

Zadanie dodatkowe - R Markdown

Maciej Beręsewicz

19 May 2015

Proszę o odtworzenie poniższego dokumentu (w HTML lub PDF). Liczby w macierzach muszą być uzależnione od danych wylosowanych ze zbioru `cars` (bez użycia funkcji `set.seed()`). Oznacza to de facto, że będą aktualizowane za każdym razem jak nastąpi kompilacja dokumentu.

Kody proszę przysyłać przez **GitHub**.

Należy użyć poniższej funkcji do wyświetlenia macierzy (Kod pochodzi z [Źródło 1](#), [Źródło 2](#)).

```
pdf.options(encoding='ISOLatin2.enc')

m2l <- function(matr) {
  printmrow <- function(x) {
    ret <- paste(paste(x, collapse = " & "), "\\")
    sprintf(ret)
  }
  out <- apply(matr, 1, printmrow)
  out2 <- paste("\\begin{bmatrix}", paste(out, collapse=' '), "\\end{bmatrix}")
  return(out2)
}
```

Regresja liniowa

Model regresji możemy zapisać następująco

$$Y = \beta X + \epsilon$$

Natomiast estymator KMNK β określony jest następująco

$$\hat{\beta} = (X^T X)^{-1} X^T Y.$$

Dane do regresji

Użyjemy danych ze zbioru `cars` i zbudujemy model regresji `dist ~ speed`. Dodatkowo, losowo (wykorzystując funkcję `sample`) wybierzemy jedynie 20 obserwacji aby dokument był czytelny.

```
car_sub <- cars[sample(1:50, 20),]
```

Macierz X wygląda następująco

$$\mathbf{X} = \begin{bmatrix} 1 & 18 \\ 1 & 18 \\ 1 & 17 \\ 1 & 11 \\ 1 & 20 \\ 1 & 20 \\ 1 & 20 \\ 1 & 7 \\ 1 & 19 \\ 1 & 25 \\ 1 & 13 \\ 1 & 10 \\ 1 & 14 \\ 1 & 12 \\ 1 & 19 \\ 1 & 15 \\ 1 & 4 \\ 1 & 22 \\ 1 & 17 \\ 1 & 13 \end{bmatrix}$$

W związku z tym $\hat{\beta}$ ma następującą postać

$$(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} = \left(\begin{bmatrix} 1 & 18 \\ 1 & 18 \\ 1 & 17 \\ 1 & 11 \\ 1 & 20 \\ 1 & 20 \\ 1 & 20 \\ 1 & 7 \\ 1 & 19 \\ 1 & 25 \\ 1 & 13 \\ 1 & 10 \\ 1 & 14 \\ 1 & 12 \\ 1 & 19 \\ 1 & 15 \\ 1 & 4 \\ 1 & 22 \\ 1 & 17 \\ 1 & 13 \end{bmatrix}^T \begin{bmatrix} 1 & 18 \\ 1 & 18 \\ 1 & 17 \\ 1 & 11 \\ 1 & 20 \\ 1 & 20 \\ 1 & 20 \\ 1 & 7 \\ 1 & 19 \\ 1 & 25 \\ 1 & 13 \\ 1 & 10 \\ 1 & 14 \\ 1 & 12 \\ 1 & 19 \\ 1 & 15 \\ 1 & 4 \\ 1 & 22 \\ 1 & 17 \\ 1 & 13 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 18 \\ 1 & 18 \\ 1 & 17 \\ 1 & 11 \\ 1 & 20 \\ 1 & 20 \\ 1 & 20 \\ 1 & 7 \\ 1 & 19 \\ 1 & 25 \\ 1 & 13 \\ 1 & 10 \\ 1 & 14 \\ 1 & 12 \\ 1 & 19 \\ 1 & 15 \\ 1 & 4 \\ 1 & 22 \\ 1 & 17 \\ 1 & 13 \end{bmatrix}^T \begin{bmatrix} 84 \\ 76 \\ 40 \\ 28 \\ 56 \\ 32 \\ 64 \\ 22 \\ 68 \\ 85 \\ 34 \\ 26 \\ 36 \\ 24 \\ 36 \\ 20 \\ 2 \\ 66 \\ 32 \\ 26 \end{bmatrix}$$

Co daje nam

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} = \left(\begin{bmatrix} 20 & 314 \\ 314 & 5446 \end{bmatrix} \right)^{-1} \begin{bmatrix} 857 \\ 15299 \end{bmatrix} = \begin{bmatrix} -13.2375048430841 \\ 3.57245253777606 \end{bmatrix}$$

Rysujemy wykres regresji

Wykres przedstawia zależność między `dist` a `speed` oraz oszacowaną przywą regresji. Należy użyć funkcji `plot` oraz `abline`.

Zależność między dystansem, a prędkością

