

Homework 7

Solution

Question 1. Finding and plotting residuals

8.2(a) (10 pts) Fit the model $E(y) = \beta_0 + \beta_1 x$ to the data.

```
# Import the data and fit the model
EX8.2 = read.csv("STAT 3113 Data Sets/EX8_2.csv")
fit = lm(Y~X, data = EX8.2)
summary(fit)

##
## Call:
## lm(formula = Y ~ X, data = EX8.2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max 
## -7.1856 -2.2261 -0.7216  3.2107  5.9046 
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) -3.179     2.747  -1.157   0.281    
## X            2.491     0.180   13.836 7.2e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.154 on 8 degrees of freedom
## Multiple R-squared:  0.9599, Adjusted R-squared:  0.9549 
## F-statistic: 191.4 on 1 and 8 DF,  p-value: 7.197e-07
```

Answer: From the R output, the Fitted model is $\hat{y} = -3.179 + 2.491x$

8.2(b) (10 pts) Calculate the residuals for the model.

```
# Find the residuals
residuals = resid(fit)
residuals

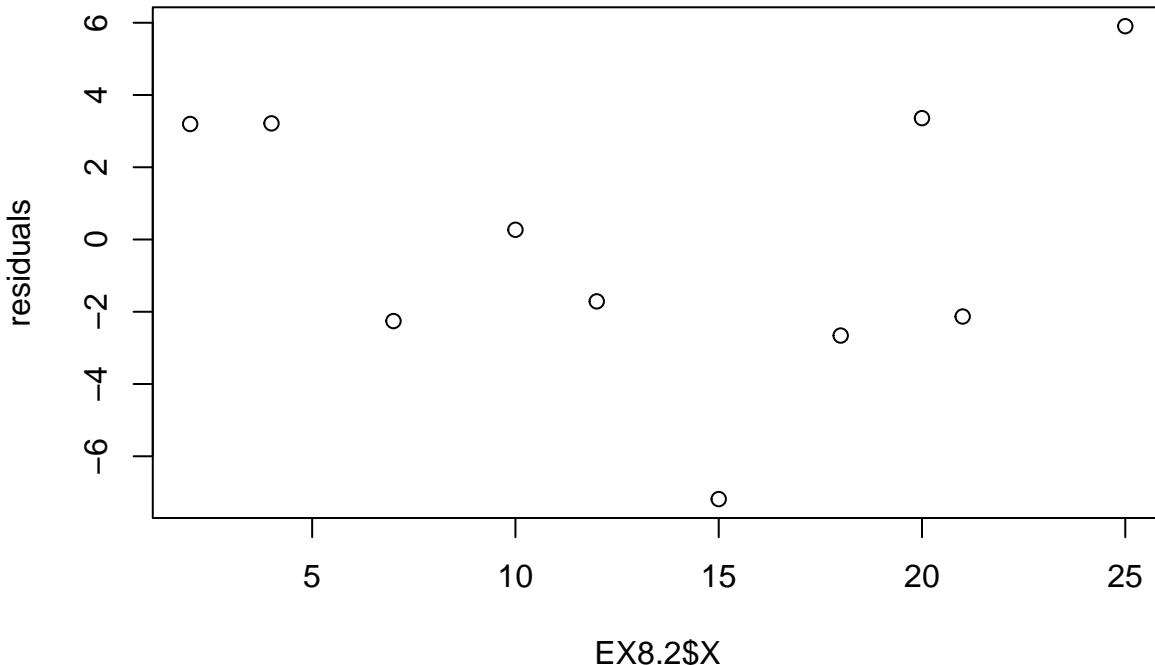
##
##      1       2       3       4       5       6       7
## 3.1972201 3.2152517 -2.2577010 0.2693464 -1.7126221 -7.1855748 -2.6585274
##      8       9      10
## 3.3595041 -2.1314801 5.9045830
```

Answer: The residuals are shown above.

8.2(c) (20 pts) Plot the residuals versus x . Do you detect any trends? If so, what does the pattern suggest about the model?

```
# Plot residuals vs x
```

```
plot(EX8.2$X, residuals)
```



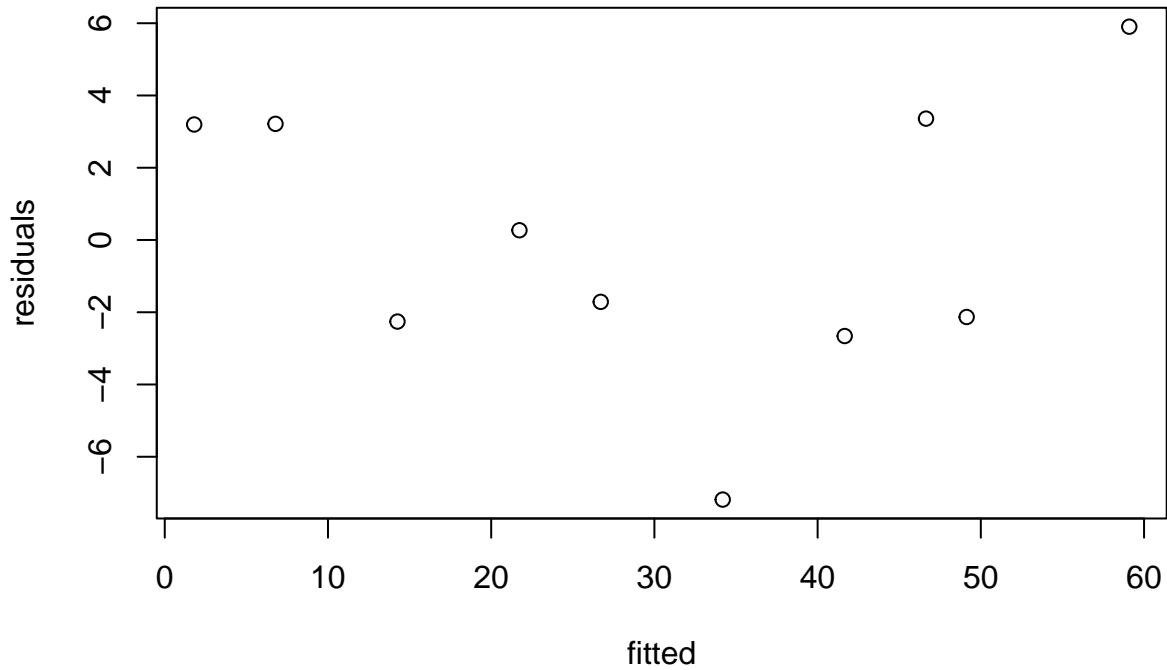
EX8.2\$X

Answer: From the residual plot, we see a quadratic U-shaped trend. For small and large values of x , the residuals are positive, while for intermediate values of x , the residuals are negative. This pattern indicates a lack-of-fit in the straight-line model and suggests to fit the quadratic model. To verify this, we would fit the quadratic model $E(y) = \beta_0 + \beta_1x + \beta_2x^2$ and test for curvature, i.e., test $H_0 : \beta_2 = 0$.

8.10 (20 pts) Plot the residuals versus \hat{y} (Note that, it is different from 8.2(c)). Do you detect any trends? If so, what does the pattern suggest about the model?

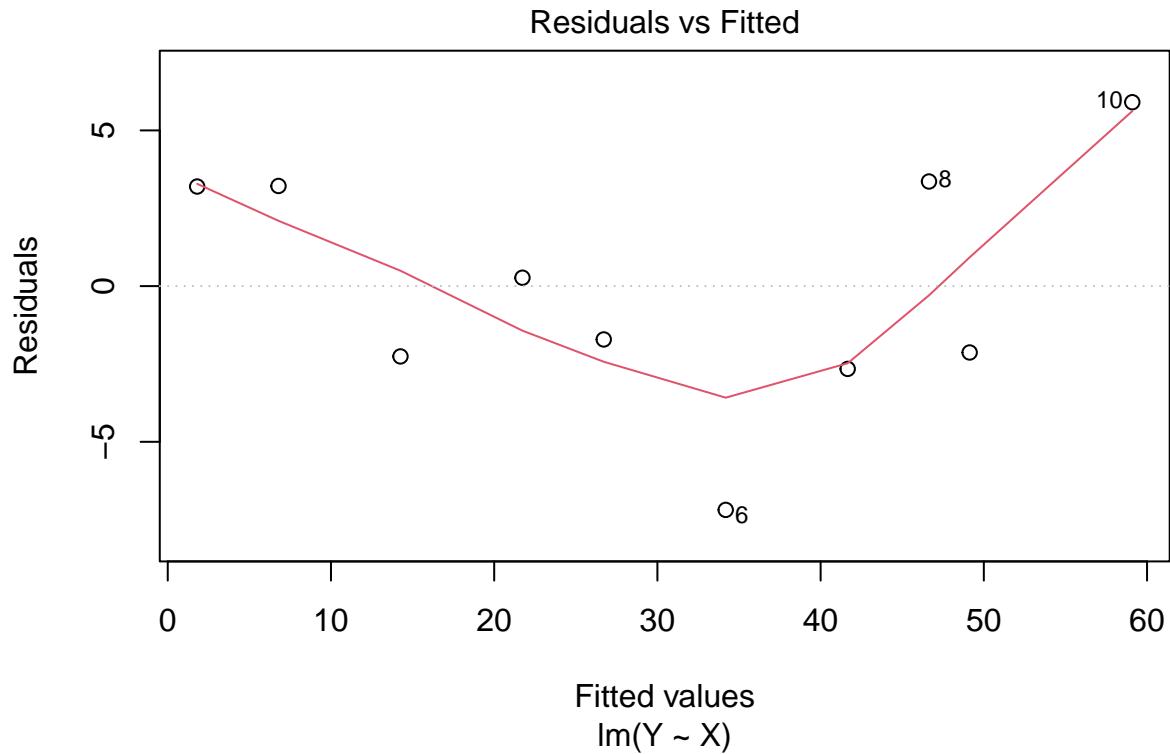
```
# Plot residuals vs y_hat
# Note that: not residual vs x as in 8.2(c)

fitted = fitted(fit)
plot(fitted, residuals)
```



```
# Or the first plot of plot(fit, which = 1)
```

```
plot(fit, which = 1)
```



Answer: Yes, there appears to be a trend. The assumption of constant variance appears to be satisfied. The pattern suggests adding curvature to the model.

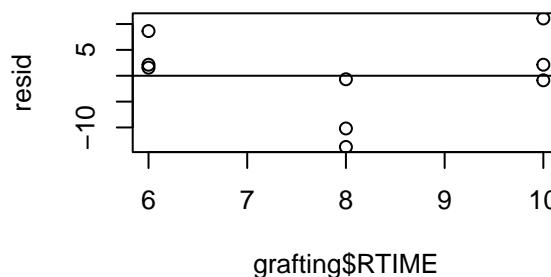
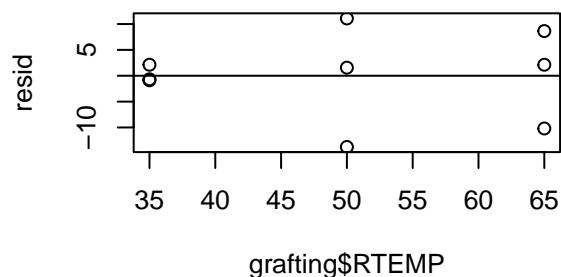
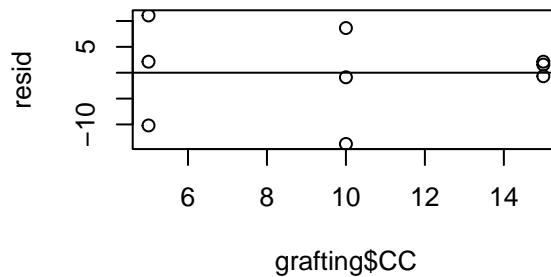
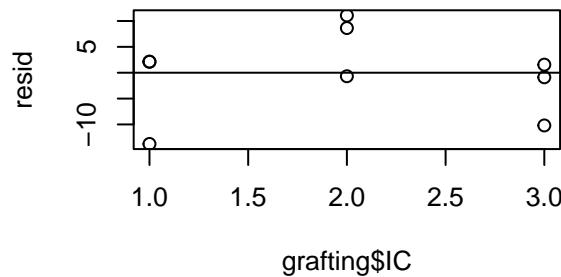
Question 2. A rubber additive made from cashew nut shells (Data set: GRAFT-ING)

8.4(a) (20 pts) Plot the model residuals against each of the four independent variables using R.

```
### Import data and plot the four residual plots
### Hint: you might need to use par(mfrow=c(2,2))
```

```
grafting = read.csv("STAT 3113 Data Sets/GRAFTING.csv")
fit_grafting = lm(GE ~ IC + CC + RTEMP + RTIME, data = grafting)
resid = resid(fit_grafting)

par(mfrow = c(2,2))
plot(grafting$IC, resid)
abline(h=0)
plot(grafting$CC, resid)
abline(h=0)
plot(grafting$RTEMP, resid)
abline(h=0)
plot(grafting$RTIME, resid)
abline(h=0)
```



```
par(mfrow=c(1,1))
```

8.4(b) (20 pts) How would you advise the researchers to modify the model? Explain

Answer: The plots of the residuals against both IC and RTIME appear to be curved. This implies that second-order terms for both IC (IC_sq) and RTIME (RTIME_sq) should be added to the model for a better fit.