymbol{\varepsilon_i}$ tidak berkorelasi dan memiliki variansi yang sama yaitu

$$cov(\boldsymbol{\varepsilon_i}, \boldsymbol{\varepsilon_j}) = \delta_{ij}\sigma^2$$

dan diasumsukan juga bahwa $var(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{I}$, akan diperlihtkan bahwa $var[\boldsymbol{Y}] = var(\boldsymbol{\varepsilon})$ yaitu sebagai berikut:

$$var[\boldsymbol{Y}] = E([\boldsymbol{Y} - E(\boldsymbol{Y})][\boldsymbol{Y} - E(\boldsymbol{Y})]^T)$$
$$= \sigma^{\mathbf{2}}\mathbf{I}$$

Selanjutnya untuk $var[\varepsilon]$, yaitu:

$$var[\varepsilon] = E([\boldsymbol{\varepsilon} - E(\boldsymbol{\varepsilon})][\boldsymbol{\varepsilon} - E(\boldsymbol{\varepsilon})]^T)$$
$$= \sigma^2\mathbf{I}$$

Oleh karena $var[\boldsymbol{Y}] = var[\boldsymbol{\varepsilon}] = \sigma^2\mathbf{I}$, sehingga variansi $(\widehat{\boldsymbol{\beta}})$ dapat dijabarkan dengan cara sebagai berikut:

$$\mathrm{Var}(\widehat{\boldsymbol{\beta}}) = \mathrm{var}[(\mathbf{X}^T\mathbf{X})^{-\mathbf{1}}\mathbf{X}^T\boldsymbol{Y}]$$

$$= E\{[(\mathbf{X}^T\mathbf{X})^{-\mathbf{1}}\mathbf{X}^T\boldsymbol{Y} - E((\mathbf{X}^T\mathbf{X})^{-\mathbf{1}}\mathbf{X}^T\boldsymbol{Y})][(\mathbf{X}^T\mathbf{X})^{-\mathbf{1}}\mathbf{X}^T\boldsymbol{Y} - E((\mathbf{X}^T\mathbf{X})^{-\mathbf{1}}\mathbf{X}^T\boldsymbol{Y})]^T\}$$

Dengan melakukan penyederhanaan secara aljabar diperoleh

$$\mathrm{Var}(\widehat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X^TX})^{-\mathbf{1}}$$

selanjutnya memperlihatkan $\hat{\beta}$ mempunyai variansi yang minimum. Untuk itu misalkan ada suatu penduga lain yaitu:

$$(\boldsymbol{\beta}^{*}) = [\{(\mathbf{X^TX})^{-\mathbf{1}}\mathbf{X^T} + \mathbf{B}\}\boldsymbol{Y}] \tag{2.5}$$

untuk **B** adalah matriks yang berukuran $p \times n$. Sehingga nilai harapannya yaitu:

$$E(\boldsymbol{\beta}^{*}) = E\{(\mathbf{X}^{\mathbf{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathbf{T}} + \mathbf{B}\}\boldsymbol{Y}$$
$$= \boldsymbol{\beta} + \mathbf{BX}\boldsymbol{\beta}$$

$\boldsymbol{\beta}^{*}$ merupakan penduga tak bias bagi β, jika $\mathbf{BX}\boldsymbol{\beta} = 0$. Sehingga $\mathbf{BX} = 0$ atau $(\mathbf{BX})^{\mathbf{T}} = 0$, sehingga $\mathbf{X}^{\mathbf{T}}\mathbf{B}^{\mathbf{T}} = 0$. Dengan demikian

$$var(\boldsymbol{\beta}^{*}) = E\{[\boldsymbol{\beta}^{*} - E(\boldsymbol{\beta}^{*})][\boldsymbol{\beta}^{*} - E(\boldsymbol{\beta}^{*})]^{T}\}$$

Dengan menggunakan sifat-sifat ekspektasi maka diperoleh

$$var(\boldsymbol{\beta}^{*}) = \mathbf{I}\boldsymbol{\sigma}^{2}\,(\mathbf{X}^{\mathbf{T}}\mathbf{X})^{-1} + \mathbf{B}\mathbf{B}^{\mathbf{T}}\mathbf{I}\boldsymbol{\sigma}^{2}$$

karena $\mathbf{BB}^{\mathbf{T}}$ adalah matriks kuadratik, maka

$$var(\boldsymbol{\beta}^{*}) \geq var(\widehat{\boldsymbol{\beta}})$$

jadi $\widehat{\boldsymbol{\beta}}$ yang diperoleh dari metode kuadrat terkecil merupakan penaksir linear tak bias yang memiliki variansi minimum untuk β.

3. Multikolinearitas

Istilah multikolinearitas atau kolinearitas ganda pertama kali diperkenalkan oleh Ragnar Frisch, yang berarti adanya hubungan linear yang sempurna di antara beberapa atau semua variabel bebas dalam model regresi. Untuk regresi *p*-variabel, mencakup variabel bebas X_1, $X_2, \cdots, X_p$. Terdapat beberapa penyebab multikolinearitas[2] diantaranya:

1. Cara pengambilan data dan kecilnya ukuran sampel.
2. Pembatasan pada model atau populasi yang disampel.
3. Spesifikasi model.
4. Model yang *overdetermined.* Model yang dimaksud memiliki lebih banyak variabel dibandingkan dengan jumlah sampel.

Ada beberapa cara untuk mengetahui terjadinya multikolinearitas dalam regrsi linear berganda:

[2]Djohart. *Multikolinearitas dan autokorelasi,* http://djonhart.economic policy.info/lecture /be/Bahan Kuliah_7.pdf (25 mei 2011)

1. Faktor variansi inflasi (VIF)
2. Menganalisis Koefisien korelasi sederhana antara variabel bebasnya
3. Nilai koefisien determinasi (R^2)

4. Metode Regresi Ridge

Cara berurutan mencari persamaan regresi yang sesuai tidak dapat digunakan bila semua peubah dalam percobaan diharuskan mempunyai peran serta dalam variabel bebas Y. Bila terdapat kolinearitas ganda yang besar maka metode kuadrat terkecil menghasilkan penaksir tak bias untuk koefisien regresi, tapi penaksir mungkin mempunyai variansi yang besar. Variansi yang besar menimbulkan dua kesulitan yaitu: penaksir mungkin tidak stabil, maksudnya peka terhadap perubahan kecil pada data yang kelihatannya tidak penting dan penaksir cenderung menghasilkan koefisien yang terlalu besar. Suatu cara menghadapi masalah ini adalah dengan tidak menggunakan metode kuadrat terkecil dan menggunakan penaksir yang bias, pada dasarnya kita bersedia menerima sejumlah bias tertentu dalam dugaan agar variansi penaksir dapat diperkecil. Penaksir yang bias diperoleh untuk mendapatkan koefisien regresi dalam model regresi yaitu pada Persamaan (1.1) dinyatakan dengan $\beta_1{}^R, \beta_2{}^R, \dots, \beta_p{}^R$ dan disebut dugaan regresi gulud.[3] Regresi gulud pertama kali diperkenalkan oleh Hoer dan R.W. Kennard (1962) yang merupakan salah satu metode untuk mengatasi multikolinearitas dengan cara memodifikasi metode kuadrat terkecil.

a. Pendugaan parameter regresi gulud.

Dugaan regresi gulud diperoleh dengan meminimumkan jumlah kuadrat galat untuk model pada persamaan

$$\boldsymbol{\varepsilon} = \boldsymbol{Y} - \mathbf{Z}\boldsymbol{\beta}$$

---

[3] Ronald Ewalpole dan Raymond H Myers, *Ilmu Peluang dan Statistika untuk Insinyur dan Ilmuwan edisi ke-4, (Bandung: ITB, 1995)*, h.510.

dengan kendala tunggal $\sum_{j=0}^{p}(\widehat{\boldsymbol{\beta}})^2 = \rho$, dimana ρ tetapan posistif yang berhingga. Berdasarkan metode pengali lagrange fungsi tujuannya yaitu:

$$F = \boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon} + c(\boldsymbol{\beta}^T\boldsymbol{\beta} - \rho)$$

$$= (\boldsymbol{Y}^T\boldsymbol{Y} - \boldsymbol{Y}^T\mathbf{Z}\boldsymbol{\beta} - \boldsymbol{\beta}^T\mathbf{Z}^T\boldsymbol{Y} + \boldsymbol{\beta}^T\mathbf{Z}^T\mathbf{Z}\boldsymbol{\beta} + c\boldsymbol{\beta}^T\boldsymbol{\beta} - c\rho) \quad (2.6)$$

selanjutnya mendiferensialkan Persamaan (4.1) terhadap $\boldsymbol{\beta}$

$$\frac{\partial F}{\partial \boldsymbol{\beta}} = \frac{\partial(\boldsymbol{Y}^T\boldsymbol{Y} - \boldsymbol{Y}^T\mathbf{Z}\boldsymbol{\beta} - \boldsymbol{\beta}^T\mathbf{Z}^T Y + \boldsymbol{\beta}^T\mathbf{Z}^T\mathbf{Z}\boldsymbol{\beta} + c\boldsymbol{\beta}^T\boldsymbol{\beta} - c\rho)}{\partial \boldsymbol{\beta}}$$

$$= \frac{\partial(\boldsymbol{Y}^T\boldsymbol{Y})}{\partial \boldsymbol{\beta}} - \frac{\partial(\boldsymbol{Y}^T\mathbf{Z}\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} - \frac{\partial(\boldsymbol{\beta}^T\mathbf{Z}^T Y)}{\partial \boldsymbol{\beta}} + \frac{\partial(\boldsymbol{\beta}^T\mathbf{Z}^T\mathbf{Z}\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} + \frac{\partial(c\boldsymbol{\beta}^T\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} - \frac{\partial(c\rho)}{\partial \boldsymbol{\beta}}$$

dengan mendisferensialkan tiap suku yaitu sebagai berikut:

$$\frac{\partial(\boldsymbol{Y}^T\boldsymbol{Y})}{\partial \boldsymbol{\beta}} = 0$$

$$\frac{\partial(-\boldsymbol{Y}^T\mathbf{Z}\boldsymbol{\beta}^T)}{\partial \boldsymbol{\beta}} = -\boldsymbol{Y}^T\mathbf{Z}$$

Nilai suatu koefisien regresi jika didiferensialkan, maka hasilnya berupa dugaan terhadap estimator untuk koefisien regresi. sehingga

$$\frac{\partial(\boldsymbol{\beta}^T\mathbf{Z}^T\mathbf{Z}\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = 2\boldsymbol{\beta}^R\mathbf{Z}^T\mathbf{Z}$$

$$\frac{\partial(-c\rho)}{\partial \boldsymbol{\beta}} = 0$$

Setelah didiferensialkan terhadap $\boldsymbol{\beta}$, sehingga diperoleh hasil yaitu:

$$\frac{\partial F}{\partial \boldsymbol{\beta}} = -\boldsymbol{Y}^T\mathbf{Z} - \mathbf{Z}^T Y + 2\boldsymbol{\beta}^R\mathbf{Z}^T\mathbf{Z} + 2c\boldsymbol{\beta}^R \quad (2.7)$$

kemudian menyamakan Persamaan (4.2) dengan nol yaitu:

$$-2\mathbf{Z}^T\boldsymbol{Y} + 2\boldsymbol{\beta}^R\mathbf{Z}^T\mathbf{Z} + 2c\boldsymbol{\beta}^R = 0$$

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

dari penjabaran tersebut sehingga didapat estimator regresi gulud untuk koefisien regresi yang nilainya dipengaruhi oleh besarnya nilai c yaitu:

$$\boldsymbol{\beta}^R(c) = (\mathbf{Z}^T\mathbf{Z} + c\mathbf{I})^{-1}\,\mathbf{Z}^T\boldsymbol{Y} \quad (2.8)$$

Sifat-sifat pendugaan regresi gulud:

a. Bias

Berdasarkan Persamaan (4.3) dimana $\mathbf{Q}$ merupakan matriks konstant, sehingga bias dari estimator gulud dapat diperoleh sebagai berikut:

$$E\left(\boldsymbol{\beta}^R(k)\right) = E((\mathbf{Z}^T\mathbf{Z} + k\mathbf{I})^{-1}\,\mathbf{Z}^T\boldsymbol{Y})$$
$$= E(\mathbf{Q}\hat{\beta})$$
$$= \mathbf{Q}\beta \qquad (2.9)$$

Berdasarkan Persamaan (2.9) maka regresi gulud bukan merupakan penduga yang tak bias, tetapi merupakan penduga yang bias.

b. Variansi minimum

Variansi minimum dapat diperoleh berdasarkan Persamaan (2.9)

$$Var\left(\beta^R(c)\right) = Var(\mathbf{Q}\hat{\beta})$$
$$= \sigma^2\mathbf{Q}\left(\mathbf{Z}^\mathbf{T}\mathbf{Z}\right)^{-\mathbf{1}}\mathbf{Q}^T \qquad (2.10)$$

jadi $\beta^R(c)$ yang diperoleh dari metode regresi gulud merupakan penaksir linear bias yang memiliki variansi minimum untuk β.

Penduga regresi gulud diperoleh dengan memasukkan suatu konstanta pembiasan ke dalam persamaan normal kuadrat terkecil yaitu:

$$(\mathbf{Z}^\mathrm{T}\mathbf{Z} + \mathrm{c}\mathbf{I})^{-\mathbf{1}}\mathbf{Z}^\mathrm{T}\mathbf{Z}(\mathbf{Z}^\mathrm{T}\mathbf{Z} + \mathrm{c}\mathbf{I})^{-1} \qquad (2.11)$$

Pemilihan nilai c merupakan hal yang perlu diperhatikan. Untuk komponen bias di dalam 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 relatif lebih kecil dibandingkan penduga metode kuadrat terkecil. Kesulitannya adalah nilai c yang optimum itu bervariasi dan penerapan satu kepenerapan lainnya tidak[4].

---

[4]Bambang sumantri, *Model Linear Terapan Buku II Analisis Regresi Ganda* (Bogor: FMIPA-IPB.1997), h. 175-181

C. HASIL PENELITIAN

Pada contoh kasus ini, data yang digunakan adalah data simulasi yang dibangkitkan dengan program Microsoft excel. Simulasi data yang dilakukan adalah untuk model regresi linear ganda dengan empat variabel bebas (X_1, X_2, X_3, X_4) dan peubah terikat Y dengan banyaknya pengamatan 30. Adapun datanya diberikan pada Lampiran I.

1. Menentukan koefisien regresi dengan menggunakan metode kudrat terkecil. Dengan menggunakan metode LSE di

$$\hat{\beta} = \begin{bmatrix} 24{,}898 \\ 0{,}462 \\ 0{,}007 \\ 0{,}155 \\ 0{,}342 \end{bmatrix}$$

Sehingga diperoleh persamaan regresi yaitu:

$$\hat{\boldsymbol{Y}}_i = 24{,}898 + 0{,}462\boldsymbol{X}_{1i} + 0{,}007\boldsymbol{X}_{2i} + 0{,}155\mathbf{X}_{3i} + 0{,}034\mathbf{X}_{4i} \qquad (3.1)$$

Uji Signifikansi pada koefisien regresi dengan hipotesis sebagai berikut:

H_0 : $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$

H_1: tidak semua $\beta_j = 0$, $j = 1, 2, \dots, p$

$$F_{hitung} = \frac{MSR}{MSE} = \frac{68847}{9723} = 44{,}26$$

Adapun Tabel ANAVAR untuk metode OLS berdasarkan hasil analisis SPSS yang ada pada Lampiran V yaitu sebagai berikut:

Tabel 4.1 ANAVAR untuk metode kuadrat terkecil

Sumber variasi	SS	Db	MS	F hitung	F
Regresi Error Total	68847 9723 78569	4 25 29	17212 389	44,247	2,76

Karena $F_{\text{hitung}} > F_{\text{tabel}}$ yaitu 44,26 > 2,76 maka gagal menerima H_0. Artinya, bahwa ada variabel bebas yang memiliki pengaruh terhadap variabel terikat. Adapun nilai koefisien determinasi yang dihitung dengan menggunakan persamaan (2.10), $R^2 = 87{,}6\%$. Artinya bahwa variansi total yang dijelaskan variabel bebas terhadap variabel terikat sebesar 87,6%.

a. Uji asumsi

Uji asumsi yang diperhatikan dalam tulisan ini adalah uji multikolinearitas, yang dalam hal ini diperhatikan berdasarkan nilai VIF

a. Variansi Inflasi Faktor (VIF)

Nilai VIF yang diperoleh yaitu:

Tabel 4.4 Variansi Inflasi Faktor

X12	X13	X14	X23	X24	X34
11,619	250,250	50,251	11,125	11,619	71,679

Dari Tabel (4.2) di atas, semua variabel menunjukkan nilai VIF yang lebih besar dari 10, sehingga disimpulkan terjadi multikolinearitas antar varibel bebas.

Karena terjadi multikolinnearitas pada setiap variabel bebas. Sehingga untuk menyelesaikannya digunakanlah regresi gulud. adapun langkah-langkah sebagai berikut:

1. Trasnsformasi (*centered and scaling*)

Setelah diketahui bahwa terjadi multikolinearitas antar variabel bebas langkah selanjutnya adalah mentrasformasi data dengan dengan metode *centered* dan *Scaling.* Adapun hasilnya setelah ditrasformasikan adalah sebagai berikut:

Tabel 4.6 Data Hasil *Centered* dan *Scaling*

N0	Y**	Z_1	Z_2	Z_3	Z_4
1	-0,25996	-0,31102	-0,34202	-0,32803	-0,32075
2	0,246637	0,221796	0,284156	0,230705	0,249741
29	-0,2992	-0,29573	-0,32043	-0,30229	-0,29611
30	-0,30277	-0,31102	-0,21031	-0,30229	-0,29611

2. Penentuan nilai c

Bentuk matriks dari data setelah ditransformasikan adalah sebagai berikut:

$$\mathbf{Z} = \begin{bmatrix} -0{,}311024 & -0{,}342024 & -0{,}328033 & -0{,}320751 \\ 0{,}221796 & 0{,}284156 & 0{,}230705 & 0{,}249741 \\ 0{,}249839 & 0{,}113576 & 0{,}248729 & 0{,}272485 \\ \vdots & \vdots & \vdots & \vdots \\ -0{,}311024 & -0{,}210310 & -0{,}302285 & -0{,}296112 \end{bmatrix}$$

$$\mathbf{Z}^T = \begin{bmatrix} -0{,}311024 & 0{,}221796 & 0{,}249839 & \cdots & -0{,}311024 \\ -0{,}342024 & 0{,}284156 & 0{,}113576 & \cdots & -0{,}210310 \\ -0{,}328033 & 0{,}230705 & 0{,}248729 & \cdots & -0{,}302285 \\ -0{,}320751 & 0{,}249741 & 0{,}272485 & \cdots & -0{,}296112 \end{bmatrix}$$

$$\mathbf{Z}^T\mathbf{Z} = \begin{bmatrix} 1{,}00000 & 0{,}95572 & 0{,}99793 & 0{,}98970 \\ 0{,}95572 & 1{,}00000 & 0{,}95427 & 0{,}94720 \\ 0{,}99793 & 0{,}95427 & 1{,}00000 & 0{,}99306 \\ 0{,}98970 & 0{,}94720 & 0{,}99306 & 1{,}00000 \end{bmatrix}$$

$$\boldsymbol{Y^*} = \begin{bmatrix} -0{,}259959 \\ 0{,}246637 \\ 0{,}232367 \\ \vdots \\ -0{,}302769 \end{bmatrix} \quad \mathbf{Z}^T\boldsymbol{Y} = \begin{bmatrix} 0{,}935849 \\ 0{,}895547 \\ 0{,}935158 \\ 0{,}928713 \end{bmatrix} \quad \boldsymbol{I} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Selanjutnya, hasil perhitungan nilai VIF dari $\beta^R(c)$ dengan berbagai nilai c dan dipilih nilai C=0,05, karena memiliki nilai MSE paling kecil diantara yang lainnya.

C	β_1^R	β_2^R	β_3^R	β_3^R	*MSE*
0,05	**0,306839**	**0,112366**	**0,282042**	**0,227159**	**0,013336**

diperoleh persamaan untuk metode regresi gulud yaitu:

$$Y = 1{,}49417Z_1 + 3{,}75167Z_2 + 1{,}01870Z_3 + 2{,}42124Z_4 \tag{3.2}$$

3. PENUTUP

Regresi gulud merupakan penduga yang dapat digunakan ketika terjadi multikolinearitas antar variabel bebas, koefisien regresi gulud diperoleh dengan menggunakan persamaan.

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

Berdasarkan data yang simulasi maka diperloleh model regresi Ridge yaitu

$$Y = 1{,}49417Z_1 + 3{,}75167Z_2 + 1{,}01870Z_3 + 2{,}42124Z_4.$$

Setelah ditransformasi ulang diperoleh penafsiran koefisien regresi yaitu:

$$\hat{Y} = 485{,}146 + 1{,}068X_1 + 2{,}271X_2 + 1{,}286X_3 + 1{,}286X_4$$

A. Saran

Banyak metode untuk mengatasi masalah multikolinearitas. Analis dapat memilih salah satu diantara semua metode yang lebih baik dari Metode Kuadrat Terkecil. Walaupun regresi gulud belum tentu dapat digunakan untuk menyelesaikan semua model yang mengandung multikolinearitas, tetapi sudah cukup bukti bahwa regresi gulud merupakan salah satu metode yang baik. Ini dikarenakan melalui model ini diusahakan memperoleh varians yang minimum dengan menentukan nilai c sehingga diperoleh keadaan yang lebih stabil.

DAFTAR PUSTAKA

Alan, Lee J dan Seber George. *Linear Regression Analisi Secon Edition*. Canada: Published Simultaneously, 2003.

Djohart. *Multikolinearitas dan autokorelasi (*http : //djonhart.economic-policy.info/ lecture/be/Bahan_Kuliah_7.pdf

Heumann Christian, *et al.,* eds., *Linear Models and Generalizations Least Squares and Alternatives Third Extended Edition.* German: Springer, 1995.

Ilyas, Baharuddin dan Muhammad Arif Tiro. *Statistika Terapan untuk Ilmu Ekonomi dan Ilmu Sosila Edisi Kedua* . Makassar: Adhira Publisher, 2002.

Iriawan, Nur dan Septin Puji Astuti, *Mengolah Data Statistik dengan mudah menggunakan minitab 14*. Yokyakarta: Andi, 2006.

Myers, Raymond H dan Ronald Ewalpole, *Ilmu Peluang dan Statistika untuk Insinyur dan Ilmuwan edisi ke-4.* Bogor: ITB, 1995.

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

Paulson, S. Daryl. *Handbook Of Regression and Modeling*. Francis: Taylor & Francis Grup, 2007.

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

Sumantri, Bambang. *Model Linear Terapan Buku II Analisis Regresi Ganda*. Bogor: FMIPA-IPB.1997.

Sunyoto, Danang, *Analisis Regresi dan Uji Hipotesis* . Yogyakarta : MedPress. 2008.

Tiro, Muhammad Arif. *Analisis Korelasi dan Regresi Edisi Kedua* .Makassar: Makassar university press, 2002).

----------------------------. *Dasar-dasar Statistika*. Makassar: State University Of makassar Press, 1999.

Weisberg, Sanford. *Applied Linear Regression*. Canada : Published Simultaneously, 1947.

Lampiran I. Data Hasil Simulasi

NO	Y	X_1	X_2	X_3	X_4
1	12	-30	-90	-35	-92
2	154	179	200	182	209
3	150	190	121	189	221
4	146	188	100	195	233
5	142	160	150	164	180
6	138	144	132	145	147
7	134	141	151	142	143
8	130	113	103	112	104
9	126	188	172	187	170
10	122	70	45	67	43
11	118	69	35	71	39
12	114	139	101	138	100
13	110	162	177	159	171
14	106	162	176	162	179
15	102	130	120	128	123
16	98	118	100	115	104
17	1	-15	-69	-14	-73
18	0	-15	-69	-14	-71
19	94	116	100	111	102
20	90	107	90	106	96
21	86	106	90	101	90
22	82	98	65	96	85
23	1	-18	-76	-12	-62
24	1	-21	-61	-11	-61
25	78	93	84	92	82

26	74	88	70	87	77
27	70	90	75	82	71
28	66	62	69	77	65
29	1	-24	-80	-25	-79
30	0	-30	-29	-25	-79