Statistika, Vol. 17 No. 1, 25 – 31
Mei 2017

Kombinasi Regresi Tak Bias *Ridge* dengan Regresi Komponen Utama untuk Mengatasi Masalah Multikolinieritas

FITRIANA NOVITASARI, SULIADI, ANNEKE ISWANI A.
Prodi Statistika, Fakultas Matematika dan Ilmu pengetahan alam, Universits islam Bandung,
Jl.Tamansari No. 1 Bandung 40116
e-mail: fitriananovitasari@gmail.com

ABSTRAK

Metode kuadrat terkecil (MKT) merupakan salah satu metode penaksir parameter regresi. Dalam metode MKT terdapat salah satu asumsi yang harus dipenuhi yaitu tidak adanya multikolinieritas. Multikolinieritas merupakan hubungan linier antara sesama variabel bebas (X) yang akan berakibat pada ragam penduga koefisisen regresi menjadi besar sehingga menyebabkan selang kepercayaan untuk parameter regresi cenderung akan lebih lebar. Salah satu metode yang dapat digunakan untuk mengatasi multikolinieritas adalah regresi MCRR, dimana metode MCRR ini merupakan kombinasi antara regresi tak bias ridge (URR) dan regresi komponen utama. Makalah ini akan membahas penanganan multikolinieritas dengan menggunakan MCRR, sebagai bahan aplikasi akan digunakan data tingkat produksi Crude Palm Oil (CPO) .

Kata Kunci: Metode kuadrat terkecil, Unbiased Ridge Regression , Regresi Komponen Utama, Penaksir kelas (r,k), Modified (r,k) Class Ridge Regression (MCRR),Multikolinieritas.

1. PENDAHULUAN

Analisis regresi merupakan suatu teknik analisis statistik untuk membuat model dan mengetahui hubungan antara dua variabel atau lebih. Salah satu dari model statistika yang sering digunakan dalam pemecahan suatu masalah adalah model regresi linier. Hubungan dalam model tersebut dapat dinyatakan dalam bentuk persamaan yang menghubungkan variabel tak bebas (Y) dengan satu atau lebih variabel bebas (X).

Salah satu cara untuk mendapatkan koefisien regresi linier berganda adalah menggunakan metode kudrat terkecil (MKT). Dalam MKT terdapat salah satu asumsi yang harus di penuhi yaitu tidak adanya multikolnieritas atau tidak ada hubungan antar variabel bebas (X). Jika antara variabel bebas (X) terdapat hubungan (mulikolinieritas) maka akan menyebabkan ragam penduga koefisien regresi menjadi besar sehingga selang kepercayaan untuk parameter regresi juga cenderung akan lebih lebar. Untuk mendeteksi apakah terdapat multikolinieritas atau tidak, dapat dilihat dengan *Variance Inflation Factors* (VIF). Jika nilai dari VIF > 10 maka dapat dinyatakan bahwa dalam variabel bebas tersebut terdapat multikolinieritas (Myers, 1990).

Untuk mengatasi masalah mutikolinieritas dapat dilakukan dengan menggunakan regresi *ridge*. Metode ini pertama kali dikemukakan oleh Hoerl dan Kennard (1970). Namun regresi *ridge* memiliki kekurangan yaitu sifat penaksir regresi *ridge* adalah bias. Oleh karena itu Crouse, *et al.*, (1995) meningkatkan kinerja regresi *ridge* yang diberi nama *Unbiased ridge regression* (URR), dimana penaksir yang dihasilkan URR merupakan penaksir yang tidak bias bagi β. Selain itu Marquardt, 1970 mengusulkan regresi komponen utama sebagai alat untuk menangani multikolinieritas, dan Baye dan Parker (1984) memodifikasi regresi komponen utama dengan menggabungkan regresi komponen utama dengan regresi *ridge* yang dinamakan penaksir (r,k) dimana r merupakan banyaknya komponen yang terpilih sedangkan k merupakan sebuah konstanta positif dari regresi *ridge*, untuk meminimumkan nilai standar error dari regresi komponen utama, Batah, *et al.* (2009) meningkatkan metode penaksir (r,k) yang di usulkan oleh Baye dan Parker (1984) yaitu dengan menggabungkan penaksir (r,k)

dengan penaksir tak bias regresi *ridge* (*unbiased ridge regression* / URR) dimana penaksir baru ini dinamakan *modified (r, k) class ridge regression* (MCRR).

Dalam penelitian ini akan membahas metode MCRR untuk menangani masalah multikolinieritas, yang akan diterapkan pada data tingkat produksi *Crude Palm Oil* (CPO) yang dipengaruhi oleh jumlah buah kelapa sawit, jumlah tenaga kerja, jam kerja mesin, penggunaaan air, penggunaan uap dan suplay listrik di Unit Adolina PT. Perkebunan Nusantara IV Sumatra Utara.

2. TINJAUAN PUSTAKA

Regresi Linier Berganda

Analisis regresi merupakan suatu metode statistik yang digunakan untuk memodelkan hubungan antara variabel tak bebas (Y) dengan variabel bebas (X). Misalkan Y_i adalah respon pengamatan ke-i, dan $X_i = \left(1, X_{i1}, X_{i2}, \dots, X_{ip}\right)^T$ adalah vektor variabel bebas pengamatan ke-i, maka hubungna antara variabel tak bebas (Y) dengan variabel bebas (X) dapat dinyatakan sebagai berikut :

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_P X_{ip} + \varepsilon_i \qquad (2.1)$$

Dalam bentuk matrik dapat dinyatakana sebegai berikut:

$$Y = X\beta + \varepsilon \qquad (2.2)$$

Dimana $Y = \left(Y_1, Y_2, \dots, Y_p \right)^T$ dan $\beta = \left(\beta_1, \beta_2, \dots, \beta_p \right)^T$ merupakan parameter koefisien regresi.

Untuk menaksir parameter-parameter dari persamaan (2.1) dapat digunakan metode kuadrat terkecil (MKT). MKT merupakan metode penaksir yang bertujuan untuk meminimumkan jumlah kuadrat kesalahan $\sum e_i^2$ sehingga nilai penaksir regresinya akan mendekati nilai yang sesungguhnya. Penaksir MKT bagi β adalah:

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^t\boldsymbol{X})^{-1}(\boldsymbol{X}^t\boldsymbol{Y}) \qquad (2.3)$$

Dalam MKT terdapat salah satu asumsi yang harus dipenuhi yaitu tidak ada hubungan linier (multikolinieritas) diantara variabel-variabel bebas. Untuk melihat apakah suatu model mengalami multikolinierits, dapat dilihat menggunakan nilai *Variance Inflation Factors* (VIF). Nilai VIF dapat diperoleh dengan persamaan :

$$VIF_j = \frac{1}{1 - R_j^2}$$

dimana R_j^2 merupakan koefisien determinasi dari regresi antara variabel bebas $\left(X_j\right)$ yang diregresikan terhadap variabel bebas lainya. Apabila Nilai VIF lebih besar dari 10 dapat dijadikan indikasi bahwa ada masalah multikolinieritas atau dapat diindikasi bahwa ada hubungan diantara variabel bebas (Myers, 1990).

Jika asumsi tidak adanya multikolinieritas ini dilanggar, terdapat beberapa alternatif untuk menangani masalah ini diantaranya adalah metode penaksir regresi *ridge* dan penaksir komponen utama, kedua penaksir ini akan di jelaskan pada sub bagian berikut ini.

Penaksir Regresi *Ridge*

Regresi *ridge* pertama kali diperkenalkan oleh Hoerl dan Kennard (1970) . Metode ini dilakukan dengan cara menambahkan konstanta yang bernilai positif λ terhadap elemen diagonal $X^{*t}X^*$. Dimana X^* merupakan nilai dari variabel bebas X yang telah di standarkan.

Penaksir parameter regresi *ridge* dilakukan dengan cara menstandarkan variabel bebas dan variabel tak bebas dengan model (Kutner, *et al.*, 2005):

$$y_i^* = \beta_1^* X_{i1}^* + \beta_2^* X_{i2}^* + \cdots + \beta_p^* X_{ip}^* + \varepsilon_i$$

Dimana X_i^* dan Y_i^* adalah X_i dan Y_i yang telah di standarkan.

Penaksir regresi *ridge* bagi $\beta_R = \left(\beta_1^*, \beta_2^*, \dots, \beta_p^*\right)^T$ adalah:

$$\widehat{\boldsymbol{\beta}}_R(\lambda) = (\boldsymbol{X}^{*t}\boldsymbol{X}^* + \lambda\boldsymbol{I})^{-1}\boldsymbol{X}^{*t}\boldsymbol{Y}^* \qquad (2.4)$$

dimana I adalah matrik identitas berukuran $(p \times p)$, dan λ adalah suatu konstanta positif dengan $\lambda > 0$.

Dalam prakteknya nilai optimal λ pada persamaan (2.4) tidak diketahui. Salah satu penaksir λ yang diusulkan oleh Hoerl dan Kennard (1970) adalah sebagai berikut:

$$\hat{\lambda} = \frac{p s_{MKT}{}^2}{\hat{\boldsymbol{\beta}}_{MKT}{}^T \hat{\boldsymbol{\beta}}_{MKT}}$$

dimana nilai $\hat{\boldsymbol{\beta}}_{MKT}$ dapat diperoleh dari persamaan (2.2), dan S^2 adalah ragam model (2.1).

$$s_{MKT}{}^2 = \frac{(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}_{MKT})^T(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}_{MKT})}{n-p}$$

untuk mengembalikan model regresi *ridge* ke model asalnya dapat di lakukan dengan persamaan berikut:

$$\hat{\beta}_0 = \bar{Y} - \left(\hat{\beta}_1\bar{X}_1 + \hat{\beta}_2\bar{X}_2 + .. + \hat{\beta}_P\bar{X}_P\right)$$

$$\hat{\beta}_j = \left(\frac{S_y}{S_j}\right)\beta_P{}^*; \quad j = 1,2,\ldots,P$$

Unbiased Ridge Regression (URR)

Penaksir yang dihasilkan oleh regresi *ridge* adalah bias, oleh karena itu Crouse *et al.*, (1995) memperkenalkan suatu metode untuk memperbaiki kinerja regresi *ridge* yaitu *unbiased ridge regression* (URR). Penaksir URR ini mampu menghasilkan penaksir yang tak bias untuk β. Penaksir URR dilakukan dengan cara mengkombinasikan nilai informasi sebelumnya (J) dengan λ.

Penaksir yang tak bias untuk β menggunakan penaksir URR adalah sebagai berikut:

$$\hat{\boldsymbol{\beta}}_{URR}(\lambda, J) = (\boldsymbol{X}^{*t}\boldsymbol{X}^* + \lambda^*\boldsymbol{I})^{-1}(\boldsymbol{X}^{*t}\boldsymbol{Y}^* + \lambda^*\boldsymbol{J}); \quad untuk\ \lambda \geq 0 \tag{2.5}$$

Dimana J adalah suatu vektor acak yang berdistribusi normal $J \sim N\left(\beta, \frac{\sigma^2}{\lambda} I\right)$ untuk $\lambda > 0$. Crouse, et al., (1995) memberikan salah satu nilai J yang dapat digunakan yaitu:

$$\boldsymbol{J} = \eta \mathbf{1}_{p\times 1}$$

dengan:

$$\eta = \frac{1}{p}\sum_{j=1}^{p}\hat{\beta}_j$$

Nilai λ^* pada persamaan (2.5) dapat diperoleh melalui persamaan sebagai berikut (Crouse, *et al.*, 1995):

$$\hat{\lambda}^* = \frac{p\sigma^2}{(\hat{\boldsymbol{\beta}}_{MKT}-\boldsymbol{J})^t(\hat{\boldsymbol{\beta}}_{MKT}-\boldsymbol{J}) - \sigma^2 tr(\boldsymbol{X}^{*t}\boldsymbol{X}^*)^{-1}}$$

Jika $(\hat{\boldsymbol{\beta}}_{MKT}-\boldsymbol{J})^t(\hat{\boldsymbol{\beta}}_{MKT}-\boldsymbol{J}) - \sigma^2 tr(\boldsymbol{X}^{*t}\boldsymbol{X}^*)^{-1} > 0$

dan

$$\hat{\lambda}^* = \frac{p\sigma^2}{(\hat{\boldsymbol{\beta}}_{MKT}-\boldsymbol{J})^t(\hat{\boldsymbol{\beta}}_{MKT}-\boldsymbol{J})} \quad \text{, untuk yang lain}$$

Jika nilai σ^2 tidak diketahui**,** maka σ^2 dapat di ganti dengan penaksir yang tak bias dari $\hat{\sigma}^2$ yaitu

$$s^2 = \frac{(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}_{MKT})^t(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}_{MKT})}{(n-p-1)}$$

Penaksir Regresi Komponen Utama (PCR)

Regresi komponen utama didasarkan pada analisis komponen utama. Melalui analisis komponen utama akan dihasilkan variabel-variabel baru (komponen utama) yang merupakan kombinasi linier dari variabel-variabel bebas asal, dan antara variabel baru ini bersifat tidak saling berkorelasi. Selanjutnya variabel bebas (komponen utama) akan diregresikan dengan variable terikat (Y). Langkah awal sebelum melakukan regresi komponen utama adalah menstandarkan variabel bebas (X) , untuk matrik variabel bebas (X) yang telah di standarkan di lambanghan X^*.

Dari variabel bebas yang telah di bakukan akan di peroleh akar ciri (δ) dan vektor ciri (v) dari $X^{*t}X^*$. Untuk menentukan akar ciri (*eigen value*) $\delta_1, \delta_2, \ldots, \delta_p$ dapat di peroleh dari penyelesaian sistem persamaan linier berikut $|\boldsymbol{X}^{*t}\boldsymbol{X}^* - \delta \mathrm{I}| = 0$. Vektor ciri v_i untuk akar ciri δ_i diperoleh dari penyelesaian $(\boldsymbol{X}^{*t}\boldsymbol{X}^* - \delta\boldsymbol{I})v = 0$. $\delta_1, \delta_1, \ldots, \delta_p$ merupakan ragama dari komponen utama ke $1,2,\ldots,p$.

Untuk penentuan banyaknya komponen yang akan digunakan yaitu dengan melihat proporsi komulatif keragaman data minimal 80% (Johnson, 2007), atau dengan menggunakan *scree plot* (Rancer, 1998).

Nilai komponen utama Z_j untuk masing-masing observasi adalah :

$$\boldsymbol{Z} = \boldsymbol{X}^*\boldsymbol{V}$$

dimana Z adalah matrik $(n \times p)$. Kolom dari matrik Z didefinisikan sebagai variabel bebas baru (komponen utama) yang saling orthogonal yaitu $Z = [Z_1, Z_2, \dots, Z_p]$. Sedangkan $\boldsymbol{V}$ merupakan matrik $(p \times p)$ dimana kolom-kolomnya merupakan vektor ciri dari $X^{*t}X^*$. Matrik $\boldsymbol{V}$ merupakan sebuah matrik orthogonal yang memenuhi persamaan $\boldsymbol{V^tV = I}$.

Oleh karena itu, $\boldsymbol{X\beta}$ pada persamaan (2.2) dapat dinyatakan sebagai :

$$\begin{aligned}\boldsymbol{X\beta} &= \boldsymbol{XVV}^t\boldsymbol{\beta}\\ &= \boldsymbol{Z\alpha}\end{aligned}$$

Maka persamaan regresi linier berganda dapat dinyatakan sebagai model regresi komponen utama seperti berikut ini (Jollife, 2002):

$$\boldsymbol{y} = \boldsymbol{Z\alpha} + \boldsymbol{\varepsilon} \tag{2.6}$$

dimana $\boldsymbol{\alpha}$ adalah vektor berukuran $(p \times 1)$ dengan bentuk $\boldsymbol{\alpha} = \boldsymbol{V\beta}$

Model umum dari regresi komponen utama adalah seperti yang di tunjukan dalam persamaan (2.6), model tersebut digunakan jika semua p variabel bebas dimasukan dalam komponen, namun jika hanya r komponen yang diambil maka model regresi komponen utama adalah :

$$\boldsymbol{y} = \boldsymbol{Z_r\alpha_r} + \boldsymbol{\varepsilon_r} \tag{2.7}$$

dimana α_r merupakan sebuah vektor dari unsur komponen yang terpilih sebanyak r yaitu bagain dari unsur V. Z_r adalah matrik berukuran $(n \times r)$ yang kolom-kolomnya merupakan bagian dari Z. Dengan menggunakan penaksir kuadrat terkecil untuk mengestimasi α di persaman (2.7) maka :

$$\begin{aligned}\hat{\boldsymbol{\alpha}}_r &= (\boldsymbol{Z}^t{}_r\boldsymbol{Z}_r)^{-1}\boldsymbol{Z}^t{}_r\boldsymbol{y}\\ &= \Lambda^{-1}\boldsymbol{Z}^t\boldsymbol{y}\end{aligned}$$

dengan : $\alpha_r = [\alpha_1, \alpha_2, \dots, \alpha_r]^t$, $Z_r = [Z_1, Z_2, \dots, Z_r]$ dan Λ merupakan matrik diagonal yang diagonal utamanya adalah nilai r akar ciri terbesar dari matrik $X^{*t}X^*$. Sedangkan matrik varian kovarian dari $\hat{\alpha}$

$$var\,(\hat{\alpha}_r) = \sigma^2\Lambda^{-1}$$

Jadi penaksir komponen utama $\hat{\beta}_{PCR}$ utama dapat ditulis :

$$\begin{aligned}\hat{\beta}_{PCR} &= \boldsymbol{V}_r\left(\boldsymbol{V}_r{}^t\boldsymbol{X}^t\boldsymbol{X}\,\boldsymbol{V}_r\right)^{-1}\boldsymbol{V}_r{}^t\boldsymbol{X}^t\boldsymbol{Y}\\ &= \boldsymbol{V}_r\Lambda_r{}^{-1}\boldsymbol{V}_r{}^t\boldsymbol{X}^t\boldsymbol{Y}\end{aligned}$$

dimana $\Lambda_r = \boldsymbol{V}_r{}^t\boldsymbol{X}^t\boldsymbol{X}\,\boldsymbol{V}_r$

Untuk meningkatkan kinerja regresi komponen utama Baye & parker 1984 melakukan peningkatan kinerja regresi komponen utama dengan menggabungkan penaksir tak bias regresi *ridge* dengan regresi komponen utama. Penaksir baru ini disebut dengan penaksir kelas (r, k) dengan penaksir sebagai berikut:

$$\begin{aligned}\hat{\boldsymbol{\beta}}_{r,k} &= \boldsymbol{V}_r\left(\boldsymbol{V}_r{}^t\boldsymbol{X}^t\boldsymbol{X}\,\boldsymbol{V}_r + \lambda\boldsymbol{I}_r\right)^{-1}\boldsymbol{V}_r{}^t\boldsymbol{X}^t\boldsymbol{Y}\\ &= \boldsymbol{V}_r(\Lambda_r + \lambda\boldsymbol{I}_r)^{-1}\boldsymbol{V}_r{}^t\boldsymbol{X}^t\boldsymbol{Y}\ \text{, untuk } \lambda > 0\end{aligned}$$

Modified $(\boldsymbol{r}, \boldsymbol{k})$ *Class Ridge Regression* (MCRR)

Penaksir MCRR diusulkan oleh Batah, *et al.*, 2009 merupakan sebuah penaksir dengan menggabungkan penaksir (r, k) dengan *unbiased ridge regression* (URR). Penaksir untuk $\hat{\beta}_{MCRR}$ adalah sebagai berikut:

$$\hat{\boldsymbol{\beta}}_{MCRR}(\lambda^*, J) = \boldsymbol{V}_r\left(\boldsymbol{V}_r{}^t\boldsymbol{X}^t\boldsymbol{X}\boldsymbol{V}_r + \lambda^*\boldsymbol{I}_r\right)^{-1}\boldsymbol{V}_r{}^t(\boldsymbol{X}^t\boldsymbol{Y} + \lambda^*\boldsymbol{J}) \tag{2.8}$$

Dari persamaan (2.8) diperoleh (Batah, 2009):

$$\begin{aligned}&Bias\,(\hat{\boldsymbol{\beta}}_{MCRR}(\lambda^*, J)) = \left[\boldsymbol{V}_r\boldsymbol{V}_r{}^t - \boldsymbol{I}_p\right]\boldsymbol{B}\\ &= -\boldsymbol{V}_{p-r}\boldsymbol{V}_{p-r}{}^t\boldsymbol{B}\end{aligned}$$

dan

$$var\,(\hat{\beta}_{MCRR}(\lambda^*,J)) = \sigma^2\boldsymbol{V}_r\left(\boldsymbol{V}_r{}^t\boldsymbol{X}^t\boldsymbol{X}\boldsymbol{V}_r + \lambda^*\boldsymbol{I}_r\right)^{-1}\boldsymbol{V}_r{}^t$$
$$= \sigma^2\boldsymbol{V}_r\boldsymbol{\Lambda}_r{}^{-1}(\lambda^*)\boldsymbol{V}_r{}^t$$

dimana $\Lambda_r(\lambda^*) = \Lambda_r + \lambda^* I_r$ dan $\Lambda_r = \boldsymbol{V}_r{}^t\boldsymbol{X}^t\boldsymbol{X}\boldsymbol{V}_r$

Dalam hal ini penaksir MCRR memiliki beberapa sifat yaitu :

1) $\hat{\beta}_p(0,J) = \hat{\beta}_{MKT} = (X^tX)^{-1}(X^tY)$
2) $\hat{\beta}_p(\lambda^*,J) = \hat{\beta}(\lambda^*,J) = (X^{*t}X^* + \lambda I)^{-1}(X^{*t}Y^* + \lambda J)$
3) $\hat{\beta}_r(0,J) = \hat{\beta}_{PCR} = T_r\left(T_r{}^tX^{*t}XT_r\right)^{-1}T_r{}^tX^{*t}Y$
4) $\lim_{\lambda^*\to 0}\hat{\beta}_{MCRR}(\lambda^*,J) = \hat{\beta}_{PCR}$
5) $\lim_{\lambda^*\to \infty}\hat{\beta}_{MCRR}(\lambda^*,J) = J$

Dimana $\hat{\beta}_p$ merupakan penaksir $\hat{\beta}_{MCRR}$ jika semua komponen dipakai dalam model.

3. BAHAN DAN METODE

Analisis yang akan digunakan untuk menangani multikolinieritas digunakan regresi MCRR, yang akan diaplikasikan pada data sekunder tingkat produksi *Crude Palm Oil* (CPO) yang dipengaruhi oleh jumlah buah kelapa sawit, jumlah tenaga kerja, jam kerja mesin, penggunaaan air, penggunaan uap dan suplay listrik Unit Adolina PT. Perkebunan Nusantara IV Sumatra Utara (Purba, 2013).

Langkah untuk menganaslisi data yang harus dilakukan adalah :

1) Mencari model MKT➔ menguji asumsi multikolinieritas ➔ jika nilai VIF >10 maka transformasi data.
2) Menetukan banyaknya komponen yang terpilih
3) Mengitung nilai η, nilai J dan λ^*
4) Subtitusikan nilai J dan λ^* kedalam persamaan (2.8) untuk menaksir regresi MCRR.

4. ANALISIS DAN PEMBAHASAN

Pemodelan Regresi Metode Kuadrat Terkecil

Hasil penaksiran parameter berdasakan metode kuadrat terkecil untuk regresi linier berganda didapatkan diperoleh model sebagai berikut :

$$\hat{Y} = 57.7267 + 0.2350\,X_1 - 0.0508\,X_2 + 0.0239\,X_3 - 0.0066\,X_4 + 0.0096\,X_5 + 0.00014\,X_6$$

Melalui persamaan model regresi dengan metode kuadrat terkecil diperoleh nilai R^2 sebesar 99.5%, namun meskipun nilai R^2 cukup tinggi setelah di uji secara parsial, banyak variabel bebas yang tidak signifikan dalam mempengaruhi variabel tak bebas, selain itu dilihat dilihat dari tanda nilai koefisien yang dihasilkan tidaklah masuk akal, jika dilihat variabel X_2 (jumlah tenaga kerja) memiliki tanda negatif hal ini bebrarti dengan bertambahnya jumlah tenaga kerja maka akan mengurangi tingkat produksi, kesimpulan yang dihailkan dalam model ini tidaklah masuk akal, sehingga hal ini mengindikasikan bahwa adanya masalah multikolinieritas. Untuk hasil uji parsial ditampilkan dalam Tabel 4.1 berikut:

Tabel 4.1 Hasil signifikansi Uji t

Variabel	**Koefisien**	**t hitung**	**P-*value***	**Keterangan**
X_1	0.2350	8.97	0.000	Signifikan
X_2	-0.0508	-0.52	0.603	Tidak signifikan
X_3	0.0239	0.11	0.914	Tidak signifikan
X_4	-0.0066	-0.36	0.719	Tidak signifikan
X_5	0.0096	1.82	0.077	Tidak signifikan
X_6	0.0001	0.63	0.529	Tidak signifikan

Untuk memastikan apakah terdapat masalah multikolinieritas dalam data tingkat produksi *crude palm oil* (CPO), maka dapat menggunakan nilai *variance inflation factors* (VIF), nilai VIF disajikan pada Tabel 4.2 berikut :

Tabel 4.2 Nilai *variance inflation factors* untuk setiap variabel bebas X

Variabel	VIF
X_1	112.437
X_2	11.303
X_3	12.096
X_4	93.446
X_5	2.420
X_6	2.676

Dari tabel di atas dapat dilihat bahwa variabel X_1, X_2, X_3 dan X_4 memiliki nilai VIF lebih besar dari 10, dengan nilai VIF masing-masing adalah 112.437, 11.303, 12.096, dan 93.446. Oleh karena itu dapat disimpulkan bahwa diantara variabel bebas terdapat masalah multikolinieritas, sehingga dapat disimpulkan bahwa dalam data tingkat produksi *crude palm oil* (CPO) terdapat kasus multikolinieritas dalam variabel bebasnya.

Penanganan multikolinieritas

Modified (r,k) class ridge regression (MCRR)

Dalam Tabel 4.2 terlihat bahwa variabel X_1, X_2, X_3 dan X_4 terindikasi multikolinieritas, oleh karena itu akan dilakukan penanganan masalah multikolinieritas dengan menggunakan metode MCRR. Sebelum menaksir menggunakan regresi MCRR, langkah pertama adalah menentukan banyaknya komponen yang akan diambil dari analisis komponen utama. Untuk menentukan banyaknya komponen yang diambil terdapat beberapa langkah diantaranya:

1) Menstandarkan varaibel bebas (X) untuk hasil dari standarisasi variabel bebas (X) diberikan dalam lampiran.
2) Menentukan nilai akar ciri dan vektor ciri. Dimana untuk nilai akar ciri disajikan dalamTabel 4.3

Tabel 4.3 Akar ciri berdasarkan analisis komponen utama

Komponen	Akar cirri	Keragaman (%)	Keragaman komulatif (%)
PC1	5.1246	85.4%	85.4%
PC2	0.4560	7.60%	93.0%
PC3	0.2791	4.70%	97.7%
PC4	0.0774	1.30%	99.0%
PC5	0.0578	1.00%	99.9%
PC6	0.0050	0.10%	100%

Dari Tabel 4.3 dapat di lihat banyaknya komponen yang akan digunakan adalah komponen dengan prosentase komulatif dari akar ciri minimal 80 %. Dalam kasus ini peneliti akan mengambil dua komponen utama yaitu komponen PC1 dan komponen PC2, karena komulatif keragaman dari kedua komponen ini sebesar 93 % > 80%.

Dalam penaksiran model regresi MCRR penetapan nilai J dan $\hat{\lambda}^*$ merupakan hal yang penting, dalam penelitian ini, dengan nilai η yang diperoleh sebesar 0.1678, untuk nilai J nilai η dikalikan dengan matrik 1_{px1},sedangkan nilai $\hat{\lambda}^*$ yang dieroleh adalah 0.0373. Sehingga diperoleh penaksir regresi MCRR dari data yang telah dikembalikan kedalam bentuk awal adalah sebagai berikut:

$$\hat{Y} = -315.776 + 0.04341\, X_1 + 0.49865\, X_2 + 1.09091\, X_3 + 0.03315\, X_4 + 0.05342\, X_5 + 0.00181\, X_6$$

Dari model MCRR diatas diperoleh nilai *mean square error* (MSE) yang dihasilkan sebesar 137423.59 dengan nilai R^2 sebesar 96.3 %. Jika dilihat dari nilai koefisien dari regresi MCRR ini semua koefisien positif artinya semua variabel berpengaruh secara positif terhadap tingkat produksi, hal ini memiliki kesimpulan yang lebih masuk akal bahwa variabel jumlah buah kelapa sawit, jumlah tenaga kerja, jam kerja mesin, penggunaaan air, penggunaan uap dan suplay listrik akan berpengaruh positif terhadap tingkat produksi. Hal ini bisa terjadi karena dalam MCRR masalah multikolinieritas sudah tidak ada.

5. KESIMPULAN

Pada model data tingkat produksi *Crude Palm Oil* (CPO) terdapat masalah multikolinieritas pada variabel jumlah buah kelapa sawit, jumlah tenaga kerja, jam kerja mesin, dan penggunaan air. Sehingga untuk menangani maslaah multikolinieritas digunakan regresi MCRR. Dengan persamaan model dari regresi MCRR adalah sebagai berikut:

$$\hat{Y} = -315.776 + 0.04341\, X_1 + 0.49865\, X_2 + 1.09091\, X_3 + 0.03315\, X_4 + 0.05342\, X_5 + 0.00181\, X_6$$

Dari hasil analisis yang telah dilakukan regresi MCRR memiliki nilai koefisien determinasi yang cukup besar dan koefisien yang dihasilkan berepengaruh secara positif terhadap variabel tingkat produksi, sehingga kesimpulan yang diambil menggunakan MCRR lebih tepat dan masuk akal.

DAFTAR PUSTAKA

Achmad, A.I., (2008). *Analisis Regresi*. Program Studi Statistika Universitas Islam Bandung.

Batah, F., Ozkale, M. R., Gore, S.D., (2009) *Combining Unbiased Ridge and Principal Component Regression Estimators. Journal of Communications in statististic- Theory and Methods* Vol: 38, No. 13: 2201-2209.

Baye, M and Parker,D. (1984). *Combinng Ridge and Principal Component Regression : a Money Demand Illustration. Journal of Communications in statististic- Theory and Methods* Vol: 13, No. 2: 2201-2209.

Crouse, r., Jin, C., Hanumara, R.C., (1995). *Unbiased Ridge Estimation ith Prior Infoermation and Ridge Trace. Journal of Communications in statististic- Theory and Methods* Vol: 24, No. 9: 2341-2354.

Hajarisman, N., (2011). *Analisis Regresi Lanjut*. Program Studi Statistika Universitas Islam Bandung.

Hoerl, A. E. dan R. W. Kennard., (1970). *Ridge Regression: Aplikation s to Nonorthogonal Problems. Technometrics.* Vol: 12 : 55-56.

Jhonson, R and Wichern, W. (2007). *Applied Multivariate Statistical Analysis*. New York: Prentice Hall, Inc.

Jollife, I.T. (2002). *Principl Component Anaysis*. Springer.

Kutner, M. H., Nachtsheim, C.J., Netter, J., and Li, W., (2005), *Applied Linear Statistical Models*, 5rd ed., McGRraw-Hill : Irwin.

Myers, R.H. (1990). *Classical and Modern Regression with Applications*. Boston, MA: Duxburry.

Netter, J. Wasserman, W. And Kutner, M.H. (1990). Applied *Linier Statistical Models*. Tokyo: Richard D. Irwin, Inc.