



PERBANDINGAN METODE *PARTIAL LEAST SQUARE (PLS)* DAN *PRINCIPAL COMPONENT REGRESSION (PCR)* UNTUK MENGATASI MULTIKOLINEARITAS PADA MODEL REGRESI LINEAR BERGANDA

Skripsi

disusun sebagai salah satu syarat
untuk memperoleh gelar Sarjana Sains

Program Studi Matematika



oleh

Eko Supriyadi

4111412023

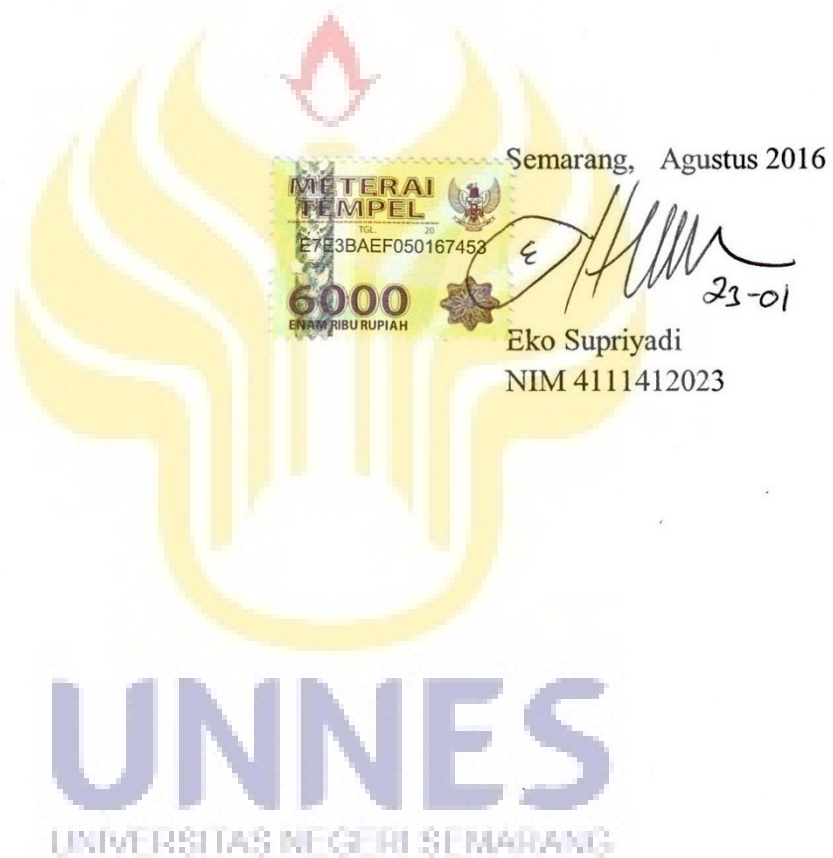
**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI SEMARANG**

2016



PERNYATAAN

Saya menyatakan bahwa skripsi ini bebas plagiat, dan apabila di kemudian hari terbukti terdapat plagiat dalam skripsi ini, maka saya bersedia menerima sanksi sesuai ketentuan peraturan perundang-undangan.



PENGESAHAN

Skripsi yang berjudul

Perbandingan Metode Partial Least Square (PLS) dan Principal Component
Regression (PCR) Untuk Mengatasi Multikolinearitas Pada Model Regresi
Linear Berganda

Disusun oleh

Eko Supriyadi

4111412023

Telah dipertahankan di hadapan sidang Panitia Ujian Skripsi FMIPA UNNES
pada tanggal 2 Agustus 2016.



Panitia,
Ketua

Prof. Dr. Zaenuri, S.E., M.Si, Akt.
NIP 196412231988031001

Sekretaris

Drs. Arief Agoestanto, M.Si.
NIP 196807221993031005

Ketua Penguji

Prof. Dr. Zaenuri, S.E., M.Si, Akt.
NIP 196412231988031001

Anggota Penguji/
Pembimbing I

Dr. Sclastika Mariani, M.Si.
NIP 196502101991022001

Anggota Penguji/
Pembimbing II

Drs. Sugiman, M.Si.
NIP 196401111989011001

MOTTO

- ❖ Sesungguhnya bersama kesulitan ada kemudahan. Maka apabila engkau telah selesai (dari suatu urusan), tetaplah bekerja keras untuk (urusan yang lain), dan hanya kepada Tuhanmulah engkau berharap. (QS. Al-Insyirah: 6-8)
- ❖ Al i'timaadu 'alan nafsi asaasun najakh (percaya diri adalah kunci kesuksesan).
- ❖ Do the best, lakukan yang terbaik.

PERSEMBAHAN

- ❖ Untuk kedua orang tua tercinta bapak Slamet Priyadi dan ibu Sukinah.
- ❖ Untuk adik-adiku tersayang, Ida Rosita dan Fathir Aghna Ahsani.
- ❖ Untuk teman-teman Prodi Matematika angkatan 2012.
- ❖ Untuk Universitas Negeri Semarang.

KATA PENGANTAR

Puji syukur kehadiran Allah SWT yang telah memberikan nikmat dan karunia-Nya serta kemudahan sehingga penulis dapat menyelesaikan skripsi yang berjudul “Perbandingan Metode Partial Least Square (PLS) dan Principal Component Regression (PCR) Untuk Mengatasi Multikolinearitas Pada Model Regresi Linear Berganda”.

Penyusunan skripsi ini dapat diselesaikan berkat kerjasama, bantuan, dan dorongan dari berbagai pihak. Oleh karena itu penulis mengucapkan terima kasih kepada :

1. Prof. Dr. Fathur Rokhman, M.Hum., Rektor Universitas Negeri Semarang.
2. Prof. Dr. Zaenuri, S.E., M.Si., Akt., Dekan FMIPA Universitas Negeri Semarang.
3. Drs. Arief Agoestanto, M.Si., Ketua Jurusan Matematika FMIPA Universitas Negeri Semarang yang telah memberikan bimbingan, pengarahan, nasihat, saran, dan dorongan selama penyusunan skripsi ini.
4. Drs. Mashuri, M.Si., Ketua Prodi Matematika FMIPA Universitas Negeri Semarang.
5. Dr. Scolastika Mariani, M.Si., Dosen Pembimbing I yang telah memberikan bimbingan, pengarahan, nasihat, saran, dan dorongan selama penyusunan skripsi ini.
6. Drs. Sugiman, M.Si., Dosen Pembimbing II yang telah memberikan bimbingan, pengarahan, nasihat, saran, dan dorongan selama penyusunan skripsi ini.

7. Prof. Dr. Zaenuri S.E, M.Si, Akt., Dosen Penguji yang telah memberikan penilaian dan saran dalam perbaikan skripsi ini.
8. Prof. Dr. St. Budi Waluya, M.Si., Dosen Wali yang telah memberikan bimbingan dan arahan selama kuliah di jurusan Matematika Universitas Negeri Semarang.
9. Dosen jurusan Matematika Universitas Negeri Semarang yang telah membekali dengan berbagai ilmu selama mengikuti perkuliahan sampai akhir penulisan skripsi ini.
10. Staf Tata Usaha Universitas Negeri Semarang yang telah banyak membantu penulis selama mengikuti perkuliahan dan penulisan skripsi ini.
11. Ayah dan ibu tercinta, Bapak Slamet Priyadi dan Ibu Sukinah yang senantiasa memberikan dukungan dan do'a yang tiada putusnya.
12. Adik-adikku tersayang, Ida Rosita dan Fathir Aghna Ahsani yang selalu memberikan motivasi, semangat, dan do'a.
13. Teman-Teman Prodi Matematika angkatan 2012 yang berjuang bersama untuk mencapai cita-cita.
14. Teman-teman kos "Tohamir" yang memberikan dukungan, semangat serta doa.
15. Ulya Nur Bayti yang selalu memberikan dorongan motivasi, semangat dan do'a.
16. Teman-teman KKN Alternatif 2A 2015 KKN 944, Srumbunggunung, Desa Poncoruso, Bawen yang memberikan semangat dan do'a.

17. Semua pihak yang tentunya tidak dapat disebutkan satu per satu yang telah memberikan bantuan.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih terdapat banyak kekurangan. Oleh karena itu, penulis mengharapkan saran dan kritik yang membangun dari pembaca.

Semarang, Juli 2016

Penulis



ABSTRAK

Supriyadi, Eko. 2016. *Perbandingan Metode Partial Least Square (PLS) dan Principal Component Regression (PCR) untuk Mengatasi Multikolinearitas Pada Model Regresi Linear Berganda*. Skripsi, Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Semarang. Pembimbing Utama Dr. Scolastika Mariani, M.Si. Dan Pembimbing Pendamping Drs. Sugiman, M.Si.

Kata Kunci: Multikolinearitas, *Partial Least Square* (PLS), *Principal Component Regression* (PCR).

Salah satu asumsi analisis regresi linear berganda yaitu tidak terjadi masalah multikolinearitas. Apabila terjadi masalah multikolinearitas, metode *Partial Least Square* (PLS) dan *Principal Component Regression* (PCR) merupakan dua metode yang dapat digunakan untuk mengatasi masalah multikolinearitas tersebut. Penelitian ini bertujuan untuk mengetahui: (1) Bagaimana model persamaan regresi dengan metode *Partial Least Square* (PLS) untuk mengatasi multikolinearitas; (2) Bagaimana model persamaan regresi dengan metode *Principal Component Regression* (PCR) untuk mengatasi multikolinearitas; (3) Metode manakah antara *Partial Least Square* (PLS) dan *Principal Component Regression* (PCR) untuk mengatasi multikolinearitas yang lebih efektif.

Metode yang digunakan dalam penelitian ini adalah *Partial Least Square* (PLS) dan *Principal Component Regression* (PCR) dengan data Anggaran Pendapatan Daerah Provinsi Jawa Tengah 2013. Metode pengumpulan data yang digunakan adalah metode dokumentasi, pustaka dan wawancara. Langkah-langkah analisis yaitu : (1) Deskripsi Data, (2) Uji Asumsi Regresi Linear, (3) Uji Asumsi Klasik, (4) Mengatasi Masalah Multikolinearitas, (5) Pemilihan Metode terbaik dengan R^2 tertinggi dan MSE terkecil dan untuk menganalisis data menggunakan program SAS.

Simpulan yang diperoleh (1) Model persamaan regresi dengan metode *Partial Least Square* (PLS) pada kasus pendapatan anggaran Daerah Provinsi Jawa Tengah yaitu $\hat{Y} = 1382126382 + 83948025.7X_1 + 40120614.88X_2 + 74135918.8X_3 + 94632319.88X_4 + 145001135.3X_5 + 59090688.22X_6$, (2) Model persamaan regresi dengan metode *Principal Component Regression* (PCR) pada kasus pendapatan anggaran Daerah Provinsi Jawa Tengah yaitu $\hat{Y} = 1382126508 + 28888566.77 X_1 + 84611231.43 X_2 - 14440705.5 X_3 - 5053009.86 X_4 - 6076336.99 X_5 - 89989904.2 X_6$, (3) Metode yang lebih efektif adalah *Partial Least Square* (PLS) dengan nilai $R^2 = 0.9170$ dan nilai MSE yang dihasilkan PLS = 1.394114E16.

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PERNYATAAN	iii
HALAMAN PENGESAHAN	iv
MOTTO DAN PERSEMBAHAN	v
KATA PENGANTAR	vi
ABSTRAK	ix
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN	xiv
 BAB 1 PENDAHULUAN	 1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	4
1.3 Rumusan Masalah	5
1.4 Batasan Masalah	6
1.5 Tujuan	6
1.6 Manfaat Penelitian	7
1.6.1 Bagi Mahasiswa	7
1.6.2 Bagi Pembaca	7
 BAB 2 TINJAUAN PUSTAKA	 8
2.1 Analisis Regresi	8
2.2 Analisis Regresi Berganda	9
2.3 Uji Asumsi Klasik	11
2.4 Multikolinearitas	12
2.4.1 Pengertian Multikolinearitas	12
2.4.2 Penyebab Terjadinya Multikolinearitas	12
2.4.3 Konsekuensi Multikolinearitas	13
2.4.4 Cara Mendeteksi Multikolinearitas	14
2.5 Metode Kuadrat Terkecil	15
2.6 Nilai Eigen dan Vektor Eigen	18
2.7 Principal Component Analisis (PCA)	19
2.8 Principal Component Regression (PCR)	19
2.9 Partial Least square (PLS)	23
2.10 Pendapatan Daerah	33
2.11 Statistical Analysis System (SAS).....	36

2.12 Kerangka Berfikir	43
BAB 3 METODE PENELITIAN	47
3.1 Fokus Penelitian	47
3.2 Studi Pustaka	47
3.3 Perumusan Masalah	48
3.4 Pengumpulan Data	48
3.5 Pemecahan Masalah	49
3.6 Menarik Kesimpulan	51
BAB 4 HASIL DAN PEMBAHASAN	52
4.1 Deskripsi Data	52
4.2 Uji Asumsi Regresi	52
4.2.1 Uji Normalitas	52
4.2.2 Model Regresi Berganda	53
4.3 Uji Asumsi Klasik	54
4.3.1 Multikolinearitas	54
4.3.2 Autokorelasi	56
4.6.3 Heteroskedastisitas	57
4.4 Mengatasi Masalah Multikolinearitas	58
4.4.1 Principal Component Regression	58
4.4.1.1 Menentukan Komponen Utama (principal Component)	58
4.4.1.2 Regresi Komponen Utama	64
4.4.2 Partial Least Square (PLS)	65
4.4.2.1 Pembentukan Komponen PLS pertama, t_1	65
4.4.2.2 Pembentukan Komponen PLS pertama, t_2	68
4.5 Pemilihan Metode Terbaik	70
4.6 Pembahasan	71
BAB 5 PENUTUP	79
5.1 Simpulan	79
5.2 Saran	80
DAFTAR PUSTAKA	81
LAMPIRAN	82

DAFTAR TABEL

Tabel	Halaman
4.1 Normalitas	53
4.2 Parameter Estimates	54
4.3 Nilai VIF dan TOL.....	55
4.4 Output Uji Autokorelasi	56
4.5 Communality	58
4.6 Eigenvalues of the Correlation Matrix	60
4.7 Skor Komponen Utama	61
4.8 Variabel Baru yang Terbentuk	62
4.9 Regresi y terhadap variabel komponen utama yang terbentuk.....	64
4.10 Uji Signifikansi y (Terstandarisasi) Terhadap Masing-Masing x	66
4.11 Komponen PLS Pertama	67
4.12 Uji signifikansi masing-masing variabel pembentukan PLS kedua	68
4.13 Regresi y Terhadap Variabel PLS yang Terbentuk.....	69
4.14 Nilai R^2 dan MSE	71

DAFTAR GAMBAR

Gambar	Halaman
2.1 Tampilan Awal Program SAS	38
2.2 Tampilan Explorer Window (Kiri) dan Result Window (Kanan)	41
2.3 Kerangka berfikir	46
3.1 Diagram Alur	50
4.1 Plot Heteroskedastisitas	57



DAFTAR LAMPIRAN

Lampiran	Halaman
1. Data Anggaran Pendapatan Daerah Provinsi Jawa Tengah 2013	66
2. Input Data dan Standarisasi Data Dengan Program SAS	86
3. Output SAS Hasil Standarisasi	87
4. Pengecekan Autokorelasi Heteroskedastisitas dan Multikolinearitas ...	88
5. Output Autokorelasi Heteroskedastisitas dan Multikolinearitas	89
6. Penanganan dengan Metode PCR	91
7. Output dengan Metode PCR	93
8. Regresi Y terhadap variabel yang terbentuk K1, K2 (PCR).....	98
9. Output Regresi Y terhadap variabel yang terbentuk K1, K2 (PCR)	99
10. Penanganan dengan Metode PLS Uji Signifikansi y' Terhadap Semua Variabel x.....	100
11. Output Penanganan dengan Metode PLS Uji Signifikansi y' Terhadap Semua Variabel x	102
12. Korelasi Tiap Variabel	103
13. Output Korelasi Tiap Variabel.....	104
14. Uji Signifikansi y' Terhadap t1 dan Masing-masing variabel x	105
15. Output Uji Signifikansi y' Terhadap t1 dan Masing-masing variabel x	106

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Analisis regresi merupakan analisis yang mempelajari bagaimana membentuk sebuah hubungan fungsional dari data untuk dapat menjelaskan atau meramalkan suatu fenomena alami atas dasar fenomena yang lain. Analisis regresi memiliki peranan yang penting dalam berbagai bidang ilmu pengetahuan. Kebanyakan analisis regresi bergantung pada metode kuadrat terkecil untuk mengestimasi parameter-parameternya dalam model regresi. Tetapi metode ini biasanya dibentuk dengan beberapa asumsi, seperti tidak ada autokorelasi, tidak terjadi multikolinearitas, homoskedastisitas, dan *error* berdistribusi normal.

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).

Pada pembentukan model regresi terdapat kemungkinan adanya hubungan antara variabel satu dengan variabel yang lain. Multikolinearitas

menyebabkan estimator mempunyai varian yang besar akibatnya interval estimasi cenderung lebih besar sehingga membuat variabel bebas secara statistika tidak signifikan padahal nilai koefisien determinasi (R^2) tinggi sehingga sulit mendapatkan estimasi yang tepat (Widarjono, 2007). Oleh karena itu perlu dilakukan tindakan lanjut untuk menangani multikolinearitas.

Efek multikolinearitas dapat menjadikan nilai model tidak dapat menjelaskan hubungan antara variabel dependent dan variabel independent secara baik. Keberadaan multikolinearitas akan menyebabkan varians parameter yang diestimasi akan menjadi lebih besar dari yang seharusnya dengan demikian tingkat akurasi dari estimasi akan menurun (Sukmono, A 2014).

Multikolinearitas dalam suatu model regresi dapat diketahui dengan menghitung nilai *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. Gejala multikolinearitas menimbulkan masalah dalam model regresi. Korelasi antar variabel bebas yang sangat tinggi menghasilkan penduga model regresi yang bias, tidak stabil, dan mungkin jauh dari nilai prediksinya (Farahani *et al*, 2010).

Multikolinearitas dapat diatasi dengan beberapa metode antara lain *Partial Least Square* (PLS) dan *Principal component Regression* (PCR) (Saika, 2014). Metode PLS merupakan metode yang mengkombinasikan sifat-sifat dari komponen utama dan regresi linear berganda. Tujuan dari metode PLS adalah mengestimasi dan menganalisis variabel terikat dari variabel-variabel bebas. Dalam hal ini, PLS mereduksi dimensi variabel-variabel bebas dengan membentuk variabel-variabel baru yang merupakan kombinasi linear dari variabel-variabel bebas dengan dimensi lebih kecil (Abdi, 2010). Sedangkan metode PCR merupakan salah satu analisis regresi yang menggunakan komponen utama untuk mengatasi masalah multikolinieritas pada regresi berganda (Tazloqoh *et al*, 2015). Komponen utama merupakan suatu teknik statistika untuk mengubah dari sebagian besar variabel asli yang digunakan yang saling berkorelasi satu dengan yang lainnya menjadi satu kumpulan variabel baru yang lebih kecil dan saling bebas (Johnson dan Wichern, 2010).

Menurut Nurhasanah (2012) hasil penelitian menunjukkan metode *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; 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 MSE dan RMSEP yang lebih rendah jika dibandingkan terhadap metode regresi komponen utama.

Dari latar belakang di atas maka penulis tertarik untuk menganalisis dengan judul “Perbandingan metode *Partial Least Square* (PLS) dan *Principal Component Regression* (PCR) untuk mengatasi multikolinearitas pada model regresi linear berganda”. Dari kedua metode *Partial Least Square* (PLS) dan *Principal component Regression* (PCR) akan membandingkan nilai koefisien determinasi R^2 dan *Mean Square Error* (MSE). Kemudian didapatkan metode mana yang lebih baik untuk mengatasi multikolinearitas.

Penelitian ini didukung dengan penggunaan paket program SAS. Paket program SAS (*Statistical Analysis System*) adalah paket program yang mendukung analisis dalam bidang statistika, riset operasi, analisis ekonomi, time series dan lain-lain.

1.2. Identifikasi Masalah

Berdasarkan latar belakang di atas, dapat diidentifikasi beberapa masalah sebagai berikut:

- 1) Bagaimana mengidentifikasi masalah multikolinearitas?
- 2) Metode apa yang dapat mengatasi masalah multikolinearitas?
- 3) Bagaimana model persamaaan dengan metode *Partial Least Square* (PLS) untuk mengatasi multikolinearitas?

- 4) Bagaimana model persamaan dengan metode *Principal Component Regression* (PCR) untuk mengatasi multikolinearitas?
- 5) Bagaimana perbandingan metode *Partial Least Square* (PLS) dan *Principal Component Regression* (PCR) untuk mengatasi multikolinearitas?
- 6) Metode manakah antara *Partial Least Square* (PLS) dan *Principal Component Regression* (PCR) untuk mengatasi multikolinearitas yang lebih baik ?

1.3. Rumusan Masalah

Berdasarkan latar belakang di atas, maka dapat dirumuskan permasalahan sebagai berikut:

- 1) Bagaimana model persamaan regresi dengan metode *Partial Least Square* (PLS) untuk mengatasi multikolinearitas?
- 2) Bagaimana model persamaan regresi dengan metode *Principal Component Regression* (PCR) untuk mengatasi multikolinearitas?
- 3) Metode manakah antara *Partial Least Square* (PLS) dan *Principal Component Regression* (PCR) untuk mengatasi multikolinearitas yang lebih baik?

1.4. Batasan Masalah

Batasan masalah yang dilakukan pada penelitian ini adalah sebagai berikut:

1. Identifikasi hanya untuk mengatasi masalah multikolinearitas.
2. Penelitian ini hanya menggunakan metode *Partial Least Square* (PLS) dan *Principal Component Regression* (PCR) untuk mengatasi multikolinearitas.
3. Paket program yang mendukung penelitian adalah SAS (*Statistical Analysis System*).
4. Data yang dipakai Anggaran Pendapatan Daerah Provinsi Jawa Tengah 2013.

1.5. Tujuan

Tujuan dari penulisan dan penelitian ini antara lain:

- 1) Mengetahui bagaimana model persamaan regresi dengan metode *Partial Least Square* (PLS) untuk mengatasi masalah multikolinearitas.
- 2) Mengetahui bagaimana model persamaan regresi dengan metode *Principal Component Regression* (PCR) untuk mengatasi masalah multikolinearitas.
- 3) Mengetahui metode yang lebih baik antara *Partial Least Square* (PLS) dan *Principal Component Regression* (PCR) untuk mengatasi multikolinearitas berdasarkan pada perbandingan nilai MSE dan R^2 .

1.6. Manfaat Penelitian

Hasil dari penelitian ini diharapkan dapat memberikan manfaat, antara lain:

1.6.1. Bagi Mahasiswa

- 1) Mahasiswa memperoleh pengetahuan tentang multikolinearitas.
- 2) Dapat membantu mahasiswa untuk menggunakan metode *Partial Least Square* (PLS) dan *Principal Component Regression* (PCR) untuk mengatasi multikolinearitas.
- 3) Dapat membantu mahasiswa untuk memilih metode yang lebih baik guna mengatasi multikolinearitas.

1.6.2. Bagi Pembaca

- 1) Diharapkan agar hasil penelitian ini dapat menambah pengetahuan pembaca mengenai topik terkait dengan penulisan ini.
- 2) Meberikan tambahan ilmu dan wawasan yang baru tentang cara mendeteksi dan mengatasi multikolinearitas dengan menggunakan metode *Partial Least Square* (PLS) dan *Principal Component Regression* (PCR).

BAB 2

TINJAUAN PUSTAKA

2.1. Analisis Regresi

Analisis regresi linear digunakan untuk menaksir atau meramalkan nilai variabel dependen bila nilai variabel independen dinaikan atau diturunkan. Analisis ini didasarkan pada hubungan satu variabel dependen dengan satu atau lebih variabel independen. Jika hanya menggunakan satu variabel independen maka disebut analisis regresi linear sederhana dan jika menggunakan lebih dari satu variabel independen maka disebut analisis regresi linear berganda (*multiple regresision*).

Asumsi yang mendasari pada analisis regresi linear adalah bahwa distribusi data adalah normal. Selain ini terdapat asumsi klasik yang biasanya digunakan pada penelitian, yaitu tidak adanya multikolinearitas, heterokedastisitas, dan autokorelasi pada model regresi (Priyanto, 2013, hal 39).

Regresi sederhana didasarkan pada hubungan fungsional ataupun kausal satu variabel independen dengan satu variabel dependen.

Persamaan umum regresi linear sederhana:

(2.1)

$$y = a + \beta x + \varepsilon$$

Dimana:

y = Subjek dalam variabel dependen yang diprediksikan.

a = Harga Y ketika harga $X = 0$ (harga konstanta)

b = Angka arah atau koefisien regresi, yang menunjukkan angka peningkatan ataupun penurunan variabel dependen yang didasarkan pada perubahan variabel independen. Bila (+) arah garis naik, dan bila (-) maka arah garis turun.

X = subjek pada variabel independen yang mempunyai nilai tertentu.

ε = adalah error (galat) pengukuran.

Manfaat dari hasil analisis regresi adalah untuk membuat keputusan apakah naik dan menurunnya variabel dependen dapat dilakukan melalui peningkatan variabel independen atau tidak. (Sugiyono, 2012, hal 260).

2.2. Analisis Regresi Berganda

Untuk memperkirakan/meramalkan nilai variabel Y , akan lebih baik apabila kita ikut memperhitungkan variabel-variabel lain yang ikut mempengaruhi Y . Dengan demikian, kita mempunyai hubungan antara satu variabel tidak bebas (*dependent variabel*) Y dengan beberapa variabel lain yang bebas (*independent variabel*) X_1, X_2, \dots, X_p .

Untuk meramalkan Y , apabila semua nilai variabel bebas diketahui, maka kita dapat mempergunakan persamaan regresi linear berganda. Hubungan Y dan X_1, X_2, \dots, X_p yang sebenarnya adalah sebagai berikut.

(2.2)

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \varepsilon_1$$

dengan X_1, X_2, \dots, X_p = variabel bebas.

Y = variabel tak bebas

β_0 = intersep

β_1 = parameter yang akan ditaksir

ε = unsur gangguan stokastik (error)

Suatu model regresi linear beganda dengan p variabel bebas dimana $\beta_j, j = 0, 1, \dots, p$ disebut bilangan pokok (koefisien) regresi. Parameter β_j mewakili perubahan yang diharapkan dalam variabel terikat Y di tiap unit berubah ke X_j ketika semua variabel bebas yang tersisa $X_i (i \neq j)$ tidak berubah.

Dan dengan bentuk umum sebagai berikut:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & X_{11} & X_{12} & \dots & X_{1p} \\ 1 & X_{21} & X_{22} & \dots & X_{2p} \\ 1 & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \dots & X_{np} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

Atau

(2.3)

$$y = x\beta + \varepsilon$$

Dimana:

y = vektor kolom $n \times 1$ variabel tak bebas Y

x = matrik $n \times (p + 1)$ dari variabel bebas X

β = vektor kolom $(p + 1) \times 1$ dari parameter yang tak diketahui $\beta_0, \beta_1, \dots, \beta_p$

ε = vektor kolom $n \times 1$ dari gangguan (*disturbance*) ε_t

Penambahan variabel bebas ini diharapkan dapat lebih menjelaskan karakteristik hubungan yang ada, walaupun masih saja ada variabel yang terabaikan.

2.3. Uji Asumsi Klasik

Analisis regresi merupakan alat analisis yang termasuk statistik parametrik. Sebagai alat statistik parametrik analisis regresi membutuhkan asumsi yang perlu dipenuhi sebelum dilakukan analisis. Analisis ini dinamakan dengan uji asumsi klasik. Asumsi klasik tersebut dapat menghilangkan estimator linear tidak bias yang terbaik dari model regresi yang diperoleh dari metode kuadrat terkecil biasa. Dengan terpenuhinya asumsi tersebut, maka hasil yang diperoleh dapat lebih akurat dan mendekati atau sama dengan kenyataan. Uji asumsi klasik dalam regresi mencakup:

a. Uji autokorelasi

Uji autokorelasi bertujuan menguji apakah dalam model regresi linear ada korelasi antara error satu dengan error yang lainnya (Sukestiyarno, 2008: 14).

b. Uji Heteroskedastisitas

Heteroskedastisitas muncul apabila *error* atau residual dari model yang diamati tidak memiliki varian yang konstan dari satu observasi ke observasi lainnya. Konsekuensi adanya heteroskedastisitas dalam model linear adalah estimator yang diperoleh (Sukestiyarno, 2008: 14).

c. Uji multikolinearitas

Uji multikolinearitas bertujuan untuk menguji apakah dalam model regresi ditemukan adanya korelasi antara variabel bebas. Jadi uji multikolinearitas terjadi hanya pada regresi ganda. Model regresi yang baik seharusnya tidak terjadi korelasi tinggi diantara variabel bebas (Sukestiyarno, 2008: 14).

2.4. Multikolinearitas

2.4.1 Pengertian multikolinearitas

Istilah multikolinearitas mula-mula ditemukan oleh Ragnar Frisch. Pada mulanya multikolinearitas berarti adanya hubungan linear yang sempurna atau pasti, diantara beberapa atau semua variabel bebas dari model regresi ganda (Gujarati, 2004:157).

2.4.2 Penyebab terjadinya Multikolinearitas

Masalah multikolinearitas dapat timbul karena berbagai sebab. Pertama, karena sifat-sifat yang terkandung dalam kebanyakan variabel ekonomi berubah bersama-sama sepanjang waktu. Besaran-besaran ekonomi dipengaruhi oleh faktor-faktor yang sama. Oleh karena itu, sekali faktor-faktor yang mempengaruhi itu menjadi operatif, maka seluruh variabel akan cenderung berubah dalam satu arah. Dalam data *time series*, pertumbuhan dan faktor-faktor kecenderungan merupakan penyebab utama adanya multikolinearitas. Kedua, penggunaan nilai lag (*lagged values*) dari variabel-variabel bebas tertentu dalam model regresi.

Mengingat sifat yang sangat mendasar dari data, multikolinearitas diperkirakan terdapat pada sebagian besar hubungan-hubungan ekonomi. Oleh karena itu, perhatian sesungguhnya bukan lagi terletak pada ada atau tidaknya multikolinearitas, tetapi lebih pada akibat-akibat yang ditimbulkan oleh adanya multikolinearitas dalam sampel (Sumodiningrat; 1998: 281- 282).

2.4.3 Konsekuensi Multikolinearitas

Jika asumsi pada model regresi linear klasik terpenuhi, maka penaksir kuadrat terkecil/*Ordinary Least Square* (OLS) dari koefisien regresi linear adalah linear, tak bias dan mempunyai varian minimum dalam arti penaksir tersebut adalah penaksir tak bias kolinear terbaik/ *Best Linear Unbiased Estimator* (BLUE), meskipun multikolinearitas sangat tinggi, penaksir kuadrat terkecil biasa masih tetap memenuhi syarat BLUE, tetapi penaksir tersebut tidak stabil. (Gujarati, 2004:162).

Dalam hal terdapat multikolinearitas sempurna, penaksir dengan kuadrat terkecil dapat menjadi tak tentu dan variansi serta standar deviasinya menjadi tak terhingga. Sedangkan jika multikolinearitas tinggi, tetapi tidak sempurna maka konsekuensinya adalah sebagai berikut:

- a. Meskipun penaksir melalui kuadrat terkecil biasa didapatkan, standar deviasinya cenderung besar jika derajat kolinearitas antara peubah bertambah.
- b. Karena standar deviasi besar, internal kepercayaan bagi parameter populasi yang relevan akan menjadi besar.
- c. Taksiran-taksiran parameter kuadrat terkecil biasa dan standar deviasi akan menjadi sangat sensitif terhadap perubahan.
- d. Jika multikolinearitas tinggi, mungkin R^2 dapat tinggi namun tidak satu pun (sangat sedikit) taksiran koefisien regresi yang signifikan secara statistik (Sumodiningrat, 1998: 287).

2.4.4 Cara Mendeteksi Multikolinearitas

Ada beberapa cara untuk mengetahui keberadaan multikolinearitas dalam suatu model regresi

1. Menganalisis matriks korelasi

Jika antara dua atau lebih variabel independen memiliki korelasi yang cukup tinggi, biasanya di atas 0,9 maka hal tersebut mengindikasikan terjadinya multikolinearitas.

2. VIF (*Variance Inflation Factor*).

VIF (*Variance Inflation Factor*) adalah salah satu cara dalam mendeteksi adanya multikolinearitas, dan dalam penulisan ini menggunakan nilai *Tolerance* atau VIF (*Variance Inflation Factor*).

(2.4)

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

dengan R_j merupakan koefisien determinasi ke-j, $j=1, 2, \dots, k$

Multikolinearitas dalam sebuah regresi dapat diketahui apabila nilai $VIF \geq 10$.

3. TOL (*Tolerance*)

Jika nilai *Tolerance* kurang dari 0,1 atau nilai VIF melebihi 10 maka hal tersebut menunjukkan bahwa multikolinearitas adalah masalah yang pasti terjadi antar variabel bebas.

2.5. Metode Kuadrat Terkecil / *Ordinary Least Square (OLS)*

Metode Kuadrat Terkecil merupakan metode yang lebih banyak digunakan dalam pembentukan model regresi atau mengestimasi parameter-parameter regresi dibandingkan dengan metode-metode lain. Metode kuadrat terkecil adalah metode yang digunakan untuk mengestimasi nilai $\hat{\beta}$ dengan cara meminimumkan jumlah kuadrat kesalahan, $s = \sum e_i^2$. Dalam notasi matriks, sama dengan meminimumkan

$$e'e = [e_1 \quad e_2 \quad \dots \quad e_n] \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} = e_1^2 + e_2^2 + \dots + e_n^2 = \sum e_i^2$$

Dengan

$$e_i = y - X\hat{\beta}$$

Sehingga

$$S = e'e$$

$$\begin{aligned}
&= (y - X\hat{\beta})'(y - X\hat{\beta}) \\
&= (y' - \hat{\beta}'X')(y - X\hat{\beta}) \\
&= y'y - \hat{\beta}'X'y - y'X\hat{\beta} + \hat{\beta}'X'X\hat{\beta} \\
&= y'y - 2\hat{\beta}'X'y + \hat{\beta}'X'X\hat{\beta}
\end{aligned}$$

Karena $\hat{\beta}'X'y = (\hat{\beta}'X'y)' = y'X\hat{\beta}$

Jika S diturunkan secara parsial terhadap parameter $\hat{\beta}$ diperoleh:

$$\frac{\partial(e'e)}{\partial\hat{\beta}} = -2X'y + 2X'X\hat{\beta}$$

Estimasi nilai β diperoleh dengan meminimumkan $\frac{\partial(e'e)}{\partial\hat{\beta}}$ maka

$$-2X'y + 2X'X\hat{\beta} = 0$$

$$X'X\hat{\beta} = X'y$$

$$(X'X)^{-1}X'X\hat{\beta} = (X'X)^{-1}X'y$$

(2.5)

$$\hat{\beta} = (X'X)^{-1}X'y$$

Jadi estimasi OLS untuk β adalah $\hat{\beta} = (X'X)^{-1}X'y$

Sifat Estimasi Kuadrat Terkecil

Jika asumsi-asumsi dasar dipenuhi maka taksiran parameter yang dihasilkan

dengan menggunakan kuadrat terkecil yaitu:

$$\hat{\beta} = (X'X)^{-1}X'y$$

Akan bersifat BLUE (Best Linear Unbiased Estimator). Sifat BLUE

ini dapat dibuktikan sebagai berikut:

i. Linear

$$\begin{aligned}
 \hat{\beta} &= (X'X)^{-1}X'y \\
 &= (X'X)^{-1}X'(X\beta + \varepsilon) \\
 &= (X'X)^{-1}X'X\beta + (X'X)^{-1}X'\varepsilon \\
 &= (\beta + (X'X)^{-1})X'\varepsilon
 \end{aligned}$$

Merupakan fungsi linear dari β dan ε .

ii. Tak bias

Dengan $(X'X)^{-1}X'X = 1$

$$\begin{aligned}
 E(\hat{\beta}) &= E[(X'X)^{-1}X'Y] \\
 &= (X'X)^{-1}X'E(Y) \\
 &= (X'X)^{-1}X'(X\beta) \\
 &= (X'X)^{-1}X'(X\beta) \\
 &= 1\beta \\
 &= \beta
 \end{aligned}$$

Jadi $E(\hat{\beta}) = \beta$ maka $\hat{\beta}$ adalah estimator yang merupakan penaksir tak bias dari β

iii. Variansi Minimum

$$\begin{aligned}
 var(\hat{\beta}) &= (X'X)^{-1}X'var(Y)X(X'X)^{-1} \\
 &= (X'X)^{-1}X'\sigma^2IX(X'X)^{-1} \\
 &= \sigma^2(X'X)^{-1}
 \end{aligned}$$

Bahwa $var(\hat{\beta}) = \sigma^2(x'x)^{-1}$ merupakan variansi terkecil dari semua penaksir linear tak bias.

Karena estimator kuadrat terkecil memenuhi sifat linear, tak bias dan mempunyai variansi minimum maka estimator kuadrat terkecil disebut bersifat BLUE (*Best Linear Unbiased Estimator*).

2.6. Nilai Eigen dan Vektor Eigen

Jika A adalah matriks $n \times n$, terdapat suatu skalar λ vektor tak nol V sehingga memenuhi persamaan berikut :

$$AV = \lambda V$$

λ adalah nilai eigen dari A dan vektor V disebut vektor eigen yang berkaitan dengan nilai eigen (λ).

$$AV = \lambda V; V \neq 0$$

$$AV = \lambda IV$$

$$(\lambda I - A) = 0$$

Supaya λ menjadi nilai eigen, maka harus ada pemecahan tak nol, Persamaan tersebut akan memiliki penyelesaian tak nol jika dan hanya jika:

$$\det(\lambda I - A) = 0 \text{ atau } |\lambda I - A| = 0$$

Nilai karakteristik λ merupakan akar polynomial derajat n . Jika $|\lambda I - A| = 0$ dengan :

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}; \lambda I = \begin{bmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{bmatrix}$$

Maka

$$\begin{aligned}
 |\lambda I - A| &= \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix} \\
 &= \begin{vmatrix} \lambda - a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & \lambda - a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix} = 0
 \end{aligned}$$

Sehingga diperoleh persamaan berikut:

$$\lambda^n + (-1)^1 M_1 \lambda^{n-1} + (-1)^2 M_2 \lambda^{n-2} + \cdots + (-1)^n M_n = 0$$

Dengan M_i adalah penjumlahan minor orde ke-I disekitar diagonal utama.

2.7. *Principal Component Analisis (PCA)*

Metode PCA bertujuan untuk menyederhanakan variabel yang diamati dengan cara mereduksi dimensinya. Hal ini dilakukan dengan cara menghilangkan korelasi diantara variabel bebas melalui transformasi variabel bebas asal ke variabel baru yang tidak berkorelasi sama sekali. Setelah beberapa komponen hasil PCA yang bebas multikolinearitas diperoleh, maka komponen-komponen tersebut menjadi variabel bebas baru yang akan diregresikan atau dianalisis pengaruhnya terhadap variabel tak bebas (Y) dengan menggunakan analisis regresi.

2.8. *Principal Component Regression (PCR)*

Principal Component Regression (PCR) merupakan salah satu metode yang telah dikembangkan untuk mengatasi masalah multikolinearitas. PCR merupakan analisis regresi dari variabel-variabel dependen terhadap komponen-komponen

utama yang tidak saling berkorelasi, dimana setiap komponen utama merupakan kombinasi linear dari semua variabel independen (Draper & Smith 1992, hal 313).

Regresi komponen utama membentuk hubungan antara variabel terikat dengan komponen utama yang dipilih dari variabel bebas (Ul-Saufie *et al.* 2011).

Principal Component Regression (PCR) merupakan suatu teknik analisis yang mengkombinasikan antara analisis regresi dengan *Principal Component Analysis* (PCA). Analisis Regresi digunakan untuk mengetahui ada tidaknya hubungan antara variabel dependen dan independen, sedangkan PCA pada dasarnya bertujuan untuk menyederhanakan variabel yang diamati dengan cara menyusutkan (mereduksi) dimensinya. Hal ini dilakukan dengan jalan menghilangkan korelasi di antara variabel melalui transformasi variabel asal ke variabel baru (merupakan kombinasi linear dari variabel-variabel asal) yang tidak saling berkorelasi. Dari p buah variabel asal dapat dibentuk p buah komponen utama, dipilih k buah komponen utama saja ($k < p$) yang telah mampu menerangkan keragaman data cukup tinggi (antara 80% sampai dengan 90%) (Johnson & Wichern, 2010, hal. 356). Komponen utama yang dipilih tersebut (k buah) dapat mengganti p buah variabel asal tanpa mengurangi informasi.

Cara pembentukan regresi komponen utama melalui analisis komponen utama ada dua cara yaitu komponen utama yang dibentuk berdasarkan matriks kovariansi dan komponen utama yang dibentuk berdasarkan matriks korelasi. Matriks korelasi dari data yang telah distandarisasi (bentuk baku Z) digunakan jika variabel yang diamati tidak memiliki satuan pengukuran yang sama. Sedangkan

Matriks varians kovarians digunakan jika semua variabel yang diamati mempunyai satuan pengukuran yang sama.

Analisis regresi komponen utama (PCR) merupakan analisis regresi variabel dependen terhadap komponen-komponen utama yang tidak saling berkorelasi, regresi komponen utama dapat dinyatakan sebagai berikut :

(2.6)

$$Y = w_0 + w_1K_1 + w_2K_2 + \dots + w_mK_m + \varepsilon$$

Dimana:

Y : variabel dependen

K : Komponen utama

W : Parameter regresi komponen utama.

$K_1, K_2, K_3, \dots, K_m$ menunjukkan komponen utama yang dilibatkan dalam analisis regresi komponen utama, dimana besaran m lebih kecil daripada banyaknya variabel independen yaitu sejumlah p , serta Y sebagai variabel dependen.

Komponen utama merupakan kombinasi linear dari variabel baku Z , sehingga:

(2.7)

$$K_1 = a_{11}Z_1 + a_{21}Z_2 + \dots + a_{p1}Z_p$$

$$K_2 = a_{12}Z_1 + a_{22}Z_2 + \dots + a_{p2}Z_p$$

⋮

$$K_m = a_{1m}Z_1 + a_{2m}Z_2 + \dots + a_{pm}Z_p$$

Apabila K_1, K_2, \dots, K_m dalam persamaan (2.7) didistribusikan kembali ke dalam persamaan regresi komponen utama, yaitu persamaan (2.6) maka diperoleh:

$$\begin{aligned}
Y &= w_0 + w_1(a_{11}Z_1 + a_{21}Z_2 + \cdots + a_{p1}Z_p) \\
&\quad + w_2(a_{12}Z_1 + a_{22}Z_2 + \cdots + a_{p2}Z_p) + \cdots \\
&\quad + w_m(a_{1m}Z_1 + a_{2m}Z_2 + \cdots + a_{pm}Z_p) + \varepsilon \\
&= w_0 + w_1a_{11}Z_1 + w_1a_{21}Z_2 + \cdots + w_1a_{p1}Z_p + w_2a_{12}Z_1 + w_2a_{22}Z_2 \\
&\quad + \cdots + w_2a_{p2}Z_p + w_ma_{1m}Z_1 + w_ma_{2m}Z_2 + \cdots \\
&\quad + w_ma_{pm}Z_p + \varepsilon \\
(2.8) \quad &= w_0 + (w_1a_{11} + w_2a_{12} + \cdots + w_ma_{1m})Z_1 + (w_1a_{21} + w_2a_{22} + \cdots \\
&\quad + w_ma_{2m})Z_2 + \cdots + (w_1a_{p1} + w_2a_{p2} + \cdots + w_ma_{pm})Z_p \\
&\quad + \varepsilon
\end{aligned}$$

Sehingga dari persamaan (2.8) diperoleh persamaan regresi dengan komponen utama sebagai berikut:

$$\begin{aligned}
(2.9) \quad Y &= b_0 + b_1Z_1 + b_2Z_2 + \cdots + b_pZ_p
\end{aligned}$$

Dengan:

$$\begin{aligned}
b_0 &= w_0 \\
b_1 &= w_1a_{11} + w_2a_{12} + \cdots + w_ma_{1m} \\
b_2 &= w_1a_{21} + w_2a_{22} + \cdots + w_ma_{2m} \\
&\vdots \\
b_p &= w_1a_{p1} + w_2a_{p2} + \cdots + w_ma_{pm}
\end{aligned}$$

Adapun algoritma *Principal Component Regression* (PCR) sebagai berikut:

1. Menghitung *eigen value* dan *eigen vector* dari matriks korelasi atau kovarians
2. Terdapat p komponen utama yang orthogonal dan tidak berkorelasi
3. Dipilih komponen yang *eigen value* > 1 atau yang mampu menerangkan keragaman cukup tinggi (80%-90%)
4. Regresi variabel dependen dengan komponen-komponen utama yang terpilih

2.9. *Partial Least Square* (PLS)

Metode *Partial Least Square* (PLS) pertama kali diperkenalkan oleh Herman Ole Andres Wold pada tahun 1960 sebagai metode alternative untuk mengatasi keterbatasan metode *Ordinary Least Square* (OLS) ketika data mengalami masalah multikolinearitas. Untuk meregresikan variabel terikat y dengan variabel x_1, x_2, \dots, x_k , metode PLS mencari komponen-komponen baru yang berperan sebagai variabel bebas untuk mengestimasi parameter regresi.

Regresi PLS ini dapat diperoleh melalui regresi sederhana maupun berganda dengan mengambil kesimpulan dari uji signifikansi. Uji signifikansi ini bertujuan untuk memilih variabel independen pembangun komponen PLS dan menentukan banyaknya komponen PLS yang terbentuk (Vinzi, Bastien, & Tenenhaus, 2004). Tujuan PLS adalah membentuk komponen yang dapat menangkap informasi dari variabel independen untuk memprediksi variabel dependen. Dalam pembentukan komponen PLS, digunakan variabel dependen y yang distandarisasi dan variabel-

variabel independen yang terpusat (Vinzi, Bastien, & Tenenhaus, 2004). Model regresi *partial least square* dengan m komponen dapat dituliskan sebagai berikut:

(2.11)

$$Y = \sum_{h=1}^m c_h t_h + \varepsilon$$

Dengan:

Y : variabel dependen

c_h : koefisien regresi Y terhadap t_h

$t_h = \sum_{j=1}^p w_{(h)j} X_j$: komponen utama ke- h yang tidak saling berkorelasi,
($h=1,2,\dots,m$)

Dengan syarat komponen PLS $t_h = \sum_{j=1}^p w_{(h)j} X_j$ orthogonal, sehingga parameter c_h dan w_h dalam persamaan (2.11) dapat diestimasi.

1. Penghitungan Komponen PLS pertama t_1

Komponen PLS pertama (t_1) adalah linear dari variabel independen X_j dengan koefisien pembobot w_1 . Persamaan komponen utama pertama dapat dituliskan sebagai berikut:

(2.12)

$$t_h = \sum_{j=1}^p w_{(h)j} X_j = w_{11}X_1 + w_{12}X_2 + \dots + w_{1p}X_p = Xw_1$$

Dengan:

t_1 : komponen PLS pertama

X_j : matriks variabel independen

w_1 :vektor koefisien bobot untuk variabel X pada komponen utama pertama.

Misalkan a_1, a_2, \dots, a_p sebagai koefisien regresi dari masing-masing variabel terpusat x_1, x_2, \dots, x_p terhadap y . komponen pertama $t_1 = Xw_1$ yang di definisikan sebagai berikut:

(2.13)

$$t_1 = \frac{1}{\sqrt{\sum_{j=1}^p a_{1j}^2}} \sum_{j=1}^p a_{1j} x_j$$

Dengan

(2.14)

$$\begin{aligned} a_{1j} &= \frac{cov(x_j, y)}{var(x_j)} \\ &= \frac{E\left(\frac{x_j y}{var(x_j)} - E(y)E\left(\frac{x_j}{var(x_j)}\right)\right)}{E\left(\frac{x_j y}{var(x_j)}\right)^2 - \left(E\left(\frac{x_j}{var(x_j)}\right)\right)^2} \\ &= \frac{\frac{1}{var(x_j)}(E(x_j y) - E(y)E(x_j))}{\left(\frac{1}{var(x_j)}\right)^2 (E(x_j)^2 - (E(x_j))^2)} \\ &= \frac{\frac{1}{var(x_j)}(E(x_j y) - E(y)E(x_j))}{\left(\frac{1}{var(x_j)}\right)^2 (var(x_j))} \end{aligned}$$

$$\begin{aligned}
&= \frac{\frac{1}{\text{var}(x_j)} (E(x_j, y) - E(y)E(x_j))}{\frac{1}{\text{var}(x_j)}} \\
&= (E(x_j, y) - E(y)E(x_j)) = \text{cov}(x_j, y)
\end{aligned}$$

Jika a_{1j} pada persamaan (2.14) disubstitusikan pada persamaan (2.13) maka persamaa tersebut dituliskan menjadi sebagai berikut:

$$\begin{aligned}
t_1 &= \frac{1}{\sqrt{\sum_{j=1}^p \left[\frac{\text{cov}(x_j, y)}{\text{var}(x_j)} \right]^2}} \sum_{j=1}^p \left[\frac{\text{cov}(x_j, y)}{\text{var}(x_j)} \right] x_j \\
&= \frac{1}{\sqrt{\sum_{j=1}^p \left[\frac{\text{cov}(x_j, y)}{\text{var}(x_j)} \right]^2}} \sum_{j=1}^p \frac{\text{cov}(x_j, y)}{\sqrt{\text{var}(x_j)}} \frac{x_j}{\sqrt{\text{var}(x_j)}} \\
&= \frac{1}{\sqrt{\text{var}(y)}} \frac{1}{\sqrt{\sum_{j=1}^p \left[\frac{\text{cov}(x_j, y)}{\text{var}(x_j)} \right]^2}} \sum_{j=1}^p \frac{\text{cov}(x_j, y)}{\sqrt{\text{var}(x_j)}} \frac{x_j}{\sqrt{\text{var}(x_j)}} \\
(2.15) \quad &= \frac{1}{\sqrt{\sum_{j=1}^p \left[\frac{\text{cov}(x_j, y)}{\text{var}(x_j)} \right]^2}} \sum_{j=1}^p \frac{\text{cov}(x_j, y)}{\sqrt{\text{var}(y)} \sqrt{\text{var}(x_j)}} \frac{x_j}{\sqrt{\text{var}(x_j)}}
\end{aligned}$$

Karena

$$\text{cov}(x_j, y) = \frac{\text{cov}(x_j, y)}{\sqrt{\text{var}(y)} \sqrt{\text{var}(x_j)}}$$

sehingga persamaan (2.15) dapat ditulis sebagai berikut:

(2.16)

$$t_1 = \frac{1}{\sqrt{\sum_{j=1}^p \text{cor}(x_j, y)^2}} \sum_{j=1}^p \text{cor}(x_j, y) x_j^*$$

Dimana:

 x_j^* : x_j yang terstandarisasi $\text{cor}(x_j, y)$: korelasi variabel x_j dengan y $\text{cov}(x_j, y)$: kovarians variabel x_j dengan y var : varians.keragaman

Variabel x_j dipilih yang berkorelasi tinggi dengan variabel dependen y , sehingga variabel x_j menjadi penting dalam pembentukan komponen t_1 . Nilai $\text{cov}(x_j, y)$ juga merupakan koefisien regresi a_{1j} dalam regresi sederhana antara y dengan variabel x_j modifikasi, yaitu $\frac{x_j}{\text{var}(x_j)}$. Regresi sederhana y dengan variabel

 x_j :

$$Y = a_{1j}x_j + \text{residu}$$

$$Y = a_{1j} \left(\frac{x_j}{\text{var}(x_j)} \right) + \text{residu}$$

Dengan:

$$a_{1j} = \frac{\text{cov} \left(\frac{y, x_j}{\text{var}(x_j)} \right)}{\text{var} \left(\frac{x_j}{\text{var}(x_j)} \right)} = \text{cov}(y, x_j)$$

Jika a_{1j} tidak signifikan maka dalam persamaan (2.15) setiap kovariansi yang tidak signifikan dapat diganti dengan nol dan artinya hubungan variabel independennya dapat diabaikan.

2. Perhitungan komponen PLS kedua t_2

Komponen PLS kedua didapatkan dengan melakukan regresi sederhana y terhadap t_1 dan masing-masing x_j terlebih dahulu kemudian regresi antara x_j terhadap t_1 . Variabel-variabel x_j yang digunakan hanya variabel yang berkontribusi secara signifikan dalam menjelaskan y pada t_1 . Model persamaan regresi keduanya adalah sebagai berikut:

(2.17)

$$y = c_1 t_1 + a_{2j} x_j + \text{residu}$$

(2.18)

$$x_j = p_{ij} t_1 + x_{1j}$$

Persamaan (2.18) didistribusikan ke (2.17) sehingga persamaan (2.17) dapat dituliskan sebagai berikut:

$$y = c_1 t_1 + a_{2j} (p_{ij} t_1 + x_{1j}) + \text{residu}$$

$$= c_1 (a_{2j} p_{ij}) t_1 + a_{2j} x_{1j} + \text{residu}$$

$$= c'_1 t_1 + a_{2j} x_{1j} + \text{residu}$$

Dengan

$$c'_1 = c_1 + a_{2j} p_{ij}$$

Dengan x_{1j} adalah residu yang dihasilkan dari regresi x_j terhadap t_1 . Maka komponen PLS kedua (t_2) dapat didefinisikan sebagai berikut:

$$\begin{aligned}
 t_2 &= \frac{1}{\sqrt{\sum_{j=1}^p a_{2j}^2}} \sum_{j=1}^p a_{2j} x_{1j} \\
 &= \frac{1}{\sqrt{\sum_{j=1}^p \text{cov}(y, x_{1j})^2}} \sum_{j=1}^p \text{cov}(y, x_{1j}) x_{1j}
 \end{aligned}$$

(2.19)

$$= \frac{1}{\sqrt{\sum_{j=1}^p \text{cor}(y, x_{1j})^2}} \sum_{j=1}^p \text{cor}(y, x_{1j}) x_{1j}^*$$

Dengan x_{1j}^* adalah residu yang telah distandarisasi dan dihasilkan dari regresi x_j terhadap t_1 . Komponen PLS t_2 ini tidak saling berkorelasi atau orthogonal dengan komponen PLS yang lain.

Nilai $\text{cov}(x_{1j}, y)$ juga merupakan koefisien regresi a_{2j} dengan regresi y pada t_1 dan variabel x_{1j} modifikasi, yaitu $\frac{x_{1j}}{\text{var}(x_{1j})}$. Regresi sederhana y dengan variabel x_{1j}

dituliskan sebagai berikut:

(2.20)

$$y = c_1 t_1 + a_{2j} \left(\frac{x_{1j}}{\text{var}(x_{1j})} \right) + \text{residu}$$

Korelasi parsial antara y dan x_j diketahui t_1 didefinisikan sebagai korelasi y dan residu x_{1j} . Karena dalam perhitungan komponen PLS $\text{cov}(y, x_{1j}) = \text{cor}(y, x_{1j})$, maka korelasi parsial antara y dan x_j yang dinyatakan dalam t_1 dapat dituliskan sebagai berikut:

$$\text{cor}(y, x_j | t_1) = \text{cor}(y, x_{1j}) \text{ atau } \text{cov}(y, x_j | t_1) = \text{cov}(y, x_{1j})$$

Oleh karena itu, komponen PLS kedua pada persamaan (2.19) dapat dituliskan sebagai berikut:

$$t_2 = \frac{1}{\sqrt{\sum_{j=1}^p \text{cor}(y, x_j | t_1)^2}} \sum_{j=1}^p \text{cor}(y, x_j | t_1) x_{1j}$$

3. Perhitungan Komponen PLS ke- h , t_h

Seperti langkah pada pembentukan komponen PLS sebelumnya, variabel yang digunakan adalah variabel-variabel yang signifikan dalam menjelaskan y pada t_1, t_2, \dots, t_{h-1} . Model regresi y terhadap t_1, t_2, \dots, t_{h-1} dan masing-masing x_j adalah sebagai berikut:

(2.21)

$$y = c_1 t_1 + c_2 t_2 + \dots + c_{h-1} t_{h-1} + a_{hj} x_j + \text{residu}$$

Untuk mendapatkan komponen t_h yang orthogonal terhadap t_{h-1} , diregresikan x_j terhadap komponen PLS yang dituliskan sebagai berikut:

(2.22)

$$x_j = p_{1j} t_1 + p_{2j} t_2 + \dots + p_{h-1j} t_{h-1} + x_{(h-1)j}$$

Dengan $x_{(h-1)j}$ adalah residu yang dihasilkan dari regresi setiap x_j terhadap t_1, t_2, \dots, t_{h-1} .

Komponen ke- h didefinisikan sebagai berikut:

$$t_h = \frac{1}{\sqrt{\sum_{j=1}^p a_{hj}^2}} \sum_{j=1}^p a_{hj} x_{(h-1)j}$$

Dengan a_{hj} adalah koefisien regresi dari $x_{(h-1)j}$ dalam regresi y pada t_1, t_2, \dots, t_{h-1} . Jika persamaan (2.22) disubstitusikan ke dalam persamaan (2.21), maka diperoleh:

$$\begin{aligned}
y &= c_1 t_1 + c_2 t_2 + \dots + c_{h-1} t_{h-1} \\
&\quad + a_{hj} (p_{1j} t_1 + p_{2j} t_2 + \dots + p_{h-1j} t_{h-1} + x_{(h-1)j}) + \text{residu} \\
&= (c_1 a_{hj} p_{1j}) + (c_2 a_{hj} p_{2j}) t_2 + \dots + (c_{h-1} a_{hj} p_{h-1j}) t_{h-1} + a_{hj} x_{(h-1)j} \\
&\quad + \text{residu} \\
&= c'_1 t_1 + c'_2 t_2 + \dots + c'_{h-1} t_{h-1} + a_{hj} x_{(h-1)j} + \text{residu}
\end{aligned}$$

Dimana $c'_{h-1} = c_{h-1} a_{hj} p_{h-1j}$

Sehingga komponen PLS ke- h dapat ditulis sebagai berikut:

$$(2.23) \quad t_h = \frac{1}{\sqrt{\sum_{j=1}^p \text{cor}(x_{(h-1)j}, y)^2}} \sum_{j=1}^p \text{cor}(x_{(h-1)j}, y) x_{(h-1)j}^*$$

Dimana $x_{(h-1)j}^*$ adalah residu standar dari regresi dan setiap x_j terhadap t_1, t_2, \dots, t_{h-1} . Perhitungan komponen PLS berhenti ketika tidak ada lagi variabel independen yang signifikan membangun komponen PLS.

4. Tranformasi Komponen PLS ke Variabel Asli

Persamaan (2.11) selanjutnya dapat ditulis ke dalam bentuk variabel aslinya, yaitu:

$$\begin{aligned}
Y &= \sum_{h=1}^m c_h t_h + \varepsilon \\
&= \sum_{h=1}^m c_h \left(\sum_{j=1}^p w_{(h)j} X_j \right) + \varepsilon \\
&= \sum_{j=1}^p \sum_{h=1}^m c_h w_{(h)j} X_j + \varepsilon
\end{aligned}$$

$$= \sum_{j=1}^p \left(\sum_{h=1}^m c_h w_{(h)j} \right) X_j + \varepsilon$$

$$Y = \sum_{j=1}^p b_j X_j + \varepsilon$$

Dimana:

Y : variabel dependen

c_h : koefisien regresi Y terhadap t_h

X_j : matriks variabel independen

$t_h = \sum_{j=1}^p w_{(h)j} X_j$: komponen utama ke- h yang tidak saling berkorelasi,
($h=1, 2, \dots, m$)

$b_j = \sum_{h=1}^m c_h w_{(h)j}$: koefisien bobot untuk variabel X_j pada komponen utama PL ke- h

ε : vektor error.

Adapun algoritma *Partial Least Square (PLS)* sebagai berikut:

1. Melakukan regresi y terhadap masing-masing x_j dan komponen ke- $(h-1)$
2. Uji signifikansi masing-masing x_j
3. Jika x_j signifikan
 - a. hitung $cor(x_{(h-1)j}, y)$ atau $cov(x_{(h-1)j}, y)$
 - b. Pembentukan komponen PLS ke- h ulangi langkah pertama sampai ketemu semua komponen PLS
4. Jika x_j tidak signifikan

Variabel tidak digunakan sebagai pembentuk komponen PLS ke- h

5. Regresi y terhadap komponen-komponen PLS yang terbentuk

2.10. Pendapatan Daerah

Pengertian pendapatan daerah berdasarkan Undang-Undang Nomor 32 Tahun 2004 tentang pemerintahan daerah pasal 1 angka 15 adalah semua hak daerah yang diakui sebagai penambah nilai kekayaan bersih dalam periode tahun anggaran bersangkutan. Pendapatan daerah berasal dari Pendapatan Asli Daerah (PAD), Dana Perimbangan serta lain-lain pendapatan yang sah.

1. Pendapatan Asli Daerah (PAD)

Berdasarkan Undang-Undang Nomor 33 Tahun 2004 tentang Perimbangan Keuangan Antara Pusat dan Daerah pasal 1 angka 18 Pendapatan Asli Daerah (PAD) merupakan pendapatan yang diperoleh daerah yang dipungut berdasarkan peraturan daerah sesuai dengan peraturan perundang-undangan. Pendapatan asli daerah berdasarkan pasal 6 Undang-Undang No.33 Tahun 2004 bersumber dari:

a. Pajak daerah

Hasil pajak daerah merupakan pungutan yang dilakukan oleh pemerintah daerah yang bersifat wajib atau memaksa berdasarkan peraturan perundangundangan yang berlaku ditetapkan melalui peraturan daerah. Pungutan ini dikenakan kepada semua objek seperti orang atau badan dan benda bergerak atau tidak bergerak, misalnya seperti pajak hotel, pajak

restoran, pajak reklame, pajak parkir, dan pajak hiburan yang nantinya akan digunakan untuk keperluan daerah dan kemakmuran rakyat.

b. Retribusi daerah

Pungutan daerah sebagai pembayaran atas jasa atau izin yang diberikan oleh pemerintah daerah untuk kepentingan pribadi atau badan. Dengan kata lain retribusi daerah adalah pungutan yang dilakukan sehubungan dengan suatu jasa atau fasilitas yang diberikan secara nyata dan langsung, seperti pelayanan kesehatan, pelayanan kebersihan, dan pelayanan pemakaman.

c. Hasil pengelolaan kekayaan daerah yang dipisahkan

Penerimaan daerah yang berasal dari hasil perusahaan milik daerah dan pengelolaan kekayaan daerah yang dipisahkan. Penerimaan ini antara lain dari BPD, perusahaan daerah, dividen BPR-BKK dan penyertaan modal daerah kepada pihak ketiga.

d. Lain-lain PAD yang sah

Penerimaan daerah yang berasal dari lain-lain milik pemerintah daerah, seperti hasil penjualan asset daerah yang tidak dipisahkan, jasa giro, dan pendapatan bunga.

2. Dana Perimbangan

Dana perimbangan bertujuan untuk menciptakan keseimbangan keuangan antara pemerintah pusat dan pemerintah daerah. Dana perimbangan yang terdiri dari Dana Bagi Hasil dari penerimaan pajak dan SDA, Dana Alokasi Umum, dan Dana Alokasi Khusus merupakan sumber

pendanaan bagi daerah dalam pelaksanaan desentralisasi, yang alokasinya tidak dapat dipisahkan satu dengan yang lain mengingat tujuan masing-masing jenis penerimaan tersebut saling mengisi dan melengkapi.

3. Lain-lain pendapatan yang sah

UU no 32/2004 maupun UU No 33/2004 tidak ada pasal yang secara tegas menetapkan aturan Dana Perimbangan dari Pemerintah Provinsi untuk Pemerintah Kabupaten/Kota. Hal yang mendasari adalah Peraturan Daerah yang dibenarkan dalam ke Undang Undang tersebut untuk mengatur adanya Dana Perimbangan, Hibah, Dana Darurat, Bagi Hasil Pajak dari Provinsi dan Pemerintah Daerah lainnya, Dana Penyesuaian dan Otonomi Khusus serta lain-lain Pendapatan Daerah yang sah.

Untuk memberi dukungan terhadap pelaksanaan otonomi daerah telah diterbitkan UU no 33/2004 tentang Perimbangan Keuangan antara Pemerintah Pusat dan Daerah. Sumber pembiayaan pemerintah daerah didalam rangka perimbangan keuangan pemerintah pusat dan daerah dilaksanakan atas dasar desentralisasi, dekonsentrasi, dan pembantuan. Berkaitan dengan perimbangan keuangan antara pemerintah pusat dan daerah, hal tersebut merupakan konsekuensi adanya penyerahan kewenangan pemerintah pusat kepada pemerintah daerah. Dengan demikian, terjadi transfer yang cukup signifikan didalam APBN dari pemerintah pusat ke pemerintah daerah, dan pemerintah daerah secara leluasa dapat menggunakan dana ini apakah untuk memberi pelayanan yang

lebih baik kepada masyarakat atau untuk keperluan lain yang tidak penting.
(Yulian & Bagio, 2015).

2.11. *Statistical Analysis System (SAS)*

SAS adalah program komputer untuk analisis statistika yang dikembangkan oleh perusahaan *SAS Institute*. Perangkat lunak ini dirancang untuk keperluan berbagai bidang dengan fitur: (1) Analisis statistika, (2) riset operasi, (3) manajemen proyek, dll. SAS diciptakan pada awal 1970-an oleh Jim Goodnight dan John Sall beserta koleganya di *North Caroline State University*, Amerika Serikat, untuk manajemen dan analisis data hasil percobaan di bidang pertanian. Pada tahun 1976, *SAS Institute* didirikan sebagai pusat pengembangan software tersebut. Ketika usaha semakin berkembang menjadi vendor *software* dan jasa analisis bisnis berskala internasional, perusahaan menghapus kata '*Institute*' dan hingga kini namanya adalah 'SAS' yang bukan akronim dari apapun. Logo SAS '*THE POWER TO KNOW*' diperkenalkan pada tahun 2000.

SAS telah merilis berbagai versi dan seri. Versi pertama yang diperkenalkan adalah 71 yang merupakan versi pertama SAS yang hanya dapat digunakan untuk analisis regresi dan ANOVA. Pada tahun 1976, *SAS Institute* merilis *Base SAS*[®]. *SAS Institute* terus melakukan inovasi dengan terus mengeluarkan SAS versi terbaru. Pada tahun 2008 SAS[®]9.2 dirilis dan memperoleh berbagai penghargaan internasional, di antaranya sebagai "*Business Intelligence Vendor of the Year in Asia Pasific*" dan "*TOP Five Companies for IT*". Satu tahun berikutnya, SAS

termasuk dalam daftar 20 perusahaan “*Best Companies to Work For*” versi majalah *FORTUNE*.

SAS lebih dari sekedar *software* analisis statistika. SAS merupakan sebuah sistem terintegrasi yang mampu memberikan solusi bagi industri, bisnis, maupun akademik. Salah satu produk dari SAS adalah SAS[®] *AnalyticsPro*. Paket SAS[®] *Analytics Pro* dikhususkan untuk keperluan analisis statistika, seperti akses, manipulasi dan analisis data serta penyajian informasi. SAS[®] *Analytics Pro* merupakan gabungan dari tiga *software* yaitu Base SAS[®], SAS/STAT[®] dan SAS/ GRAP[®] (Saefuddin, Sartono, & Setiabudi. 2010).

1. Base SAS[®]

Base SAS[®] merupakan fondasi SAS yang harus diinstal apabila akan menggunakan *software* SAS yang lain. Base SAS[®] sudah menggunakan bahasa pemrograman generasi keempat serta dilengkapi dengan program-program SAS siap pakai dan fasilitas *Macro* yang berguna untuk mengurangi kerumitan bahasa pemrograman. Base SAS dirancang untuk keperluan integrasi data, analisis data, pelaporan dan pemrograman SAS.

2. SAS/STAT[®]

SAS/STAT[®] dirancang untuk analisis statistika sederhana maupun kompleks, diantaranya ANOVA, modeling, regresi, analisis data kategorik, analisis Bayesian, analisis peubah ganda, analisis daya tahan (*survival*), analisis *nonparametric*, analisis psikometrik, analisis data survey, dan lain-lain.

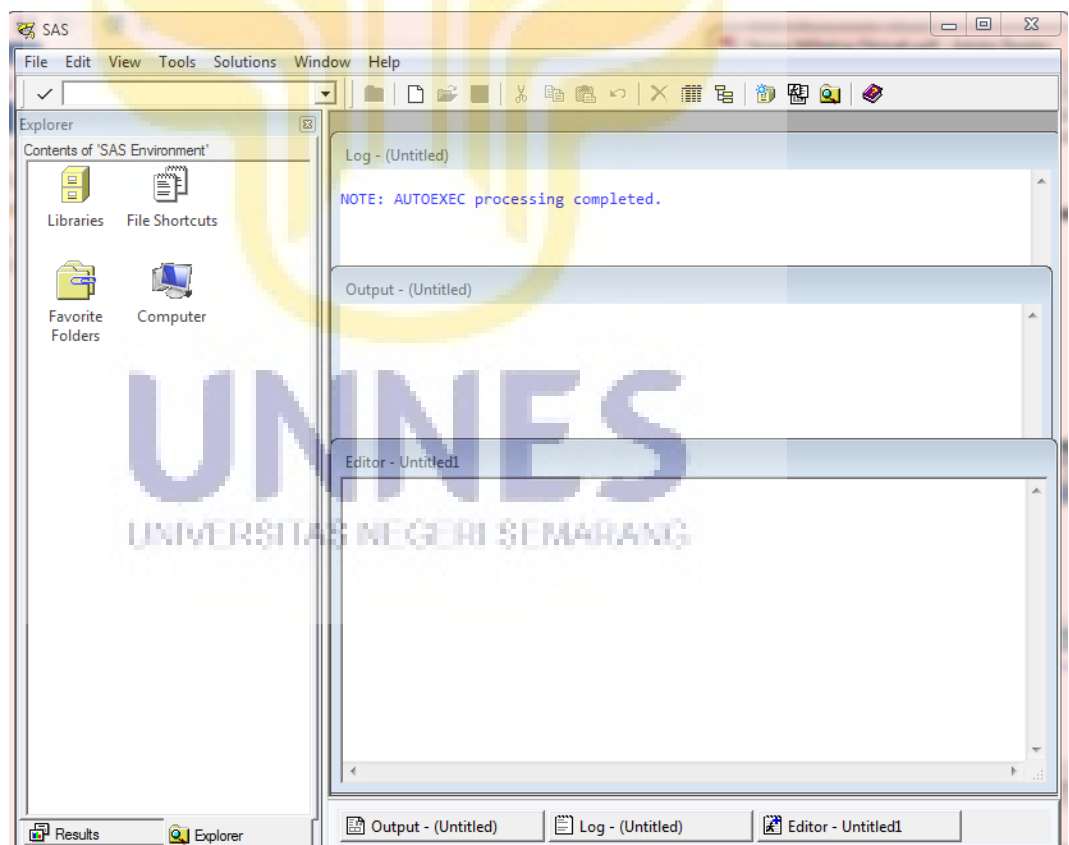
3. SAS/GRAPH®

SAS/GRAPH® merupakan *software* SAS untuk visualisasi hasil analisis data yang mengesankan dalam bentuk dua dimensi. SAS/GRAPH® menyediakan fasilitas untuk membuat berbagai macam diagram, plot dan peta.

2.9.1. Menjalankan SAS

Ada beberapa cara untuk memulai menjalankan SAS pada sistem operasi *Windows*, antara lain:

1. Klik ganda ikon SAS pada desktop, atau
2. Dari *Start menu* di desktop, Klik All Program ® SAS ® SAS 9.1 for Windows.



Gambar 2.1 Tampilan Awal Program SAS

Setelah SAS dijalankan, ada lima bagian utama pada window SAS, yaitu Program Editor, *Log*, *Output*, *Result* dan *Explorer*. Fungsi kelima windows tersebut adalah sebagai berikut:

a. Program Editor

Program editor atau *editor window* berguna untuk membuka dan mengedit program SAS yang telah ditulis sebelumnya, menuliskan program SAS baru, menjalankan program SAS, menyimpan program SAS ke dalam *file*, maupun mencetak program SAS.

Menjalankan program yang ditulis pada editor window dapat dilakukan dengan tiga cara, yaitu dengan:

- 1) Perintah ***Submit*** dari menu ***Run***,
- 2) Klik ikon pada *toolbar*, atau
- 3) Tekan tombol F8 pada *Keyboard*

Jika program dapat dibaca dan dijalankan, SAS akan menampilkan *log window*. Selain itu, jika program SAS yang dilakukan mengandung perintah untuk menghasilkan *output* tertentu, SAS juga akan menampilkan *output window*.

b. Log

Log window menampilkan pernyataan SAS (log) yang telah dijalankan, dan informasi mengenai program SAS yang dijalankan, termasuk peringatan dan pesan kesalahan (*error & warning*).

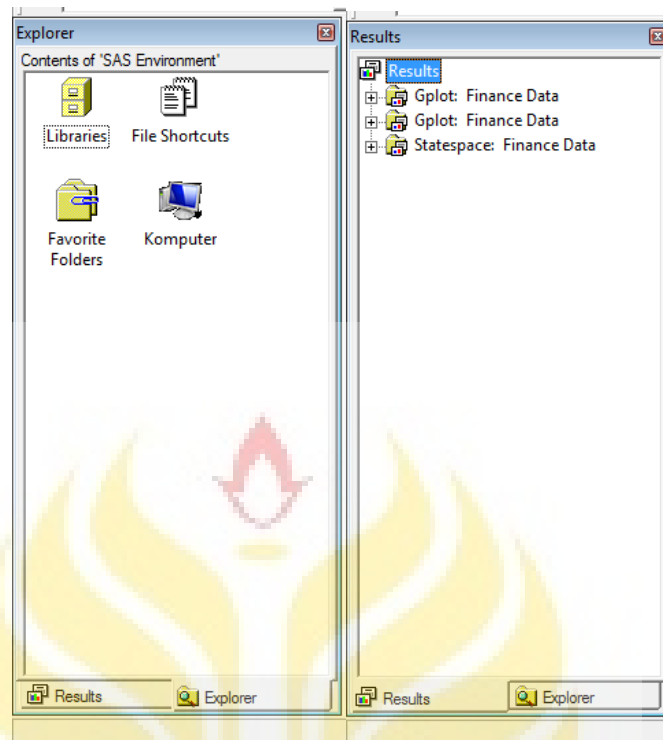
c. *Output*

Output window menampilkan keluaran dari program yang telah dijalankan. Secara *default*, *output SAS* berupa teks akan ditampilkan dalam format *listing*. SAS juga mengeluarkan *window* lain untuk menampilkan grafik ketika prosedur SAS grafik dijalankan. Isi dari log dan *output window* tidak dapat diedit. Namun *log window* dapat dikosongkan melalui perintah **Clear** dari menu **Edit**. *Output SAS* juga dapat ditampilkan dalam format HTML dengan cara klik **Tools** → **Option** → **Preferences**. Pada kotak dialog **Preferences**, aktifitas tab **Results**, dan cek pada **Create HTML**.

Apabila prosedur SAS memuat perintah untuk menampilkan grafik (diagram, plot, peta dan lain-lain), SAS akan mengeluarkan *window* baru yaitu *GRAPH viewer*.

d. *Result dan Explorer Window*

Result window menampilkan indeks dari *output SAS*. Sedangkan *explorer window* memiliki fungsi yang serupa dengan *Windows Explorer* pada *Windows*, yaitu untuk menampilkan data-set maupun library SAS. Perintah copy, cut, paste, maupun delete juga dapat dilakukan. *Result window* dan *Explorer window* tidak dapat ditampilkan bersama-sama. Jika *Result window* ditampilkan, *Explorer window* akan tersembunyi, dan demikian sebaliknya.



Gambar 2.2 Tampilan *Explorer Window* (Kiri) dan *Result Window* (Kanan)

2.10. Penelitian Terdahulu

Pada penelitian Indahwati (2014) mengenai Metode *Partial Least Square* untuk mengatasi multikolinearitas pada model regresi linear berganda memberikan kesimpulan bahwa penduga parameter regresi yang dihasilkan oleh metode PLS menjadi bias dan tidak efisien ketika digunakan untuk mengestimasi parameter regresi ketika terdapat korelasi negatif, tidak ada korelasi dan korelasi kecil antar variabel bebasnya. Hal tersebut terlihat dari nilai bias dan MSE yang dihasilkan metode PLS lebih besar dibandingkan dengan metode OLS.

Pada penelitian Tazliqoh (2015) Perbandingan regresi komponen utama dengan regresi ridge pada analisis factor-faktor pendapatan asli daerah (PAD)

provinsi jawa tengah memberikan kesimpulan bahwa Pada analisis Pendapatan Asli Daerah (PAD) dengan menggunakan metode kuadrat terkecil terdapat masalah multikolinieritas sehingga dilakukan penanganan dengan menggunakan regresi komponen utama dan regresi ridge. Berdasarkan perhitungan nilai standar *error* menunjukkan bahwa penanganan multikolinieritas menggunakan regresi komponen utama lebih baik dibandingkan dengan regresi ridge dalam analisis faktor-faktor Pendapatan Asli Daerah (PAD) Provinsi Jawa Tengah.

Pada penelitian Ifadah (2011) mengenai analisis metode *Principal Component Analysis* (komponen utama) dan ridge dalam mengatasi dampak multikolinearitas dalam analisis regresi linear berganda memberi kesimpulan Metode *Principal Component Analysis* (PCA) digunakan untuk mengatasi multikolinearitas yang bertujuan untuk menyederhanakan variabel yang diamati dengan cara mereduksi dimensinya. Hal ini dilakukan dengan cara menghilangkan korelasi diantara variabel bebas melalui transformasi variabel bebas asal ke variabel baru yang tidak berkorelasi sama sekali. Setelah beberapa komponen hasil PCA yang bebas multikolinearitas diperoleh, maka komponen-komponen tersebut menjadi variabel bebas baru yang akan diregresikan atau dianalisis pengaruhnya terhadap variabel tak bebas (*Y*) dengan menggunakan analisis regresi. Metode PCA dilakukan dengan menggunakan bantuan analisis faktor dalam SPSS. Sedangkan metode regresi ridge pada hakikatnya mengusahakan sifat-sifat jumlah kuadrat MSE menjadi lebih kecil dengan cara menambahkan suatu konstanta positif yang kecil pada diagonal matriks persamaan normal. Hal ini akan menyebabkan taksiran regresi ridge menjadi stabil walaupun menjadi bias.

Pada penelitian Nurhasanah (2012) mengenai Perbandingan metode *Partial Least Square (PLS)* dengan regresi komponen utama untuk mengatasi multikolinearitas memberikan kesimpulan Pada Metode *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; 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.

2.12. Kerangka Berpikir

Di dalam asumsi klasik terdapat uji multikolinearitas, uji multikolinearitas bertujuan untuk menguji apakah dalam model regresi ditemukan adanya korelasi antara variabel bebas. Jadi uji multikolinearitas terjadi hanya pada regresi ganda. Model regresi yang baik seharusnya tidak terjadi korelasi tinggi diantara variabel bebas.

Ada beberapa cara untuk mengatasi masalah multikolinearitas di antaranya dengan metode *Principal Component Analisis (PCA)*, Ridge Regression, Stepwise, Jackknife, *Partial Least Square (PLS)*, *Principal Component Regression (PCR)*

dan lain sebagainya. Metode *Partial Least Square* (PLS) dan *Principal Component Regression* (PCR) merupakan metode yang mengkombinasikan sifat-sifat dari *Principal Component Analysis* (PCA) dan regresi linear berganda untuk mengatasi masalah multikolinearitas.

PLS dibedakan menjadi dua macam yaitu regresi dan analisis jalur, secara umum metode PLS dapat diklasifikasikan menjadi dua bagian yaitu PLS-R dan PLS-PM. PLS-PM atau PLS-SEM masih bias diklasifikasikan menjadi beberapa pengembangan. Demikian juga untuk PLS-R atau PLS-GLR karena dapat dikembangkan untuk GLR seperti logistic, regresi cox dan keluarga eksponensial dalam GLM. Dalam hal ini PLS diklasifikasi menjadi dua kategori yaitu untuk variabel dependen Y kuantitatif dan Y kualitatif. Penerapan PLS untuk variabel dependen kualitatif/kategori disebut PLS-DA (PLS discriminant analysis).

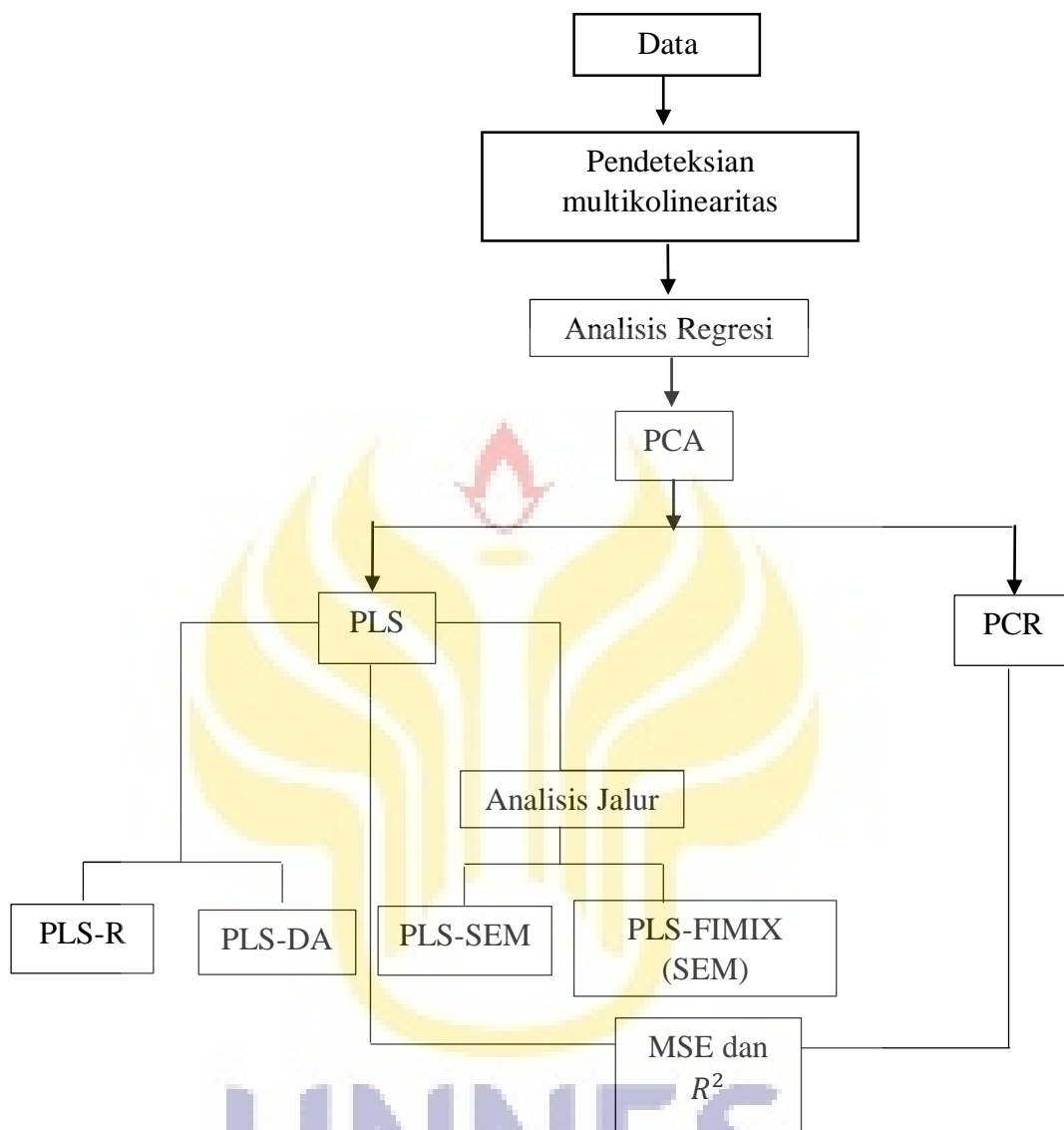
Metode PLS merupakan metode yang mengkombinasikan sifat-sifat dari komponen utama dan regresi linear berganda. Tujuan dari metode PLS adalah mengestimasi dan menganalisis variabel terikat dari variabel-variabel bebas. Dalam hal ini, PLS mereduksi dimensi variabel-variabel bebas dengan membentuk variabel-variabel baru yang merupakan kombinasi linear dari variabel-variabel bebas dengan dimensi lebih kecil.

Analisis Regresi digunakan untuk mengetahui ada tidaknya hubungan antara variabel dependen dan independen, sedangkan PCA pada dasarnya bertujuan untuk menyederhanakan variabel yang diamati dengan cara menyusutkan (mereduksi) dimensinya. Hal ini dilakukan dengan jalan menghilangkan korelasi di antara variabel melalui transformasi variabel asal ke variabel baru (merupakan

kombinasi linear dari variabel-variabel asal) yang tidak saling berkorelasi. Dari p buah variabel asal dapat dibentuk p buah komponen utama, dipilih k buah komponen utama saja ($k < p$) yang telah mampu menerangkan keragaman data cukup tinggi (antara 80% sampai dengan 90%).

Cara pembentukan regresi komponen utama melalui analisis komponen utama ada dua cara yaitu komponen utama yang dibentuk berdasarkan matriks kovariansi dan komponen utama yang dibentuk berdasarkan matriks korelasi. Matriks korelasi dari data yang telah distandarisasi (bentuk baku Z) digunakan jika variabel yang diamati tidak memiliki satuan pengukuran yang sama. Sedangkan Matriks varians kovarians digunakan jika semua variabel yang diamati mempunyai satuan pengukuran yang sama.

Penelitian ini memfokuskan pada metode PLS dan PCR. Dari kedua metode tersebut dibandingkan sehingga diperoleh metode terbaik. Setelah selesai menggunakan metode PLS dan PCR membandingkan nilai koefisien determinasi R^2 dan nilai Mean Square Error (MSE) untuk memperoleh metode terbaik. Untuk mempermudah analisis ini, peneliti menggunakan bantuan *software* SAS. Diagram alur kerangka berfikir Gambar 2.3 sebagai berikut.



Gambar 2.3 Kerangka Berfikir

BAB 5

PENUTUP

5.1. Kesimpulan

Berdasarkan hasil analisis dan pembahasan pada bab 4, maka dapat diperoleh beberapa kesimpulan sebagai berikut:

1. Model persamaan dengan metode *Partial Least Square* (PLS) pada kasus pendapatan anggaran Daerah Provinsi Jawa Tengah sebagai berikut:

$$\begin{aligned}\hat{Y} = & 1382126382 + 83948025.7X_1 + 40120614.88X_2 \\ & + 74135918.8X_3 + 94632319.88X_4 + 145001135.3X_5 \\ & + 59090688.22X_6\end{aligned}$$

2. Model persamaan dengan metode *Principal Component Regression* (PCR) pada kasus pendapatan anggaran Daerah Provinsi Jawa Tengah sebagai berikut:

$$\begin{aligned}\hat{Y} = & 1382126508 + 28888566.77 X_1 + 84611231.43 X_2 \\ & - 14440705.5X_3 - 5053009.86X_4 - 6076336.99X_5 \\ & - 89989904.2X_6\end{aligned}$$

3. Perbandingan metode manakah antara *Partial Least Square* (PLS) dan *Principal Component Regression* (PCR) untuk mengatasi multikolinearitas yang lebih baik dilihat dari nilai R^2 dan nilai MSE. dapat dilihat bahwa nilai R^2 yang dihasilkan PLS = 0.7752 lebih besar dibandingkan dengan PCR = 0.5652 dan nilai MSE yang dihasilkan PLS = 3.660671E16 lebih kecil dibandingkan dengan PCR = 7.301565E16 maka dapat disimpulkan bahwa

metode PLS lebih baik dibandingkan dengan metode PCR dalam mengatasi masalah multikolinearitas.

5.2. Saran

Berdasarkan simpulan di atas peneliti memberikan saran sebagai berikut:

1. Jika pada suatu model regresi terjadi penyimpangan asumsi multikolinearitas. Maka harus dilakukan tindakan perbaikan untuk menghilangkan multikolinearitas tersebut, sebaiknya menggunakan metode *Partial Least Square* (PLS) karena terbukti lebih baik dibandingkan dengan metode *Principal Component Regression* (PCR).
2. Masalah multikolinearitas dapat diatasi dengan berbagai cara. Metode PLS terbukti dapat mengatasi multikolinearitas lebih baik daripada metode PCR. Penelitian selanjutnya dapat menggunakan metode PLS dengan metode yang lain untuk mengatasi masalah multikolinearitas.

5.3. Kelemahan Penelitian

Pada data terjadi sebuah *error* karena penjumlahan antar variabel X_1, X_2, X_3, X_4, X_5 sampai dengan X_6 tidak sesuai dengan nilai variabel Y hal ini terjadi karena *human error* pada saat perhitungan di data BPS Jawa Tengah.

DAFTAR PUSTAKA

- Abdi, H. 2010. Partial Least Square (PLS) Regression and Projection On Latent Structure Regression. *Encyclopedia of Social Sciences Research Methods*.
- Draper, H., & Smith, H. 1992. *Applied Regression Analysis, 2nd. (Analisis Regresi Terapan Edisi Ke-2)*. Penerjemah : Bambang Sumantri. Jakarta:PT. Gramedia Pustaka Utama.
- Farahani, A, H. Rahiminezhed, A. Same, L, & Immannezhed, K. 2010. A Comparison of Partial Least Square (PLS) and Ordinary Least Square (OLS) regressions in predicting of couples mental health based on their communicational patterns. *Procedia*.
- Gujarati, D. 2004. *Ekonometrika Dasar*. Sumarno Zain Penerjemah. Jakarta: Erlangga. Terjemahan dari: *Basic Econometrics*.
- Ifadah, Ana. 2011. *Analisis Metode Principal Component Analisis (Komponen Utama) dan Regresi Ridge Dalam Mengatasi Dampak Multikolinearitas analisis regresi linear berganda*. Skripsi, Program Studi Matematika, Fakultas MIPA UNNES.
- Indahwati, R. Kusnandar, D. & Sulistianingsih, E. 2014. Metode *Partial Least Square* untuk Mengatasi Multikolinearitas pada Model Regresi Linear Berganda. Bimaster
- Johnson, R. A., Wichern, D. W. 2010. *Applied Multivariate Statistical Analysis*. Sixth edition, Prentice Hall. New Jersey.
- Nurhasanah, Subianto, M. & Fitriani, R. 2012. Perbandingan Metode Partial Least Square (PLS) dengan Regresi Komponen Utama untuk Mengatasi Multikolinearitas. *Jurusan Matematika FMIPA UNSYIAH, Statistika*, 12 (1): 33 – 42.
- Pryantono, Duwi. 2013. *Analisis Korelasi, Regresi dan Multivariate dengan SPSS*. Yogyakarta: Gava Media.
- Saefuddin, A., B. Sartono, & N.A. Setiabudi. 2010. *Pengenalan Umum Analisis Statistika dengan SAS Seri 1: Peringkasan dan Penyajian Data*. Bogor: IPB Press.
- Saikia, B. & Singh, R. 2014. Estimation of Principal Components Regression Coefficients. *International Journal of Science and Research (IJSR)*.
- Soemartini. 2008. *Penyelesaian Multikolinearitas Melalui Ridge Regression*. Tesis. Universitas Padjajaran.
- Sugiyono. 2012. *Statistika untuk Penelitian*. Bandung: ALFABETA.

- Sukestiyarno. 2008. *Workshop Olah Data Penelitian dengan SPSS*. Diklat Mata Kuliah Model Linear.
- Sukmono, A. 2014. *Penggunaan Partial Least Square Regression (PLSR) untuk Mengatasi Multikolinearitas Dalam Estimasi Klorofil Daun Tanaman Padi dengan Citra Hiperspektral*. Program Studi Teknik Geodesi, Fakultas Teknik, Universitas Diponegoro.
- Sumodiningrat, G. 1998. *Ekonometrika Pengantar*. Yogyakarta: BPFE.
- Tazliqoh, A, H. Rahmawati, H, & Safitri, D. 2015. Perbandingan Regresi Komponen Utama Regresi Ridge Pada Analisis Faktor-Faktor Pendapatan Asli Daerah (PAD) Provinsi Jawa Tengah. *Jurnal Gaussian*.
- Widarjono, A. 2007. *Ekonometrika Teori dan Aplikasi untuk Ekonomi dan Bisnis*. Yogyakarta: Ekonisia FE UII.
- Vinzi, V. Bastien, P. & Tenenhaus, M. 2004. Partial Least Square Generalized Linear Regression. *Computational statistics & Data Analysis* (2005).

