PERBANDINGAN PERFORMA *PRINCIPAL COMPONENT REGRESSION* (PCR) DAN *LATENT ROOT REGRESSION* (LRR) DALAM MENANGANI MASALAH MULTIKOLINEARITAS

(Skripsi)

Oleh

SISKA AZZIRA AMANDA

JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2025

ABSTRACT

COMPARISON THE PERFORMANCE OF PRINCIPAL COMPONENT REGRESSION (PCR) AND LATENT ROOT REGRESSION (LRR) FOR HANDLING MULTICOLLINEARITY PROBLEMS

By

SISKA AZZIRA AMANDA

Regression analysis is a method used to model the relationship between independent variables and a dependent variable. One of the assumptions that must be met in regression analysis is the absence of a relationship among the independent variables, which is referred to as multicollinearity. This study aims to examine the application of the Principal Component Regression (PCR) and Lasso Ridge Regression (LRR) methods in addressing multicollinearity under various data conditions. The study uses simulated data with sample sizes of (n = 20, 60, 100) and correlation coefficients (ρ = 0.3, 0.6, 0.9, 0.95), repeated 1000 times. The results of this study show that PCR performs better in handling multicollinearity, as it produces the smallest Mean Squared Error (MSE) compared to LRR.

Keyword: Principal Component Regression, Latent Root Regression, Multicollinearity

ABSTRAK

PERBANDINGAN PERFORMA *PRINCIPAL COMPONENT REGRESSION* (PCR) DAN *LATENT ROOT REGRESSION* (LRR) DALAM MENANGANI MASALAH MULTIKOLINEARITAS

Oleh

SISKA AZZIRA AMANDA

Analisis regresi adalah metode yang mampu memodelkan hubungan yang terjadi antara variabel bebas dan variabel terikatnya. Asumsi yang harus dipenuhi dalam analisis regresi adalah tidak adanya hubungan antara variabel bebas disebut sebagai multikolinearitas. Penelitian ini bertujuan untuk mengkaji penerapan metode PCR dan LRR dalam menangani multikolinearitas pada beberapa kondisi data. Penelitian ini menggunakan data simulasi dengan ukuran sampel $(n = 20, 60, 100)$ serta koefisien korelasi $(\rho = 0.3, 0.6, 0.9, 0.95)$ dan diulang sebanyak 1000 kali. Hasil pada penelitian ini menunjukkan bahwa PCR lebih baik dalam mengatasi masalah multikolinearitas karena menghasilkan nilai MSE terkecil dibandingkan dengan LRR.

Kata kunci: *Principal Component Regression*, *Latent Root Regression*, Multikolinearitas

PERBANDINGAN PERFORMA *PRINCIPAL COMPONENT REGRESSION* (PCR) DAN *LATENT ROOT REGRESSION* (LRR) DALAM MENANGANI MASALAH MULTIKOLINEARITAS

Oleh
SISKA AZZIRA AMANDA
2117031099

Skripsi
Sebagai Salah Satu Syarat untuk Mencapai Gelar
SARJANA MATEMATIKA

Pada

Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Lampung

JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2025

Judul Skripsi : **PERBANDINGAN PERFORMA *PRINCIPAL COMPONENT REGRESSION* (PCR) DAN *LATENT ROOT REGRESSION* (LRR) DALAM MENANGANI MASALAH MULTIKOLINEARITAS**

Nama Mahasiswa : Siska Azzira Amanda

Nomor Pokok Mahasiswa : **2117031099**

Program Studi : **Matematika**

Fakultas : **Matematika dan Ilmu Pengetahuan Alam**

MENYETUJUI

1. Komisi Pembimbing

Widiarti, S.Si., M.Si.
NIP. 198005022005012003

Dina Eka Nurvazly, S.Pd., M.Si.
NIP. 199311062019032018

2. Ketua Jurusan Matematika

Dr. Aang Nuryaman, S.Si., M.Si.
NIP. 197403162005011001

MENGESAHKAN

1. Tim Penguji

Ketua	: **Widiarti, S.Si., M.Si.**	………
Sekretaris	: **Dina Eka Nurvazly, S.Pd., M.Si.**	………
Penguji Bukan Pembimbing	: **Prof. Drs. Mustofa Usman, M.A., Ph.D.**	………

2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

Dr. Eng. Heri Satria, S.Si., M.Si.
NIP. 197110012005011002

Tanggal Lulus Ujian Skripsi : **22 Mei 2025**

PERNYATAAN SKRIPSI MAHASISWA

Yang bertanda tangan di bawah ini :

Nama	: **Siska Azzira Amanda**
Nomor Pokok Mahasiswa	: **2117031099**
Jurusan	: **Matematika**
Judul Skripsi	: **PERBANDINGAN PERFORMA *PRINCIPAL COMPONENT REGRESSION* (PCR) DAN *LATENT ROOT REGRESSION* (LRR) DALAM MENANGANI MASALAH MULTIKOLINEARITAS**

Dengan ini menyatakan bahwa penelitian ini adalah hasil pekerjaan saya sendiri. Apabila di 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, 22 Mei 2025

Yang menyatakan

Siska Azzira Amanda

NPM. 2117031099

RIWAYAT HIDUP

Penulis bernama lengkap Siska Azzira Amanda dilahirkan di Bekasi pada tanggal 07 Juni 2004. Penulis merupakan anak kedua dari pasangan Bapak Johana dan Ibu Mardianah.

Penulis mengawali pendidikan Taman kanak-kanak di TK Dinul Islam pada tahun 2008 – 2009, Sekolah Dasar di SDIT Assalam pada tahun 2009 – 2015, Sekolah Menengah Pertama di MTs YPPA Cipulus pada tahun 2015 – 2018 dan melanjutkan Sekolah Menengah Atas di MA YPPA Cipulus pada tahun 2018 – 2021. Setelah itu, penulis diterima sebagai mahasiswi Program Studi S1 Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung melalui jalur Seleksi Bersama Masuk Perguruan Tinggi Negeri (SBMPTN) pada tahun 2021.

Selama masa perkuliahan, penulis aktif di beberapa kegiatan di antaranya aktif dalam kepengurusan organisasi HIMATIKA FMIPA Unila sebagai Anggota Biro Dana dan Usaha pada tahun 2022, menjadi Bendahara Pelaksan dalam kepanitiaan Dies Natalis Jurusan Matematika ke-24 (DINAMIKA XXIV) pada tahun 2023, dan melakukan kerja praktik di Badan Perencanaan Pembangunan dan Penelitian Pengembangan Daerah Kota Bekasi.

KATA INSPIRASI

"Karena sesungguhnya sesudah kesuliatn itu ada kemudahan, sesungguhnya setelah kesulitan ada kemudahan."

(Q.S. Al Insyirah: 5-6)

"Allah tidak akan membebani seseoran melainkan sesuai dengan kesanggupannya."

(Q.S. Al Baqarah: 286)

"Dan barang siapa yang bertakwa kepada Allah, niscaya Allah akan menjadikan baginya kemudahan dalam urusannya."

(Q.S. At Talaq: 4)

"Nobody comes to save you now. You've got something they don't. You've just gotta keep your eyes open."

(Taylor Swift)

PERSEMBAHAN

Dengan megucapkan Alhamdulillah, puji dan syukur kehadirat Allah SWT yang telah melimpahkan rahmat serta hidayah-Nya sehingga skripsi ini dapat diselesaikan dengan baik. Tak lupa sholawat serta salam selalu tercurahkan kepada junjungan besar Nabi Muhammad SAW. Dengan penuh ketulusan hati, penulis mempersembahkan karya ini untuk:

Kedua Orang Tua

Cinta pertamaku Ayahanda Johana serta pintu surgaku Ibunda Mardianah, yang selalu menjadi sumber semangat. Tiada henti kasih sayang dan dukungan yang kalian berikan dengan penuh cinta serta doa-doa tulus, yang menjadi kekuatan utama dalam setiap langkah perjalanan ini. Terima kasih atas setiap doa, dukungan, kasih sayang tanpa batas, serta setiap pengorbanan yang tidak terhitung.

Kedua saudara kandung

Muhammad Aziz Kurniawan dan Muhammad Aditya Fajar Nugraha

Dosen Pembimbing dan Pembahas

Terima kasih kepada dosen pembimbing dan pembahas yang telah membantu, memberikan motivasi, memberikan kritik dan saran serta ilmu yang sangat amat berharga.

Keluarga Besar dan Sahabat Terbaik

Almamater tercinta, Universitas Lampung

SANWACANA

Puji syukur atas kehadirat Allah SWT atas segala limpahan rahmat, karunia serta nikmat hidayah yang telah diberikan-Nya, sehingga penulis dapat menyelesaikan skripsi dengan judul **"Perbandingan Performa *Principal Component Regression* (PCR) dan *Latent Root Regression* (LRR) dalam Menangani Masalah Multikolinearitas"**. Shalawat serta salam juga tak lupa selalu tercurahkan kepada junjungan kita nabi agung, Nabi Muhammad SAW yang telah membawa kita meninggalkan zaman jahiliyah yang gelap gulita menuju zaman yang terang benderang dan semoga kita mendapatkan syafaatnya di yaumul akhir kelak. Aamiin. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Matematika pada Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

Dalam penyusunan skripsi ini tentunya tidak terlepas dari dukungan, bimbingan, serta masukan dari berbagai pihak yang telah membantu. Untuk itu penulis ingin mengucapkan rasa terima kasih kepada:

1. Ibu Widiarti, S.Si., M.Si., selaku Dosen Pembimbing Akademik dan Dosen Pembimbing I, yang senantiasa memberikan bimbingan dengan penuh kesabaran, memberikan bantuan dan arahan yang bijaksana, motivasi yang tidak pernah pudar, serta saran yang membangun sehingga penulis dapat menyelesaikan skripsi ini.
2. Ibu Dina Eka Nurvazly, S.Pd., M.Si., selaku Dosen Pembimbing II atas bantuan dan bimbingannya yang begitu berarti dalam penyusunan skripsi ini.
3. Bapak Prof. Drs. Mustofa Usman, M.A., Ph.D., selaku Dosen Pembahas yang telah memberikan kritik dan saran yang membangun selama proses penyusunan skripsi ini.

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, staff, karyawan Jurusan Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
7. Bapak, Mamah, Aa, Ade, Nenek, Teh Ria, Daffa, dan Salsa yang selalu mendoakan dengan tulus, tiada henti memberikan dukungan, motivasi, pengorbanan, serta kasih sayang tiada tara demi kesuksesan penulis.
8. Sahabat-sahabat terbaik penulis (Nesa, Margel, Vara, Arsie, Yunda Callista, Yunda Claudya, dan Yunda Asti) yang senantiasa mendukung dan hadir dalam setiap proses yang penulis lalui.
9. Teman-teman seperbimbingan yang telah membersamai berjuang saat bimbingan berlangsung.
10. Seluruh pihak terkait yang tidak dapat disebutkan satu persatu yang telah membantu penulis dalam menyelesaikan skripsi ini.

Penulis menyadari skripsi ini masih jauh dari kesempurnaan, sehingga masih terdapat banyak kekurangan baik dalam penyajian maupun pada penulisan yang murni karena ketidaksempurnaan penulis. Oleh karena itu, penulis mengharapkan saran dan kritik yang membangun agar lebih baik.

Bandar Lampung, 22 Mei 2025
Penulis

Siska Azzira Amanda

DAFTAR ISI

DAFTAR TABEL

Tabel Halaman

DAFTAR GAMBAR

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

Analisis regresi adalah metode yang mampu memodelkan hubungan yang terjadi antara variabel bebas dan variabel terikatnya (Kadir, 2019). Variabel-variabel regresi yang berhubungan secara linear disebut regresi linear. Regresi linear yang menghubungkan satu variabel terikat dengan satu variabel bebas disebut dengan regresi linear sederhana, sedangkan regresi linear yang menghubungkan satu variabel terikat dengan dua atau lebih variabel bebas disebut regresi linear berganda (Kutner, *et al*., 2005). Analisis regresi berganda merupakan suatu metode statistik yang digunakan untuk menguji adanya hubungan variabel bebas dan variabel terikat. Analisis regresi dalam bentuk linear dengan lebih dari satu variabel bebas disebut dengan analisis regresi linear berganda (Kartiningrum, dkk, 2022).

Asumsi yang harus dipenuhi dalam analisis regresi adalah tidak adanya hubungan antara variabel bebas disebut sebagai multikolinearitas. Multikolinearitas dapat dideteksi dengan nilai VIF. Jika nilai $VIF > 10$ maka terjadi multikolinearitas pada data. Gejala multikolinearitas menimbulkan masalah dalam model regresi. Korelasi antar variabel bebas yang sangat tinggi menghasilkan penduga model regresi yang berbias, tidak stabil, dan mungkin jauh dari nilai prediksinya. Beberapa metode yang dapat digunakan untuk menangani masalah multikolinearitas yaitu *Principal Component Regression* (PCR), *Latent Root Regression* (LRR), *Ridge Regression*, *Partial Least Square* (PLS), dan *Least Absolute Shrinkage and Selection Operator* (LASSO) (Draper & Smith, 1998).

Principal Component Regression (PCR) adalah analisis regresi dari variabel-variabel terikat terhadap komponen-komponen utama yang tidak saling berkorelasi, dengan setiap komponen utama merupakan kombinasi linear dari semua variabel bebas. PCR digunakan untuk meminimumkan masalah multikolinearitas tanpa harus mengeluarkan variabel bebas yang terindikasi berkorelasi (Draper & Smith, 1998). *Latent Root Regression* (LRR) adalah metode perluasan dari metode regresi komponen utama (*Principal Component Regression*) yang menyatukan matriks data dari variabel bebas dan variabel terikatnya yang sudah dibakukan. Komponen utama ini yang akan diregresikan dengan variabel terikat sehingga korelasinya dapat dihilangkan dan masalah multikolinearitas dapat teratasi (Untari & Susanti, 2017).

Penelitian terkait PCR dan LRR yang telah dilakukan sebelumnya yaitu penelitian dari Candra, dkk. (2024), yang mengkaji tentang perbandingan antara *Latent Root Regression* dan *Ridge Regression* dalam mengatasi multikolinearitas. Hasil penelitian ini menunjukkan bahwa *Latent Root Regression* dapat mengatasi masalah multikolinearitas dengan lebih baik daripada *Ridge Regression* dengan membandingkan nilai VIF karena menghasilkan nilai VIF sama dengan 1. Hidayatullah, dkk. (2024), mengkaji tentang analisis regresi komponen utama dan penelitian ini menunjukkan bahwa *Principal Component Regression* mampu menghasilkan model terbaik dengan nilai koefisien determinasi (R^2) sebesar $0{,}834$. Amelia & Putra (2023) mengkaji tentang Regresi Komponen Utama dalam mengatasi multikolinearitas pada faktor yang memengaruhi pendapatan asli daerah di Sumatera Barat. Hasil penelitian ini menunjukkan bahwa komponen utama $W1$ dan $W2$ dapat menjelaskan 30% dan $16{,}9\%$ dari keragaman data. Alfah, dkk. (2022) melakukan penelitian mengenai teknologi NIRS untuk memprediksi jumlah pencampuran minyak sawit dalam minyak nilam menggunakan metode PCR. Hasil penelitian ini menunjukkan bahwa model terbaik adalah *Principal Component Regression* (PCR) dengan *pretreatment derivative* 1 ($D1$), dan model ini memiliki nilai koefisien determinasi (R^2) sebesar $0{,}92$. Sahriman & Yulianti (2023) melakukan penelitian mengenai penerapan model *Statistical Downscaling* (SD) dengan PCR dan LRR untuk peramalan curah hujan di Kabupaten Pangkep. Hasil penelitian ini menunjukkan bahwa model LRR lebih baik dalam menjelaskan keragaman data curah hujan Kabupaten Pangkep dibandingkan model PCR berdasarkan nilai korelasi tertinggi sebesar 0,97.

Berdasarkan uraian di atas, penulis tertarik untuk membandingkan metode *Principal Component Regression* (PCR) dan *Latent Root Regression* (LRR). Perbandingan kedua metode tersebut dilakukan secara simulasi dengan berbagai kondisi yaitu ukuran data, jumlah variabel, dan besar korelasi yang berbeda-beda.

1.2 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah untuk mengkaji penerapan metode *Principal Component Regression* (PCR) dan *Latent Root Regression* (LRR) dalam menangani multikolinearitas pada beberapa kondisi data.

1.3 Manfaat Penelitian

Manfaat dari penelitian ini adalah mendapatkan informasi mengenai metode dengan performa terbaik dari metode *Principal Component Regression* (PCR) dan *Latent Root Regression* (LRR) dalam menangani multikolinearitas.

II. TINJAUAN PUSTAKA

2.1 Analisis Regresi

Analisis regresi adalah metode statistik yang digunakan untuk menilai hubungan antara satu atau lebih variabel bebas dan variabel terikat. Tujuan utama analisis regresi adalah untuk memahami dan mengukur sejauh mana perubahan dalam satu variabel dapat memprediksi perubahan dalam variabel lainnya. Hasil analisis regresi dapat memberikan informasi tentang kekuatan, arah, dan signifikansi statistik dari hubungan antar variabel tersebut (Montgomery, *et al*., 2012).

2.2 Analisis Regresi Berganda

Analisis regresi dalam bentuk linear dengan lebih dari satu variabel bebas disebut dengan analisis regresi linear berganda (Kartiningrum, dkk., 2022). Menurut Draper & Smith (1998) model regresi linear berganda dapat dituliskan sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + \varepsilon_i;\ i = 1,2,\ldots,n \tag{2.1}$$

Keterangan:

Y_i : Variabel terikat.

β_0 : Parameter konstanta/intersep regresi yang tidak diketahui nilainya.

β_k : Parameter koefisien regresi yang tidak diketahui nilainya.

X_{ki} : Variabel bebas.

ε_i : Variabel galat.

n : Banyaknya amatan.

k : Banyaknya variabel bebas.

2.3 Estimasi Parameter

Menurut Kutner, *et al*. (2005) salah satu metode estimasi parameter yang digunakan dalam regresi linear adalah metode kuadrat terkecil atau sering juga disebut dengan metode *Ordinary Least Square* (OLS). Tujuan dari metode OLS adalah untuk mengestimasi parameter dengan meminimalkan Jumlah Kuadrat Galat (JKG) yang dapat dituliskan dalam persamaan berikut:

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \tag{2.2}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{1,1} & \cdots & x_{1,k} \\ 1 & x_{2,1} & \cdots & x_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n,1} & \cdots & x_{n,k} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Keterangan:

$\boldsymbol{Y}$: Vektor variabel terikat berukuran $(n \times 1)$.

$\boldsymbol{X}$: Matriks variabel bebas berukuran $(n \times p)$.

$\boldsymbol{\beta}$: Vektor pendugaan parameter berukuran $(p \times 1)$.

Berdasarkan persamaan (2.2) nilai duga $\boldsymbol{\beta}$ dapat dihitung menggunakan metode *Ordinary Least Square* dengan meminimumkan jumlah kuadrat *error*, langkah pertama yaitu mendapatkan nilai sisaan dengan persamaan:

$$\boldsymbol{\varepsilon} = \boldsymbol{Y} - \boldsymbol{X\beta} \tag{2.3}$$

Kemudian dihitung jumlah kuadrat sisaan sebagai berikut:

$$\begin{aligned} \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{\varepsilon}_i^2 &= \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\boldsymbol{Y} - \boldsymbol{X\beta})'(\boldsymbol{Y} - \boldsymbol{X\beta}) \\ &= \boldsymbol{Y}'\boldsymbol{Y} - \boldsymbol{X}'\boldsymbol{\beta}'\boldsymbol{Y} - \boldsymbol{X\beta}\boldsymbol{Y}' + \boldsymbol{X}'\boldsymbol{\beta}'\boldsymbol{X\beta} \\ &= \boldsymbol{Y}'\boldsymbol{Y} - 2\boldsymbol{X}'\boldsymbol{\beta}'\boldsymbol{Y} + \boldsymbol{X}'\boldsymbol{\beta}'\boldsymbol{X\beta} \end{aligned} \tag{2.4}$$

Selanjutnya meminimumkan persamaan jumlah kuadrat sisaan dengan menurunkan secara parsial terhadap parameter $\boldsymbol{\beta}$ dan disamakan dengan nol. Berikut adalah persamaan jumlah kuadrat sisaan yang diturunkan:

$$\begin{aligned}
\frac{\partial \sum_{i=1}^{n} \boldsymbol{\varepsilon}_i^2}{\partial \boldsymbol{\beta}} = \frac{\partial(\boldsymbol{Y}'\boldsymbol{Y} - 2\boldsymbol{X}'\boldsymbol{\beta}'\boldsymbol{Y} + \boldsymbol{X}'\boldsymbol{\beta}'\boldsymbol{X}\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} &= 0 \\
-2\boldsymbol{X}'\boldsymbol{Y} + 2\boldsymbol{X}'\boldsymbol{X}\widehat{\boldsymbol{\beta}} &= 0 \\
(\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{X}\widehat{\boldsymbol{\beta}} &= (\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{Y} \\
\boldsymbol{I}\widehat{\boldsymbol{\beta}} &= (\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{Y} \\
\widehat{\boldsymbol{\beta}} &= (\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{Y}
\end{aligned} \tag{2.5}$$

Sehingga didapat persamaan $\widehat{\boldsymbol{\beta}}$ seperti berikut:

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{X}'\boldsymbol{Y} \tag{2.6}$$

Keterangan:

$\widehat{\boldsymbol{\beta}}$: Vektor pendugaan parameter berukuran $(p \times 1)$.

$\boldsymbol{X}$: Matriks variabel bebas berukuran $(n \times p)$.

$\boldsymbol{Y}$: Matriks variabel terikat berkuran $(n \times 1)$.

2.4 Multikolinearitas

Ragnar Frisch pertama kali memperkenalkan istilah multikolinearitas pada tahun 1934, yang berarti adanya korelasi di antara variabel-variabel bebas dari model regresi. Multikolinearitas adalah terjadinya hubungan linear antara variabel bebas dalam suatu model regresi linear berganda (Gujarati & Porter, 2009). Menurut Best & Wolf (2005), multikolinearitas adalah kondisi di mana terdapat korelasi yang signifikan antara satu atau lebih pasang variabel bebas. Ketika terdapat korelasi yang kuat antara variabel bebas, hal ini dapat menyebabkan masalah dalam analisis regresi linear. Multikolinearitas dapat dideteksi dengan melihat nilai *Variance Inflation Factor* (VIF). Jika nilai $VIF > 10$ maka terjadi multikolinearitas pada data. Statistik uji untuk VIF adalah:

$$VIF_j = \frac{1}{1 - R_j^2} \tag{2.7}$$

Keterangan:

R_j^2 : Koefisien determinasi antar variabel bebas.

2.5 *Principal Component Regression* (PCR)

Terdapat dua tahap dalam *Principal Component Regression* yaitu melakukan *principal component analysis* dan melakukan regresi variabel terikat dengan komponen utama. PCR merupakan analisis regresi variabel terikat terhadap komponen-komponen utama yang tidak saling berkorelasi, dengan setiap komponen utama merupakan kombinasi linear dari semua variabel bebas (Draper & Smith, 1998). Berikut adalah persamaan PCR:

$$Y = w_0 + w_1 K_1 + w_2 K_2 + \cdots + w_m K_m + \varepsilon \tag{2.8}$$

Keterangan:

Y : Variabel terikat.

w_0 : Konstanta *principal component regression*.

w_m : Parameter *principal component regression*.

K_m : Komponen utama.

$m; m \leq k$.

Komponen utama merupakan kombinasi linear dari variabel standar yaitu:

$$\begin{aligned} K_1 &= a_{11} Z_1 + a_{21} Z_2 + \cdots + a_{k1} Z_k \\ K_2 &= a_{12} Z_1 + a_{22} Z_2 + \cdots + a_{k2} Z_k \\ &\vdots \\ K_m &= a_{1m} Z_1 + a_{2m} Z_2 + \cdots + a_{km} Z_k \end{aligned} \tag{2.9}$$

Keterangan:

Z_k : Variabel bebas yang telah distandarisasi

Jika persamaan (2.9) disubstitusikan ke persamaan (2.8) maka diperoleh:

$$\begin{aligned} Y &= w_0 + w_1(a_{11}Z_1 + a_{21}Z_2 + \cdots + a_{k1}Z_k) + w_2(a_{12}Z_1 + a_{22}Z_2 + \cdots \\ &\quad + a_{k2}Z_k) + w_m(a_{1m}Z_1 + a_{2m}Z_2 + \cdots + a_{km}Z_k) + \varepsilon \\ &= w_0 + w_1a_{11}Z_1 + w_1a_{21}Z_2 + \cdots + w_1a_{k1}Z_k + w_2a_{12}Z_1 + w_2a_{22}Z_2 \\ &\quad + \cdots + w_2a_{k2}Z_k + w_ma_{1m}Z_1 + w_ma_{2m}Z_2 + w_ma_{km}Z_k + \varepsilon \\ &= w_0 + (w_1a_{11} + w_2a_{12} + \cdots + w_ma_{1m})Z_1 + (w_1a_{21} + w_2a_{22} \ldots + \\ &\quad w_ma_{2m})Z_2 + (w_1a_{k1} + w_2a_{k2} + \cdots + w_ma_{km})Z_k + \varepsilon \end{aligned} \tag{2.10}$$

Sehingga diperoleh persamaan PCR sebagai berikut:

$$Y = b_0 + b_1Z_1 + b_2Z_2 + \cdots + b_kZ_k + \varepsilon \tag{2.11}$$

dengan

$$\begin{aligned} b_0 &= w_0 \\ b_1 &= w_1a_{11} + w_2a_{12} + \cdots + w_ma_{1m} \\ b_2 &= w_1a_{21} + w_2a_{22} + \cdots + w_ma_{2m} \\ b_k &= w_1a_{k1} + w_2a_{k2} + \cdots + w_ma_{km} \end{aligned}$$

2.6 *Latent Root Regression (LRR)*

Metode *Latent Root Regression* merupakan perluasan dari PCR. Perbedaannya terletak pada nilai akar laten yang dihasilkan dari matriks korelasinya. Pada LRR, matriks korelasi diperoleh dari penggabungan variabel terikat dan variabel bebas yang telah dibakukan, yang dapat ditulis sebagai berikut (Draper & Smith, 1998):

$$\boldsymbol{Z}^* = [\boldsymbol{Z}_y, \boldsymbol{Z}] \tag{2.12}$$

Dengan $\boldsymbol{Z}_y$ merupakan matriks **Y** dan $\boldsymbol{Z}$ merupakan matriks **X** yang sama-sama telah dibakukan. Setelah matriks **Y** dan **X** dibakukan, maka matriks **X** dan **Y** akan menjadi seperti berikut:

$$\boldsymbol{Z}^* = \begin{bmatrix} Z_{1y} & Z_{11} & \ldots & Z_{1k} \\ Z_{2y} & Z_{21} & \ldots & Z_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{ny} & Z_{n1} & \ldots & Z_{nk} \end{bmatrix} \tag{2.13}$$

Setelah didapat matriks $\boldsymbol{Z}^*$ maka dapat dihitung matriks korelasi $\boldsymbol{Z}^{*\prime}\boldsymbol{Z}^*$. Akar laten (λ_j) dan vektor laten (γ_j) pada LRR diperoleh dari matriks korelasi gabungan yang memenuhi persamaan berikut:

$$\left|\boldsymbol{Z}^{*\prime}\boldsymbol{Z}^* - \lambda_j \boldsymbol{I}\right| = 0; \lambda_1, \lambda_2, \dots, \lambda_k \tag{2.14}$$

dan

$$(\boldsymbol{Z}^{*\prime}\boldsymbol{Z}^* - \lambda_j \boldsymbol{I})\boldsymbol{\gamma_j} = 0; \gamma_{1j}, \gamma_{2j}, \dots, \gamma_{kj} \tag{2.15}$$

Setelah mendapatkan akar laten dan vektor laten dari matriks korelasi gabungannya, selanjutnya dilakukan analisis komponen utama. Komponen utama regresi akar laten dapat diperoleh dengan menggunakan rumus berikut:

$$C_j = \gamma_{0j} Z_y + \gamma_j^0 Z \tag{2.16}$$

dengan:

$$\gamma_j' = (\gamma_{0j}, \gamma_{1j}, \gamma_{2j}, \dots, \gamma_{kj} \text{ dan } \gamma_j^0 = (\gamma_{1j}, \gamma_{2j}, \dots, \gamma_{kj})$$

(Vigneau & Qannari, 2002)

Selanjutnya, komponen utama yang akan digunakan pada tahap analisis diperoleh dengan membuang komponen utama yang bersesuaian dengan nilai akar laten $\lambda_j \leq 0.05$ dan elemen pertama vektor laten $\left|\gamma_{0j}\right| < 0.10$ (Webster, *et al*., 1974). Setelah membuang variabel mana yang dipertahankan, selanjutnya menghitung vektor koefisien kuadrat terkecil termodifikasi dengan rumus berikut:

$$\boldsymbol{\beta}^* = \begin{bmatrix} \beta_1^* \\ \beta_2^* \\ \vdots \\ \beta_k^* \end{bmatrix} = c \sum_{j}^{k} \gamma_{0j} \lambda_j^{-1} \begin{bmatrix} \gamma_{1j} \\ \gamma_{2j} \\ \vdots \\ \gamma_{rj} \end{bmatrix} \tag{2.17}$$

dengan c adalah konstanta yaitu:

$$c = -\left\{\sum_{j}^{k} \gamma_{0j} \gamma_j^{-1}\right\}^{-1} \left\{\sum_{i=1}^{n} (Y_i - \bar{Y})^2\right\}^{\frac{1}{2}}$$

Keterangan:

λ_j : Akar laten ke-j dari matriks $Z^{*\prime}Z^*$.

γ_j : Elemen vektor laten ke-j.

γ_{0j} : Elemen pertama dari vektor laten ke-j.

j : $0,1,2,\ldots,k$.

Menurut Draper & Smith (1998), penduga koefisien regresi pada variabel asal diperoleh dengan membagi penduga koefisien regresi pada variabel yang telah dibakukan dengan S_j sehingga diperoleh:

$$\hat{\beta}_j = \frac{\beta_j^*}{S_j} \tag{2.18}$$

dengan:

$$S_j = \sqrt{\sum_{j=1}^{k} (X_j - \bar{X}_j)^2}$$

Sedangkan untuk perhitungan koefisien regresi β_0 diperoleh berdasarkan rumus:

$$\beta_0 = \bar{Y} - \beta_1\bar{X}_1 - \beta_2\bar{X}_2 - \cdots - \beta_k\bar{X}_k \tag{2.19}$$

2.7 Simulasi Monte Carlo

Salah satu model simulasi pengendalian persediaan yang paling populer adalah simulasi Monte Carlo. Simulasi Monte Carlo merupakan teknik sampling statistik yang digunakan untuk memperkirakan solusi untuk masalah kuantitatif. Model simulasi Monte Carlo adalah suatu bentuk simulasi probabilistik dengan solusi suatu masalah diberikan berdasarkan proses pengacakan. Simulasi Monte Carlo sangat efektif saat digunakan untuk memodelkan suatu aliran antrian dalam sebuah kegiatan, evolusi sebuah epidemi penyakit berdasarkan ruang dan waktu, pengujian statistik, serta prediksi harga (Haryanti, dkk, 2023).

2.8 *Mean Squared Error* (MSE)

Mean Squared Error (MSE) merupakan salah satu ukuran evaluasi yang paling umum digunakan dalam analisis statistik dan pemodelan prediktif. MSE mengukur rata-rata kuadrat selisih antara nilai yang diprediksi oleh model dan nilai aktual (observasi) (Draper & Smith, 1998). Adapun persamaan MSE sebagai berikut:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - y)^T(y_i - y) \tag{2.20}$$

Keterangan:

n : Jumlah pengulangan data simulasi.

y : Nilai sebenarnya.

y_i : Nilai prediksi untuk data simulasi ke-l.

III. METODE PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilaksanakan pada semester ganjil tahun akademik 2024/2025 bertempat di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

3.2 Data Penelitian

Data dalam penelitian ini berupa data simulasi untuk melihat di beberapa kondisi data. Pada simulasi dibangkitkan data menggunakan *software* RStudio dengan variabel bebas X sebanyak lima variabel $(k = 5)$ dan galat berdistribusi normal $N(0,1)$. Masing-masing data dibangkitkan dengan ukuran sampel $(n = 20, 60, 100)$ untuk melihat kondisi data pada saat n kecil, sedang, dan besar serta koefisien korelasi $(\rho = 0.3, 0.6, 0.9, 0.95)$ untuk melihat kondisi data pada saat ρ rendah, sedang, dan tinggi dan diulang sebanyak 1000 kali. Untuk mendapatkan data multikolinearitas pada setiap himpunan data, X_{ik} dibangkitkan menggunakan simulasi Monte Carlo berdasarkan McDonald & Galarneau (1975) dengan persamaan sebagai berikut:

$$X_{ik} = \sqrt{1-\rho^2}Q_{ik} + \rho Q_{i(k+1)}; i = 1,2,\ldots,n \text{ dan } j = 1,2,\ldots,k \tag{3.1}$$

3.3 Metode Penelitian

Penelitian ini membandingkan metode *Principal Component Regression* (PCR) dan *Latent Root Regression* (LRR) pada data simulasi. Adapun langkah-langkah analisis data yang dilakukan pada penelitian ini dibagi menjadi beberapa tahap yaitu:

3.3.1 Proses Pembangkitan Data

Adapun langkah-langkah membangkitkan data sebagai berikut:

1. Menetapkan parameter model regresi yaitu $\beta_0 = 0$, $\beta_1 = \beta_2 = \cdots = \beta_p = 1$, dengan k adalah jumlah variabel.
2. Menetapkan jumlah variabel dan ukuran sampel
   Pada penelitian ini, digunakan variabel bebas sebanyak lima variabel $(k = 5)$ dan ukuran sampel $(n = 20, 60, 100)$.
3. Menetapkan koefisien korelasi $(\rho = 0.3, 0.6, 0.9, 0.95)$.
4. Ulangi langkah 1 sampai 4 sebanyak 1000 kali iterasi.

3.3.2 *Principal Component Regression* (PCR)

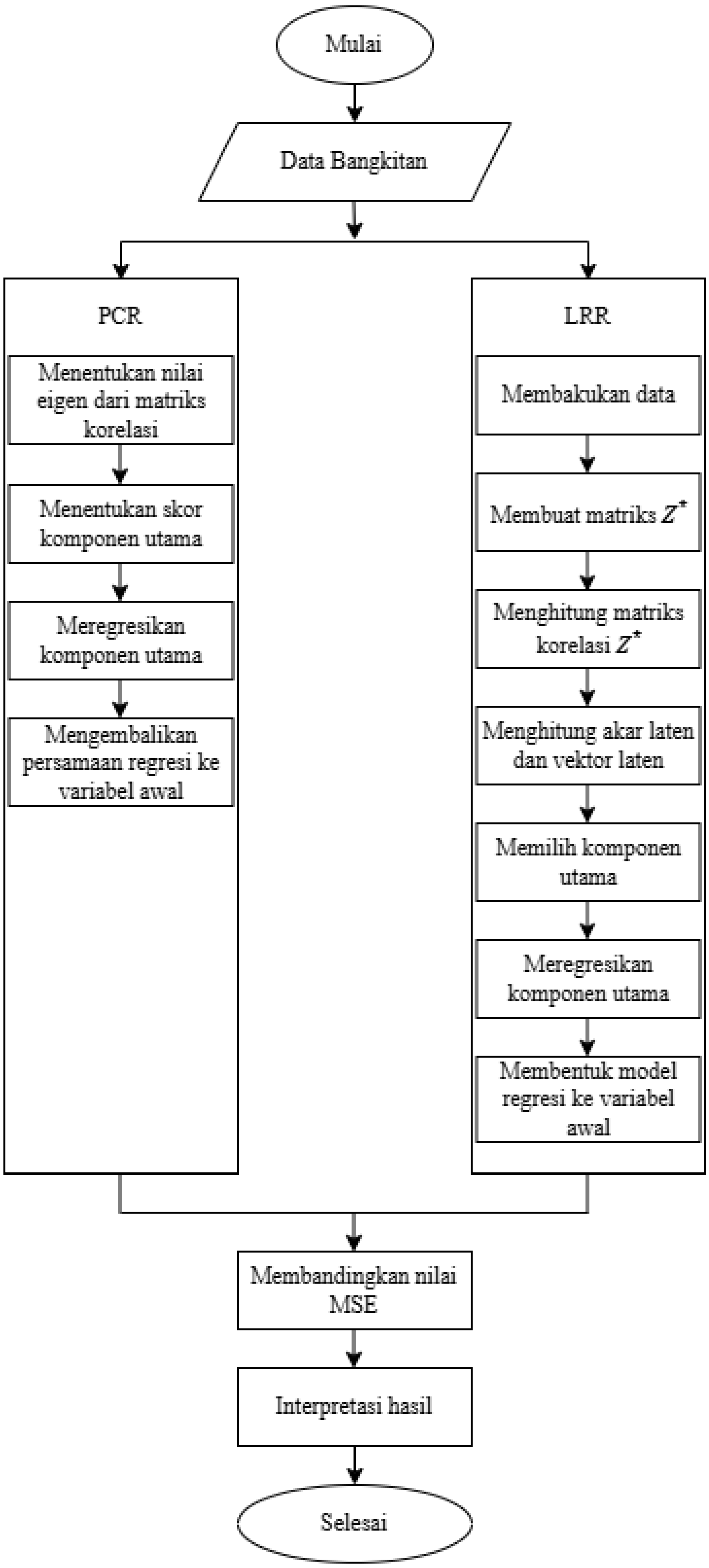
1. Menentukan nilai eigen (λ) dari matriks korelasi $(\boldsymbol{\rho})$.
2. Menentukan vektor eigen atau skor komponen utama (K) untuk nilai eigen lebih dari satu.
3. Meregresikan variabel dependen (Y) terhadap komponen utama (K).
4. Mentransformasi persamaan regresi linear berganda ke dalam bentuk variabel standar (Z).
5. Mentransformasi persamaan regresi linear berganda dengan variabel standar (Z) ke dalam model regresi linear berganda dengan variabel asal (X).

6. Menguji kebaikan model regresi yang diperoleh.

3.3.3 *Latent Root Regression* (LRR)

1. Membakukan data (variabel terikat Y dan variabel bebas X).
2. Membuat matriks gabungan dari variabel yang telah dibakukan, dilambangkan dengan $\boldsymbol{Z}^* = [\boldsymbol{Z_y}\colon \boldsymbol{Z_x}]$.
3. Menghitung matriks korelasi gabungan dari matriks $\boldsymbol{Z}^*$.
4. Menghitung akar laten dan vektor laten berdasarkan matriks korelasi gabungan.
5. Memilih komponen utama yang digunakan dengan membuang komponen utama yang mempunyai akar laten $\lambda_j \leq 0{,}05$ dan elemen pertama dari vektor laten $|\gamma_{0j}| < 0{,}10$.
6. Komponen utama yang terpilih kemudian diregresikan dengan variabel terikatnya.
7. Melakukan pendeteksian terhadap multikolinearitas pada model regresi akar laten yang dihasilkan berdasarkan nilai VIF.
8. Membentuk model regresi ke dalam variabel asal.

Berikut merupakan diagram alir dari penelitian ini:



Gambar 1. Diagram Alir Penelitian

V. KESIMPULAN

Berdasarkan hasil dan pembahasan sebelumnya, diperoleh kesimpulan sebagai berikut:

1. Data yang mengalami multikolinearitas dihasilkan pada studi simulasi dengan variasi koefisien korelasi ($\rho = 0.9, 0.95$) dan variasi ukuran sampel ($n = 20, 60, 100$), sedangkan pada studi simulasi dengan variasi koefisien korelasi ($\rho = 0.3, 0.6$) dan variasi ukuran sampel ($n = 20, 60, 100$) dihasilkan data yang tidak mengalami multikolinearitas.
2. Pada studi simulasi dengan variasi koefisien korelasi ($\rho = 0.9, 0.95$) dan variasi ukuran sampel ($n = 20, 60, 100$), metode *Principal Component Regression* (PCR) menghasilkan nilai MSE terkecil dibandingkan dengan metode *Latent Root Regression* (LRR).
3. Berdasarkan nilai MSE, metode *Principal Component Regression* (PCR) lebih baik dalam mengatasi masalah multikolinearitas dibandingkan dengan metode *Latent Root Regression* (LRR).

DAFTAR PUSTAKA

Alfah, N., Munawar, A. A., & Zulfahrizal. 2022. Teknologi NIRS untuk Memprediksi Jumlah Pencampuran Minyak Sawit dalam Minyak Nilam menggunakan Metode PCR. *Jurnal Ilmiah Mahasiswa Pertanian*. **7**(4): 2615-2878.

Amelia, S. & Putra, A. A. 2023. Regresi Komponen Utama dalam Mengatasi Multikolinearitas pada Faktor yang Mempengaruhi Pendapatan Asli Daerah di Sumatera Barat. *Jurnal Pendidikan Tambusai*. **7**(2): 10906-10914.

Best, H. & Wolf, C. 2005. *Regression Analysis and Causal Inference*. Sage Publications, Croydon.

Candra, P. A., Sukarsa, I. K. G., & Gandhiadi, G. K. 2024. Perbandingan Antara Latent Root Regression dan Ridge Regression dalam Mengatasi Multikolinearitas. *INNOVATIVE: Journal of Social Science Research*. **4**(1): 10300-10312.

Draper, N. & Smith, H. 1998. *Applied Regression Analysis* (3rd ed.). John Wiley & Sons, New York.

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

Haryanti, D. A., Nugraha, N., Andriani, F., Lestari, D. P., & Susanti, B. 2023. *Prediksi Nilai Nisab Zakat Dengan Pendekatan Model Stokastik*. Uwais Inspirasi Indonesia, Jawa Timur.

Hidayatullah, A. F., Saputra, D., Inarah, F., Evita, I., Fadillah, M., & Harsyiah, L. 2024. Analisis Regresi Komponen Utama untuk Mengatasi

Multikolinearitas pada Faktor-Faktor yang Mempengaruhi Indeks Pembangunan Manusia. *Jurnal Sains Natural*. **2**(1): 19-24.

Kadir. 2019. *Statistika Terapan: Konsep, Contoh, dan Analisis Data dengan Program SPSS/Lisrel dalam Penelitian* (3rd ed.). RajaGrafindo Persada, Jakarta.

Kartiningrum, E. D., Notobroto, H. B., Otok, B. W., Kumatijati, N. E., & Yuswatiningsih, E. 2022. *Aplikasi Regresi dan Korelasi dalam Analis Data Hasil Penelitian*. STIKes Majapahit, Mojokerto

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

McDonald, G. C., & Galarneau, D. I. 1975. A Monte Carlo Evaluation of Some Ridge-Type Estimators. *Journal of American Statistical Association*. **70**(350): 407-416.

Montgomery, D. C., Peck, E. A., & Vining, G. G. 2012. *Introduction to Linear Regression Analysis*. John Wiley & Sons, New Jersey.

Sahriman, S. & Yulianti, A, S. 2023. Statistical Downscaling with Principal Component Regression and Latent Root Regression to Forecast Rainfall in Pangkep Regency. *Journal of Mathematics and Its Applications*. **17**(1): 401-410.

Untari, D. P. & Susanti, M. 2017. Latent Root Regression untuk Mengatasi Multikolinearitas. *Pythagoras: Jurnal Pendidikan Matematika*. **12**(1): 23-32

Vigneau, E., & Qannari, E. M. 2002. A New Algorithm for Latent Root Regression Analysis. *Computational Statistics & Data Analysis*. **41**(1): 231-242.

Webster, J. T., Gunst, R. F., & Mason, R. L. 1974. Latent Root Regression Analysis. *Technometrics*. **16**(4): 513-522.