ANALISIS METODE *GENERALIZED TWO STAGE RIDGE REGRESSION* (GTSRR) UNTUK MENGATASI MASALAH MULTIKOLINEARITAS DAN AUTOKORELASI

(Skripsi)

Oleh

ANIS WATUN HASANAH

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2023

ABSTRACT

ANALYSIS OF *GENERALIZED TWO STAGE RIDGE REGRESSION* (GTSRR) METHOD TO OVERCOME MULTICOLINEARITY AND AUTOCORELATION PROBLEMS

By

ANIS WATUN HASANAH

Multicollinearity, or the presence of relatively high correlation between independent variables, is a common issue in linear regression models. On the other hand, autocorrelation occurs when there is a correlation between the actual observation values and the previous observation values. Autocorrelation is commonly observed in time series data. The GTSRR method is used as an alternative to address the issue of multicollinearity in estimating the parameters of a linear regression model. Therefore, the objective of this study is to evaluate the effectiveness of the GTSRR method in addressing multicollinearity and autocorrelation issues through data simulation with sample sizes of $n = 50, 75$, 100, using 6 independent variables that are highly correlated with a correlation coefficient of 0.99, and an autocorrelation coefficient of 0.15. The results of this study indicate that a sample size of $n = 100$ yields better results compared to sample sizes of $n = 50$ and 75. This is evidenced by the smaller values of MSE and AMSE. Therefore, in modeling data that contains multicollinearity and autocorrelation issues, using a sample size of n=100 will produce better results.

Keywords: *Generalized Two Stage Ridge Regression* (GTSRR), Multicollinearity, Autocorrelation

ABSTRAK

ANALISIS METODE *GENERALIZED TWO STAGE RIDGE REGRESSION* (GTSRR) UNTUK MENGATASI MASALAH MULTIKOLINEARITAS DAN AUTOKORELASI

Oleh

ANIS WATUN HASANAH

Multikolinearitas atau adanya korelasi yang relatif tinggi antara variabel independen, merupakan masalah umum dalam model regresi linier. Di sisi lain, autokorelasi terjadi ketika terdapat korelasi antara nilai pengamatan aktual dengan nilai pengamatan sebelumnya. Gejala autokorelasi umumnya terlihat pada data deret waktu. Metode GTSRR digunakan sebagai alternatif untuk mengatasi masalah multikolinearitas dalam estimasi parameter model regresi linier. Oleh karena itu, tujuan dari penelitian ini adalah untuk mengevaluasi sejauh mana metode GTSRR efektif dalam mengatasi masalah multikolinearitas dan autokorelasi melalui simulasi data dengan ukuran $n = 50, 75, 100$, serta menggunakan 6 variabel independen yang saling berkorelasi dengan tingkat korelasi sebesar 0.99, dan autokorelasi sebesar 0.15. Hasil penelitian ini menunjukkan bahwa ukuran sampel $n = 100$ memberikan hasil yang lebih baik dibandingkan dengan ukuran sampel $n = 50$ dan 75. Hal ini dibuktikan dengan nilai MSE dan AMSE yang semakin kecil. Oleh karena itu, dalam pemodelan data yang mengandung masalah multikolinearitas dan autokorelasi, menggunakan ukuran sampel n=100 akan menghasilkan hasil yang lebih baik.

Kata Kunci: *Generalized Two Stage Ridge Regression* (GTSRR), Multikolinearitas, Autokorelasi

ANALISIS METODE *GENERALIZED TWO STAGE RIDGE REGRESSION* (GTSRR) UNTUK MENGATASI MASALAH MULTIKOLINEARITAS DAN AUTOKORELASI

Oleh

ANIS WATUN HASANAH

Skripsi

Sebagai Salah Satu Syarat untuk Mencapai Gelar
SARJANA MATEMATIKA

Pada

Jurusan Matematika
Fakultas Matematika Dan Ilmu Pengetahuan Alam

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2023

Judul Skripsi : **ANALISIS METODE *GENERALIZED TWO STAGE RIDGE REGRESSION* (GTSRR) UNTUK MENGATASI MASALAH MULTIKOLINEARITAS DAN AUTOKORELASI**

Nama Mahasiswa : ***Anis Watun Hasanah***

Nomor Pokok Mahasiswa : **1917031009**

Jurusan : **Matematika**

Fakultas : **Matematika dan Ilmu Pengetahuan Alam**

MENYETUJUI

1. Komisi Pembimbing

Dr. Ir. Netti Herawati, M.Sc.	**Agus Sutrisno, S.Si., M.Si.**
NIP 19650125 199003 2 001	NIP 19700831 199903 1 002

2. Ketua Jurusan Matematika

Dr. Aang Nuryaman. S.Si.,M.Si.
NIP. 19740316 200501 1 001

MENGESAHKAN

1. Tim Penguji

Ketua : **Dr. Ir. Netti Herawati, M.Sc.** ....

Sekretaris : **Agus Sutrisno, S.Si., M.Si.** ....

Penguji
Bukan Pembimbing : **Dr. Khoirin Nisa, S.Si., M.Si.** ....

2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

Dr. Eng. Heri Satria, S.Si., M.Si.
NIP 19711001 200501 1 002

Tanggal Lulus Ujian Skripsi : **9 Juni 2023**

PERNYATAAN SKRIPSI MAHASISWA

Yang bertanda tangan di bawah ini:

Nama	: **Anis Watun Hasanah**
Nomor Pokok Mahasiswa	: **1917031009**
Jurusan	: **Matematika**
Judul Skripsi	: **ANALISIS METODE *GENERALIZED TWO STAGE RIDGE REGRESSION* (GTSRR) UNTUK MENGATASI MASALAH MULTIKOLINEARITAS DAN AUTOKORELASI**

Dengan ini menyatakan bahwa skripsi ini adalah hasil pekerjaan saya sendiri. Dan apabila kemudian hari terbukti bahwa skripsi ini merupakan hasil salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku.

Bandar Lampung, Juni 2023
Yang menyatakan,

Anis Watun Hasanah

RIWAYAT HIDUP

Penulis bernama lengkap Anis Watun Hasanah. Penulis lahir di Lampung pada tanggal 15 Maret 2001, sebagai anak kedua dari dua bersaudara. Perjalanan pendidikan penulis dimulai di SDN Marga Kaya pada tahun 2007 hingga 2013. Setelah menyelesaikan pendidikan dasar, penulis melanjutkan ke jenjang Sekolah Menengah Pertama di SMPN 3 Jatiagung dari tahun 2013 hingga 2016. Kemudian, penulis melanjutkan pendidikan di jenjang Sekolah Menengah Atas di SMAS Al-Huda dari tahun 2016 hingga 2019.

Pada tahun 2019, penulis berhasil terdaftar sebagai mahasiswi Program Studi S1 Matematika di Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung melalui jalur Seleksi Nasional Masuk Perguruan Tinggi Negeri (SNMPTN). Selama menjadi mahasiswi, penulis aktif mengikuti berbagai kegiatan. Penulis juga berkesempatan mengikuti perkuliahan di Institut Teknologi Sepuluh Nopember selama satu semester pada tahun 2021 melalui program Kredensial Mikro Mahasiswa Indonesia (KMMI). Pada tahun 2023, penulis juga mengikuti kegiatan Kampus Mengajar Angkatan 5 yang ditempatkan di SDN 5 Karang Anyar.

Pada bulan Januari hingga Februari 2022, penulis melaksanakan Kerja Praktik di Badan Pusat Statistik Lampung Selatan sebagai bentuk penerapan ilmu yang telah penulis peroleh selama kuliah. Pada bulan Juni hingga Juli 2022, penulis juga melaksanakan Kuliah Kerja Nyata (KKN) di Desa Gunung Mekar, Kecamatan Jabung, Lampung Timur sebagai bentuk pengabdian mahasiswi dan pelaksanaan Tri Dharma Perguruan Tinggi.

Dengan pengalaman dan pengetahuan yang penulis peroleh selama masa kuliah, penulis berharap dapat terus mengembangkan diri dan memberikan kontribusi positif dalam bidang studi penulis serta masyarakat di sekitar.

KATA INSPIRASI

"Dan barangsiapa yang bertakwa kepada Allah, niscaya Dia akan mencarikan jalan keluar baginya (dari kesulitan), dan memberinya rezeki dari arah yang tidak disangka-sangka."

(Surat Ath-Thalaq, 65:2-3)

"Sesungguhnya, setelah kesulitan ada kemudahan."

(Surat Al-Insyirah, 94:5-6)

"Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya."

(Surat Al-Baqarah, 2:286)

"Ilmu pengetahuan tidak memiliki batasan. Semakin banyak yang kita ketahui, semakin banyak yang kita sadari bahwa masih ada begitu banyak hal yang perlu kita jelajahi."

(Stephen Hawking)

"Ketika Anda berhenti belajar, Anda mulai menua."

(Albert Einstein)

"Kesuksesan tidak datang dengan sendirinya, tetapi melalui kerja keras, ketekunan, dan tekad yang kuat. Teruslah berjuang dan jangan pernah menyerah."

PERSEMBAHAN

Dengan segala kerendahan hati, penulis persembahkan karya sederhana ini kepada:

Keluargaku Tercinta

Dengan penuh rasa cinta dan terima kasih, persembahan ini aku dedikasikan kepada orang tua tercinta, kakak, dan adikku yang selalu ada di setiap langkah hidupku. Kalian adalah pilar yang kuat, memberikan cinta, dukungan, dan pengorbanan tanpa batas. Terima kasih telah mendidik dan membimbingku dengan penuh kasih sayang dan keteladanan. Terima kasih atas cinta, dukungan, dan kebaikan yang kalian berikan

Dosen Pembimbing dan Pembahas

dengan rasa hormat dan penuh rasa terima kasih telah memberikan arahan, bimbingan, dan wawasan berharga selama perjalanan penulisan skripsi ini. Tanpa dukungan dan kerjasama, pencapaian ini tidak akan pernah terwujud.

Sahabatku

Terima kasih, sahabat-sahabatku, karena kalian telah menjadi bagian tak terpisahkan dari pencapaian ini. Kebersamaan kita adalah harta yang tak ternilai, dan saya berharap hubungan kita tetap langgeng dan abadi.

SANWACANA

Puji syukur kehadirat Allah SWT. yang telah melimpahkan segala rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul "Analisis Metode *Generalized Two Stage Ridge Regression* (GTSRR) untuk Mengatasi Masalah Multikolinearitas dan Autokorelasi".

Penulisan skripsi ini dapat terlaksana karena bantuan dan dukungan dari berbagai pihak baik secara langsung maupun tidak langsung, sehingga pada kesempatan ini penulis mengucapkan terima kasih kepada:

1. Ibu Dr. Ir. Netti Herawati, M.Sc., selaku pembimbing utama dan pembimbing akademik atas kesediaan waktu dan pemikiran nya dalam memberikan bimbingan dan arahan yang membangun dalam proses penyusunan skripsi ini.
2. Bapak Agus Sutrisno, S.Si., M.Si., selaku pembimbing kedua yang telah memberikan arahan dalam proses penyusunan skripsi ini.
3. Ibu Dr. Khoirin Nisa, S. Si., M.Si., selaku dosen penguji yang telah memberikan evaluasi dan saran bagi perbaikan skripsi penulis.
4. Bapak Dr. Aang Nuryaman, S.Si., M.Si., selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
5. Bapak Dr. Eng. Heri Satria, S.Si., M.Si., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
6. Seluruh dosen dan staf Jurusan Metematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
7. Kedua orang tua beserta kakak dan adik tercinta yang selalu menjadi rumah tanpa riuh serta menjadi motivasi terbesar penulis untuk menyelesaikan perkuliahan ini.

8. Sahabatku yaitu Ahmad Rizki Wiranto, Ameliana Wijayanti, Elsa Rahmawati, Herlina Juwita Sukma, Novi Darina, Risma Nurul Hidayati serta teman-teman Jurusan Matematika Angkatan 2019, terima kasih untuk canda tawa, dukungan dan bantuan kalian selama ini.
9. Teman-teman Kampus Mengajar Batch 5 April, Dinda, Rena, Figo, serta seluruh guru dan staf SDN 5 Karang Anyar yang telah memberikan pengalaman berharga.
10. Kakak-kakak dan adik-adik tingkat di Jurusan Matematika serta semua pihak terkait lainnya yang telah memberikan dukungan dalam penyusunan tugas akhir ini.

Penulis menyadari bahwa tugas akhir ini masih jauh dari kesempurnaan. Oleh karena itu, penulis mengharapkan kritik serta saran yang bersifat membangun dari pembaca. Semoga tugas akhir ini dapat berguna dan bermanfaat bagi kita semua.

Bandar Lampung,   Juni 2023
Penulis,

Anis Watun Hasanah

DAFTAR ISI

Halaman

DAFTAR TABEL

Tabel Halaman

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

Deret waktu adalah suatu rangkaian data yang dikumpulkan berdasarkan urutan waktu dengan interval yang sama. *Autoregressive* (AR) merupakan bentuk regresi deret waktu yang menghubungkan nilai pengamatan saat ini dengan nilai pengamatan sebelumnya. Model ini menggambarkan suatu proses acak yang berlangsung dalam variasi waktu tertentu, di mana variabel keluaran tergantung secara linear pada nilai proses sebelumnya.

Analisis regresi digunakan untuk memodelkan dan menganalisis hubungan antara satu atau lebih variabel independen terhadap variabel dependen. Model regresi linear dapat diperoleh dengan melakukan estimasi terhadap parameter modelnya. Umumnya, estimasi parameter regresi linear dilakukan dengan menggunakan Metode Kuadrat Terkecil (MKT). MKT dapat digunakan jika beberapa asumsi terpenuhi, seperti asumsi ragam berdistribusi normal, tidak terjadi masalah multikolinearitas, kehomogenan ragam sisaan, dan tidak terjadi autokorelasi.

Salah satu masalah yang sering muncul dalam model regresi linear adalah multikolinearitas, yaitu terjadinya korelasi yang tinggi antara variabel-variabel independen, dan autokorelasi yaitu terjadinya korelasi antara pengamatan pada X_t dan pengamatan pada $X_{t-1}, X_{t-2}, \dots, X_{t-k}$. Gejala autokorelasi kemungkinan besar terjadi pada data deret waktu.

Metode *Generalized Ridge Regression* (GRR) adalah sebuah metode alternatif untuk memperkirakan parameter dalam model regresi linear yang diperkenalkan oleh Hoerl & Kennard pada tahun 1970. GRR merupakan pengembangan dari metode *Ridge Regression* (RR). Perbedaan utama antara metode *Ridge Regression* (RR) dan GRR terletak pada jumlah nilai konstanta bias k. Pada metode *Ridge Regression* (RR), hanya terdapat satu nilai *k* untuk konstanta bias, sedangkan pada metode GRR, terdapat nilai *k* yang sesuai dengan jumlah *p* variabel independen dalam model regresi linear.

Penanganan autokorelasi dapat dilakukan dengan menggunakan metode *Two Stages Least Square* (TSLS). Jika koefisien autokorelasi diketahui maka digunakan *Generalized Least Square* (GLS), namun ketika koefisien autokorelasi tidak diketahui maka digunakan *Feasible Generalized Least Square* (FGLS).

Eledum & Alkhaifa (2012) mengenalkan suatu metode inovatif dalam regresi Ridge yang disebut *Generalized Two Stage Ridge Regression* (GTSRR). GTSRR merupakan penggabungan dari metode *Two Stages Least Squares* (TSLS) dan *Generalized Ridge Regression* (GRR).

Berdasarkan latar belakang diatas, maka penulis tertarik untuk meneliti tentang "Analisis Metode *Generalized Two Stage Ridge Regression* (GTSRR) untuk Mengatasi Masalah Multikolinearitas dan Autokorelasi".

1.2 Tujuan Penelitian

Tujuan dari penelitian ini meliputi:

1. Menguji efektivitas metode *Generalized Two Stage Ridge Regression* (GTSRR) dalam menangani masalah multikolinearitas dan autokorelasi.
2. Mengidentifikasi langkah-langkah yang digunakan dalam melakukan estimasi menggunakan *Generalized Two Stage Ridge Regression* (GTSRR).

3. Menerapkan metode *Generalized Two Stage Ridge Regression* (GTSRR) pada data deret waktu yang ada.

1.3 Manfaat Penelitian

Manfaat dari penelitian ini meliputi:

1. Menyediakan metode alternatif yang efektif dalam menangani masalah multikolinearitas dan autokorelasi.
2. Memperluas pemahaman tentang penerapan metode *Generalized Two Stage Ridge Regression* (GTSRR) pada data deret waktu.
3. Meningkatkan wawasan dan pengetahuan tentang penggunaan metode *Generalized Two Stage Ridge Regression* (GTSRR).

II. TINJAUAN PUSTAKA

2.1 Deret waktu

Deret waktu merujuk pada kumpulan data yang diurutkan berdasarkan waktu dengan interval yang konsisten. Analisis deret waktu adalah metode kuantitatif yang digunakan untuk menganalisis pola data yang dikumpulkan dalam urutan waktu selama periode waktu tertentu, yang dikenal sebagai data deret waktu. Ketika waktu dianggap sebagai variabel diskret (meskipun waktu sebenarnya dapat bersifat kontinu), frekuensi pengumpulan data biasanya tetap. Dalam konteks diskret, frekuensi pengumpulan dapat berupa detik, menit, jam, hari, minggu, bulan, atau tahun, tergantung pada kebutuhan dan skala waktu yang relevan (Montgomery, 2008).

2.2 *Autoregressive* (AR)

Autoregressive (AR) adalah jenis model regresi deret waktu yang menghubungkan nilai pengamatan saat ini dengan nilai pengamatan sebelumnya. Model ini digunakan untuk menggambarkan proses acak yang berubah seiring waktu, di mana variabel keluaran hanya bergantung secara linear pada nilai proses sebelumnya. AR (p) adalah salah satu bentuk model *autoregressive* di mana nilai proses acak saat ini hanya dipengaruhi oleh nilai proses acak dalam p periode waktu sebelumnya. Dengan kata lain, AR (p) mempertimbangkan pengaruh nilai-nilai sebelumnya sebanyak p periode waktu untuk memprediksi nilai saat ini.

Menurut Rosadi (2021), proses *autoregressive* orde p AR(p) didefenisikan dengan persamaan:

$$Y_t = \varphi_0 + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \cdots + \varphi_p Y_{t-p} + \varepsilon_t \quad (2.2.1)$$

dengan:

Y_t	:	Variabel dependen pada waktu ke-*t*
$Y_{t-1}, Y_{t-2}, Y_{t-p}$	:	Nilai masa lalu yang bersangkutan pada t-1, t-2, dan t-*p*
$\varphi_0, \varphi_1, \varphi_2, \varphi_p$	:	Parameter autoregressive , i=1,2,…,p
ε_t	:	Residual pada waktu ke-t

2.3 Regresi Linear

Menurut Draper & Smith (1992), regresi linear adalah metode statistika yang digunakan untuk menganalisis data dan mengambil kesimpulan yang signifikan tentang hubungan antara variabel independen dan variabel dependen. Regresi linear berganda adalah suatu teknik statistik yang digunakan untuk menganalisis dan memodelkan pengaruh antara dua atau lebih variabel independen terhadap variabel dependen. Dengan demikian, model regresi linear berganda merujuk pada model regresi linear yang menggambarkan hubungan antara dua atau lebih variabel independen dan variabel dependen.

Analisis regresi dapat dibagi menjadi dua jenis berdasarkan jumlah variabel independen yang digunakan, yaitu analisis regresi linear sederhana dan analisis regresi linear berganda. Analisis regresi linear sederhana adalah metode regresi yang memodelkan hubungan antara satu variabel independen dengan variabel dependen. Dalam hal ini, satu variabel independen digunakan untuk memprediksi atau menjelaskan variasi dalam variabel dependen. Di sisi lain, analisis regresi linear berganda adalah metode regresi yang memodelkan hubungan antara dua atau lebih variabel independen dengan variabel dependen.

Menurut Draper & Smith (1992), model regresi linear berganda dengan *p* variabel independen, sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_p X_{ip} + \varepsilon_i \tag{2.3.1}$$

Keterangan:

N	:	Banyaknya pengamatan
p	:	Banyaknya variabel independen dalam model
Y_i	:	Variabel dependen pada pengamatan ke-*i*
β_j	:	Parameter regresi atau koefisien regresi independen ke-*j*
X_{ij}	:	Variabel independen ke-*j* pada pengamatan ke-*i*
I	:	1, 2, …,*n*
J	:	1, 2, …, *p*

Jika terdapat *n* pengamatan, model regresi linear berganda menjadi:

$$\begin{aligned} Y_1 &= \beta_0 + \beta_1 X_{11} + \beta_2 X_{12} + \cdots + \beta_p X_{1p} + \varepsilon_1 \\ Y_2 &= \beta_0 + \beta_1 X_{21} + \beta_2 X_{22} + \cdots + \beta_p X_{2p} + \varepsilon_2 \\ &\vdots \\ Y_n &= \beta_0 + \beta_1 X_{n1} + \beta_2 X_{n2} + \cdots + \beta_p X_{np} + \varepsilon_n \end{aligned} \tag{2.3.2}$$

Maka, dalam bentuk matriks model regresi linear berganda dapat dituliskan sebagai berikut:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1p} \\ 1 & X_{21} & X_{22} & \cdots & X_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \tag{2.3.3}$$

Sehingga,

$$\begin{aligned} \boldsymbol{Y}_{n\times 1} &= \boldsymbol{X}_{n\times(p+1)} \boldsymbol{\beta}_{(p+1)\times 1} + \boldsymbol{\varepsilon}_{n\times 1} \\ \boldsymbol{Y} &= \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \end{aligned} \tag{2.3.4}$$

Keterangan:

$\boldsymbol{Y}$	:	Vektor pengamatan berukuran $n \times 1$
$\boldsymbol{X}$	:	Matriks variabel independen berukuran $n \times (p+1)$
$\boldsymbol{B}$	:	Vektor parameter regresi berukuran $(p+1) \times 1$
$\boldsymbol{\varepsilon}$	:	Vektor random sisaan berukuran $n \times 1$

2.4 Asumsi-Asumsi Model Regresi

Menurut Gujarati (2003), dalam analisis regresi harus memenuhi asumsi-asumsi, sebagai berikut:

1. Model regresi linearnya merupakan regresi linear dalam parameter, dapat dituliskan sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \tag{2.4.1}$$

2. Nilai yang diambil oleh variabel independen X dianggap tetap dalam sampel berulang. Sehingga, X diasumsikan nonstokastik.
3. Nilai ekspektasi dari vektor residualnya adalah 0.

$$E(\varepsilon_i|X_i) = 0 \tag{2.4.2}$$

dimana, $i = 1, 2, \dots, n$

4. Homoskedastisitas (variansi dari residualnya adalah konstan).

$$\begin{aligned} var(\varepsilon_i|X_i) &= E[\varepsilon_i - E(\varepsilon_i|X_i)]^2 \\ &= E(\varepsilon_i^2|X_i) \\ &= \sigma^2 \end{aligned} \tag{2.4.3}$$

dimana, $i = 1, 2, \dots, n$

5. Tidak terjadi autokorelasi antar residual.

$$\begin{aligned} cov(\varepsilon_i, \varepsilon_j|X_i, X_j) &= E\{[\varepsilon_i - E(\varepsilon_i)]|X_i\}\{[\varepsilon_j - E(\varepsilon_j)]|X_j\} \\ &= E(\varepsilon_i|X_i)(\varepsilon_j|X_j) \\ &= 0 \end{aligned} \tag{2.3.4}$$

dimana, $i, j = 1, 2, \dots, n$ dan $i \neq j$

6. Residual berdistribusi normal.
7. Jumlah *n* observasi harus lebih besar dari jumlah estimasi parameter.
8. Nilai X dalam sampel tidak boleh semuanya sama.
9. Tidak ada bias dalam model yang digunakan.
10. Tidak terjadi multikolinieritas (tidak terdapat hubungan linier antara variabel independen dengan yang lainnya).

2.5 Metode Kuadrat Terkecil (MKT)

Metode Kuadrat Terkecil yang disebut juga *Ordinary Least Square* (OLS) merupakan sebuah metode yang dapat digunakan untuk menaksir nilai rata-rata dari variabel random. Metode Kuadrat Terkecil pertama kali diperkenalkan oleh Carl Friedreich Gauss, seorang ahli matematikawan Jerman untuk menghitung masalah astronomi. Menurut Gujarati & Porter (2009), Metode Kuadrat Terkecil merupakan suatu metode pendugaan β dengan meminimumkan jumlah kuadrat dari *error*.

Menurut Sembiring (2003), untuk menentukan estimasi parameter β dengan meminumkan bentuk kuadrat sebagai berikut:

$$\begin{aligned} J &= (Y - X\beta)^T(Y - X\beta) \\ &= Y^TY - 2Y^TX\beta + \beta X^TX\beta \end{aligned} \tag{2.5.1}$$

Pandang kembali persamaan $(2.3.1)$ maka,

$$J = \sum \varepsilon_i^2 = \sum\left(y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2} - \cdots - \beta_p x_{ip}\right)^2 \tag{2.5.2}$$

Selanjutnya, *J* akan diturunkan secara parsial terhadap $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ dan samakan dengan nol.

Jadi,

$$\frac{\partial J}{\partial \beta_0} = -2\sum\left(y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2} - \cdots - \beta_p x_{ip}\right)$$
$$\frac{\partial J}{\partial \beta_1} = -2\sum\left(y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2} - \cdots - \beta_p x_{ip}\right) x_{i1}$$
$$\frac{\partial J}{\partial \beta_2} = -2\sum\left(y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2} - \cdots - \beta_p x_{ip}\right) x_{i2}$$
$$\vdots$$
$$\frac{\partial J}{\partial \beta_p} = -2\sum\left(y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2} - \cdots \beta_p x_{ip}\right) x_{ip} \qquad (2.5.3)$$

Setelah disusun kembali dan ganti semua parameter dengan estimasinya, sistem persamaan ini dapat dituliskan sebagai berikut:

$$\sum y_i = nb_0 + b_1 \sum x_{i1} + b_2 \sum x_{i2} + \cdots + b_p \sum x_{ip}$$
$$\sum y_i x_{i1} = b_0 \sum x_{i1} + b_1 \sum x_{i1}^2 + b_2 \sum x_{i2}\, x_{i1} + \cdots + b_p \sum x_{ip} x_{i1}$$
$$\sum y_i x_{i2} = b_0 \sum x_{i2} + b_1 \sum x_{i1}\, x_{i2} + b_2 \sum x_{i2}^2 + \cdots + +b_p \sum x_{ip} x_{i2}$$
$$\vdots$$
$$\sum y_i x_{ip} = b_0 \sum x_{ik} + b_1 \sum x_{i1} x_{ik} + b_2 \sum x_{i2}\, x_{ik} + \cdots + +b_p \sum x_{ip} \qquad (2.5.4)$$

Persamaan diatas merupakan persamaan normal. Jika ditulis dalam bentuk matriks maka, dapat dituliskan sebagai berikut:

$$(X^T X)\beta = X^T Y$$
$$\beta = (X^T X)^{-1} X^T Y \qquad (2.5.5)$$

Persamaan diatas dapat disebut sebagai penduga parameter β.

Berikut ini merupakan sifat-sifat dari penduga parameter Metode Kuadrat Terkecil (MKT), yaitu:

1. Tak bias

$$\begin{aligned} E(\beta) &= E((X^T X)^{-1} X^T Y) \\ &= (X^T X)^{-1} X^T E(Y) \\ &= (X^T X)^{-1} X^T X \beta \\ &= \beta \end{aligned} \qquad (2.5.6)$$

Sehingga, b merupakan penduga takbias dari β.

2. Variansi minimum

$$\begin{aligned} \textstyle\sum_b &= (X^T X)^{-1} X^T \textstyle\sum_Y X (X^T X)^{-1} \\ &= (X^T X)^{-1} X^T \sigma^2 I X (X^T X)^{-1} \\ &= \sigma^2 (X^T X)^{-1} \end{aligned} \tag{2.5.7}$$

Dimana $\sigma^2(X^T X)^{-1}$ merupakan matriks simetris berukuran $p \times p$. Pada praktiknya nilai σ^2 belum diketahui, sehingga nilai σ^2 harus diduga menggunakan data terlebih dahulu.

$$\begin{aligned} SS_{RES} = J &= (Y - X\beta)^T (Y - X\beta) \\ &= Y^T Y - 2Y^T X\beta + \beta X^T X\beta \\ &= Y^T Y - \beta X^T Y \end{aligned} \tag{2.5.8}$$

Sehingga,

$$MS_{RES} = \frac{SS_{RES}}{n-p} \tag{2.5.9}$$

Dimana n merupakan banyaknya pengamatan dan p merupakan banyaknya variabel bebas. Jadi, estimasi dari σ^2 dapat dituliskan sebagai berikut:

$$\begin{aligned} \sigma^2 &= MS_{RES} \\ &= \frac{SS_{RES}}{n-p} \end{aligned} \tag{2.5.10}$$

2.6 Multikolinearitas

Menurut Gujarati (2003), multikolinearitas merupakan adanya hubungan linear yang kuat antara beberapa atau seluruh variabel independen dalam model regresi. Berdasarkan sifat hubungan linear yang terjadi antara variabel independen, multikolinearitas dapat dibedakan menjadi dua jenis:

1. Multikolinearitas Sempurna:

   Multikolinearitas sempurna terjadi ketika terdapat hubungan linear yang tepat antara variabel-variabel independen, dan kondisi berikut terpenuhi:

- Salah satu variabel independen dapat diungkapkan sebagai kombinasi linear dari variabel lainnya.
- Terdapat persamaan linier yang sempurna antara variabel-variabel independen.

$$\lambda_1 X_1 + \lambda_2 X_2 + \cdots + \lambda_p X_p = 0 \tag{2.6.1}$$

dimana,

$\lambda_1, \lambda_2, \ldots, \lambda_p$ adalah konstan. Sehingga tidak semuanya bernilai nol secara bersamaan. Dalam kasus multikolinearitas sempurna, tidak mungkin memperoleh estimasi unik untuk koefisien regresi. Hal ini mengakibatkan ketidakstabilan numerik dalam perhitungan dan mempengaruhi interpretasi hasil analisis regresi.

2. Multikolinearitas Tidak Sempurna:

Multikolinearitas tidak sempurna terjadi ketika terdapat hubungan linear yang kuat antara variabel-variabel independen, tetapi tidak ada hubungan linear yang tepat seperti pada multikolinearitas sempurna. Dalam kasus ini, variabel-variabel independen masih saling berkorelasi, tetapi tidak ada kombinasi linear yang presisi antara variabel-variabel tersebut. Hubungan linear yang tidak sempurna terjadi jika berlaku suatu hubungan sebagai berikut:

$$\lambda_1 X_1 + \lambda_2 X_2 + \cdots + \lambda_p X_p + \varepsilon_i = 0 \tag{2.6.2}$$

dimana, ε_i merupakan galat sisa (*error*).

Untuk mengetahui perbedaan antara multikolinearitas sempurna dan kurang sempurna, maka dimisalkan $\lambda_2 \neq 0$, sehingga persamaan X_{2i} dapat dituliskan sebagai berikut:

$$X_{2i} = -\frac{\lambda_1}{\lambda_2} X_{1i} - \frac{\lambda_3}{\lambda_2} X_{3i} - \cdots - \frac{\lambda_p}{\lambda_2} X_{pi} - \frac{1}{\lambda_2} \varepsilon_i \tag{2.6.3}$$

Yang menunjukkan bahwa X_2 tidak berhubungan linear sempurna dengan sisa variabel lainnya, sebab tergantung pada nilai ε_i. Multikolinearitas, baik sempurna maupun tidak sempurna, dapat mempengaruhi interpretasi dan keandalan estimasi koefisien regresi, serta mempersulit identifikasi pengaruh individu dari setiap variabel independen dalam model regresi.

2.6.1 Akibat Multikolinearitas

Menurut Gujarati & Porter (2009), akibat yang akan ditimbulkan jika terjadi maslah multikolinearitas pada analisis regresi sebagai berikut:

1. Estimasi parameter yang diperoleh menggunakan Metode Kuadrat Terkecil (MKT) menunjukkan variasi yang belum mencapai titik minimum.
2. Selang kepercayaan cenderung memiliki lebar yang signifikan dikarenakan adanya nilai *standard error* yang tinggi.
3. Rasio t atau statistik uji t tidak signifikan karena nilai *standard error* yang besar.
4. Ukuran goodness of fit suatu model (R^2) secara keseluruhan sangat tinggi, namun nilai statistik uji t tidak signifikan.

2.6.2 Mendeteksi Adanya Multikolinearitas

Menurut Gujarati & Porter (2009), terdapat beberapa cara untuk mendeteksi ada tidaknya multikolinearitas pada analisis regresi sebagai berikut:

1. Meskipun nilai R^2 cukup tinggi, terdapat sedikit sekali nilai statistik uji *t* yang signifikan.
2. Terdapat korelasi yang kuat antara variabel prediktor.
3. *Variance Inflating Factor* (VIF)

$$VIF = \frac{1}{1 - R_j^2} \tag{2.6.4}$$

Dimana, R_j^2 adalah koefisien determinasi dari variabel independen X_j yang saat diregresikan terhadap $(p - 1)$ variabel independen lainnya. Sebagai aturan umum, jika nilai VIF suatu variabel melebihi 5, ini menandakan adanya multikolinearitas, yang berarti variabel tersebut memiliki hubungan yang kuat dengan variabel lainnya. Secara khusus, jika nilai R_j^2 melebihi 0.9, ini mengindikasikan adanya korelasi yang kuat atau hubungan saling ketergantungan antar variabel, yang memperkuat keberadaan multikolinearitas.

2.6.3 Cara Mengatasi Multikolinearitas

Menurut Montgomery et al. (2012), terdapat beberapa teknik yang umumnya digunakan untuk mengatasi masalah multikolinearitas, yaitu:

1. Menambahkan Sampel Baru

   Dengan menambahkan sampel baru ke dalam data, informasi tambahan pada variabel independen dapat diperoleh. Hal ini dapat mengurangi kemungkinan terjadinya multikolinearitas antara variabel independen.

2. Menghilangkan Satu atau Beberapa Variabel Independen

   Multikolinearitas sering terjadi ketika dua atau lebih variabel independen memiliki korelasi yang kuat dalam model. Dalam situasi ini, multikolinearitas dapat diatasi dengan menghilangkan satu atau beberapa variabel yang memiliki korelasi tinggi dengan variabel lainnya. Dengan menghapus variabel-variabel yang saling berkorelasi tinggi, dapat mengurangi dampak multikolinearitas dan meningkatkan stabilitas serta interpretabilitas model.

3. Menggunakan Metode Alternatif Selain Metode Kuadrat Terkecil (MKT):

   Masalah multikolinearitas dapat menyebabkan Metode Kuadrat Terkecil (MKT) tidak memberikan estimasi parameter yang akurat. Oleh karena itu, digunakan metode alternatif seperti regresi ridge. Pada regresi ridge, kendala ditambahkan ke dalam MKT dengan menambahkan konstanta pada diagonal utama dari matriks $X^T X$. Hal ini membantu mengurangi multikolinearitas dan menghasilkan estimasi parameter yang lebih baik.

2.7 Autokorelasi

Autokorelasi adalah fungsi yang mengukur sejauh mana terdapat korelasi (hubungan linier) antara pengamatan pada waktu t, yang direpresentasikan sebagai X_t, dengan pengamatan pada waktu-waktu sebelumnya, yang direpresentasikan sebagai $(X_{t-1}, X_{t-2}, \ldots, X_{t-k})$. Menurut Gujarati (2003), autokorelasi dapat diartikan sebagai korelasi diantara sisaan. Autokorelasi umumnya terjadi pada data deret waktu, karena pada data deret waktu mengikuti urutan waktu sehingga antar observasi saling berhubungan. Autokorelasi untuk lag *k* dari data deret waktu dinyatakan sebagai berikut:

$$r_k = \frac{\sum_{t=1}^{n-k}(X_t - \overline{X_t})(X_{t+k} - \overline{X_t})}{\sum_{t=1}^{n}(X_t - \overline{X_t})^2} \qquad (2.7.1)$$

Dalam analisis autokorelasi, hubungan antara autokorelasi dan lag (keterlambatan) dikenal sebagai fungsi autokorelasi. Fungsi autokorelasi yang dinyatakan sebagai ρk, menggambarkan hubungan atau korelasi antara variabel pada waktu *t* dengan variabel pada waktu lag *k*. Fungsi autokorelasi memungkinkan untuk mengukur dan mengidentifikasi pola ketergantungan dalam data deret waktu serta melihat keberadaan struktur autokorelasi yang mungkin mempengaruhi analisis dan prediksi data tersebut (Rosadi, 2021).

2.7.1 Akibat Autokorelasi

Dampak dari adanya autokorelasi adalah sebagai berikut (Gujarati, 2003):

1. Estimasi Parameter yang Tidak Efisien

   Estimasi parameter dengan menggunakan Metode Kuadrat Terkecil (MKT) tetap linear dan takbias, namun estimasi yang dihasilkan menjadi tidak efisien. Hal ini berarti estimasi parameter tidak optimal dan memiliki ragam yang lebih besar dari yang seharusnya.

2. Ketidaksignifikan Selang Kepercayaan dan Uji Hipotesis

   Kehadiran autokorelasi mengakibatkan pelanggaran asumsi klasik dalam analisis regresi. Akibatnya, selang kepercayaan dan uji hipotesis yang didasarkan pada statistik uji t dan statistik uji F tidak lagi signifikan. Dalam konteks ini, kesimpulan tentang signifikansi koefisien regresi tidak dapat diambil secara valid.

3. Pengaruh pada Interpretasi Hasil

   Adanya autokorelasi dapat mengaburkan pengaruh sebenarnya dari variabel independen terhadap variabel dependen. Hal ini membuat interpretasi hasil regresi menjadi sulit, karena koefisien regresi yang seharusnya signifikan dapat dianggap tidak signifikan dan sebaliknya.

2.7.2 Cara Mendeteksi Autokorelasi

Pengujian autokorelasi yang paling popular yang digunakan untuk mendeteksi adanya autokorelasi adalah uji *Durbin-Watson* (DW), yang dikembangkan oleh Durbin dan Watson.

Berikut ini adalah langkah-langkah uji *Durbin-Watson* (DW):

1. Hipotesis

   $H_0 : \rho = 0$ (tidak ada autokorelasi)

   $H_1 : \rho \neq 0$ (ada autokorelasi)

2. Taraf nyata $\alpha = 0.05$

3. Statistik Uji

$$d = \frac{\sum_{t=2}^{n}(\hat{\varepsilon}_t \hat{\varepsilon}_{t-1})^2}{\sum_{t=2}^{n} \hat{\varepsilon}_t^{\,2}} \qquad (2.7.2)$$

   Nilai statistik d, akan dibandingkan dengan nilai dU dan dL pada tabel Durbin-Watson.

4. Kriteria keputusan

   Jika d > dU, maka gagal tolak H_0

   Jika d< dL, maka H_0 ditolak

   Jika dL $\leq d \leq$ dU, maka tidak bisa ditarik kesimpulan

   dengan dU merupakan batas atas *Durbin Watson* dan dL merupakan batas bawah *Durbin Watson.*

2.7.3 Cara Mengatasi Autokorelasi

Penanganan autokorelasi dapat dilakukan dengan menggunakan metode *Two Stage Least Square* (TSLS). Apabila koefisien autokorelasi diketahui maka akan digunakan *Generalized Least Square* (GLS). *Generalized Least Square* (GLS) merupakan salah satu metode yang dapat digunakan untuk mengatasi masalah autokorelasi (Gujarati & Porter, 2010).

Menurut Gujarati (2003), metode iterasi *Cochrane-Orcutt* merupakan salah satu metode yang dapat digunakan untuk mengatasi adanya masalah pada regresi linear. Metode ini sangat efisien digunakan untuk mengatasi terjadinya pelanggaran asumsi autokorelasi yang terjadi pada model regresi (Wibowo, 2020). Metode iterasi *Cochrane-Orcutt* dilakukan dengan meregresikan residual ε_t dengan ε_{t-1} untuk AR (1) sehingga diperoleh nilai koefisien autokorelasi yang konstan. Sebagai ilustrasi, pertimbangkan model regresi berikut ini:

$$Y_t = \beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \cdots + \beta_p X_{p,t} + \varepsilon_t \qquad (2.7.3)$$

dimana ε_t mengikuti AR (1) dan nilai ρ telah diketahui.

$$\varepsilon_t = \rho\varepsilon_{t-1} + v_t \qquad (2.7.4)$$

Tulis kembali persamaan (2.7.3) untuk periode $t - 1$.

$$Y_{t-1} = \beta_0 + \beta_1 X_{1,t-1} + \beta_2 X_{2,t-1} + \cdots + \beta_p X_{p,t-1} + \varepsilon_{t-1} \qquad (2.7.5)$$

Selanjutnya, persamaan (2.7.5) dikalikan dengan ρ, sehingga didapat:

$$\rho Y_{t-1} = \rho\beta_0 + \rho\beta_1 X_{1,t-1} + \rho\beta_2 X_{2,t-1} + \cdots + \rho\beta_p X_{p,t-1} + \rho\varepsilon_{t-1} \qquad (2.7.6)$$

Kemudian eliminasi persamaan (2.7.3) dan (2.7.6).

$$Y_t - \rho Y_{t-1} = (1-\rho)\,\beta_0 + (X_t - \rho X_{t-1})\beta_1 + (X_{2t} - \rho X_{2t-1})\beta_2 + \cdots + \left(X_{pt} - \rho X_{pt-1}\right)\beta_p + (\varepsilon_t - \rho\varepsilon_{t-1}) \qquad (2.7.7)$$

Persamaan (2.7.7) dapat dituliskan sebagai berikut:

$$Y_t^* = \beta_0^* + \beta_p^* X_{p,t}^* + \varepsilon_t \qquad (2.7.8)$$

dengan,

$Y_t^* = Y_t - \rho Y_{t-1}$

$\beta_0^* = (1-\rho)\beta_0$

$\beta_p^* = \beta_p$

$X_{p,t}^* = X_{p,t} - \rho X_{p,t-1}$

$\varepsilon_t = \varepsilon_t - \rho\varepsilon_{t-1}$

2.8 *Generalized Two Stage Ridge Regression* (GTSRR)

Metode *Generalized Two Stage Ridge Regression* (GTSRR) merupakan metode alternatif yang dapat digunakan untuk mengatasi masalah multikolinearitas pada estimasi parameter model regresi linear. GTSRR merupakan kombinasi antara metode *Two Stages Least Squares* (TSLS) dan *Generalized Ridge Regression* (GRR). Metode *Generalized Ridge Regression* (GRR) merupakan pengembangan dari metode ridge, dimana pada metode ridge hanya dibutuhkan satu konstanta bias *k* untuk semua variabel independen, sedangkan pada Metode GRR dibutuhkan konstanta bias k sebanyak *p* variabel independen.

Menurut Utami, dkk. (2018), metode *Generalized Ridge Regression* adalah salah satu metode alternatif yang sangat efektif dalam mengatasi masalah multikolinearitas. Metode ini telah terbukti berhasil mengurangi efek multikolinearitas, seperti yang dapat dilihat dari nilai *Variance Inflation Factor* (VIF) yang lebih kecil dari 5 untuk setiap variabel independen. Dengan menggunakan *Generalized Ridge Regression*, dapat memperoleh estimasi parameter yang lebih stabil, serta meningkatkan interpretabilitas model dalam kasus adanya multikolinearitas.

Suatu persamaan regresi linear berganda dapat dilihat sebagai berikut:

$$Y_i^T = \beta_0 + \beta_1 X_{i1}^T + \beta_2 X_{i2}^T + \cdots + \beta_p X_{ip}^T + \varepsilon_i$$

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon_i} \quad (2.8.1)$$

Keterangan:

Y	:	Vektor pengamatan berukuran $n \times 1$
X	:	Matriks variabel independen berukuran $n \times p$
β	:	Vektor parameter berukuran $p \times 1$
ε_i	:	Vektor random sisaan

Varibael yang akan digunakan dalam analisis dengan menggunakan Metode *Generalized Two Stage Ridge Regression* (GTSRR) adalah variabel yang telah dilakukan transformasi. Misalkan P merupakan matriks ortogonal berordo p dengan kolom-kolomnya adalah vektor eigen dari matriks $X^T X$ maka:

$$\begin{aligned} \boldsymbol{\Lambda} &= \boldsymbol{P^T}(\boldsymbol{X^T X})\boldsymbol{P} \\ &= \boldsymbol{P^T X^T X P} \\ &= (\boldsymbol{XP})^{\boldsymbol{T}} \boldsymbol{XP} \\ &= (\boldsymbol{V})^{\boldsymbol{T}} \boldsymbol{V} \end{aligned} \quad (2.8.2)$$

Karena matriks P merupakan matriks ortogonal, dimana $P^T = P^{-1}$.

Maka, $P^T P = P P^T = I$.

Dengan demikian, persamaan regresi linear dapat ditulis sebagai berikut:

$$\begin{aligned}\boldsymbol{Y} &= \boldsymbol{X\beta} + \boldsymbol{\varepsilon}\\ &= \boldsymbol{X}(\boldsymbol{P^T P})\boldsymbol{\beta} + \boldsymbol{\varepsilon}\\ &= (\boldsymbol{XP})(\boldsymbol{P^T\beta}) + \boldsymbol{\varepsilon}\\ &= \boldsymbol{V\alpha} + \boldsymbol{\varepsilon}\end{aligned} \tag{2.8.3}$$

dimana, $V = XP$ dan $\alpha = P^T\beta$.

Dengan menggunakan Metode Kuadrat Terkecil (MKT), penduga bagi α adalah:

$$\widehat{\boldsymbol{\alpha}} = (\boldsymbol{V^T V})^{-1}\boldsymbol{V^T Y} \tag{2.8.4}$$

Selanjutnya persamaa (2.9.4) dapat dibentuk menjadi:

$$\begin{aligned}\widehat{\boldsymbol{\alpha}} &= (\boldsymbol{V^T V})^{-1}\boldsymbol{V^T Y}\\ &= ((\boldsymbol{XP})^{\boldsymbol{T}}\boldsymbol{XP})^{-1}(\boldsymbol{XP})^{\boldsymbol{T}}\boldsymbol{X}\widehat{\boldsymbol{\beta}}\\ &= (\boldsymbol{P^T X^T XP})^{-1}\boldsymbol{X^T P^T X}\widehat{\boldsymbol{\beta}}\\ &= (\boldsymbol{P^T X^T XP})^{-1}\boldsymbol{P^T X^T X}\widehat{\boldsymbol{PP^T}}\boldsymbol{\beta}\\ &= (\boldsymbol{P^T X^T XP})^{-1}(\boldsymbol{P^T X^T XP})\boldsymbol{P^T}\widehat{\boldsymbol{\beta}}\\ &= \boldsymbol{P^T}\widehat{\boldsymbol{\beta}}\end{aligned}$$

Dengan kedua ruas dikalikan dengan P, maka akan diperoleh:

$$\begin{aligned}\boldsymbol{P}\widehat{\boldsymbol{\alpha}} &= \boldsymbol{PP^T}\widehat{\boldsymbol{\beta}}\\ \widehat{\boldsymbol{\beta}} &= \boldsymbol{P}\widehat{\boldsymbol{\alpha}}\end{aligned} \tag{2.8.5}$$

Menurut Hoerl & Kennard (1970), dengan menggunakan pengganda *lagrange* (F), dimana $\hat{\alpha}(K)$ merupakan nilai yang meminimumkan fungsi dengan syarat $\hat{\alpha}(K)^{\,T}\hat{\alpha}(K) \leq c^2$ dimana *c* merupakan nilai konstanta. Maka akan diperoleh persamaan sebagai berikut:

$$\boldsymbol{F} = (\boldsymbol{\rho Y^*} - \boldsymbol{\rho X^*}\widehat{\boldsymbol{\alpha}}(\boldsymbol{K})\,)^{\boldsymbol{T}}\left(\boldsymbol{\rho Y^*} - \boldsymbol{\rho X^*}\widehat{\boldsymbol{\alpha}}(\boldsymbol{K})\right) + \boldsymbol{K}(\widehat{\boldsymbol{\alpha}}(\boldsymbol{K})^{\boldsymbol{T}}\widehat{\boldsymbol{\alpha}}(\boldsymbol{K}) - \boldsymbol{c^2})$$

$$\boldsymbol{F} = \left(\boldsymbol{Y^*}^{\boldsymbol{T}}\boldsymbol{\rho^T} - \widehat{\boldsymbol{\alpha}}(\boldsymbol{K})^{\boldsymbol{T}}\boldsymbol{X^*}^{\boldsymbol{T}}\boldsymbol{\rho^T}\,\right)(\boldsymbol{\rho Y^*} - \boldsymbol{\rho X^*}\widehat{\boldsymbol{\alpha}}(\boldsymbol{K})\,) + \boldsymbol{K}(\widehat{\boldsymbol{\alpha}}(\boldsymbol{K})^{\boldsymbol{T}}\widehat{\boldsymbol{\alpha}}(\boldsymbol{K}) - \boldsymbol{c^2})$$

$$\begin{aligned}\boldsymbol{F} = {} & \boldsymbol{Y^*}^{\boldsymbol{T}}\boldsymbol{\rho^T\rho Y^*} - \boldsymbol{Y^*}^{\boldsymbol{T}}\boldsymbol{\rho^T\rho X^*}\widehat{\boldsymbol{\alpha}}(\boldsymbol{K}) - \widehat{\boldsymbol{\alpha}}(\boldsymbol{K})^{\boldsymbol{T}}\boldsymbol{X^*}^{\boldsymbol{T}}\boldsymbol{\rho^T\rho Y^*} +\\ & \widehat{\boldsymbol{\alpha}}(\boldsymbol{K})^{\boldsymbol{T}}\boldsymbol{X^*}^{\boldsymbol{T}}\boldsymbol{\rho^T\rho X^*}\widehat{\boldsymbol{\alpha}}(\boldsymbol{K}) + \boldsymbol{K}(\widehat{\boldsymbol{\alpha}}(\boldsymbol{K})^{\boldsymbol{T}}\widehat{\boldsymbol{\alpha}}(\boldsymbol{K}) - \boldsymbol{c^2})\end{aligned}$$

Karena $\hat{\alpha}(K)^T X^{*^T} \rho^T \rho Y^*$ merupakan skalar, maka dengan menggunakan sifat $\left(\hat{\alpha}(K)^T X^{*^T} \rho^T \rho Y^*\right)^T = \hat{\alpha}(K) \rho X^* Y^{*^T} \rho^T$, sehingga persamaan tersebut menjadi:

$$\boldsymbol{F} = \boldsymbol{Y}^{*^T} \boldsymbol{\rho}^T \boldsymbol{\rho} \boldsymbol{Y}^* - 2\hat{\boldsymbol{\alpha}}(\boldsymbol{K})^T \boldsymbol{X}^{*^T} \boldsymbol{\rho}^T \boldsymbol{\rho} \boldsymbol{Y}^* + \hat{\boldsymbol{\alpha}}(\boldsymbol{K})^T \boldsymbol{X}^{*^T} \boldsymbol{\rho}^T \boldsymbol{\rho} \boldsymbol{X}^* \hat{\boldsymbol{\alpha}}(\boldsymbol{K}) + (\hat{\boldsymbol{\alpha}}(\boldsymbol{K})^T \hat{\boldsymbol{\alpha}}(\boldsymbol{K}) - \boldsymbol{c}^2) \tag{2.8.6}$$

Nilai minimum F diperoleh jika $\frac{\partial F}{\partial \hat{\alpha}(K)} = 0$, maka akan diperoleh persemaan sebagai berikut:

$$\begin{aligned}
-2\boldsymbol{X}^{*^T} \boldsymbol{\rho}^T \boldsymbol{\rho} \boldsymbol{Y}^* + 2\boldsymbol{X}^{*^T} \boldsymbol{\rho}^T \boldsymbol{\rho} \boldsymbol{X}^* \hat{\boldsymbol{\alpha}}(\boldsymbol{K}) + 2\boldsymbol{K}\left(\hat{\boldsymbol{\alpha}}(\boldsymbol{K})\right) &= \mathbf{0} \\
-\boldsymbol{X}^{*^T} \boldsymbol{\rho}^T \boldsymbol{\rho} \boldsymbol{Y}^* + \boldsymbol{X}^{*^T} \boldsymbol{\rho}^T \boldsymbol{\rho} \boldsymbol{X}^* \hat{\boldsymbol{\alpha}}(\boldsymbol{K}) + \boldsymbol{K}\left(\hat{\boldsymbol{\alpha}}(\boldsymbol{K})\right) &= \mathbf{0} \\
-\boldsymbol{X}^{*^T} \boldsymbol{\rho}^T \boldsymbol{\rho} \boldsymbol{Y}^* + \left(\boldsymbol{X}^{*^T} \boldsymbol{\rho}^T \boldsymbol{\rho} \boldsymbol{X}^* + \boldsymbol{K}\right)\left(\hat{\boldsymbol{\alpha}}(\boldsymbol{K})\right) &= \mathbf{0} \\
\left(\boldsymbol{X}^{*^T} \boldsymbol{\rho}^T \boldsymbol{\rho} \boldsymbol{X}^* + \boldsymbol{K}\right)\left(\hat{\boldsymbol{\alpha}}(\boldsymbol{K})\right) &= \boldsymbol{X}^{*^T} \boldsymbol{\rho}^T \boldsymbol{\rho} \boldsymbol{Y}^* \\
\hat{\boldsymbol{\alpha}}(\boldsymbol{K}) &= \left(\boldsymbol{X}^{*^T} \boldsymbol{\rho}^T \boldsymbol{\rho} \boldsymbol{X}^* + \boldsymbol{K}\right)^{-1} \boldsymbol{X}^{*^T} \boldsymbol{\rho}^T \boldsymbol{\rho} \boldsymbol{Y}^* \\
\hat{\boldsymbol{\alpha}}(\boldsymbol{K}) &= \left(\boldsymbol{X}^{*^T} \boldsymbol{\Omega} \boldsymbol{X}^* + \boldsymbol{K}\right)^{-1} \boldsymbol{X}^{*^T} \boldsymbol{\Omega} \boldsymbol{Y}^* \qquad (2.8.7)
\end{aligned}$$

Dimana *K* merupakan matriks diagonal dengan diagonal utama berisi nilai konstanta $k_1,\ k_2, \ldots, k_p$.

$$K = \begin{bmatrix} k_1 & 0 & \cdots & 0 \\ 0 & k_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & k_p \end{bmatrix} \quad , \quad k > 0$$

Jadi, penduga bagi metode *Generalized Two Stage Ridge Regression* (GTSRR) sebagai berikut:

$$\begin{aligned}
\hat{\boldsymbol{\beta}}_{\boldsymbol{GTSRR}} &= \hat{\boldsymbol{\alpha}}(\boldsymbol{K}) \\
\hat{\boldsymbol{\beta}}_{\boldsymbol{GTSRR}} &= \left(\boldsymbol{X}^{*^T} \boldsymbol{\Omega} \boldsymbol{X}^* + \boldsymbol{K}\right)^{-1} \boldsymbol{X}^{*^T} \boldsymbol{\Omega} \boldsymbol{Y}^* \qquad (2.8.8)
\end{aligned}$$

2.9 Penentuan Nilai Konstanta Bias K

Menurut Hoerl & Kennard (1970), untuk memilih nilai konstanta bias (K_j) pada metode *Generalized Two Stage Ridge Regression* (GTSRR) dengan menggunakan rumus sebagai berikut:

$$K_j = \frac{\hat{\sigma}^2}{\hat{\alpha}_j^{\,2}} \tag{2.9.1}$$

dengan *j=1, 2, …, p*

Keterangan:

K_j	:	Nilai konstanta bias untuk parameter ke-j
$\hat{\sigma}^2$	:	Kuadrat tengah sisaan dari persamaan $(2.7.3)$
$\hat{\alpha}_j$	:	Nilai penduga parameter ke-j yang didapatkan dari persamaan $(2.7.4)$

2.10 *Mean Squared Error* (MSE) dan *Average Mean Squared Error* (AMSE)

Mean Squared Error adalah salah satu cara untuk mengukur kesalahan yang terjadi pada metode yang digunakan. MSE merupakan rata-rata kuadrat kesalahan antara nilai yang diramalkan dan nilai yang diamati (Ghazali, 2017). Apabila peneliti menggunakan data simulasi yang direplikasi maka salah satu cara yang digunakan untuk membandingkan hasil prediksi yaitu AMSE. Menurut Montgomery, *et al*. (2012), semakin kecil nilai *Mean Squared Error* (MSE) dan *Average Mean Squared Error* (AMSE) menunjukkan bahwa suatu metode atau model prediksi semakin baik. MSE mengukur rata-rata dari kuadrat kesalahan antara nilai yang diramalkan dan nilai yang diamati. Semakin kecil nilai MSE, semakin dekat nilai prediksi dengan nilai sebenarnya, menunjukkan kualitas yang lebih baik dalam melakukan prediksi atau estimasi. AMSE, yang merupakan rata-rata dari MSE di seluruh pengamatan yang berguna untuk mengevaluasi kinerja secara keseluruhan. Semakin kecil nilai AMSE, semakin akurat dan baik metode atau model dalam meramalkan atau memprediksi secara umum. Dalam prakteknya, MSE dan AMSE sering digunakan untuk membandingkan dan memilih model prediksi yang terbaik di antara beberapa model yang tersedia. Model dengan nilai MSE dan AMSE terendah akan dianggap sebagai model yang paling baik atau optimal. Nilai MSE dan AMSE dapat dihitung dengan menggunakan persamaan sebagai berikut:

$$MSE = \sum_{i=1}^{n} \frac{(\hat{Y}_i - Y_i)^2}{n} \qquad , i = 1,2,\ldots,n \qquad (2.10.1)$$

$$AMSE = \sum_{i=1}^{100} \frac{MSE}{m} \qquad , m = 100 \qquad (2.10.2)$$

dengan:

n	:	Jumlah data
$\hat{Y}_i$	:	Nilai yang diramalkan
Y_i	:	Nilai yang diamati
m	:	Banyaknya ulangan (replikasi)

III. METODOLOGI PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada semester genap tahun akademik 2022/2023 di Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung.

3.2 Data Penelitian

Dalam penelitian ini, akan menggunakan data simulasi yang dihasilkan menggunakan perangkat lunak R-Studio. Pada penelitian ini data yang akan digunakan adalah data simulasi dengan menggunakan *software R-Studio*. Data yang dimaksud kemudian didesain untuk memenuhi asumsi multikolinearitas dan autokorelasi pada GTSRR. Adapun jumlah sampel yang akan digunakan adalah $n = 50, 75, 100$ dan variabel independen yang akan digunakan sebanyak 6 variabel dengan tingkat korelasi antar variabel independen sebesar 0.99 serta terdapat autokorelasi sebesar 0.15 dengan nilai $\beta_1 = \beta_2 = \cdots = \beta_p = 1$.

Data simulasi yang dimaksudkan diatas pada setiap himpunan data X_{ij} dibangkitkan dengan menggunakan persamaan sebagai berikut:

$$X_{ij} = (1-\rho^2)^{1/2} Z_{ij} + \rho Z_{i(p+1)} \tag{3.2.1}$$

dimana $i = 1, 2, .., n$, $j = 1, 2, .., p$, dan $Z_{ij} \sim N(0, 1)$ serta $\rho = 0.99$.

Simulasi data X_{ij} sebagai berikut:

$X_{i1} = (1-\rho^2)^{1/2} z_{i1} + \rho z_{i7}$

$X_{i2} = (1-\rho^2)^{1/2} z_{i2} + \rho z_{i7}$

$X_{i3} = (1-\rho^2)^{1/2} z_{i3} + \rho z_{i7}$

$X_{i4} = (1-\rho^2)^{1/2} z_{i4} + \rho z_{i7}$

$X_{i5} = (1-\rho^2)^{1/2} z_{i5} + \rho z_{i7}$

$X_{i6} = (1-\rho^2)^{1/2} z_{i6} + \rho z_{i7}$

Sedangkan pengaruh autokorelasi pada model deret waktu ditetapkan menyesuaikan persamaan AR(2), dengan variabel dependen yang saling berkorelasi sesuai dengan persamaan (3.2.1) sebagai berikut:

$$\tilde{Y} = X_{i1} + X_{i2} + X_{i3} + X_{i4} + X_{i5} + X_{i6} \qquad ; i = 1, 2, .., n \qquad (3.2.2)$$

Kemudian membangkitkan galat berdistribusi normal $\varepsilon_t \sim N(0,1)$.
Untuk memodelkan data dengan masalah multikolinearitas dan autokorelasi, selanjutnya dibangkitkan data menggunakan persamaan AR(2) dengan $\varphi_1 = 1$ dan $\varphi_2 = 1$, sebagai berikut:

$$Y_t = \varphi_0 + \varphi_1 \tilde{Y}_{t-1} + \varphi_2 \tilde{Y}_{t-2} + \varepsilon_t \qquad (3.2.3)$$

Untuk kelayakan suatu model regresi dapat dilihat dari nilai MSEdan AMSE. Nilai MSE dan AMSE didapatkan dari persamaan berikut:

$$MSE = \sum_{i=1}^{n} \frac{(\hat{Y}_i - Y_i)^2}{n} \qquad , i = 1,2, \dots, n \qquad , n = 50, 75, 100$$

$$AMSE = \sum_{i=1}^{100} \frac{MSE}{m} \qquad , m = 100$$

dengan m merupakan banyaknya ulangan (replikasi) dalam simulasi.

3.3 Metode Penelitian

Untuk mengetahui data bangkitan memenuhi asumsi multikolinearitas dan autokorelasi akan dilakukan uji nilai VIF dan uji *Durbin Watson* sebelum simulasi menggunakan MKT dan setelah simulasi menggunakan GTSRR.

Data pada penelitian ini diolah dengan bantuan *software* R. Adapun langkah-langkah simulasi sebagai berikut:

1. Membangkitkan data simulasi $n = 50,75,100$.
2. Melakukan pendugaan parameter menggunakan iterasi *Cochrane Orcutt Iterative Procedure* untuk mengatasi masalah autokorelasi.
3. Melakukan analisis menggunakan *Generalized Two Stage Ridge Regression* (GTSRR) dengan langkah-langkah sebagai berikut:
   a. Menentukan nilai konstanta bias k.
   b. Menentukan penduga β GTSRR.
4. Menghitung nilai MSE.
5. Mengulang langkah 1 sampai langkah 4 sebanyak 100 kali.
6. Menghitung nilai AMSE.
7. Menganalisis hasil simulasi.

Untuk melihat kelayakan model tersebut maka dilakukan uji efektivitas pada salah satu data bangkitan dengan melihat nilai MSE dan AMSE.

V. KESIMPULAN

Kesimpulan dari penelitian ini adalah sebagai berikut:

1. Metode GTSRR baik digunakan untuk mengatasi adanya masalah multikolinearitas dan autokorelasi pada data simulasi deret waktu dengan $n = 50, 75, 100,$ dibuktikan dengan nilai VIF setelah dilakukan analisis dengan metode GTSRR kurang dari 5, nilai *Durbin Watson* yang lebih besar dari batas atas pada tabel *Durbin Watson*, dan nilai MSE dan AMSE yang semakin kecil setelah dilakukan analisis dengan menggunakan metode GTSRR.
2. Berdasarkan hasil yang diperoleh dari ukuran sampel $n = 100$ lebih baik dari ukuran sampe $n = 50$ dan $75.$ Dibuktikan dengan nilai MSE dan AMSE yang semakin kecil sehingga dalam menentukan nilai pemodelan dari data yang mengandung masalah multikolinearitas dan autokorelasi akan semakin baik.

DAFTAR PUSTAKA

Aprianto, A., Debataraja, N. N., & Imro'ah, N. 2020. Metode Cochrane-Orcutt Untuk Mengatasi Autokorelasi Pada Estimasi Parameter Ordinary Least Squares. *Buletin Ilmiah Mat, Stat, dan Terapannya (Bimaster)*, **9** (1): 95-102.

Draper, N. R., & Smith, H. 1992. *Applied Regression Analysis*. 3th Edition. John Wiley and Sons, New York.

Eledum, H. Y. A., & Alkhaifa, A. A. 2012. Generalized Two Stages Ridge regression Estimator GTR for Multicollinearity and Autocorrelated Errors. *Canadian Journal on Science and Engineering Mathematics,* **3** (3): 79-85.

Gujarati, D. N. 2003. *Basic Econometrics*. 4th Edition. McGraw-Hill, New York.

Gujarati, D. N., & Porter, D. C. 2009. *Basic Econometrics*. 5th Edition. McGraw-Hill, New York.

Hoerl, A. E., & Kennard, R. W. 1970. Ridge Regression Biased Estimation For Non Ortogonal Problems. *Technometrics*, **12** (1): 55-67.

Kutner, M. H., Neter, J., & Wasserman, W. 1983. *Applied Linear Regression Models*. Richard D. Irwin, United States.

Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. 2005. *Applied Linear Statistical Models*. 5th Edition. McGraw-Hill, New York.

Montgomery, D. C., Peck, E. A., & Vining, G. G. 2012. *Introduction to Linear Regression Analysis*. 5th Edition. John Wiley and Sons, New York.

Montgomery, D.C. 2008. *Introduction to Time Series Analysis and Forecasting.* John Wiley & Sons. Inc., New Jersey.

Rosadi, D. 2021. *Analisis Runtun Waktu*. UGM, Yogyakarta.

Sembiring, R. K. 2003. *Analisis Regresi*. Edisi ke-2. ITB, Bandung.

Utami, N. K. T., Sukarsa, I. K. G., & Kencana, I. P. E. N. 2018. Penerapan Metode Generalized Ridge Regression dalam Mengatasi Masalah Multikolinearitas. *E-Jurnal Matematika*, **2** (1): 54-59.

Wibowo, A. 2020. Perbaikan Asumsi Autokorelasi Menggunakan metode Cochrane-Orcutt (Studi Apakah Minyak Dan Gas Merupakan Berkah atau Musibah). *Jurnal Matematika dan Pendidikan Matematika*, **5** (1): 23-30.