1. done

2. done

3. i j sigmahat R-squared

Model 2 3 31.52520 0.4795617

Model 2 6 32.74106 0.4386434

Model 3 5 32.82083 0.4359044

Model 2 5 33.27873 0.4200548

Model 1 3 34.74315 0.3678912

Model 1 6 35.33417 0.3462026

Model 3 6 35.43275 0.3425494

Model 1 5 35.59213 0.3366215

Model 5 6 35.77746 0.3296949

Model 1 2 37.15697 0.2770069

Model 4 5 37.64381 0.2579374

Model 3 4 38.23286 0.2345320

Model 4 6 38.57122 0.2209233

Model 2 4 38.64980 0.2177458

Model 1 4 40.92204 0.1230636

The first model[2,3] is the best by far with the lowest sigmahat and the highest r-squared. models [2,6] and [3,5] are very close together with the sigmahat being separated by about .1 and the r-squared being almost exactly alike.

4.

Call:

lm(formula = y ~ xmat[, 2] + xmat[, 3])

Residuals:

Min 1Q Median 3Q Max

-87.377 -12.918 4.627 17.161 67.539

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 50.4469 15.8250 3.188 0.00255 \*\*

xmat[, 2] -0.7736 0.1640 -4.718 2.17e-05 \*\*\*

xmat[, 3] -0.9313 0.1879 -4.956 9.74e-06 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 31.53 on 47 degrees of freedom

Multiple R-squared: 0.4796, Adjusted R-squared: 0.4574

F-statistic: 21.65 on 2 and 47 DF, p-value: 2.161e-07

It is significant because the P-value is less than 0.5 so the 95% confidence interval will not contain 0.

5.Coefficients:

(Intercept) xmat[, 2] xmat[, 3]

50.4469 -0.7736 -0.9313

95% is -1.112 to -.33.

Confidence Intevals are important because they tell us how useful the data is. The higher the confidence on a smaller interval, the better the data is.

6.

lmbest3 <- function(y,xmat){

#

# This returns R-squared and sigmahat for all linear

# models of size 2 (2 x's). The results are sorted on R-squared.

#

p <- length(xmat[1,])

coll <- c()

for(i in 1:(p-2)){

for(j in (i+1):(p-1)){

for(k in (j+1):p){

fit <- lm(y ~ xmat[,i] + xmat[,j] + xmat[,k])

sfit <- summary(fit)

tmp <- c(i,j,k,sfit$sigma,sfit$r.squared)

coll <- rbind(coll,tmp)

}

}

}

ind <- order(coll[,4],decreasing=TRUE)

collf <- coll[ind,]

colnames(collf) <- c("i","j","k","sigmahat","R-squared")

ic <- length(collf[,1])

rownames(collf) <- rep("Model",ic)

return(collf)

}

7.

lmbest3(y,xmat)

i j k sigmahat R-squared

Model 1 2 4 36.99539 0.2985307

Model 4 5 6 36.13877 0.3306392

Model 3 4 6 35.79407 0.3433475

Model 1 4 6 35.69916 0.3468250

Model 1 4 5 35.27705 0.3621802

Model 1 3 4 34.82294 0.3784953

Model 1 5 6 33.16619 0.4362267

Model 2 4 6 33.09418 0.4386720

Model 2 4 5 33.05304 0.4400666

Model 3 4 5 32.97633 0.4426626

Model 3 5 6 32.40143 0.4619263

Model 1 2 5 31.91208 0.4780561

Model 1 3 6 31.69097 0.4852639

Model 2 3 4 31.66212 0.4862009

Model 2 5 6 30.59784 0.5201615

Model 1 2 6 30.54667 0.5217651

Model 1 3 5 29.65990 0.5491284

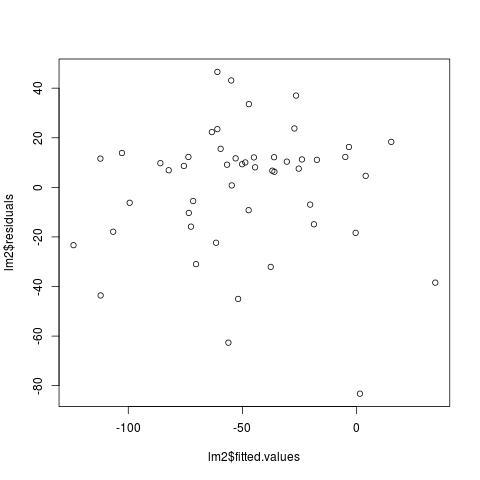
Model 1 2 3 29.03445 0.5679434

Model 2 3 6 28.31877 0.5889808

Model 2 3 5 26.31224 0.6451630

Model 2 3 5 is the best model for lmbest3(y,xmat). It has the lowest sigmahat and highest r-squared. Compared to the the best model of lmbest2(y,xmat) it is a significant improvement. The sigma hat is lower by almost 5 and the r-squared is larger by .2

8.



We’re not completely sure what to comment on this, but we did notice quite a lot of points are heavily negative.

