



IPB University
— Bogor Indonesia —

Study Program
Statistics and Data Science
Department of Statistics

PRAKTIKUM 8

ANALISIS REGRESI

Metode dan Kriteria Pemilihan Peubah Penjelas Terbaik

Oleh:

Nabil Bintang Prayoga – G1401221017

Selasa, 8 April 2025



BACKWARD SELECTION

- Dimulai dengan **model penuh** berisi semua k peubah penjelas
- Secara iteratif, peubah penjelas yang **tidak signifikan** akan **dikeluarkan satu per satu** (paling tidak signifikan dari model yang ada)
- Proses **berhenti** ketika **tidak ada** lagi peubah penjelas yang **tidak signifikan**
- **Keuntungan: Persamaan** regresi dengan **semua peubah penjelas** dapat diketahui dan **menghemat waktu** dibandingkan metode *selection* lain
- **Kerugian:** Peubah yang **dihapus tidak dapat** dipertimbangkan kembali untuk masuk ke dalam model
- **Banyaknya model** yang terbentuk adalah $1 + \frac{k(k+1)}{2}$ serta memiliki syarat bahwa $n > k$

FORWARD SELECTION

- Dimulai dengan **model kosong** tanpa peubah penjelas
- Secara iteratif, peubah penjelas yang **paling signifikan** akan **dimasukkan satu per satu** ke dalam model
- Proses **berhenti** ketika **tidak ada** lagi peubah penjelas yang **signifikan** di antara peubah penjelas yang **tersisa**
- **Keuntungan: Satu-satunya** metode *subset selection* yang layak ketika **banyak peubah penjelas sangat besar** yaitu $n < k$
- **Kerugian:** Peubah yang **dimasukkan tidak dapat** dikeluarkan dari model sehingga **bisa saja bukan model terbaik**
- **Banyaknya model** yang terbentuk adalah $1 + \frac{k(k+1)}{2}$

STEPWISE SELECTION / hybrid / bertatar

- **Gabungan** antara metode *forward* dan *backward*
- Dimulai dengan **model kosong** tanpa peubah penjelas
- **Proses dilakukan seperti *forward*** tetapi peubah penjelas yang masuk dapat **dikeluarkan seperti *backward***
- Proses **berhenti** ketika **tidak ada** lagi peubah penjelas yang **tidak signifikan dalam model** dan **tidak ada** peubah penjelas yang **signifikan** di antara peubah penjelas yang **tidak masuk model**
- **Keuntungan:** Pendekatan metode ***subset selection* terbaik**

BEST SUBSET SELECTION

- Melakukan pemilihan model terbaik dengan mencoba **setiap kemungkinan kombinasi dari k peubah penjelas**
- Model yang dibentuk **sebanyak 2^k** yang terdiri dari k buah model berisi satu peubah penjelas, $\binom{k}{2}$ model berisi tepat dua peubah penjelas, dan seterusnya sampai ada model dengan semua peubah penjelas.
- **Keuntungan: Sederhana dan menarik** secara konseptual
- **Kerugian: Mahal** dalam komputasinya sehingga tidak dapat diterapkan pada k peubah penjelas yang sangat besar

KRITERIA PEMILIHAN MODEL

- *Adjusted R-Square* terbesar
- MSE (*Mean Squared Error*) terkecil
- AIC (*Akaike's Information Criterion*) terkecil
- BIC (*Bayesian Information Criterion*) terkecil
- C_p mallow terkecil dan terdekat dengan banyaknya parameter ($k+1$)
- PRESS (*Prediction Sum of Squares*) terkecil



SHRINKAGE METHODS / penyusutan

Menambahkan **pinalti/kendala** untuk parameter model dalam pendugaan MKT/OLS

a. Ridge Regression

- Menduga koefisien regresi ketika terdapat **multikolinearitas tinggi**
- Digunakan pada pemodelan regresi dengan **peubah penjelas sangat banyak** ($k \gg n$)
- Dugaan parameter **berbias** dan selalu **mempertahankan semua** peubah penjelasnya

b. Lasso Regression

- **Menyeleksi** peubah penjelas dalam model dengan $k \gg n$
- Dugaan parameter **berbias** dan memungkinkan **tidak semua** peubah penjelas dipilih (memiliki koefisien yang tidak sama dengan nol)

TERIMA KASIH



IPB University
— Bogor Indonesia —

Department of Statistics
Jl. Meranti W22 L4
Kampus IPB Dramaga Bogor 16680
Telp.: 0251-8624535
E-mail: statistika@apps.ipb.ac.id