Statistika, Vol. 12 No. 1, 33 – 42
Mei 2012

Perbandingan Metode *Partial Least Square* (PLS) dengan Regresi Komponen Utama untuk Mengatasi Multikolinearitas

NURHASANAH, MUHAMMAD SUBIANTO, RIKA FITRIANI

Jurusan Matematika FMIPA UNSYIAH
Jl. Syech Abdul Rauf No.3 Darussalam, Banda Aceh
E-mail : nurhasanah.math@gmail.com

ABSTRAK

Dalam mengatasi multikolinearitas pada suatu data, ada beberapa metode yang dapat digunakan, diantaranya yaitu metode *Partial Least Square* (PLS) dan metode regresi komponen utama (RKU). Data yang digunakan dalam penulisan ini adalah data sekunder yang diperoleh dari Jurnal *Technometrics* (Naes, 1985). Hasilnya menunjukkan bahwa metode PLS lebih baik dari pada RKU berdasarkan nilai koefisien determinasi (R^2) yang tinggi, nilai *Mean Square Error Prediction* (MSEP) dan nilai *Root Mean Square Error Prediction* (RMSEP) yang minimum.

Kata kunci: multikolinearitas, metode Partial Least Square (PLS), regresi komponen utama (RKU), R^2*, MSEP, RMSEP.*

1. PENDAHULUAN

Analisis regresi linear berganda yang mempunyai banyak variabel bebas, sering timbul masalah karena terjadinya hubungan antara dua atau lebih variabel bebasnya. Variabel bebas yang saling berkorelasi disebut multikolinearitas (*multicollinearity*). Salah satu dari asumsi model regresi linear adalah bahwa tidak terdapat multikolinearitas diantara variabel bebas yang termasuk dalam model. Multikolinearitas terjadi apabila terdapat hubungan atau korelasi diantara beberapa atau seluruh variabel bebas (Gonst and Mason, 1977 dalam Soemartini, 2008).

Untuk mengetahui adanya multikolinearitas yaitu dengan menghitung koefisien korelasi sederhana antara sesama variabel bebas, jika terdapat koefisien korelasi sederhana yang hampir mendekati ± 1 maka hal tersebut menunjukkan terjadinya masalah multikolinearitas dalam regresi (Walpole, 1988). Selain itu, salah satu alat untuk mengukur adanya multikolinearitas adalah *Variance Inflation Factor* (VIF). VIF adalah suatu faktor yang mengukur seberapa besar kenaikan ragam dari koefisien penduga regresi dibandingkan terhadap variabel bebas yang orthogonal jika dihubungkan secara linear. Nilai VIF akan semakin besar jika terdapat korelasi yang semakin besar diantara variabel bebas. Nilai VIF > 10 dapat digunakan sebagai petunjuk adanya multikolinearitas pada data. Gejala multikolinearitas menimbulkan masalah dalam model regresi. Korelasi antar variabel bebas yang sangat tinggi menghasilkan penduga model regresi yang berbias, tidak stabil, dan mungkin jauh dari nilai prediksinya (Bilfarsah, 2005).

Salah satu cara untuk mendapatkan koefisien regresi pada persamaan regresi linear berganda adalah melalui metode kuadrat terkecil. Metode ini menghasilkan penaksir terbaik (tak bias dan bervarians minimum) jika saja tidak ada korelasi antar variabel bebas. Namun jika hal itu terjadi, ada beberapa cara atau metode yang dapat digunakan untuk mengatasi masalah multikolinearitas yaitu regresi komponen utama, regresi ridge, metode kuadrat terkecil parsial (*partial least square*) dan bebrapa metode lainnya. Dalam penulisan ini hanya membandingkan metode *Partial Least Square* (PLS) dan regresi komponen utama. Metode *Partial Least Square* (PLS) merupakan proses pendugaan yang dilakukan secara iteratif dengan melibatkan struktur keragaman variabel bebas dan variabel tak bebas. Metode kedua yang dikaji dalam penelitian ini adalah regresi komponen utama yaitu regresi dengan mengambil komponen utama sebagai variabel bebas. Koefisien penduga dari metode ini diperoleh melalui penyusutan dimensi variabel penduga komponen utama, dimana subset komponen utama yang dipilih harus tetap mempertahankan keragaman yang besar terhadap variabel tak bebasnya (Herwindiati, 1997).

Dari pengkajian kedua metode tersebut akan dihitung nilai R^2, *Mean Square Error Prediction* (MSEP), dan *Root Mean Square Error Prediction* (RMSEP) dan kemudian didapatkan metode mana yang lebih baik diantara kedua metode tersebut dengan melihat nilai R^2 yang lebih tinggi dan nilai MSEP dan RMSEP yang lebih rendah.

2. TINJAUAN PUSTAKA

Partial Least Square (PLS)

Metode *Partial Least Square* (PLS) merupakan *soft* model yang dapat menjelaskan struktur keragaman data. *Partial Least Square* (PLS) dapat dilihat sebagai bentuk yang saling berkaitan dengan *Prinsip Component Regression* (PCR). Model yang dihasilkan oleh metode *Partial Least Square* (PLS) mengoptimalkan hubungan antara dua kelompok variabel. Pendugaan model hubungan *Y* dengan *X* dan pendugaan nilai *Y* tertentu menggunakan suatu algoritma. Proses penentuan model dilakukan secara iterasi dengan melibatkan keragaman pada variabel *X* dan *Y*. Struktur ragam dalam *Y* mempengaruhi perhitungan komponen kombinasi linear dalam *X* dan sebaliknya, struktur ragam dalam *X* berpengaruh terhadap kombinasi linear dalam *Y* (Bilfarsah, 2005).

Pada dasarnya *Partial least square* (PLS) memodelkan hubungan variabel *Y* dengan variabel *X* berdasarkan variabel internal. Variabel *X* dibagi ke dalam skor t_h dan *loading* p_h, yang dinyatakan sebagai:

$$X = t_1 p_1' + t_2 p_2' + t_3 p_3' + \ldots + t_h p_h' + E_h \qquad (2.1)$$

dimana: *X* = variabel bebas

t_h = vektor skor (*score vector*) variabel *X*

p_h = vektor muatan (*loading vector*) variabel *X*

E_h = matriks sisaan variabel X

Variabel *Y* juga dibagi dalam skor u_h dan *loading* q_h yang dinyatakan sebagai:

$$Y = u_1 q_1' + u_2 q_2' + u_3 q_3' + \ldots + u_h q_h' + F_h \qquad (2.2)$$

dimana: *Y* = variabel tak bebas

u_h = vektor skor (*score vector*) variabel *Y*

q_h = vektor muatan (*loading vector*) variabel *Y*

F_h = matriks sisaan variabel *Y*

(Wigena dan Aunuddin, 1998).

Pemodelan *Partial Least Square* (PLS) ditempuh melalui hubungan variabel u_h dan t_h yang konvergen. Jika proses konvergensi dari skor variabel *X* (t_h) dan skor variabel tak bebas *Y* (u_h) dihitung secara terpisah, maka model yang dihasilkan mempunyai hubungan yang lemah. Untuk memperbaiki kondisi tersebut, proses konvergensi dari u_h dan t_h dilakukan secara bersama-sama dengan cara melibatkan skor *Y* pada perhitungan *loading X* :

$$u_{awal} = Y_j \qquad (2.3)$$

$$p' = \frac{u' X}{u' u} \quad (p \text{ sebagai fungsi dari } u) \qquad (2.4)$$

serta melibatkan skor *X* pada perhitungan *loading Y* :

$$q' = \frac{t'Y}{t't} \qquad (q \text{ sebagai fungsi dari } t) \qquad (2.5)$$

Melalui cara tersebut, akan mempercepat proses konvergensi, tetapi masih ada beberapa kelemahan antara lain skor *X* (t_h) yang dihasilkan ternyata tidak orthogonal. Jika t_h tidak orthogonal akan terjadi korelasi yang cukup besar antara variabel bebas *X*. Untuk mengatasi kendala tersebut, skor *X* perlu diskalakan lagi dengan suatu pembobot *w* (*loading*).

Regresi Komponen Utama

Regresi komponen utama merupakan metode yang cukup baik untuk memperoleh koefisien penduga pada persamaan regresi yang mempunyai masalah multikolinearitas. Variabel bebas pada regresi komponen utama berupa hasil kombinasi linear dari variabel asal Z, yang disebut sebagai komponen utama. Koefisien penduga dari metode ini diperoleh melalui penyusutan dimensi komponen utama, dimana subset komponen utama yang dipilih harus tetap mempertahankan keragaman yang sebesar-besarnya. Dimana Z adalah hasil normal baku dari variabel X. Adapun hasil normal baku yang dimaksud adalah dengan mengurangkan setiap variabel bebas asal X_j dengan rata-rata dan dibagi dengan simpangan baku, dinotasikan:

$$Z_{ij} = \frac{\left(X_j - \bar{X}\right)}{s_j} \qquad (2.6)$$

Cara penghapusan komponen utama dimulai dari prosedur seleksi akar ciri dari suatu persamaan:

$$\left| \, AX - \lambda I \, \right| = 0 \qquad (2.7)$$

Jika akar ciri λ_j diurutkan dari nilai terbesar sampai nilai terkecil, maka pengaruh komponen utama W_j berpadanan dengan pengaruh λ_j. Ini berarti bahwa komponen-komponen tersebut menerangkan proporsi keragaman terhadap variabel tak bebas Y yang semakin lama semakin kecil.

Komponen utama W_j saling orthogonal sesamanya dan dibentuk melalui suatu hubungan:

$$W_j = v_{1j}\; Z_1 + v_{2j}\; Z_2 + v_{31j}\; Z_{31} + \ldots + v_{pj}\; Z_p \qquad (2.8)$$

Vektor ciri v_j diperoleh dari setiap akar ciri λ_j yang memenuhi suatu sistem persamaan homogen:

$$\left| \, AX - \lambda_j I \, \right| v_j = 0 \qquad (2.9)$$

dimana $v_j = \left(v_{1j}\,, v_{2j}, v_{3j}, \ldots, v_{pj}\right)$

Jika terdapat *m* subset komponen utama yang akan masuk dalam persamaan regresi, maka persamaan tersebut dapat ditulis sebagai:

$$Y = W_m\; \hat{\beta}_m + \varepsilon \qquad (2.10)$$

Perhitungan koefisien penduga regresi komponen utama $\hat{\beta}$ dapat dilakukan secara analog dengan penduga metode kuadrat terkecil, yaitu:

$$\hat{\beta} = \left(W'\,W\right)^{-1} W'\; y \qquad (2.11)$$

Untuk mendapatkan nilai t_{hitung} pada nilai koefisien regresi komponen utama dapat dilakukan dengan menghitung nilai simpangan baku untuk masing-masing koefisien regresi dengan menggunakan persamaan sebagai berikut:

$$s\left(\gamma_i\right) = \sqrt{\operatorname{var}\left(\gamma_i\right)} \qquad , i = 1, 2, \ldots \qquad (2.12)$$

dimana:

$$\operatorname{var}(\gamma_i) = \frac{s^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}\left(\frac{a_{i1}^2}{\lambda_1} + \frac{a_{i2}^2}{\lambda_2} + \cdots + \frac{a_{ij}^2}{\lambda_j}\right) \quad (2.13)$$

dimana: s^2 = varians

y_i = variabel tak bebas

$\bar{y}$ = nilai rata-rata dari variabel tak bebas

$a_{i1}, a_{12}, \ldots, a_{ij}$ = nilai vektor ciri yang terpilih

$\lambda_1, \lambda_2, \ldots, \lambda_j$ = nilai akar ciri yang terpilih

Untuk mendapatkan nilai t hitung dari koefisien baku regresi komponen utama dapat dihitung dengan persamaan sebagai berikut:

$$t_{hitung} = \frac{koefisien\ penduga(\gamma_i)}{s(\gamma_i)} \quad (2.14)$$

(Jollife, 1986 dalam Herwindiati, 1997).

Seleksi Model Terbaik

Pemilihan model terbaik dapat dilakukan dengan melihat nilai determinasi (R^2). Model dikatakan baik jika nilai R^2 tinggi, nilai R^2 berkisar antara 0 sampai 1. Adapun cara untuk memperoleh nilai R^2 adalah sebagai berikut:

$$R^2 = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2} \quad (2.15)$$

dimana: R^2 = Koefisien determinasi

$\hat{y}_i$ = variabel tak bebas dugaan

$\bar{y}$ = nilai rata-rata dari variabel tak bebas

(Sembiring, 1995).

Menurut (Andriyanto dan Basith, 1999), pemilihan model terbaik juga dapat dilakukan dengan melihat nilai *Mean Square Error Prediction* (MSEP) dan *Root Mean Square Error Prediction* (RMSEP). Metode terbaik adalah metode dengan nilai MSEP dan nilai RMSEP terkecil. Kriteria MSEP dan RMSEP dapat ditentukan dengan cara:

$$MSEP = \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n} \quad (2.16)$$

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \quad (2.17)$$

dimana:

$\hat{y}_i$ = variabel tak bebas dugaan

y_i = variabel tak bebas sebenarnya

n = banyaknya data

3. HASIL DAN PEMBAHASAN

Penyelesaian dengan Metode *Partial Least Square* (PLS)

Perhitungan vektor pembobot w_h, vektor skor (*score vector*) dari variabel *X* dan variabel *Y* serta vektor muatan (*loading vector*) dari variabel *X* dan variabel *Y* merupakan nilai- nilai yang diperlukan untuk menduga koefisien pada metode *Partial Least Square* (PLS) .

Selanjutnya perhitungan nilai R^2, *Mean Square Error Prediction* (MSEP), dan *Root Mean Square Error Prediction* (RMSEP) pada setiap komponen dari metode *Partial Least Square* (PLS) disajikan pada Tabel 3.1 berikut:

Tabel 3.1. Nilai R^2, MSEP dan RMSEP pada *Partial Least Square* (PLS)

Komponen ke	R^2	MSEP	RMSEP
1	0.07197	12.6394	3.5552
2	0.63231	5.0077	2.2378
3	0.69262	4.1863	2.0461
4	0.80321	2.6801	1.6371
5	0.93115	0.9377	0.9683
6	0.94733	0.7174	0.8470
7	0.95499	0.6131	0.7830
8	0.95649	0.5926	0.7698
9	0.95659	0.5912	0.7689

Berdasarkan hasil nilai R^2, MSEP dan RMSEP yang tercantum pada Tabel 3.1 terlihat bahwa pada komponen kelima sudah tercapai kondisi konvergen karena pada komponen tersebut nilai R^2 mulai stabil, yang berarti bahwa pada komponen kelima sampai komponen kesembilan nilai R^2 nya sudah relatif sama. Jadi komponen yang diambil sebagai penduga model adalah komponen kelima. Nilai R^2, MSEP dan RMSEP pada komponen kelima yaitu sebagai berikut:

Tabel 3.2. Nilai R^2 MSEP dan RMSEP pada Komponen Kelima

Komponen Ke	Nilai		
	R^2	MSEP	RMSEP
5	0.93115	0.9377	0.9683

Berdasarkan hasil yang diperoleh, koefisien penduga dan nilai t_{hitung} pada komponen kelima adalah sebagai berikut:

Tabel 3.3. Nilai Koefisien Penduga dan t_{hitung} pada Komponen Kelima

Variabel	Koefisien Penduga	Nilai Koefisien Penduga	t_{hitung}
X_1	β_1	69.7400	1.6000
X_2	β_2	-23.7629	-1.8356 *
X_3	β_3	-98.5289	-3.6448 *
X_4	β_4	-73.2028	-2.8396 *
X_5	β_5	0.1985	0.0039
X_6	β_6	74.7836	2.7270 *
X_7	β_7	1.7561	0.9783
X_8	β_8	85.2408	5.6926 *
X_9	β_9	-12.5974	-0.3367

Keterangan : * Signifikan pada $\alpha = 5\%$

Berdasarkan Tabel 3.3 menunjukkan bahwa koefisien penduga pada metode *Partial Least Square* (PLS) tidak semuanya berpengaruh nyata pada taraf nyata 0.05. Variabel-variabel yang berpengaruh nyata adalah variabel X_2, X_3, X_4, X_6, dan X_8, sedangkan variabel-variabel lainnya tidak berpengaruh nyata. Hal ini ditunjukkan dengan melihat nilai t_{hitung} pada masing-masing variabel X_2, X_3, X_4, X_6, dan X_8 yang lebih besar dari nilai $t_{\text{tabel}} = t_{(0.95,36)} = 1.70$. Berdasarkan pengujian disimpulkan bahwa H_0 ditolak karena nilai $|t_{\text{hitung}}| > |t_{\text{tabel}}|$ sehingga pengujian nilai statistik uji t untuk regresi adalah nyata pada taraf nyata 0.05.

Penyelesaian dengan Metode Regresi Komponen Utama

Dalam analisis regresi komponen utama hal yang terlebih dahulu dilakukan menormal bakukan variabel-variabel X menjadi Z, kemudian menentukan nilai akar ciri dan vektor ciri dari matriks Z dapat dilihat pada Tabel 3.4 dan Tabel 3.5 berikut:

Tabel 3.4. Nilai Akar Ciri λ_j

Akar Ciri (λ_j)	Proporsi	Komulatif
7.1795	0.7977	0.7977
1.3862	0.1540	0.9517
0.2987	0.0332	0.9849
0.1048	0.0116	0.9966
0.0226	0.0025	0.9991
0.0066	0.0007	0.9999
0.0016	0.0002	0.9999
$7.8816\text{x}10^{-5}$	$8.7573\text{x}10^{-6}$	0.9999
$3.4793\text{x}10^{-5}$	$3.8658\text{x}10^{-6}$	1.0000

Tabel 3.5. Nilai Vektor Ciri

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9
Z_1	0.3688	-0.0847	-0.0423	0.2806	-.02761	-0.6861	0.1092	0.3750	0.2770
Z_2	0.3701	-0.0689	-0.250	0.2692	-0.2939	-0.1104	0.0796	0.6595	-0.4953
Z_3	0.3712	-0.0494	-0.0179	0.2119	-0.1934	0.5400	0.0500	0.1979	0.6663
Z_4	0.3714	-0.0554	-0.0185	0.1845	-0.0933	0.4553	-0.0365	0.6127	-0.4815
Z_5	0.2144	0.5961	0.7696	-0.0683	0.0136	-0.0383	0.0011	-1.2×10^{-5}	0.0021
Z_6	0.3687	-0.0879	0.0566	-0.0756	0.7095	-0.0393	0.5834	0.0453	-0.0088
Z_7	-0.0796	-0.7778	0.6227	-0.0277	-0.0073	0.0083	-0.0006	0.0007	0.0037
Z_8	0.3702	-0.0836	-0.0513	-0.0020	0.4376	-0.1211	-0.7983	-0.0879	0.0459
Z_9	0.3550	-0.0865	-0.1048	-0.8710	-0.3100	-0.0209	0.0154	0.0027	-0.0033

Berdasarkan Tabel 3.4 menunjukkan bahwa akar ciri pertama menjelaskan sekitar 79.77% dari keragaman yang terjadi, dan akar ciri yang kedua menjelaskan 15.4% dan akar ciri yang berikutnya hanya menjelaskan sekitar 0.33% dan 0.11% saja. Berdasarkan Tabel 3.5 menunjukkan bahwa dari sembilan komponen utama yang diturunkan dari matriks korelasi antar variabel bebas, ada dua komponen utama yang memegang peranan penting dalam menerangkan keragaman total data, yaitu komponen utama pertama dan kedua atau dilihat dari nilai akar ciri yang lebih besar dari 1. Dengan demikian komponen utama pertama (W_1) dan komponen utama kedua (W_2) yang merupakan kombinasi linear dari Z dapat dinyatakan dalam persamaan berikut:

$$W_1 = 0.3688\ Z_1 + 0.3701\ Z_2 + 0.3712\ Z_3 + 0.3714\ Z_4 + 0.2144\ Z_5 + 0.3687\ Z_6 - 0.0796\ Z_7 + 0.3702\ Z_8 + 0.3550\ Z_9$$

$$W_2 = -0.0847\ Z_1 - 0.0689\ Z_2 - 0.0494\ Z_3 - 0.0554\ Z_4 + 0.5961\ Z_5 - 0.0879\ Z_6 + 0.7778\ Z_7 - 0.0836\ Z_8 - 0.0865\ Z_9$$

Matriks W_j berisi skor komponen utama yang diperoleh dari persamaan W_1 dan W_2. Selanjutnya Y diregresikan terhadap skor komponen utama W_1 dan W_2, hasilnya dapat dilihat pada Tabel 3.6 berikut:

Tabel 3.6. Penduga Parameter Regresi Komponen Utama

Variabel	DB	Pendugaan	t_{hitung}	P Value	Nilai VIF
Konstanta	1	40.6361	105.850	2×10^{-16}	
W_1	1	1.0481	7.213	2.86×10^{-8}	1
W_2	1	-0.8985	-2.717	0.0104	1

Tabel 3.7 Tabel Sidik Ragam Regresi Komponen Utama

Sumber Keragaman	DB	Jumlah Kuadrat	Kuadrat Tengah	F_{hitung}	P Value
Model	2	315.214	157.607	29.71	4.179^{-8}
Galat	33	175.089	5.306		
Total	35	490,303			

$R^2 = 0.6429$ adj $R^2 = 0.6213$

Berdasarkan Tabel 3.6 dan Tabel 3.7 terlihat bahwa nilai R^2 yang dihasilkan oleh koefisien penduga model regresi komponen utama sangat tinggi yaitu sebesar 64.29%, kemudian nilai F_{hitung} sebesar 29.71 dan F_{tabel} = $F_{0.05}(9,33)$ = 2.17. Berdasarkan pengujian disimpulkan bahwa H_0 ditolak karena nilai $F_{hitung} > F_{tabel}$, sehingga pengujian nilai statistik uji F untuk regresi adalah nyata pada taraf nyata 0.05 dengan nilai VIF sebesar 1.

Model yang sudah didapat selanjutnya ditransformasikan kembali ke variabel asal Z, sehingga diperoleh persamaan regresi dalam variabel baku sebagai berikut:

$$\hat{Y} = 40.6361 + 1.0481\,W_1 - 0.8985\,W_2 \tag{3.1}$$

$$\hat{Y} = 40.3631 + 0.4626\,Z_1 + 0.4498\,Z_2 + 0.4334\,Z_3 + 0.4390\,Z_4 + 0.3109\,Z_5 + 0.4654\,Z_6 + 1.0869\,Z_7 + 0.4631\,Z_8 + 0.4498\,Z_9 \tag{3.2}$$

Nilai simpangan baku dan nilai t_{hitung} untuk masing-masing koefisien regresidapat dilihat pada Tabel 3.8 berikut:

Tabel 3.8. Analisis Signifikansi Koefisien Regresi Parsial

Variabel	Nilai Koefisien	Simpangan Baku	t_{hitung}
Z_1	0.4626	0.0162	28.6329 *
Z_2	0.4498	0.0156	28.8235 *
Z_3	0.4334	0.1506	28.7819 *
Z_4	0.4390	0.0152	28.8293 *
Z_5	0.3109	0.0533	5.8305 *
Z_6	0.4654	0.0163	28.5772 *
Z_7	1.0869	0.0688	15.7995 *
Z_8	0.4631	0.0162	28.6576 *
Z_9	0.4498	0.0158	28.5408 *

Keterangan : * Signifikan pada $\alpha = 5\,\%$

Analisis signifikansi koefisien parsial baku pada regresi komponen utama disajikan dalam Tabel 3.8 yang menunjukkan bahwa semua koefisien regresi nyata secara statistik pada taraf nyata 0.05. Hal ini ditunjukkan dengan melihat nilai t_{hitung} pada masing-masing variabel dengan nilai $t_{tabel} = t_{(0.95,36)}$ = 1.70. Berdasarkan pengujian disimpulkan bahwa H_0 ditolak karena nilai $|t_{hitung}| > |t_{tabel}|$ sehingga pengujian nilai statistik uji t untuk regresi adalah nyata pada taraf nyata 0.05.

Untuk memperoleh persamaan penduga dengan menggunakan variabel asli, ditransformasikan kembali ke model regresi $\hat{Y} = f\,(w)$ ke variabel asalnya $\hat{Y} = f\,(x)$, yaitu:

$$\hat{Y} = 40.3631 + 0.4626\,Z_1 + 0.4498\,Z_2 + 0.4334\,Z_3 + 0.4390\,Z_4 + 0.3109\,Z_5 + 0.4654\,Z_6 + 1.0869\,Z_7 + 0.4631\,Z_8 + 0.4498\,Z_9$$

$$\hat{Y} = 2.7413 + 3.7641\,X_1 + 3.9860\,X_2 + 4.2835\,X_3 + 4.3283\,X_4 - 0.1689\,X_5 + 5.1074\,X_6 + 5.7933\,X_7 + 4.8193\,X_8 + 7.7344\,X_9 \tag{3.3}$$

Tabel 3.9 berikut ini merupakan nilai R^2, MSEP dan RMSEP dari regresi komponen utama :

Tabel 3.9. Nilai R^2, MSEP dan RMSEP dari Metode Regresi Komponen Utama

Metode Analisa	Nilai		
	R^2	MSEP	RMSEP
Regresi Komponen Utama	0.6429	5.2513	2.2916

Perbandingan Metode

Dari hasil seluruh pembahasan, nilai koefisien penduga dan nilai t hitung dari kedua metode dapat dilihat pada Tabei 3.10 berikut:

Tabel 3.10. Nilai Koefisien Penduga dan Nilai t hitung dari Kedua Metode

Koefisien Penduga	Nilai Koefisien Penduga		Nilai t hitung	
	Partial Least Square (PLS)	Regresi Komponen Utama	*Partial Least Square* (PLS)	Regresi Komponen Utama
β_1	69.7400	0.4626	1.6000	28.6329 *
β_2	-23.7629	0.4498	-1.8356 *	28.8235 *
β_3	-98.5289	0.4334	-3.6448 *	28.7819 *
β_4	-73.2028	0.4390	-2.8396 *	28.8293 *
β_5	0.1985	0.3109	0.0039	5.8305 *
β_6	74.7836	0.4654	2.7270 *	28.5772 *
β_7	1.7561	1.0869	0.9783	15.7995 *
β_8	85.2408	0.4631	5.6926 *	28.6576 *
β_9	-12.5974	0.4498	-0.3367	28.5408 *

Keterangan : * Signifikan pada $\alpha = 5\%$

Berdasarkan Tabel 3.10 menunjukkan bahwa koefisien penduga pada metode *Partial Least Square* (PLS) tidak semuanya berpengaruh nyata pada taraf nyata 0.05. Variabel-variabel yang berpengaruh nyata adalah variabel X_2, X_3, X_4, X_6, dan X_8, sedangkan variabel-variabel lainnya tidak berpengaruh nyata. Hal ini ditunjukkan dengan melihat nilai t_{hitung} pada masing-masing variabel X_2, X_3, X_4, X_6, dan X_8 yang lebih besar dari nilai $t_{\text{tabel}} = t_{(0.95,36)} = 1.70$. Berdasarkan pengujian disimpulkan bahwa H_0 ditolak karena nilai $|t_{\text{hitung}}| > |t_{\text{tabel}}|$ sehingga pengujian nilai statistik uji t untuk regresi adalah nyata pada taraf nyata 0.05. Sedangkan pada regresi komponen utama menunjukkan bahwa semua koefisien regresi nyata secara statistik pada taraf nyata 0.05. Hal ini ditunjukkan dengan melihat nilai t_{hitung} pada masing-masing variabel dengan nilai $t_{\text{tabel}} = t_{(0.95,36)} = 1.70$. Berdasarkan pengujian disimpulkan bahwa H_0 ditolak karena nilai $|t_{\text{hitung}}| > |t_{\text{tabel}}|$ sehingga pengujian nilai statistik uji t untuk regresi adalah nyata pada taraf nyata 0.05.

Untuk mengetahui metode mana yang lebih baik, perlu dikaji nilai R^2, MSEP dan RMSEP nya. Nilai R^2, MSEP dan RMSEP yang diperoleh dari kedua metode dapat dilihat pada Tabel 3.11 berikut ini:

Tabel 3.11. Nilai R^2, MSEP dan RMSEP dari Kedua Metode

Metode Analisa	Nilai		
	R^2	MSEP	RMSEP
Partial Least Square (PLS)	0.93115	0.9377	0.9683
Regresi Komponen Utama	0.6429	5.2513	2.2916

Berdasarkan Tabel 3.11, nilai R^2 dari metode *Partial Least Square* (PLS) memberikan nilai yang lebih besar dibandingkan dengan metode regresi komponen utama. Hal ini berarti metode *Partial Least Square* (PLS) memberikan ketepatan model yang lebih baik dari pada metode regresi komponen utama. Begitu juga jika ditinjau dari nilai MSEP dan nilai RMSEP, metode *Partial Least Square* (PLS) mempunyai nilai yang lebih rendah dari pada metode regresi komponen utama sehingga metode *Partial Least Square* (PLS) memberikan ketepatan model yang lebih baik dari pada metode regresi komponen utama.

4. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang dilakukan terhadap kedua metode untuk mengatasi multikolinear yaitu metode *Partial Least Square* (PLS) dan metode regresi komponen utama dapat disimpulkan bahwa: (1) Pada Metod*e Partial Least Square* (PLS) nilai koefisien penduga pada masing-masing variabel tidak semuanya berpengaruh nyata pada taraf nyata 0.05, sedangkan pada regresi komponen utama semua nilai koefisien penduga pada masing-masing variabel semuanya berpengaruh nyata pada taraf nyata 0.05; (2) Metode *Partial Least Square* (PLS) memberikan hasil yang lebih baik jika dibandingkan dengan metode regresi komponen utama. Hal ini dapat disimpulkan dengan melihat nilai R^2, *Mean Square Error Prediction* (MSEP), dan *Root Mean Square Error Prediction* (RMSEP). Metode *Partial Least Square* (PLS) mempunyai nilai R^2 yang lebih tinggi dan mempunyai nilai MSEP dan RMSEP yang lebih rendah jika dibandingkan terhadap metode regresi komponen utama.

Peneliti yang berkeinginan melanjutkan pengembangan tulisan ini diharapkan dapat menggunakan metode yang berbeda misalnya dengan menggunakan metode regresi *ridge* sebagai metode pembanding serta menggunakan data riil.

DAFTAR PUSTAKA

[1]. Bilfarsah, A. 2005. Efektifitas Metode Aditif Spline Kuadrat Terkecil Parsial Dalam Pendugaan Model Regresi. *Makara, Sains*, 9 (1) : 28 - 33.

[2]. Herwindiati, D.E. 1997. Pengkajian Regresi Komponen Utama, Regresi Ridge, dan Regresi Kuadrat Terkecil Parsial untuk Mengatasi Kolinearitas. *Tesis*. Institut pertanian Bogor. Bogor.

[3]. Montgomery, D.C dan Peck, E.A. 1992. *Introduction to Linier Regression Analysis*. John Willey & Sons. New York.

[4]. Naes, T. 1985. Multivariate Calibration When The Error Covariance Matrix is Structured. *Technometrics*, 27 (3) : 301 – 311.

[5]. Noryanti. 2009. Pengaruh Hasil-hasil Ujian di Sekolah Terhadap Hasil Ujian Nasional di SMU Negeri 1 Limboto Kabupaten Gorontalo. *Teknologi Technoscientia*. 2 (1): 85 – 95.

[6]. Sembiring, R. K. 1995. *Analisis Regresi*. ITB Bandung. Bandung.

[7]. Soemartini. 2008. Penyelesaian Multikolinearitas Melalui Metode Ridge Regression. *Tesis*. Universitas Padjajaran.

[8]. Wigena, A. H dan Aunuddin. 1998. Metode PLS untuk Mengatasi Kolinearitas dalam Kalibrasi Ganda. *Forum Statistika dan Komputasi*. 3 (1) : 17-19.