

# Handling Multicollinearity Problems in Indonesia's Economic Growth Regression Modeling Based on Endogenous Economic Growth Theory

Penanganan Masalah Multikolinieritas pada Pemodelan Pertumbuhan Ekonomi Indonesia Berdasarkan Teori Pertumbuhan Ekonomi *Endogenous*

Aldino Yanke<sup>1,2</sup>, Nofrida Elly Zendrato<sup>1,2</sup>, Agus M Soleh<sup>1‡</sup>

<sup>1,2</sup>Department of Statistics, IPB University, Indonesia

<sup>2</sup>Badan Pusat Statistik, Indonesia

<sup>‡</sup>corresponding author: [agusms@apps.ipb.ac.id](mailto:agusms@apps.ipb.ac.id)

Copyright © 2022 Aldino Yanke, Nofrida Elly Zendrato, and Agus M Soleh. This is an open-access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Abstract

One of the multiple linear regression applications in economics is Indonesia's economic growth model based on the theory of endogenous economic growth. Endogenous economic theory is the development of classical theory which cannot explain how the economy grows in the long run. The regression model based on the theory of endogenous economic growth used many independent variables, which caused multicollinearity problems. In this study, the multiple linear regression model using the least-squares estimation method and some methods to handle the multicollinearity problem was implemented. Variable selection methods (backward, forward, and stepwise), principal component regression (PCR), partial least square (PLS), and regularization methods (Ridge, Lasso, and Elastic Net) were applied to solve the multicollinearity problem. Variable selection method with backward, forward, and stepwise has not been able to overcome the problem of multicollinearity. In contrast, Principal Component Regression, PLS regression, and regularization regression methods overcame the multicollinearity problem. We used "leave one out cross-validation" (LOOCV) to determine the best method for handling multicollinearity problems with the smallest mean square of error (MSE). Based on the MSE value, the best method to overcome the multicollinearity problem in the economic growth model based on endogenous economic growth theory was the Lasso regression method.

**Keywords:** endogenous economic growth, Lasso, multicollinearity

---

\* Received: Jan 2021; Reviewed: Jul 2022; Published: Aug 2022

## 1. Pendahuluan

Analisis regresi merupakan alat statistika untuk mengevaluasi hubungan antara satu peubah dengan peubah lain atau satu peubah dengan beberapa peubah lainnya. Hubungan antar peubah tersebut bisa merupakan hubungan sebab-akibat ataupun bukan merupakan hubungan sebab-akibat melainkan hanya hubungan kualitatif dan kuantitatif antar peubah. Ada dua tujuan suatu peneliti melakukan analisis regresi yaitu untuk tujuan memprediksi suatu peubah berdasarkan peubah yang lain atau peneliti ingin mengetahui hubungan antara suatu peubah terhadap peubah yang lain. Dalam analisis regresi berganda yang melibatkan lebih dari satu peubah penjelas, permasalahan yang sering dihadapi adalah masalah multikolinieritas atau adanya korelasi yang kuat antar peubah penjelas. Masalah multikolinieritas pada regresi linier berganda dengan metode estimasi Metode Kuadrat Terkecil (MKT) akan menimbulkan masalah yaitu estimasi parameter model regresi tidak stabil disebabkan nilai varians kovarians koefisien regresi mempunyai nilai yang besar (Ng, 2013).

Penerapan pemodelan regresi linier dapat dilakukan pada berbagai bidang keilmuan, termasuk dalam bidang ekonomi. Dalam bidang ekonomi sangat menarik untuk melakukan pemodelan regresi pada indikator-indikator penting perekonomian. Salah satu indikator penting perekonomian suatu negara adalah pertumbuhan ekonomi yang menunjukkan perkembangan ekonomi suatu negara.

Ada banyak teori mengenai pertumbuhan ekonomi yang dapat dimodelkan dan dianalisis lebih jauh. Salah satu teori mengenai pertumbuhan ekonomi adalah teori pertumbuhan endogeneous. Teori ini dilatarbelakangi dengan teori pertumbuhan klasik yang tidak dapat menjelaskan bagaimana pertumbuhan ekonomi pada jangka panjang. Berdasarkan teori pertumbuhan ekonomi klasik, perubahan teknologi, saving, dan pertumbuhan penduduk merupakan faktor eksogen. Sedangkan teori pertumbuhan endogeneous memosisikan perubahan teknologi, saving, dan pertumbuhan penduduk sebagai faktor endogen.

Model pertumbuhan endogeneous mendefinisikan bahwa pertumbuhan ekonomi digerakkan oleh tiga faktor yaitu faktor pertama adalah pertumbuhan penduduk dan akumulasi modal manusia (Sumber Daya Manusia-SDM), faktor kedua adalah perubahan teknologi, dan faktor ketiga adalah peran pemerintah. Selain faktor yang dijelaskan dalam teori tersebut, pada analisis data ekonomi biasanya terkait dengan peubah ekonomi lainnya. Adapun peubah yang erat kaitannya dengan pertumbuhan ekonomi adalah data Pendidikan, inflasi, ekspor-impor, perdagangan luar negeri, pertumbuhan investasi, dan nilai tukar rupiah.

Pemodelan terhadap pertumbuhan ekonomi berdasarkan teori pertumbuhan endogeneous yang juga dipengaruhi oleh peubah ekonomi lainnya mengidentifikasi banyak peubah bebas dalam model. Banyaknya peubah bebas dalam model menyebabkan model sangat rentan tidak memenuhi asumsi model regresi yaitu asumsi multikolinieritas.

Masalah kolinieritas akan berdampak pada estimasi parameter model dengan metode MKT, sehingga diperlukan penanganan terhadap masalah multikolinieritas. Terdapat banyak cara untuk menangani masalah multikolinieritas diantaranya dengan

seleksi peubah (*Backward, Forward, dan Stepwise*), regresi regularisasi (Regresi Gulud, Lasso dan Elastic Net), Regresi Komponen Utama dan Regresi Partial Least Square (PLS). (Apriansyah, 2018) menggunakan *stepwise regression* untuk mempermudah membuat asosiasi dengan mengurangi dimensi peubah SNP (*single nucleotide polymorphism*). (Fanny et al., 2018) menangani permasalahan multikolinieritas pada pemodelan produktivitas penangkapan bagan perahu dengan metode regresi regularisasi. Dengan metode yang sama, (Kusuma & Wulansari, 2019) menerapkannya pada pemodelan kemiskinan dan kerentanan kemiskinan. (Agus M Soleh et al., 2015) membandingkan regresi komponen utama dan Lasso pada Statistical Downscaling pemodelan curah hujan bulanan. Sedangkan (Korkmazoglu & Kemalbay, 2012) menerapkan metode regresi PLS untuk menangani multikolinieritas pada pemodelan pertumbuhan ekonomi Turki. Dalam penelitian ini akan diterapkan beberapa metode penanganan masalah multikolinieritas data yaitu seleksi peubah, regresi regularisasi, regresi komponen utama, dan regresi PLS pada kasus pemodelan pertumbuhan ekonomi endogen Indonesia yang melibatkan beberapa peubah penjelas serta menentukan metode terbaik berdasarkan nilai Kuadrat Tengah Galat (KTG) validasi silang.

## 2. Tinjauan Pustaka

### 2.1. Teori Pertumbuhan Ekonomi *Endogenous*

Pada tahun 1980an, Paul Romer dan Robert Lucas menggagas suatu teori pertumbuhan ekonomi *endogenous* yang merupakan salah satu perkembangan dari teori pertumbuhan ekonomi yang masuk dalam neo klasik. Teori ini memasukkan perubahan teknologi, *saving* dan pertumbuhan penduduk yang merupakan faktor endogen pertumbuhan ekonomi.

Beberapa ekonom memberikan gambaran karakteristik dari teori pertumbuhan ekonomi *endogenous*. (Romer, 1994) menyebutkan bahwa pertumbuhan ekonomi merupakan hasil dari dalam sistem ekonomi itu sendiri bukan berasal dari luar. Sedangkan (Pomfret, 1997) menyatakan bahwa teori pertumbuhan ekonomi *endogenous* berusaha menjelaskan sumber-sumber pertumbuhan dari peubah di dalam model.

Secara teoritis, model pertumbuhan ekonomi merupakan fungsi produksi suatu daerah yang didalamnya terdapat unsur modal (kapital) dan tenaga kerja. Selain fungsi tersebut juga terdapat fungsi *saving*. Sedangkan secara empirik, model pertumbuhan ekonomi *endogenous* telah dikembangkan pada beberapa penelitian. (Taşel & Bayarcelik, 2013) mengembangkan model pertumbuhan ekonomi dipengaruhi oleh indikator modal manusia (SDM) yaitu melalui tingkat partisipasi sekolah. (Zareen & Qayyum, 2015) mengembangkan model pengaruh konsumsi pemerintah, barang modal, inflasi dan pengangguran terhadap pertumbuhan ekonomi di Pakistan. (Korkmazoglu & Kemalbay, 2012) mengembangkan model pertumbuhan ekonomi dengan peubah bebas tiga faktor utama teori pertumbuhan ekonomi *endogenous*, yaitu peubah yang berkaitan dengan penduduk dan ketenagakerjaan, perubahan teknologi, dan peran pemerintah dalam perekonomian. Sedangkan (Neog & Gaur,

2020) mengembangkan model pengaruh pajak terhadap pertumbuhan ekonomi di Negara Bagian Indian.

## 2.2. Analisis Regresi Linier Berganda

Menurut Montgomery *et al.* (2015) analisis regresi merupakan teknik statistika untuk menginvestigasi dan memodelkan hubungan antara satu peubah dengan peubah lainnya. Model regresi yang melibatkan lebih dari satu peubah bebas disebut dengan model Regresi Linier Berganda (RLB). Adapun model persamaan dari RLB adalah sebagai berikut:

$$y_i = \sum_{j=1}^p x_{ij}\beta_j + \varepsilon_i; \quad i = 1, 2, \dots, n; \quad j = 1, 2, \dots, p \quad (1)$$

Untuk memperoleh model RLB diperlukan tahapan estimasi parameter model dan metode yang populer digunakan adalah MKT. Estimasi parameter model RLB dengan MKT sebagai berikut:

$$\hat{\beta}^{\text{MKT}} = \arg \min_{\beta} \{(\mathbf{y} - \mathbf{X}\beta)'(\mathbf{y} - \mathbf{X}\beta)\} \quad (2)$$

Permasalahan yang sering terjadi dalam pemodelan RLB adalah masalah multikolinieritas atau terjadi kolerasi antar peubah bebas. Adanya masalah multikolinieritas akan berpengaruh pada matriks varians kovarians parameter model dan menyebabkan estimasi parameter model regresi tidak konsisten.

Ada beberapa metode yang bisa digunakan untuk mengatasi masalah multikolinieritas pada model RLB. Metode yang dapat diterapkan untuk mengatasi multikolinieritas adalah metode seleksi peubah (*backward, forward, and stepwise selection*), regresi komponen utama, regresi PLS, serta regresi regularisasi (Regresi Gulud, Lasso, dan *Elastic Net*).

## 2.3. Seleksi Peubah Model RLB

Salah satu solusi ketika ada korelasi antar peubah penjelas adalah melakukan seleksi peubah dimana dengan seleksi peubah akan diperoleh model dengan peubah penjelas yang lebih sedikit sehingga diharapkan tidak terjadi multikolinieritas. Seleksi peubah merupakan tahapan yang akan menghasilkan peubah penjelas yang sedikit dalam model RLB tanpa mengurangi maksud dari tujuan penelitian. Peubah yang terseleksi merupakan peubah yang paling berpengaruh terhadap model. Menurut (Montgomery *et al.*, 2015b) proses seleksi peubah dilakukan dengan cara mengevaluasi hanya sebagian dari peubah bebas pada model regresi dengan memasukkan dan mengeluarkan peubah bebas secara bersamaan dalam satu waktu. Terdapat tiga metode dasar dalam seleksi peubah yaitu *Backward Elimination, Forward Selection, dan Stepwise Selection*. Estimasi parameter hasil seleksi peubah menggunakan MKT.

Prosedur eliminasi langkah mundur (*Backward Elimination Procedure*) merupakan salah satu metode seleksi peubah dalam regresi linier yaitu dengan cara mengeluarkan peubah satu per satu dari model awal. Untuk peubah penjelas yang telah dikeluarkan dari model, tidak akan dimasukkan kembali ke dalam model.

Pemilihan peubah yang dikeluarkan dari model berdasarkan apakah peubah tersebut berpengaruh nyata atau tidak dalam model pada nilai  $\alpha$  tertentu. Proses berakhir jika tidak ada lagi peubah bebas yang tidak nyata diantara yang tersisa.

Prosedur eliminasi langkah maju (*Forward Elimination Procedure*) menyeleksi peubah dengan memasukkan peubah penjelas yang paling berpengaruh kedalam persamaan regresi kemudian menambahkan satu per satu peubah penjelas berikutnya yang paling berpengaruh berikutnya, demikian seterusnya. Pengaruh suatu peubah bebas terhadap peubah respon ditentukan pada tingkat  $\alpha$  tertentu yang ditentukan peneliti. Peubah penjelas yang telah dimasukkan kedalam model tidak dapat dikeluarkan kembali dari model.

Sedangkan prosedur regresi bertatar (*Stepwise Regression*) merupakan proses seleksi peubah yang menggabungkan prosedur *backward* dan *forward*. Prosedur dalam penyeleksian peubah dimulai dari tanpa peubah penjelas, kemudian menyeleksi peubah penjelas yang paling berpengaruh (seperti prosedur pada *Forward*), namun peubah yang telah masuk kedalam model dapat dikeluarkan dari model (seperti prosedur pada *Backward*). Proses akan berhenti ketika tidak ada lagi peubah penjelas yang tidak nyata diantara peubah penjelas yang tersisa dalam model dan tidak ada lagi peubah penjelas yang nyata diantara peubah penjelas yang tidak masuk kedalam model. Saat ini, para peneliti dapat mengontrol proses secara detail dengan mengatur tingkat signifikansi dan manipulasi peubah (Wang & Chen, 2016).

#### 2.4. Regresi Komponen Utama

Regresi Komponen Utama merupakan salah satu teknik dalam mengatasi permasalahan multikolonieritas. Metode ini merupakan gabungan antara regresi linier dengan analisis komponen utama. Menurut (Artigue & Smith, 2019) Regresi Komponen Utama mengurangi jumlah peubah bebas pada model regresi ke dalam komponen utama yang lebih sedikit dengan menghilangkan korelasi peubah melalui transformasi peubah asal ke peubah baru yang tidak berkorelasi. (Marcus et al., 2012). Ide dari Regresi Komponen Utama adalah Analisis Komponen Utama pada Matriks  $\mathbf{X}$  (peubah penjelas) dan menggunakan hanya  $k$  komponen utama sebagai peubah bebas dalam model regresi.

Tujuan dari Analisis Komponen Utama adalah mencari matriks *orthogonal*  $\mathbf{P}$  dengan ordo  $m \times m$  yang mampu mengubah peubah  $\mathbf{X}$ .

$$\mathbf{T} = \mathbf{XP} \quad (3)$$

Dimana peubah baru  $t_1, \dots, t_m$  tidak saling berkorelasi dan terurut keragamannya dari tertinggi ke terendah. Dengan kata lain, diinginkan matriks kovarians  $\text{cov}(\mathbf{T}, \mathbf{T})$  merupakan matriks diagonal dengan urutan diagonal terurut. Matriks  $\mathbf{P}$  disebut komponen utama dan  $\mathbf{T}$  disebut skor.

Seringkali, sebagian besar keragaman dari  $\mathbf{X}$  dapat tercermin dari beberapa komponen utama pertama. Misalnya diambil  $\mathbf{P}_{1k}$  merupakan komponen utama pertama dimana  $k \leq m$ . Sehingga terbentuk aproksimasi dari  $\mathbf{X}$  sebagai berikut:

$$\mathbf{T}_{|k} = \mathbf{X}\mathbf{P}_{|k} \quad (4)$$

Dimana  $\mathbf{T}_{|k}$  merupakan matriks berordo  $n \times k$  merupakan pemampatan dari matriks berordo  $n \times m$  yang menangkap sebagian besar keragaman dari  $\mathbf{X}$ . Ide dari regresi komponen utama adalah meregresikan  $\mathbf{T}_{|k}$  terhadap  $\mathbf{Y}$  dimana antar kolom matriks  $\mathbf{T}_{|k}$  tidak saling berkorelasi.

## 2.5. Regresi PLS

Menurut (Zifarelli et al., 2020) regresi PLS merupakan metode yang membangun model yang bisa diterapkan ketika peubah penjelas banyak dan memiliki korelasi yang kuat. Proses pada Regresi PLS adalah mengurangi jumlah peubah penjelas menjadi set peubah yang lebih sedikit yang tidak saling berkorelasi dan melakukan regresi kuadrat terkecil pada komponen tersebut. Idanya adalah melakukan dekomposisi pada matriks  $\mathbf{X}$  dan matriks respon  $\mathbf{Y}$  (kasus umumnya untuk respon juga multivariat) dengan metode yang sejalan dengan analisis komponen utama.

$$\mathbf{X} = \mathbf{T}\mathbf{P}' + \mathbf{E} \quad (5)$$

$$\mathbf{Y} = \mathbf{T}\mathbf{Q}' + \mathbf{F} \quad (6)$$

$$\mathbf{T} = \mathbf{X}\mathbf{W}(\mathbf{P}'\mathbf{W})^{-1} \quad (7)$$

Dimana  $\mathbf{T}$ ,  $\mathbf{P}$ , dan  $\mathbf{E}$  adalah matriks skor, loading, dan sisaan dari matriks  $\mathbf{X}$ . Matriks  $\mathbf{Q}$  dan  $\mathbf{F}$  adalah matriks loading dan sisaan untuk  $\mathbf{Y}$  serta  $\mathbf{W}$  adalah matriks penimbang. Jumlah komponen utama diperoleh dengan k-fold cross validation. Dekomposisi matriks  $\mathbf{X}$  dan  $\mathbf{Y}$  dapat dilakukan dengan proses iterasi melalui algoritma *Nonlinear Iterative Partial Least Square* (NIPALS). Sehingga penduga parameter model regresi PLS dalam (Liu & Li, 2017) didefinisikan:

$$\hat{\boldsymbol{\beta}}^{\text{PLS}} = \mathbf{W}(\mathbf{P}'\mathbf{W})^{-1}\mathbf{Q}' \quad (8)$$

## 2.6. Regresi Regularisasi

Regresi Regularisasi (*regularization method*) merupakan metode penyusutan estimasi atau prediksi pada regresi yang digunakan ketika terjadi pelanggaran asumsi multikolinieritas. Metode ini memberikan batasan atau penalisasi pada koefisien regresi.

### 2.6.1 Regresi Gulud

(Hoerl & Kennard, 1970) memperkenalkan metode Regresi Gulud yaitu salah satu metode regresi *shrinkage* yang dapat mengatasi multikolinieritas dengan tidak melakukan seleksi peubah. Metode ini memodifikasi MKT, dengan memperbolehkan penduga yang bias pada parameter model. Penduga yang memiliki bias yang kecil dan memiliki ragam yang jauh lebih kecil lebih dipilih dibandingkan penduga tak bias yang memiliki ragam yang cukup besar untuk mendekati nilai parameter yang sebenarnya. Menurut (Arashi et al., 2021) dalam Regresi Gulud dilakukan pengaturan terhadap parameter dan kombinasi liniernya untuk memperoleh penduga dengan variabilitas lebih kecil.

Regresi Gulud meminimumkan JKG dengan batasan  $L_2$ -norm kuadrat dari koefisien regresi. Sehingga estimasi parameter pada model Regresi Gulud adalah:

$$\hat{\beta}^{\text{Gulud}} = \arg \min_{\beta} \{(\mathbf{y} - \mathbf{X}\beta)'(\mathbf{y} - \mathbf{X}\beta) + \lambda_2 \beta' \beta\} \quad (9)$$

Dengan meminimumkan JKG Regresi Gulud diperoleh persamaan:

$$\mathbf{X}'\mathbf{y} = (\mathbf{X}'\mathbf{X} + \lambda_2 \mathbf{I})\beta \quad (10)$$

Ketika  $\lambda_2 > 0$  maka  $(\mathbf{X}'\mathbf{X} + \lambda_2 \mathbf{I})$  berpangkat penuh. Sehingga solusi dari persamaan (8) dapat diperoleh:

$$\hat{\beta}^{\text{Gulud}} = (\mathbf{X}'\mathbf{X} + \lambda_2 \mathbf{I})^{-1} \mathbf{X}'\mathbf{y} \quad (11)$$

### 2.6.2 Lasso

Metode *Least Absolute Shrinkage and Selection Operator* (Lasso) yang merupakan salah satu upaya memperbaiki hasil dugaan parameter regresi MKT diperkenalkan oleh (Tibshirani, 1996). (Agus Mohamad Soleh & others, 2013) menyatakan bahwa Lasso merupakan alternatif bagi penyeleksian peubah dan pendugaan koefisien regresi pada kondisi peubah bebas tidak benar-benar saling bebas. Hal ini yang menjadi pembeda dengan Regresi Gulud, karena pada Regresi Gulud secara simultan koefisien yang diduga tidak mungkin bernilai nol.

Lasso meminimumkan JKG dengan batasan  $L_1$ -norm pada koefisien regresi. Sehingga estimasi parameter pada model regresi dengan metode Lasso adalah:

$$\hat{\beta}^{\text{LASSO}} = \arg \min_{\beta} \{(\mathbf{Y} - \mathbf{X}\beta)'(\mathbf{Y} - \mathbf{X}\beta) + \lambda_1 \mathbf{1}_p |\beta|\} \quad (12)$$

Untuk menyelesaikan persamaan di atas digunakan pemrograman kuadrat. Namun, dengan metode ini Lasso memiliki kelemahan yaitu ketika peubah penjelasnya ( $p$ ) lebih besar dari jumlah amatan ( $n$ ) maka penduga parameter regresi menjadi tidak terdefinisi. Oleh karena itu dikembangkan regresi Lasso dengan algoritma *Least Angle Regression* (LARS) oleh Hastie et al. (2008).

### 2.6.3 Elastic Net

*Elastic Net* merupakan metode estimasi parameter regresi linear yang dapat mengatasi multikolinieritas dengan banyak peubah penjelas ( $p$ ) lebih besar dari jumlah amatan ( $n$ ). Metode ini merupakan gabungan dari batasan Regresi Gulud dan Lasso. Metode ini dapat menyusutkan koefisien regresi sehingga dapat melakukan seleksi peubah serta dapat memberikan efek pengelompokan peubah ketika peubah-peubah penjelas memiliki korelasi yang kuat satu sama lain.

Untuk setiap  $\lambda_1$  dan  $\lambda_2$ , (Zou & Hastie, 2005) mendefinisikan ukuran standar dari *Naïve Elastic Net*, yaitu:

$$L(\lambda_1, \lambda_2, \beta) = \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j)^2 + \lambda_2 \sum_{j=1}^p \beta_j^2 + \lambda_1 \sum_{j=1}^p |\beta_j| \quad (13)$$

Sehingga penduga parameter model *Elastic Net* menurut Liu et al. (2017) adalah:

$$\hat{\beta}^{\text{enet}} = \left(1 + \frac{\lambda_2}{n}\right) \arg \min_{\beta_0, \beta} \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j)^2 + \lambda_2 \sum_{j=1}^p \beta_j^2 + \lambda_1 \sum_{j=1}^p |\beta_j| \quad (14)$$

Dimana L1-norm ( $\lambda_1$ ) adalah batasan Lasso yang mengecilkan koefisien regresi tepat nol dan L2-norm ( $\lambda_2$ ) adalah batasan Gulud. Untuk menduga parameter regresi *Elastic Net* digunakan dengan algoritma LARS-EN. Berdasarkan Kayanan *et al.* (2019), LARS-EN merupakan modifikasi algoritma LARS yang digunakan untuk memperoleh solusi dari *Elastic Net*.

### 3. Metodologi

#### 3.1. Data

Dalam penelitian ini dilakukan studi empiris analisis regresi linier berganda yang diterapkan pada bidang ekonomi yaitu model teori ekonomi *endogenous*. Untuk estimasi parameter model regresi digunakan metode MKT serta diterapkan metode penanganan permasalahan multikolinieritas. Metode penanganan multikolinieritas yang digunakan dalam penelitian ini adalah metode seleksi peubah (*forward, backward, stepwise*), regresi komponen utama, regresi PLS, serta regresi regularisasi (Regresi Gulud, Lasso, dan *Elastic Net*). Selanjutnya dipilih metode penanganan multikolinieritas terbaik berdasarkan nilai Kuadrat Tengah Galat (KTG) validasi silang terkecil.

Penelitian ini melakukan pemodelan untuk tahun data 2000-2019. Data yang digunakan bersumber dari Badan Pusat Statistik (BPS). Berikut adalah daftar peubah dalam penelitian ini:

Tabel 1. Daftar Peubah Penelitian

No	Peubah	Satuan	Transformasi
Peubah Respon			
1	Produk Domestik Bruto Atas Dasar Harga Konstan (ADHB)	Miliar rupiah	Logaritma Natural
Peubah Bebas			
2	Jumlah Penduduk	Ribu jiwa	Logaritma Natural
3	Angka Partisipasi Angkatan Kerja	Persen	-
4	Tingkat Pengangguran	Persen	-
5	Persentase Pertanian dalam PDB	Persen	-
6	Impor Barang dan Jasa	Miliar rupiah	Logaritma Natural
7	Ekspor Barang dan Jasa	Miliar rupiah	Logaritma Natural
8	Persentase Industri Pengolahan dalam PDB	Persen	-
9	Persentase Industri Migas dalam PDB	Persen	-
10	Pembentukan Modal Tetap Bruto (PMTB)	Miliar rupiah	Logaritma Natural
11	Angka Partisipasi Sekolah Dasar	Persen	-
12	Angka Partisipasi Sekolah Menengah	Persen	-
13	Angka Partisipasi Sekolah Atas	Persen	-
14	Konsumsi Pemerintah	Miliar rupiah	Logaritma Natural
15	Pendapatan Pajak	Miliar rupiah	Logaritma Natural



### 3.2. Metode Penelitian

Tahapan analisis dalam penelitian ini adalah :

1. Melakukan eksplorasi data.  
 Dalam tahapan ini, dilakukan analisis deskriptif. Alat analisis deskriptif yang digunakan adalah dengan metode grafik dimana pada penelitian ini dilakukan analisis dalam bentuk diagram garis perkembangan data dari tahun ke tahun. Selain itu dalam tahapan eksplorasi data juga dilakukan analisis awal sebelum masuk analisis regresi dengan melakukan analisis korelasi antar peubah respon dan penyerta untuk mengetahui hubungan antar peubah dalam penelitian ini.
2. Mengestimasi parameter model regresi dengan MKT  
 Melakukan pemodelan regresi dengan peubah respons yang ditentukan dalam hal ini In PDB terhadap peubah bebas yang ditentukan berdasarkan studi literatur. Untuk memperoleh dugaan model regresi dilakukan estimasi parameter model dengan metode MKT.
3. Melakukan pengecekan multikolinieritas  
 Salah satu kondisi yang harus terpenuhi pada pemodelan RLB dengan metode estimasi parameter MKT adalah kebebasan antar peubah bebas (non-multikolinieritas). Salah satu cara pengecekannya adalah dengan *Variance Inflation Factor* (VIF). Montgomery *et al.* (2012) menyatakan bahwa terjadi masalah multikolinieritas jika nilai VIF lebih dari 5 atau 10.
4. Melakukan penanganan masalah multikolinieritas: metode seleksi peubah (*forward, backward, stepwise*), regresi komponen utama, regresi PLS, serta regresi regularisasi (Regresi Gulud, Lasso, dan *Elastic Net*).
5. Memilih model terbaik  
 Pemilihan model terbaik dilakukan dengan menggunakan indikator rata-rata KTG yang diperoleh dari proses validasi silang dengan metode LOOCV. Metode validasi silang dengan LOOCV dilakukan dengan mengeluarkan satu amatan lalu melakukan pemodelan berdasarkan data selain data yang dikeluarkan. Setelah itu dilakukan pendugaan respons dari data yang dikeluarkan dan dibandingkan dengan data peubah respon. Kuadrat dari selisih antara data dugaan dan data asli merupakan nilai KTG. Proses ini dilakukan untuk semua data, selanjutnya KTG untuk masing-masing data dirata-ratakan.

Adapun tahapan untuk masing-masing metode untuk mengatasi masalah multikolinieritas, sebagai berikut:

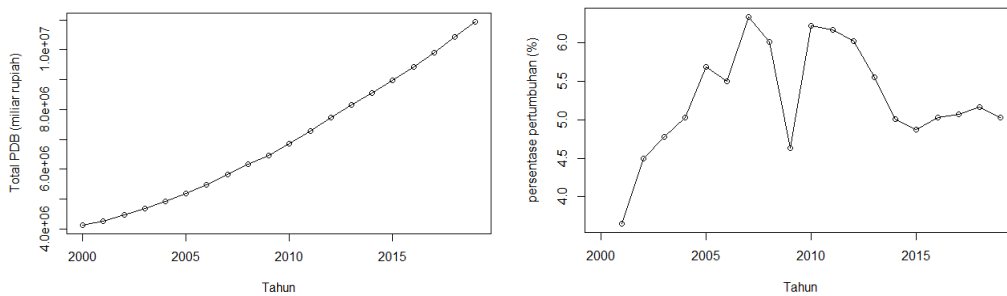
- Prosedur *Backward Elimination*:
  1. Melakukan pemodelan RLB dengan semua peubah penjelas
  2. Melakukan seleksi peubah dengan mengeluarkannya satu persatu berdasarkan nyata atau tidaknya berpengaruh terhadap peubah respon pada taraf nyata ( $\alpha$ ) tertentu. Peubah bebas yang paling tidak nyata dikeluarkan dari model.
  3. Selanjutnya melakukan pemodelan setelah peubah yang paling nyata dikeluarkan. Dan melakukan pengecekan nyata atau tidak semua peubah bebas terhadap peubah respon. Jika ada yang tidak nyata maka peubah yang paling tidak nyata dikeluarkan dari model.

4. Proses akan berhenti jika tidak ada lagi peubah bebas yang tidak nyata dalam model.
- Prosedur *Forward Selection*:
    1. Melakukan pemodelan tanpa peubah penjelas
    2. Melakukan pemodelan dengan memasukkan satu peubah penjelas. Peubah yang berpengaruh nyata dan memiliki nilai F parsial terbesar terpilih masuk dalam model
    3. Melakukan pemodelan dengan satu peubah penjelas sesuai langkah kedua.
    4. Melakukan pemodelan dengan memasukkan peubah bebas lainnya dengan asumsi dalam model sudah ada peubah penjelas pertama. Ketika ada peubah bebas lainnya yang nyata dan nilai F parsial terbesar terpilih masuk dalam model.
    5. Proses akan berakhir jika tidak ada lagi peubah bebas yang nyata mempengaruhi peubah respon dengan asumsi semua peubah yang nyata sudah masuk dalam model.
  - Prosedur *Stepwise Selection*:
    1. Melakukan tahapan memasukkan peubah penjelas dalam model sesuai dengan *Forward Selection*.
    2. Selanjutnya ketika peubah penjelas masuk model dilakukan pengecekan terhadap peubah yang ada dalam model. Jika ada peubah bebas yang tidak nyata mempengaruhi peubah respon maka dikeluarkan dari model.
  - Prosedur Regresi Komponen Utama:
    1. Melakukan analisis komponen utama pada matriks peubah penjelas dan mendapatkan komponen utama dan skor dari komponen utama.
    2. Memilih jumlah komponen utama yang akan menjadi peubah penjelas model. Jumlah komponen utama yang masuk dalam model memperhatikan total keragaman yang dapat dijelaskan oleh komponen utama.
    3. Melakukan pemodelan RLB dengan peubah bebas komponen utama terpilih dengan metode estimasi MKT.
  - Prosedur Regresi PLS:
    1. Menentukan jumlah komponen dalam model berdasarkan proses validasi silang. Jumlah komponen terpilih berdasarkan nilai Kuadrat Tengah Galat Prediksi (KTGP).
    2. Dengan jumlah komponen yang ditentukan dilakukan estimasi parameter model regresi PLS
  - Prosedur Regresi Regularisasi:
    1. Melakukan validasi silang untuk memperoleh nilai *tuning parameter*. Nilai *tuning parameter* dipilih berdasarkan nilai KTG terkecil. (*tuning parameter* Regresi Gulud:  $\lambda_2$ , Lasso:  $\lambda_1$ , *Elastic Net*:  $\lambda_1$  dan  $\lambda_2$ )
    2. Melakukan estimasi parameter model regresi regularisasi dengan *tuning parameter* yang diperoleh.

#### 4. Hasil dan Pembahasan

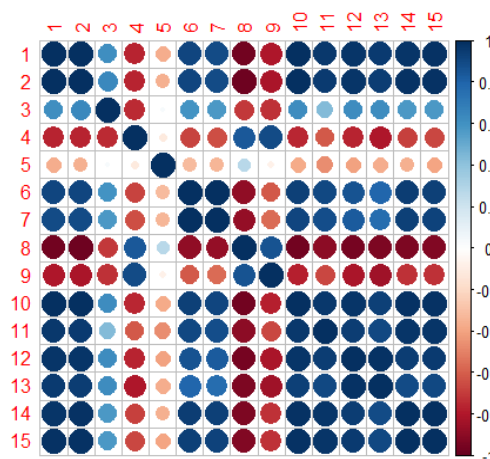
Pada penelitian ini dilakukan pemodelan pertumbuhan ekonomi Indonesia berdasarkan teori ekonomi *endogenous*. Dalam pemodelan pada bidang ekonomi biasanya menggunakan banyak peubah bebas yang mempengaruhi peubah respon. Berdasarkan studi teoritis dan studi empirik diidentifikasi beberapa peubah yang

mempengaruhi PDB sebagai indikator pertumbuhan ekonomi. Sebelum masuk dalam pemodelan, disajikan karakteristik dari PDB dan pertumbuhan ekonomi di Indonesia.



Gambar 1. (a) Pe (a) bangun PDB Indonesia, (b) Laju Pertur (b) an Ekonomi Indonesia Tahun 2000-2019

Berdasarkan Gambar 1 dapat diperoleh informasi bahwa nilai PDB Indonesia terus mengalami peningkatan sejalan dengan berjalannya waktu. Namun terlihat bahwa perubahan PDB dari tahun ke tahun yang direpresentasikan dalam pertumbuhan ekonomi cenderung berfluktuasi selama tahun 2001 hingga 2019. PDB dan pertumbuhan ekonomi merupakan indikator yang sangat penting dalam perekonomian suatu negara karena kedua indikator tersebut mewakili keseluruhan kegiatan ekonomi suatu negara. Sehingga pemodelan terhadap indikator PDB dan pertumbuhan ekonomi menjadi penting supaya tercipta kestabilan ekonomi suatu negara. Salah satu model mengenai pertumbuhan ekonomi adalah model pertumbuhan ekonomi *endogenous* yang memodelkan beberapa peubah ekonomi yang terkait dengan kapital, tenaga kerja (termasuk kualitas sumber daya manusia), dan *saving* terhadap pertumbuhan ekonomi.



Gambar 2. Matriks Korelasi antar Peubah Penjelas

Sebelum melakukan pemodelan, dilakukan identifikasi awal membentuk matriks korelasi antar peubah. Berdasarkan nilai korelasi hubungan peubah bebas terhadap peubah respons PDB, sebagian besar peubah bebas memiliki hubungan linier positif yang cukup kuat dimana peubah yang memiliki korelasi positif lebih dari 0,9 adalah jumlah penduduk, impor, PMTB, angka partisipasi dasar, menengah dan atas,

pengeluaran konsumsi pemerintah, dan pajak. Beberapa peubah bebas menunjukkan korelasi negatif terhadap PDB, yaitu pengangguran, *share* sektor pertanian terhadap PDB, *share* sektor industri pengolahan terhadap PDB, dan persentase industri migas. Sedangkan peubah lainnya memiliki korelasi positif namun besarnya moderat.

Tabel 2. Estimasi Parameter Model RLB dengan MKT dan nilai VIF

Koefisien Regresi	Estimasi Parameter	p-value	VIF
(Intercept)	-27,5400	0,0005*	
ln jumlah penduduk	3,2890	0,0002*	1214,3759
angka partisipasi angkatan kerja	-0,0226	0,0018*	18,7048
Pengangguran	-0,0091	0,0672	74,7266
<i>share</i> pertanian thd PDB	-2,9930	0,0036*	25,9252
ln impor	0,0121	0,6583	427,3265
ln ekspor	-0,0465	0,2670	374,7210
<i>share</i> industri pengolahan thd PDB	0,0298	0,0012*	152,9755
persen industri migas	-0,0212	0,0542	99,1917
ln pmtb	0,4215	0,0012*	872,8569
angka partisipasi dasar	-0,0478	0,0176*	335,9881
angka partisipasi menengah	-0,0002	0,9486	542,2368
angka partisipasi atas	-0,0006	0,7948	485,3764
ln pengeluaran pemerintah	0,4218	0,0004*	405,2071
ln pajak	-0,1972	0,0006*	607,0123

\*) signifikan pada  $\alpha=5$  persen

Estimasi parameter model RLB pertumbuhan ekonomi berdasarkan teori pertumbuhan ekonomi *endogenous* dapat dilihat pada Tabel 2. Penggunaan peubah bebas yang banyak dalam pemodelan RLB akan menyebabkan masalah multikolinieritas dalam model regresi atau antar peubah bebas saling berkorelasi. Hal ini dapat dikonfirmasi dari matriks korelasi antar peubah dan nilai VIF model regresi semua peubah penjelas ada dalam model dengan metode estimasi MKT yang dapat dilihat pada Tabel 2. Nilai VIF semua peubah bebas bernilai lebih dari 10 yang menunjukkan bahwa terjadinya masalah multikolinieritas.

Tabel 3. Peubah Bebas Hasil Seleksi Peubah

<i>Backward</i>	Peubah Bebas Hasil Seleksi				VIF
	VIF	<i>Forward</i>	VIF	<i>Stepwise</i>	
Jumlah		Jumlah		Jumlah	
Penduduk	374,359	Penduduk	145,226	Penduduk	145,226
Impor	138,427	Impor	10,979	Impor	10,979
Ekspor	97,149	PMTB	191,051	PMTB	191,051
<i>Share</i> Industri					
Pengolahan	60,479				
PMTB	247,966				
Pengeluaran					
Pemerintah	148,944				
Pajak	312,534				

Pada penelitian ini akan dilakukan beberapa metode menangani masalah multikolinieritas. Metode pertama yang digunakan adalah dilakukan seleksi peubah dengan metode *backward*, *forward*, dan *stepwise* dimana dalam penelitian ini digunakan  $\alpha$  enter dan *remove* sebesar 0,15 dan diperoleh hasil pada Tabel 3. Berdasarkan hasil seleksi peubah tersebut masih mempunyai nilai VIF yang tinggi dan lebih besar dari 10 sehingga model hasil seleksi peubah dengan metode estimasi MKT masih mengalami masalah multikolinieritas.

Metode selanjutnya upaya dalam mengatasi masalah multikolinieritas adalah metode regresi komponen utama. Dalam regresi komponen utama, dilakukan analisis komponen utama pada peubah bebas dan diperoleh bahwa dengan menggunakan dua komponen utama pertama telah mampu menjelaskan 98,24 persen dari keragaman total peubah bebas. Sehingga dalam regresi komponen utama digunakan dua komponen utama sebagai peubah bebas dalam model regresi. Penggunaan komponen utama sebagai peubah bebas dalam model regresi menjamin tidak terjadinya masalah multikolinieritas dalam model regresi. Kelemahan dari metode regresi komponen utama adalah dalam hal intepretasi hasil model yang diperoleh karena peubah bebas yang digunakan merupakan skor komponen utama yang merupakan kombinasi linier dari peubah bebas asal. Model yang diperoleh dengan metode regresi komponen utama adalah:

$$\ln \widehat{PDB} = 15,714 + 0,028 * KU1 - 0,010KU2 \quad (13)$$

Setelah diperoleh nilai dugaan parameter model regresi komponen utama, perlu dilakukan transformasi balik ke peubah bebas penelitian. Koefisien model hasil transformasi balik dapat dilihat pada Tabel 4.

Tabel 4. Koefisien Model Transformasi Balik Regresi Komponen Utama

Peubah	Koefisien Model	Peubah	Koefisien Model
intercept	15.714	persen industri migas	-0.0052
ln jumlah penduduk	0.0002	ln pmtb	0.0008
angka partisipasi angkatan kerja	0.0048	angka partisipasi dasar	0.0007
pengangguran	-0.0112	angka partisipasi menengah	0.0124
share pertanian thd PDB	0.0031	angka partisipasi atas	0.0219
ln impor	0.0012	ln pengeluaran pemerintah	0.0005
ln ekspor	0.0007	ln pajak	0.0011
share industri pengolahan thd PDB	-0.0077		

Model regresi PLS merupakan salah satu metode yang mampu mengatasi masalah multikolinieritas. Berdasarkan proses validasi silang diperoleh bahwa jumlah komponen yang memiliki nilai KTGP terkecil adalah 7 komponen dengan KTGP 0,039.

Sehingga pemodelan regresi PLS dilakukan dengan 7 komponen dan diperoleh hasil estimasi parameter yang dapat dilihat pada Tabel 5.

Tabel 5. Estimasi Parameter Model Regresi PLS

Peubah	Estimasi Parameter PLS	Peubah	Estimasi Parameter PLS
In jumlah penduduk	0,0114	persen industri migas	-0,0847
angka partisipasi angkatan kerja	0,0066	In pmtb	0,0571
pengangguran	-0,0248	angka partisipasi dasar	0,0981
share pertanian thd PDB	-0,0594	angka partisipasi menengah	-0,0112
In impor	0,0267	angka partisipasi atas	0,0009
In ekspor	0,0185	In pengeluaran pemerintah	0,0413
share industri pengolahan thd PDB	0,0087	In pajak	0,1143

Alternatif metode yang populer digunakan untuk menangani masalah multikolinieritas adalah dengan penerapan regresi regularisasi. Dalam regresi regularisasi ada 3 metode yaitu Regresi Gulud, Lasso, dan *Elastic Net*. Tahapan awal dalam penerapan regresi regularisasi adalah penentuan *tuning parameter* masing-masing model yang meminimumkan KTG. Terdapat satu *tuning parameter* pada model Regresi Gulud dan Lasso, sedangkan pada *Elastic Net* terdapat dua *tuning parameter* karena *Elastic Net* merupakan gabungan antara Regresi Gulud dan Lasso. Berdasarkan hasil validasi silang diperoleh *tuning parameter* untuk Regresi Gulud adalah 0,00533, untuk Lasso 0,001, sedangkan untuk *Elastic Net* diperoleh *tuning parameter*  $\lambda_1=0,93083$  dan  $\lambda_2=0,00367$ . Hasil estimasi parameter untuk ketiga model tersebut sebagai berikut:

Tabel 6. Hasil Estimasi Parameter Model Regresi Regularisasi

Peubah	Regresi Gulud	Lasso	<i>Elastic Net</i>
(intercept)	0,399	-9,050	-7,932
In jumlah penduduk	0,717	1,469	1,286
angka partisipasi angkatan kerja	-0,007	-0,004	
pengangguran	-0,015	-0,004	
share pertanian thd PDB	-0,035	-0,014	-0,003
In impor	0,009		
In ekspor	0,009		
share industri pengolahan thd PDB	0,000		
persen industri migas	-0,047	-0,026	-0,018
In pmtb	0,141	0,338	0,363
angka partisipasi dasar	0,032	0,024	0,025
angka partisipasi menengah	-0,001		
angka partisipasi atas	0,001		0,002
In pengeluaran pemerintah	0,122		
In pajak	0,053		

Selanjutnya dari seluruh metode untuk mengatasi masalah multikolinieritas dipilih metode terbaik berdasarkan nilai dari KTG terkecil validasi silang dengan metode LOOCV. Diperoleh nilai rata-rata KTG sebagai berikut:

Tabel 7. Rataan KTG validasi silang

Metode	Rataan KTG
MKT	0,0003196
<i>Backward</i>	0,0002624
<i>Forward</i>	0,0002227
<i>Stepwise</i>	0,0002227
regresi PCA	0,0080485
regresi PLS	0,0014848
regresi Gulud	0,0005093
regresi lasso	0,0001172
<i>Elastic Net</i>	0,0001762

Berdasarkan nilai rataan KTG validasi silang diperoleh model terbaik untuk menangani masalah multikolinieritas adalah Regresi Lasso. Regresi Lasso merupakan regresi regularisasi yang merupakan metode yang juga melakukan seleksi peubah. Berdasarkan hasil Regresi Lasso, berikut model yang diperoleh:

$$\ln \widehat{PDB} = -9,05 + 1,47 \ln jml \text{ penduduk} - 0,004 \text{ APAK} - 0,004 \text{ pengangguran} - 0,014 \text{ share pertanian} - 0,026 \text{ persen industri migas} + 0,338 \ln \text{ PMTB} + 0,024 \text{ angka partisipasi dasar} \quad (15)$$

Pada model Regresi Lasso tersebut, PDB dipengaruhi jumlah penduduk, angka partisipasi Angkatan kerja, tingkat pengangguran, persentase pertanian, persentase industry migas, PMTB, dan angka partisipasi dasar. Regresi Lasso merupakan salah satu metode regularisasi. Menurut Simeon et.al. (2021), regresi regularisasi memperbaiki kesalahan prediksi model dengan mengurangi variabilitas dalam estimasi koefisien regresi dengan menyusutkan koefisien regresi menuju nol. Dengan diterapkan factor regularisasi/penalty akan menyesuaikan efek multikolinieritas dalam model.

## 5. Kesimpulan

Pemodelan pertumbuhan ekonomi Indonesia berdasarkan teori pertumbuhan ekonomi *endogenous* dengan RLB metode estimasi MKT mengalami masalah multikolinieritas. Dalam penelitian ini dilakukan beberapa upaya penanganan multikolinieritas. Metode pertama, dengan metode seleksi peubah (*backward*, *forward*, dan *stepwise*). Metode ini tidak mampu menangani masalah multikolinieritas. Selanjutnya dilakukan beberapa alternatif pemodelan yang dapat menangani masalah multikolinieritas yaitu Regresi Komponen Utama, Regresi PLS, serta Regresi Regularisasi yang terdiri dari Regresi Gulud, Lasso, dan Elastic Net. Berdasarkan KTG validasi silang terkecil, metode penanganan masalah multikolinieritas terbaik dalam pemodelan pertumbuhan ekonomi Indonesia berdasarkan model teori pertumbuhan ekonomi *endogenous* adalah metode regularisasi Lasso.

**Daftar Pustaka**

- Apriansyah, D. (2018). Asosiasi Single Nucleotide Polymorphism dan Fenotipe pada Penyakit Diabetes Mellitus Tipe 2 Menggunakan Stepwise Regression.
- Arashi, M., Roozbeh, M., Hamzah, N. A., & Gasparini, M. (2021). Ridge regression and its applications in genetic studies. *Plos One*, 16(4): e0245376.
- Artigue, H., & Smith, G. (2019). The principal problem with principal components regression. *Cogent Mathematics & Statistics*, 6(1): 1622190.
- Fanny, R., Djuraidah, A., & Alamudi, A. (2018). Pendugaan Produktivitas Bagan Perahu dengan Regresi Gulud, LASSO dan Elastic-net. *Xplore: Journal of Statistics*, 2(2): 7–14.
- Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction (Vol. 2)*. Springer.
- Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1): 55–67.
- Kayanan, M., & Wijekoon, P. (2019). Performance of LASSO and Elastic net estimators in Misspecified Linear Regression Model.
- Korkmazoglu, O. B., & Kemalbay, G. (2012). Econometrics application of partial least squares regression: an endogeneous growth model for Turkey. *Procedia-Social and Behavioral Sciences*, 62: 906–910.
- Kusuma, G. W., & Wulansari, I. Y. (2019). Analisis kemiskinan dan kerentanan kemiskinan dengan Regresi Ridge, LASSO, dan Elastic-Net di Provinsi Jawa Tengah tahun 2017. *Seminar Nasional Official Statistics*, 2019(1), 503–513.
- Liu, W., & Li, Q. (2017). An efficient elastic net with regression coefficients method for variable selection of spectrum data. *PloS One*, 12(2): e0171122.
- Marcus, G. L., Wattimanela, H. J., & Lesnussa, Y. A. (2012). Analisis Regresi Komponen Utama Untuk Mengatasi Masalah Multikolinieritas Dalam Analisis Regresi Linier Berganda. *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, 6(1): 31–40.
- Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015a). *Introduction to time series analysis and forecasting*. John Wiley & Sons.
- Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015b). *Introduction to time series analysis and forecasting*. John Wiley & Sons.
- Neog, Y., & Gaur, A. K. (2020). Tax structure and economic growth: a study of selected Indian states. *Journal of Economic Structures*, 9(1): 1–12.
- Ng, K. S. (2013). *A simple explanation of partial least squares*. The Australian National University, Canberra, 1–10.
- Pomfret, R. (1997). Growth and transition: why has China's performance been so different? *Journal of Comparative Economics*, 25(3): 422–440.
- Romer, P. M. (1994). The origins of endogenous growth. *Journal of Economic Perspectives*, 8(1): 3–22.
- Soleh, Agus M, Wigena, A. H., Djuraidah, A., & Saefuddin, A. (2015). Statistical downscaling to predict monthly rainfall using linear regression with L1 regularization (LASSO). *Applied Mathematical Sciences*, 9(108): 5361–5369.
- Soleh, Agus Mohamad & others. (2013). LASSO: SOLUSI ALTERNATIF SELEKSI PEUBAH DAN PENYUSUTAN KOEFISIEN MODEL REGRESI LINIER. *Forum Statistika Dan Komputasi*, 18(1).
- Taşel, F., & Bayarcelik, E. B. (2013). The effect of schooling enrolment rates on economic sustainability. *Procedia-Social and Behavioral Sciences*, 99: 104–111.



- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1): 267–288.
- Wang, K., & Chen, Z. (2016). Stepwise regression and all possible subsets regression in education. *Electronic International Journal of Education, Arts, and Science (EIJEAS)*, 2.
- Zareen, S., & Qayyum, A. (2015). An analysis of the impact of government size on economic growth of Pakistan: an endogenous growth. *Research Journal Social Sciences (RJSS)*, 4(1).
- Zifarelli, A., Giglio, M., Menduni, G., Sampaolo, A., Patimisco, P., Passaro, V. M., ... Spagnolo, V. (2020). Partial least-squares regression as a tool to retrieve gas concentrations in mixtures detected using quartz-enhanced photoacoustic spectroscopy. *Analytical Chemistry*, 92(16): 11035–11043.
- Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2): 301–320.