

Komputasi Statistika

Johannes Pande Manurung

Membuat Fungsi Regresi Linear dengan Class System S3

Penjelasan

Fungsi “reglinear” didefinisikan dengan membutuhkan tiga atribut, yaitu y (peubah respons), x (peubah penjelas), dan metode. Mode data yang digunakan pada y (peubah respons) dan x (peubah penjelas) adalah numerik. Peubah penjelas pada fungsi ini dapat menggunakan lebih dari satu peubah (dapat menjalankan analisis regresi linear berganda). Untuk menggunakan lebih dari satu peubah penjelas, merge/cbind data antara setiap peubah penjelas dan didefinisikan menjadi x . Metode pendugaan menggunakan teknik optimasi. Terdapat dua metode pendugaan parameter regresi yang tersedia, yaitu “metode kuadrat terkecil” (metode default) dan “metode kemungkinan”.

```
reglinear <- function(y,x,metode = "metode kuadrat terkecil") {  
  obj <- list(y = y, x = x, metode = metode)  
  y <- obj$y  
  n <- length(y)  
  obj$x <- cbind(1, obj$x)  
  x <- obj$x  
  m <- ncol(x)  
  metode <- obj$metode  
  
  if (!is.numeric(x) || !is.numeric(y) || !all(is.finite(x)) ||  
  !all(is.finite(y)))  
    stop("Input harus numerik")  
  if (nrow(x) != length(y))  
    stop("Panjang x dan y berbeda")  
  
  class(obj) <- "reglinear"  
  
  # Pemilihan metode  
  if(metode == "metode kuadrat terkecil"){  
    JKSisaan <- function(beta) {  
      sum((y - x %*% beta)^2)  
    }  
    hasil <- optim(rep(0,m), JKSisaan, method = "BFGS")  
  }  
  else if(metode == "metode kemungkinan"){  
    lossLikelihood <- function(beta) {  
      mu <- x %*% beta  
      -1*sum(dnorm(y, mean = mu, sd = sd(y - mu), log = TRUE))  
    }  
    hasil <- optim(rep(0,m), lossLikelihood, method = "BFGS")  
  }  
}
```

```

    }
  else{
    stop("Metode tidak diketahui.
          Gunakan 'metode kuadrat terkecil' atau 'metode kemungkinan'.")
  }

  obj$coefficients <- hasil$par
  obj$fittedvalues <- x %*% obj$coefficients
  obj$residuals <- (y - obj$fittedvalues)
  obj

}

```

Fungsi summary untuk class “reglinear”

Penjelasan

Fungsi summary didefinisikan untuk menampilkan ringkasan hasil model. Terdapat dua tabel yang ditampilkan, yaitu tabel koefisien dan tabel anova. Tabel koefisien menampilkan hasil pendugaan parameter/koefisien, galat baku/standard error, nilai uji-t statistik, dan p-value. Tabel anova menampilkan hasil anova regresi yang terdiri dari derajat bebas, jumlah kuadrat, kuadrat Tengah, nilai uji-F statistik, dan p-value. Selain itu, ditampilkan nilai R-Square dan Adjusted R-Squared.

```

summary.reglinear <- function(obj) {
  coefficients <- obj$coefficients
  fittedvalues <- obj$fittedvalues
  residuals <- obj$residuals

  n <- length(obj$y)
  m <- length(coefficients)

  # Membuat Tabel Anova
  y <- obj$y
  ybar <- mean(obj$y)
  dbRegresi <- m-1
  dbTotal <- n-1
  dbGalat <- dbTotal - dbRegresi
  JKTotal <- sum((y-ybar)^2)
  JKRegresi <- sum((fittedvalues-ybar)^2)
  JKGalat <- JKTotal-JKRegresi
  KTRegresi <- JKRegresi/dbRegresi
  KTGalat <- JKGalat/dbGalat
  Fhit <- KTRegresi/KTGalat
  pvalue_anova <- pf(Fhit, dbRegresi, dbGalat, lower.tail = FALSE)

  rowregresi <- as.numeric(c(dbRegresi, JKRegresi, KTRegresi, Fhit,

```

```

pvalue_anova))
rowgalat <- as.numeric(c(dbGalat, JKGalat, KTGalat, NA, NA))
rowtotal <- as.numeric(c(dbTotal, JKTotal, NA, NA, NA))
Tabel_Anova <- rbind(rowregresi, rowgalat, rowtotal)
colnames(Tabel_Anova) <- c("db", "SS", "MS", "F", "p-value")
rownames(Tabel_Anova) <- c("Regresi", "Error", "Total")

Rsq <- JKRegresi/JKTotal
Adj_Rsq <- 1-((1-Rsq)*(n-1)/(n-(m-1)-1))

# Membuat Tabel Koefisien
x <- obj$x
c <- solve(t(x)%*%x)
se_b <- suppressWarnings(sqrt(KTGalat*c))

# Mencari standard error
se <- c(rep(0,m))
for (i in 1:m) {
  se[i] <- se_b[i,i]
  i <- i+1
}

# Mencari t-value koefisien
tvalue_koef <- c(rep(0,m))
for (i in 1:m) {
  tvalue_koef[i] <- (coefficients[i]-0)/se[i]
  i <- i+1
}

# Mencari p-value koefisien
pvalue_koef <- c(rep(0,m))
for (i in 1:m) {
  pvalue_koef[i] <- 2*pt(-abs(tvalue_koef[i]),df <- n-m)
  i <- i+1
}

Tabel_koef <- cbind(coefficients, se, tvalue_koef, pvalue_koef)
colnames(Tabel_koef) <- c("Koefisien", "Std. Error", "t-value", "p-value")
rownames(Tabel_koef) <- c(rep(0,m))
for (i in 1:m) {
  if (i==1) rownames(Tabel_koef)[i] <- "Intersep"
  else rownames(Tabel_koef)[i] <- paste("X", (i-1), sep="")
  i <- i+1
}

cat("Tabel Koefisien:\n")
print(Tabel_koef)
cat("\n")

```

```

cat("Tabel Anova:\n")
print(Tabel_Anova)
cat("\n")
cat("Residual standard error:", sqrt(KTGalat), "on", dbGalat, "degrees of
freedom\n")
cat("Multiple R-squared:", Rsq, "\n")
cat("R-squared Adj:", Adj_Rsq, "\n")
cat("F-statistic:", Fhit, "on", dbRegresi, "and", dbGalat, "DF, p-value:",
pvalue_anova, "\n")
}

```

Fungsi plot untuk class “reglinear”

Penjelasan

Fungsi plot didefinisikan untuk menampilkan dua buah plot, yaitu plot quantile normal dari error dan plot fitted vs error.

```

plot.reglinear <- function(obj) {
  fittedvalues <- obj$fittedvalues
  residuals <- obj$residuals

  par(mfcol = c(1, 2))

  # Normal Q-Q plot
  qqnorm(residuals, main = "Residuals Q-Q Plot")
  qqline(residuals)

  # Fitted vs Residuals plot
  plot(fittedvalues, residuals,
        main = "Fitted vs Residuals",
        xlab = "Fitted Values",
        ylab = "Residuals")
  abline(h = 0, col = "black")
}

```

Interface dengan Shiny

Penjelasan

Interface didefinisikan menggunakan shiny. Panel interface tedapat disebelah kiri dan menerima file input berupa “.csv” dan “.xlsx”. Pengguna dapat memilih satu peubah respons (y) pada kolom peubah respons dan beberapa peubah penjelas (x) pada kolom peubah penjelas. Terdapat dua metode yang dapat dipilih, yaitu “metode kuadrat terkecil” dan “metode kemungkinan”. Tombol “Analisis” dapat ditekan untuk memulai analisis. Hasil yang ditampilkan berupa output fungsi “summary” dan “plot” sehingga dihasilkan tabel koefisien (berisi hasil pendugaan parameter/koefisien, galat baku/standard error,

nilai uji-t statistik, dan p-value), tabel anova, nilai R-Square dan Adjusted R-Squared, serta plot quantile normal dari error dan plot fitted vs error.

```
library(shiny)
ui <- fluidPage(
  titlePanel("Analisis Regresi"),
  sidebarLayout(
    sidebarPanel(
      fileInput("file", "Pilih File (CSV/Excel)", accept = c(".csv",
                                                               ".xlsx")),
      uiOutput("var_select"),
      selectInput("Metode", "Metode", choices = c("metode kuadrat terkecil",
                                                    "metode kemungkinan")),
      actionButton("analisis", "Analisis")
    ),
    mainPanel(
      verbatimTextOutput("summary"),
      plotOutput("plot")
    )
  )
)

server <- function(input, output, session){
  data <- reactive({
    req(input$file)
    ext <- tools:::file_ext(input$file$datapath)
    library(readxl)
    switch(ext,
      csv = read.csv(input$file$datapath),
      xlsx = read_excel(input$file$datapath),
      stop(" Harap unggah file CSV atau Excel."))
  })
}

output$var_select <- renderUI({
  req(data())
  colnames <- names(data())
  tagList(
    selectInput("respons", "Peubah Respons", choices = colnames),
    selectInput("penjelas", "Peubah Penjelas", choices = colnames, multiple
= TRUE)
  )
})

model <- reactive({
  req(data(), input$respons, input$penjelas)
  y <- data()[[input$respons]]
  x <- as.matrix(data()[, input$penjelas, drop = FALSE])
  reglinear(y, x, metode = input$Metode)
```

```
)  
  
output$summary <- renderPrint({  
  req(input$analisis)  
  summary(model())  
})  
  
output$plot <- renderPlot({  
  req(input$analisis)  
  plot(model())  
})  
}  
  
shinyApp(ui = ui, server = server)
```