Week 2 Homework

Ben Arancibia

June 20, 2015

HA 4.1 Electricity consumption was recorded for a small town on 12 randomly chosen days. The following maximum temperatures (degrees Celsius) and consumption (megawatt-hours) were recorded for each day.

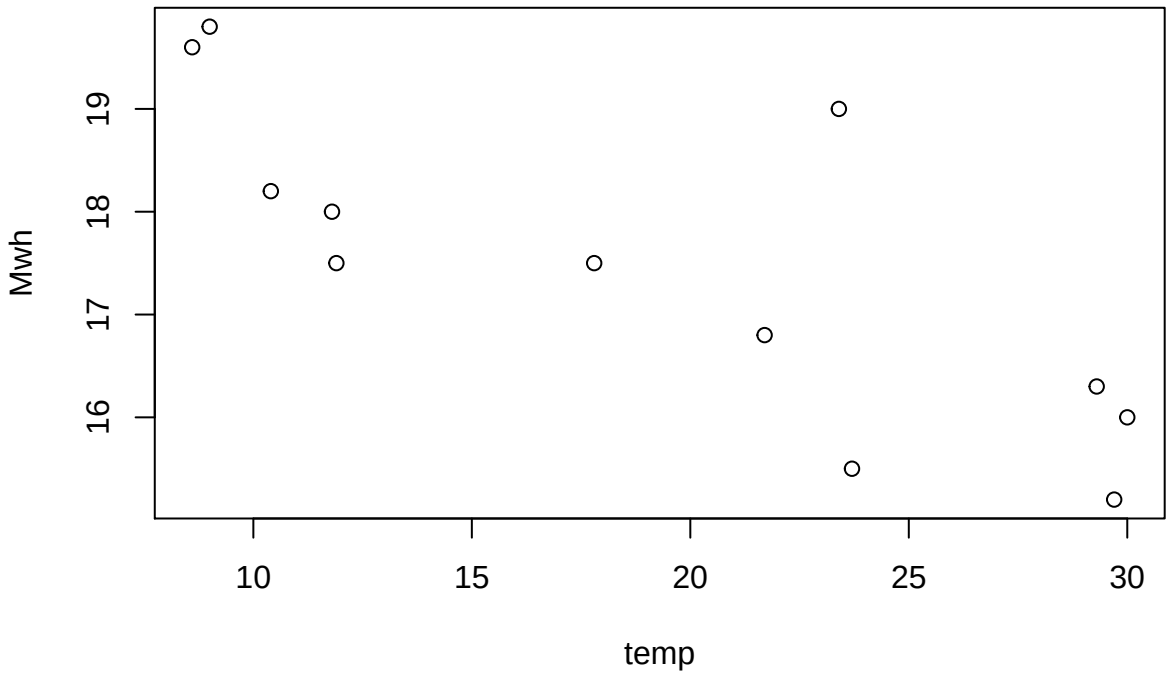
```
library(fpp)
```

```
knitr::kable(econsumption)
```

Mwh	temp
16.3	29.3
16.8	21.7
15.5	23.7
18.2	10.4
15.2	29.7
17.5	11.9
19.8	9.0
19.0	23.4
17.5	17.8
16.0	30.0
19.6	8.6
18.0	11.8

a) Plot the data and find the regression model for Mwh with temperature as an explanatory variable. Why is there a negative relationship?

```
plot(Mwh ~ temp, data=econsumption)
```



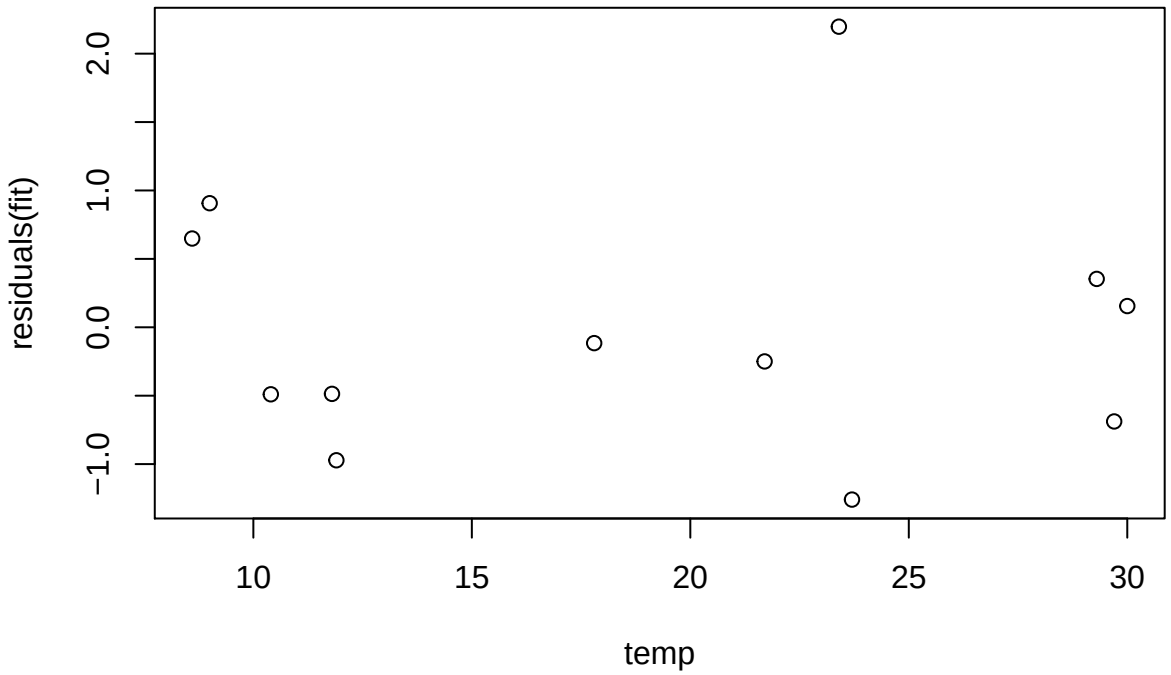

```
fit  <- lm(Mwh ~ temp, data=econsumption)
summary(fit)
```

```
##
## Call:
## lm(formula = Mwh ~ temp, data = econsumption)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2593 -0.5395 -0.1827  0.4274  2.1972
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 20.19952    0.73040   27.66 8.86e-11 ***
## temp        -0.14516    0.03549   -4.09  0.00218 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9888 on 10 degrees of freedom
## Multiple R-squared:  0.6258, Adjusted R-squared:  0.5884
## F-statistic: 16.73 on 1 and 10 DF,  p-value: 0.00218
```

It seems that there is a negative relationship because a simple linear model is not appropriate. A non-linear model will be necessary for the data.

b) Produce a residual plot. Is the model adequate? Are there any outliers or influential observations?

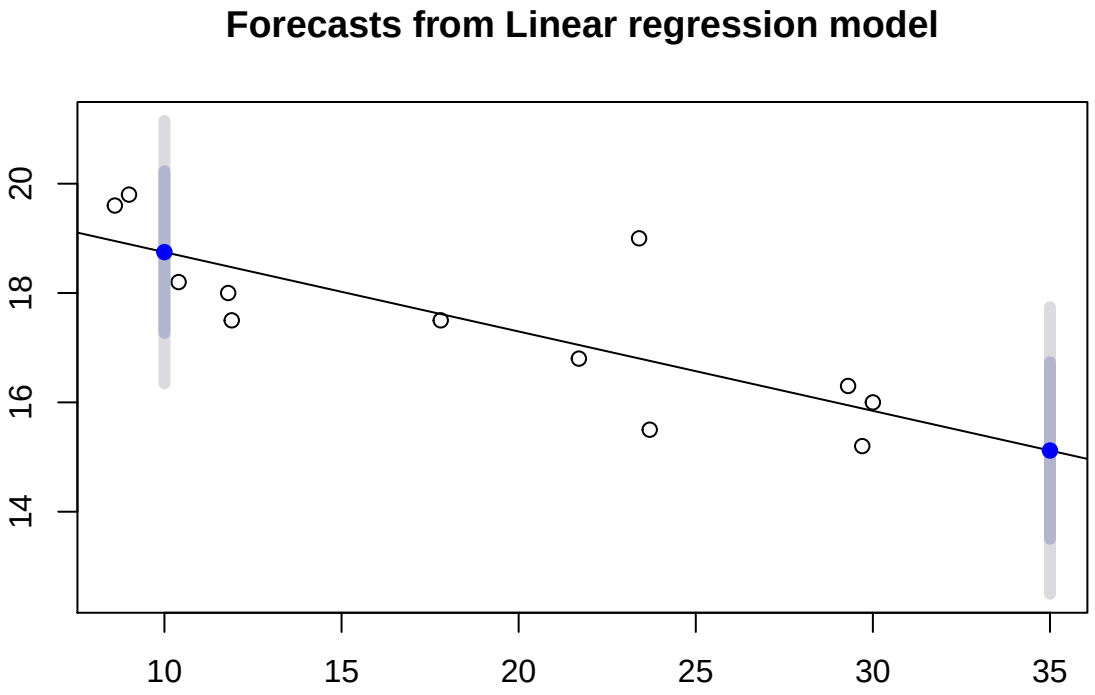
```
plot(residuals(fit)~temp, data=econsumption)
```



The model appears to be adequate, there is no apparent pattern. The only thing that might cause an issues is 1 outlier at temperature 23.4.

c) Use the model to predict the electricity consumption that you would expect for a day with maximum temperature 10 and a day with maximum temperature 35. Do you believe these predictions?

```
fcast <- forecast(fit, newdata=data.frame(temp=c(10,35)))
plot(fcast)
```



Based on the predictive intervals and the nearby data points, the predictions are believable.

d) Give prediction intervals for your forecasts. The following R code will get you started:

```
summary(fit)
```

```
##
## Call:
## lm(formula = Mwh ~ temp, data = econsumption)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2593 -0.5395 -0.1827  0.4274  2.1972
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 20.19952    0.73040   27.66 8.86e-11 ***
## temp        -0.14516    0.03549   -4.09  0.00218 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9888 on 10 degrees of freedom
## Multiple R-squared:  0.6258, Adjusted R-squared:  0.5884
## F-statistic: 16.73 on 1 and 10 DF,  p-value: 0.00218
```

```
forecast(fit, newdata=data.frame(temp=c(60)))
```

```
##   Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## 1       11.49008 9.041979 13.93819 7.514874 15.46529
```