Zadanie 1:
Drużyna informatyków testuje nowy program rozciągania pleców w przerwach od programowania zespołowego, by ograniczyć liczbę dolegliwości. Poniższa tabelka pokazuje codzienną liczbę minut rozciągania się przez uczestników i powiązaną z nimi liczbę dolegliwości przez cały semestr.

(a) Znajdź prostą regresji, pokazującą zależność ilości dolegliwości od czasu rozciągania.

(b) O ile spada ilość dolegliwości z każdą minutą rozciągania?

(c) Ile minut rozciągania jest potrzebne, by zawodnik uniknął dolegliwości?

Rozwiązanie:

(a)

Y - ilość dolegliwości

X - czas rozciągania

n = 8

$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{160}{8} = 20$

$\overline{Y} = \frac{\sum_{i=1}^{n} Y_i}{n} = \frac{14}{8} = 1.75$

$\beta_1 = \frac{\sum_{i=1}^{n}(X_i - \overline{X}) * (Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2} = \frac{-120}{1500} = -0.08$

$\beta_0 = \overline{Y} - \beta_1 * \overline{X} = 1.75 - (-0.08) * 20 = 3.35$

Prosta regresji to

$Y = \beta_0 + X\beta_1$, czyli

$Y = 3.35 - 0.08 * X$

(b) Z każdą minutą rozciągania ilość dolegliwości spada średnio o 0,08

(c)

$0 = 3.35 - 0.08X$

$X = 41.875 \approx 42$

Żeby zawodnik uniknął dolegliwości powinien się rozciągać ok 42 min

Zadanie 2:
Pokaż, że jeżeli $n \geq p$ oraz $rank(X) = n$ to estymator β metodą najmniejszych kwadratów

$\widehat{\beta} = argmin\|Y - X\beta\|^2$,

jest postaci

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

Dodatkowo jeżeli $\varepsilon \sim N(0, \sigma^2 Id)$ to pokaż że $\widehat{\beta}$ jest estymatorem największej wiarygodności oraz znajdź estymator największej wiarygodności σ^2.

Rozwiązanie:

X- macierz planu $\in R^{nxp}$

$\beta \in R^p$ -wektor nieznanych współczynników

$Y \in R^n$ - zmienna objaśniana

$Y = X\beta + \epsilon$

$rank(X) = n$ czyli macierz jet pełnego rzędu

$RSS = argmin\|Y - X\beta\|^2 = (Y - X\beta)^T(Y - X\beta) = YY^T - 2Y^TX\beta + \beta^TX^TX\beta$

$\frac{\delta RSS}{\delta\beta} = -2X^TY + 2X^TX\beta = -2X^T(Y - X\beta)$

$X^T(Y - X\beta) = 0$

$\hat{\beta} = (X^TX)^{-1}X^TY$

Jeśli $\varepsilon \sim N(0, \sigma^2 Id)$ to $Y \sim N(X\beta, I\sigma^2)$

$L = \frac{1}{(2\pi)^{\frac{n}{2}}(I\sigma^2)^{\frac{n}{2}}} e^{-\frac{1}{2}*(Y-\beta X)\sigma^{-2}I(Y-\beta X)} = \frac{1}{(2\pi)^{\frac{n}{2}}(\sigma^2)^{\frac{n}{2}}} e^{-\frac{1}{2\sigma^2}*\|Y-X\beta\|^2}$

$ln(L) = ln(\frac{1}{(2\pi)^{\frac{n}{2}}}) + ln(\frac{1}{(\sigma^2)^{\frac{n}{2}}}) + ln(e^{-\frac{1}{2\sigma^2}*\|Y-X\beta\|^2}) =$

$= -\frac{n}{2}ln(2\pi) - \frac{n}{2}ln(\sigma^2) - \frac{1}{2\sigma^2} * \|Y - X\beta\|^2$

Od β zależy jedynie $\frac{1}{2\sigma^2} * \|Y - X\beta\|^2$, które przyjmuje największe wartości dla $\beta = (X^TX)^{-1}X^TY$

$z = \sigma^2$

$\frac{\delta ln(L)}{\delta z} = -\frac{n}{2z} + \frac{1}{2z^2} * \|Y - X\beta\|^2$

$-\frac{n}{2z} + \frac{1}{2z^2} * \|Y - X\beta\|^2 = 0$

$-nz + \|Y - X\beta\|^2 = 0$

$z = \frac{\|Y - X\beta\|^2}{n}$

$\sigma^2 = \frac{\|Y - X\beta\|^2}{n}$ - estymator największej wiarygodności σ^2

Zadanie 3:

Oblicz wartość oczekiwaną i wariancję $\hat{\beta}$. Jaki ma rozkład $\hat{\beta}$, jeśli $\epsilon \sim N(0, \sigma^2 Id)$?

Rozwiązanie: Obliczam wartość oczekiwaną i wariancję, korzystam z $\mathbb{E}Y = X\beta$:

$$\begin{aligned}
\mathbb{E}\hat{\beta} &= \mathbb{E}\left[(X^TX)^{-1}X^TY\right] = (X^TX)^{-1}X^T\mathbb{E}Y = (X^TX)^{-1}X^TX\beta = I\beta = \beta \\
\operatorname{Var}\hat{\beta} &= \mathbb{E}(\hat{\beta} - \beta)^2 = \mathbb{E}(\hat{\beta} - \beta)^T(\hat{\beta} - \beta) \\
&= \mathbb{E}\left(\left(X^TX\right)^{-1}X^TY - \left(X^TX\right)^{-1}X^T\mathbb{E}Y\right)\left(\left(X^TX\right)^{-1}X^TY - \left(X^TX\right)^{-1}X^T\mathbb{E}Y\right)^T \\
&= \mathbb{E}\left(\left(X^TX\right)^{-1}X^T(Y - \mathbb{E}Y)\right)\left(\left(X^TX\right)^{-1}X^T(Y - \mathbb{E}Y)\right)^T \\
&= \left(X^TX\right)^{-1}X^T\mathbb{E}\left[(Y - \mathbb{E}Y)(Y - \mathbb{E}Y)^T\right]\left(\left(X^TX\right)^{-1}X^T\right)^T \\
&= \left(X^TX\right)^{-1}X^T\sigma^2 I_n\left(\left(X^TX\right)^{-1}X^T\right)^T = \sigma^2\left(X^TX\right)^{-1}\left(X^TX\right)\left(X^TX\right)^{-1} \\
&= \sigma^2\left(X^TX\right)^{-1}
\end{aligned}$$

Traktujemy Y jako zmienną losową, ma ona rozkład normalny, tak jak szum ϵ (bo X oraz β są ustalone). Z zadania 2. $\hat{\beta}$ jest funkcją liniową od Y, więc również ma rozkład normalny. Średnią i wariancję już znamy: $\hat{\beta} \sim N(\beta, \sigma^2(X^TX)^{-1})$.

Zadanie 4:
Pokaż, że predykcja $\hat{Y} = X\hat{\beta}$ jest rzutem prostopadłym Y na przestrzeń liniową rozpiętą przez kolumny macierzy X.

Rozwiązanie:

$\hat{Y} = X\hat{\beta}$

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

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

Wektor $\hat{Y}$ jest rzutem ortogonalnym na przestrzeń liniową rozpiętą przez wektory będące kolumnami X

A więc ta predykcja jest rzutem prostokątnym na przestrzeń liniową rozpiętą przez kolumny macierzy X.