ISSN: 2339-2541
JURNAL GAUSSIAN, Volume 4, Nomor 3, Tahun 2015, Halaman XX - XX
Online di: http://ejournal-s1.undip.ac.id/index.php/gaussian

ESIMASI PARAMETER REGRESI *RIDGE* MENGGUNAKAN ITERASI HOERL, KENNARD, DAN BALDWIN (HKB) UNTUK PENANGANAN MULTIKOLINIERITAS

Nur Aeniatus Solekakh[1], Dwi Ispriyanti [2], Sudarno[3]
[1]Mahasiswa Jurusan Statistika FSM Universitas Diponegoro
[2,3]Staff Pengajar Jurusan Statistika FSM Universitas Diponegoro
Email: nuraeniatus30@gmail.com ispriyanti.dwi@gmail.com dsghani@gmail.com

ABSTRACT

Regression analysis is statistical method used to analyze the dependence of respond variables to predictor variable. In multiple linear regression analysis, there are assumptions that must be met, they are normality, homoscedasticity, absence of multicollinearity, and absence of autocorrelation. One of assumption frequently found is multicollinearity. If multicollineraity is exist between predictor variables, then regression analysis with ordinary least square is no longer used. *Ridge* regression is regression method to handle multicollinearity. The *ridge* estimator involves adding biasing constant *(k)* to each diagonal element of **X'X**. Biasing constant *(k)* is determined by Hoerl, Kennard, and Baldwin (HKB) iteration method. This regression can be applied to inflation rate in Indonesia data and the factors that influence, they are BI rate, money supply, and exchange rate of rupiah. *Ridge* regression analysis, the VIF (Variance Inflation Factor) values for each predictor variables BI rate, money supply, and exchange rate of rupiah are 1.6637, 3.2712, and 4.3309. Since VIF values are not exceed to 10, then there is no multicollinearity in *ridge* regression model.

Keywords: *Inflation, Multikollinearity, Ridge Regression, HKB Iteration, VIF*

1. PENDAHULUAN

1.1. Latar Belakang

Indonesia merupakan salah satu negara berkembang di kawasan asia. Perkembangan dan kemajuan suatu negara dapat dilihat dari faktor perekonomian, karena perekonomian menjadi salah satu pondasi utama kekuatan suatu negara. Salah satu indikator utama yang digunakan untuk melihat perkembangan perekonomian suatu negara adalah tingkat inflasi.

Inflasi merupakan suatu situasi dimana harga sebuah barang dan jasa meningkat, misalnya penyewaan apartemen, harga buku, dan hampir semua hal dan semakin lama cenderung mengalami peningkatan (Ristono dan Puryani, 2011). Inflasi sebenarnya dipengaruhi oleh banyak faktor, bahkan politik dan keamanan. Dalam penulisan kali ini, penulis ingin meneliti faktor-faktor yang mempengaruhi laju inflasi, yaitu BI rate, jumlah uang beredar, dan nilai tukar rupiah.

Secara teori, ketiga faktor tersebut memang mempengaruhi laju inflasi, misalkan BI rate yang ditetapkan oleh Bank Indonesia. BI rate yang ditetapkan rendah, masyarakat enggan untuk menyimpan uang di bank dan memilih untuk menggunakan uangnya untuk kegiatan ekonomi yang lain sehingga terjadi inflasi. Selanjutnya, semakin banyak uang beredar semakin tinggi angka inflasi karena menyebabkan daya beli masyarakat menjadi rendah dan harga-harga kebutuhan menjadi naik. Ketika nilai tukar rupiah terhadap dolar rendah, hal ini akan mengakibatkan kenaikan harga barang-barang kebutuhan dan dapat meningkatkan laju inflasi

Untuk melihat pengaruh BI rate, jumlah uang beredar, dan nilai tukar rupiah terhadap tingkat inflasi di Indonesia digunakan analisis regresi. Analisis regresi merupakan metode statistika yang digunakan untuk menganalisis ketergantungan dari variabel respon terhadap variabel prediktor (Suliyanto, 2011). Analisis regresi yang menganalis lebih dari satu variabel prediktor terhadap variabel responnya disebut analis regresi linier berganda. Pada analisis regresi linier berganda, terdapat beberapa asumsi yang harus dipenuhi, yaitu normalitas residual, homoskesdastisitas, tidak adanya autokorelasi, dan tidak adanya multikolinieritas antar variabel prediktor. Pada asumsi-asumsi tersebut, parameter dapat diestimasi dengan menggunakan metode kuadrat terkecil.

Bila salah satu asumsi tidak terpenuhi, yaitu multikolinieritas, maka analisis regresi linier berganda tidak dapat digunakan. Untuk itu digunakan analisis regresi *ridge* untuk mengatasi multikolinieritas tersebut. Regresi *ridge* merupakan metode regresi yang menghasilkan estimator yang bias karena penambahan konstanta bias *(k)* pada modelnya, namun varian yang dihasilkan lebih kecil (Weisberg, 1985). Salah satu kesulitan dalam metode ini adalah menentukan konstanta bias (k) yang tepat. Terdapat beberapa metode untuk menentukan konstanta bias (k), namun pada penulisan ini akan digunakan iterasi HKB (Hoerl, Kennard, dan Balwin). Metode iterasi ini merupakan metode yang diajukan oleh Hoerl dan Kennard (1976) dalam Motgomery dan peck (1992) dengan mengambil nilai k awal dari k_{HKB}. Pemilihan konstanta bias (k) dengan iterasi ini merupakan metode yang tidak subjektif seperti *ridge trace*, karena menggunakan cara analitik.

1.2. Tujuan

Tujuan dalam penelitian ini adalah:

1. Mengestimasi parameter dari regresi ridge menggunakan iterasi HKB berdasarkan konstanta bias *(k)*.
2. Menentukan model regresi *ridge*.
3. Memprediksi nilai inflasi pada periode bulan Maret sampai Agustus 2015.

2. TINJAUAN PUSTAKA

2.1. Analisis Regresi Berganda

Kajian yang digunakan untuk menganalisis hubungan antara satu variabel prediktor dan satu variabel respon disebut analisis regresi linier sederhana. Sedangkan untuk mengkaji hubungan antara satu variabel respon dengan lebih dari satu variabel prediktor disebut analisis regresi linier berganda.

Pada regresi liner berganda, persamaan yang menyatakan variabel respon sebagai fungsi linier dari p variabel prediktor didekati dengan menggunakan data yang telah diobservasi. Model tersebut ditentukan oleh persamaan linier sebagai berikut:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \ldots + \beta_p x_{ip} + \varepsilon_i \quad (1)$$

Untuk mengestimasi $\boldsymbol{\beta}$ yang ditunjukkan dengan $\hat{\boldsymbol{\beta}}$, dimana model metode kuadrat terkecil dapat ditulis dengan:

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \hat{\beta}_3 x_{i3} + \ldots + \hat{\beta}_p x_{ip} + e_i \quad (2)$$

Bila ditulis dalam bentuk matriks,

$$\mathbf{y} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e} \quad (3)$$

Sehingga untuk mengestimasi $\boldsymbol{\beta}$ yang ditunjukkan dengan $\hat{\boldsymbol{\beta}}$, adalah:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-1}\mathbf{X'y} \quad (4)$$

2.2. Multikolinieritas

Multikolinieritas pertama kali diperkenalkan oleh Ragnar Frisch yang berarti adanya hubungan linier diantara beberapa atau semua variabel prediktor pada model regresi linier berganda (Gujarati, 2002). Dalam regresi linier berganda seharusnya tidak terjadi multikolinieritas. Jika multikolinieritas terjadi, maka dapat memberikan dampak variansi yang besar sehingga sulit mendapatkan estimasi yang tepat (Widarjono, 2005).

Dalam bentuk matriks, multikolinearitas adalah kondisi buruk (*ill condition*) dari matriks **X'X** yaitu kondisi yang tidak memenuhi asumsi klasik (Montgomery dan Peck, 1992). Matriks **X'X** akan menjadi matriks singular, sehingga nilai perkiraan koefisien regresi menjadi tidak berhingga dan tak mungkin menduganya (Gujarati, 2002).

2.3. Pendeteksian Multikolinieritas

Terdapat beberapa metode yang dapat digunakan untuk mendeteksi multikolinieritas, salah satunya yaitu dengan menggunakan nilai VIF (*Variance Inflation Factor*) dari masing-masing variabel prediktor terhadap variabel responnya.

Jika nilai VIF tidak lebih dari 10, maka model dinyatakan tidak mengandung multikolinieritas (DeMaris, 2004). Nilai VIF dapat diperoleh dengan rumus:

$$VIF_j = \frac{1}{(1-R_j^2)} \tag{5}$$

dengan R_j^2 adalah koefisien determinasi yang diperoleh jika X_j diregresikan dengan (*p-1*) variabel lain.

2.4. Metode *Centering* dan *Rescaling*

Centering dan *rescaling* merupakan bagian dari membakukan variabel (Kutner *et al*, 2005). Prosedur centering membuat perhitungan regresi menjadi lebih sederhana dengan menghilangkan $\hat{\beta}_0$ (intersep).

Prosedur *rescaling* dengan mentransformasikan variabel respon y dan variabel prediktor x dalam bentuk (Kutner *et al*, 2005):

$$x_{ij}^* = \frac{1}{\sqrt{n-1}}\left(\frac{x_{ij}-\bar{x}_j}{S_{x_j}}\right), i = 1,2,...,n\,;\ j = 1,2,...,p \tag{6}$$

$$y_i^* = \frac{1}{\sqrt{n-1}}\left(\frac{y_i-\bar{y}}{S_y}\right) \quad , i = 1,2,...,n \tag{7}$$

dengan: $x_{ij}^* = \text{variabel } x \text{ dalam bentuk baku}$

$y_j^* = \text{variabel } y \text{ dalam bentuk baku}$

$\bar{x}_j = \text{rata - rata dari } x_j$

$\bar{y} = \text{rata - rata dari } y$

$S_{x_j} = \text{standar deviasi dari } x_j$

$S_y = \text{standar deviasi dari } y$

2.5. Regresi *Ridge*

Dalam Dunteman (1984) menyatakan bahwa Hoerl dan Kennard (1970) memperkenalkan estimasi parameter dengan metode yang berbeda yaitu regresi *ridge*. Metode ini digunakan untuk menangani masalah multikolinieritas dengan memberikan estimator yang bias namun memberikan varian yang lebih kecil daripada metode kuadrat terkecil. Model regresi *ridge* didasarkan pada penambahan konstanta bias *(k)* pada diagonal matriks **X'X**, sehingga:

$$\hat{\boldsymbol{\beta}}^{\mathbf{R}} = (\mathbf{X}^{*'}\mathbf{X}^{*} + k\mathbf{I})^{-1}\mathbf{X}^{*'}\mathbf{y}^{*} \tag{8}$$

2.6. Metode Memilih *Biasing Constant* (*k*)

Menentukan nilai *k* merupakan hal yang sulit untuk mencapai koefisien regresi yang stabil. Terdapat beberapa metode untuk menentukan kontanta bias *k*, yaitu :

1. *Ridge Trace*

Ridge trace adalah plot antara koefisien regresi *ridge* terhadap nilai-nilai *k*, dengan $k \in [0,1]$. Untuk nilai *k* yang berbeda-beda dapat ditentukan $\hat{\boldsymbol{\beta}}^{*}$. Pada grafik, *k* sebagai sumbu horizontal dan $\hat{\boldsymbol{\beta}}^{*}$ sebagai sumbu vertikal (Montgomery dan Peck, 1992). Melalui *ridge trace*, tujuannya adalah memilih *k* yang bernilai kecil, dimana pada *k* tersebut dianggap bahwa koefisien regresi mulai stabil. Metode ini, dianggap terlalu subjektif karena memerlukan keputusan peneliti untuk menentukan nilai *k* yang akan dipilih.

2. Menghitung k_{HKB}

Dalam Montgomery dan Peck (1992) terdapat metode lain yang lebih analitik untuk memilih konstanta bias (*k*), metode ini diperkenalkan oleh Hoerl, Kennard dan Balwin (1975) dengan menghitung nilai *k* dengan menggunkan rumus HKB.

$$k_{HKB} = \frac{p\hat{\sigma}^{2}}{\hat{\boldsymbol{\beta}}^{*}{'}\hat{\boldsymbol{\beta}}^{*}} \tag{9}$$

dimana $k_{HKB} =$ konstanta bias oleh Hoerl, Kennard, dan Baldwin

$p =$ banyaknya variabel bebas

$$\hat{\sigma}^{2} = \frac{(\mathbf{y}^{*} - \mathbf{X}^{*}\hat{\boldsymbol{\beta}}^{*})'(\mathbf{y}^{*} - \mathbf{X}^{*}\hat{\boldsymbol{\beta}}^{*})}{n-p-1}$$

$\hat{\boldsymbol{\beta}}^{*} =$ estimasi parameter dari metode kuadrat terkecil

3. Iterasi

Dalam Khalaf daan Iguernane (2014) menyatakan bahwa Hoerl dan Kennard (1976) mengusulkan sebuah metode iterasi untuk memilih *k*. Metode ini didasarkan pada persamaan (9).

Adapun langkah-langkahnya sebagai berikut:

1) Nilai awal dari konsatanta bias (k_i) diperoleh dari persamaan (9) yang mana parameternya diperoleh dari metode kuadrat terkecil.
2) Selanjutnya mengestimasi parameter regresi *ridge* dari nilai *k* awal, yaitu $\hat{\boldsymbol{\beta}}^{\mathbf{R}} = (\mathbf{X}^{*'}\mathbf{X}^{*} + k_i\mathbf{I})^{-1}\mathbf{X}^{*'}\mathbf{y}^{*}$.
3) Untuk memperoleh nilai konstanta bias yang baru (k_{i+1}), digunakan parameter regresi *ridge* dari langkah (b), yaitu $k_{(i+1)} = \dfrac{p\hat{\sigma}^{2}}{\hat{\boldsymbol{\beta}}^{R}{'}\hat{\boldsymbol{\beta}}^{R}}$.
4) Jika $k_{(i+1)} - k_i \approx 10^{-10}$ maka iterasi berakhir, namun jika tidak maka $k_{(i+1)}$ akan menjadi konstanta bias awal k_i dan kembali ke langkah (b).

2.7. Uji Regresi

Uji regresi dilakukan untuk mengetahui apakah variabel-variabel prediktor dalam model regresi secara signifikan mempengaruhi variabel respon.

a. Uji Secara Simultan

Uji secara simultan digunakan untuk menguji pengaruh secara simultan variabel prediktor terhadap variabel responnya. Rumusan uji simultan model regresi adalah sebagai berikut:

H_0 : $\beta_1 = \beta_2 = \cdots = \beta_p = 0$

H_1 : minimal ada satu j dengan $\beta_j \neq 0 \qquad j = 1,2,\ldots,p$

Statistik Uji : $F_{hitung} = \dfrac{R^2/p}{(1-R^2)/(n-p-1)}$

dengan : $R^2 = 1 - \dfrac{\sum (y-\hat{y})^2}{\sum (y-\bar{y})^2}$

Kriteria uji : tolak H_0 jika nilai $F_{hitung} > F_{(\alpha;p;n-p-1)}$

Penolakan H_0 menunjukkan bahwa variabel prediktor memiliki pengaruh secara simultan terhadap variabel respon.

b. Uji Individu

Uji individu digunakan untuk menguji pengaruh masing-masing variabel prediktor secara individu terhadap variabel respon. Rumusan uji individu model regresi adalah sebagai berikut:

H_0 : $\beta_j = 0, \qquad j = 1,2,\cdots,p$

H_1 : $\beta_j \neq 0, \qquad j = 1,2,\cdots,p$

Statistik Uji : $t_{hitung} = \dfrac{\hat{\beta}_j}{se(\hat{\beta}_j)}$

dengan, $se(\hat{\beta}_j) = \sqrt{\hat{\sigma}^2 c_{jj}}$

c_{jj} = elemen diagonal matriks $\mathbf{(X'X)^{-1}}$

Kriteria uji : tolak H_0 jika nilai $|t_{hitung}| > t_{(\alpha/2;n-p-1)}$

Penolakan H_0 menunjukkan bahwa variabel penjelas memberikan kontribusi terkait hubungan linear terhadap variabel respon.

3. METODOLOGI PENELITIAN

3.1. Jenis dan Sumber Data

Penelitian menggunakan data sekunder yang diambil dari website Bank Indonesia (www.bi.go.id), yaitu data inflasi, BI rate, dan nilai tukar rupiah terhadap dolar, serta dari website BPS (www.bps.go.id), yaitu data jumlah uang beredar dari bulan Januari 2011 sampai Februari 2015 dengan jumlah data sebanyak 50 data.

3.2. Tahapan Analisis Data

Adapun tahapan analisis yang dilakukan dalam penelitian ini adalah sebagai berikut:

1. Input data variabel respon inflasi (y) dan variabel-variabel prediktor BI rate (x_1), jumlah uang beredar (x_2), dan nilai tukar rupiah (x_3).
2. Menstandarkan data dengan centering dan rescaling seperti pada persamaan 6 dan 7.
3. Menentukan estimasi parameter dengan OLS.
4. Pengujian asumsi multikolinieritas.

a. Jika tidak terdapat multikolinieritas, maka proses langsung berakhir.
b. Jika terdapat multikolinieritas, maka digunakan analisis regresi *ridge*. Langkah awal regresi *ridge* adalah menentukan konstanta bias (*k*) menggunakan iterasi HKB.

5. Jika konstanta bias (*k*) sudah ditentukan, maka dapat diperoleh parameter regresi *ridge* dan model regresi *ridge*.
6. Model yang dihasilkan lalu dijuji signifikansinya.

4. HASIL DAN PEMBAHASAN

4.1. *Centering* dan *Rescaling*

Pada data yang akan digunakan terdapat perbedaan satuan, untuk itu diperlukan pembakuan variabel yaitu dengan *centering* dan *rescaling*. Setelah data ditransformasi, dihasilkan data yang memiliki satuan yang setara.

4.2. Estimasi Parameter dengan Metode OLS

Data yang telah dibakukan siap untuk diestimasi parameternya. Diperoleh persamaan regresi adalah $y^*=0{,}212x_1^* - 0{,}884\, x_2^* + 1{,}23\, x_3^*$.

4.3. Uji Asumsi

1. Normalitas

Pada uji normalitas, diperoleh P-Value>0,150. Jadi dapat disimpulkanresidual berdistribusi normal

2. Multikolinieritas

Diperoleh nilai VIF untuk masing-masing variabel predictor BI rate, jumlah uang beredar, dan nilai tukar rupiah yaitu 3,027; 11,410; dan 16,936. Karena terdapat nilai VIF yang melebihi 10, maka diindikasikan terdapat multikolinieritas pada model.

3. Heterokedastisitas

Untuk mengetahui model mengandung heterokedastitas, diakukan uji glejser. Melalui uji glejser, diperoleh hasi bahwa tidak terdapat gejala heterokodastisitas pada model karena tidak terdapat pengaruh yang signifikan variabel prediktor terhadap nilai mutlak residual.

4. Autokorelasi

Uji autokorelasi menggunakan nilai Durbin-Watson, dan diperoleh nilai Durbin-Watson sebesar 1,689015. Pada tabel Durbin-Watson dengan n=50 dan p=3, maka akan diperoleh nilai dL=1,421 dan dU=1,674. Sehingga nilai 4-dU=2,326 dan 4-dL=2,579. Karena nilai Durbin-Watson terletak antara dU dan 4-dU, maka dapat disimpulkan bahwa model persaman regresi tidak mengandung autokorelasi.

4.4. Regresi *Ridge*

Pada regresi *ridge* menggunakan iterasi HKB, diperoleh konstanta bias awal (k_1)=0,0149. Selanjutnya dapat diestimasi parameter regresi *ridge* dari konstanta bias awal (k_1), yaitu:

$$\hat{\boldsymbol{\beta}}^{\mathbf{R}} = \begin{bmatrix} 0{,}2954 \\ -0{,}6181 \\ 0{,}9021 \end{bmatrix}$$