Soluzione al Problema 0

Aldo Solari, matricola 2575

Risposte

La risposta alla domanda 0 è volutamente prolissa per esigenze esemplicative.

0.

Il modello di regressione polinomiale di terzo grado stimato con le 30 osservazioni del training set fornisce le seguenti stime dei coefficienti:

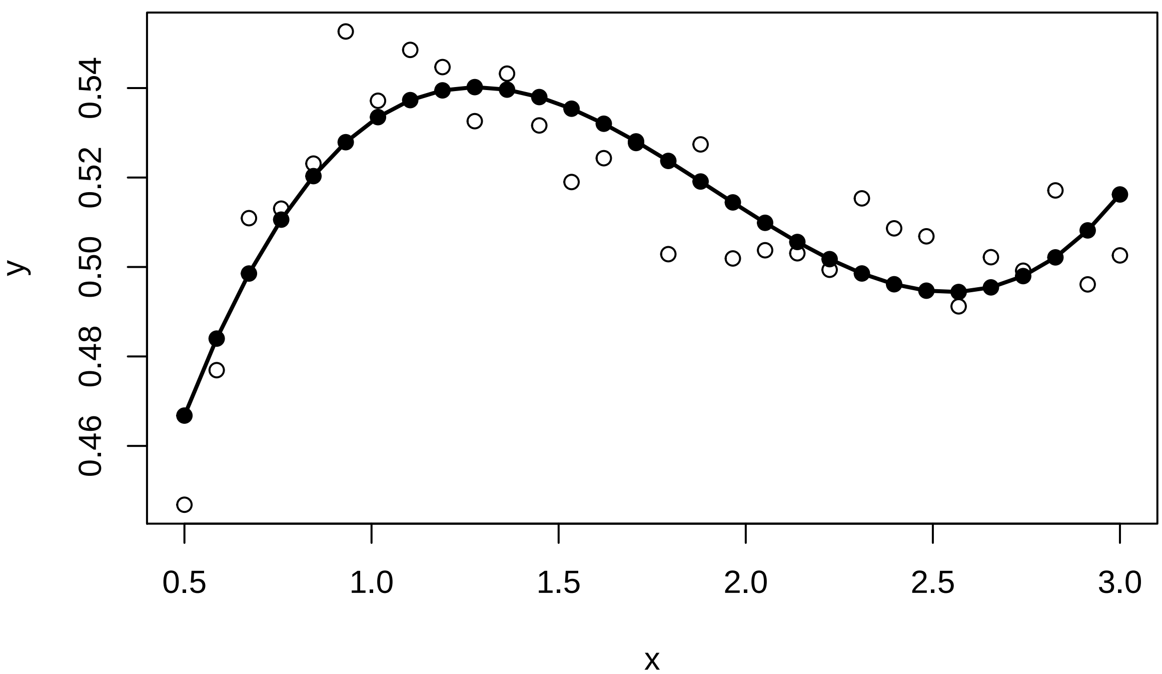
```
##    hatbeta1    hatbeta2    hatbeta3    hatbeta4
##  0.30546280  0.44075802 -0.25878583  0.04509449
```

Dato un qualunque $x \in \mathbb{R}$, il modello prevede

$$\hat{y} = \hat{\beta}_1 + \hat{\beta}_2 x + \hat{\beta}_3 x^2 + \hat{\beta}_4 x^3$$

Nel seguito indicheremo con $\hat{y}_i$ i valori previsti per x_i, $i = 1, \ldots, n$.

Il seguente grafico rappresenta il diagramma di dispersione delle osservazioni del training set (x_i, y_i) (pallini vuoti), e le mette a confronto con i valori previsti dal modello $(x_i, \hat{y}_i)$ (pallini pieni).



Infine, $\mathrm{MSE}_{\mathrm{Tr}} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = 1.2469566 \times 10^{-4}$.

I risultati ottenuti si possono riprodurre con il seguente codice R:

```
library(readr)
dat <- read_table2("http://azzalini.stat.unipd.it/Libro-DM/ieri-domani.dat")[-31,]
train <- data.frame(x=dat$x, y=dat$y.domani)
d <- 3
fit <- lm(y ~ poly(x, degree=d, raw=T), train)
betahat <- coef(fit)
names(betahat) <- paste("hatbeta", 1:(d+1), sep="")
yhat <- fitted(fit)
plot(y ~ x, train)
points(yhat ~ x, train, pch=19)
lines(yhat ~ x, train, lwd=2)
MSE.tr <- mean((train$y - yhat)^2)
```

1.

...