Quadratic terms in regression models

We will look at the nutrition data example from class and look at how height and weight are associated.

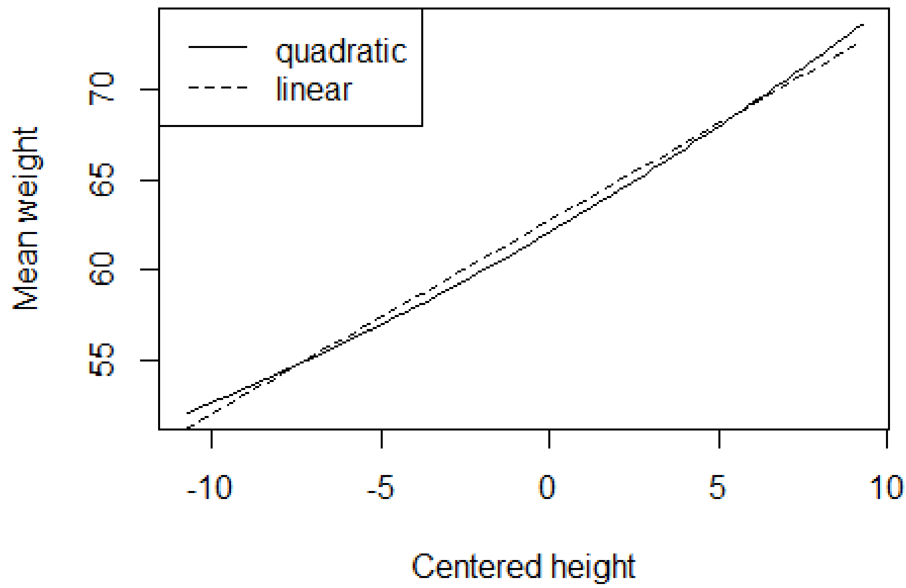
```
nutrition.data <- data.frame(wgt=c(64,71,53,67,55,58,77,57,56,51,76,68),
                             hgt=c(57,59,49,62,51,50,55,48,42,42,61,57),
                             age=c(8,10,6,11,8,7,10,9,10,6,12,9))


nutrition.data$hgt.ctr <- nutrition.data$hgt-mean(nutrition.data$hgt)
# height as just a linear term
mod.hgt1 <- lm(wgt ~ hgt.ctr, data=nutrition.data)
# adding quadratic term
mod.hgt5 <- lm(wgt ~ hgt.ctr + I(hgt.ctr^2),data=nutrition.data)
summary(mod.hgt5)

##
## Call:
## lm(formula = wgt ~ hgt.ctr + I(hgt.ctr^2), data = nutrition.data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.6054 -3.5705 -0.6399  2.0158 12.3692
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  62.07513    2.45272  25.309 1.13e-09 ***
## hgt.ctr       1.10027    0.26385   4.170  0.00241 **
## I(hgt.ctr^2)  0.01581    0.04247   0.372  0.71830
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.723 on 9 degrees of freedom
## Multiple R-squared:  0.6681, Adjusted R-squared:  0.5944
## F-statistic: 9.059 on 2 and 9 DF,  p-value: 0.006989
```

The slope estimate for the centered height term is `1.10027`, so for someone with the average height (i.e., `hgt.ctr==0`), mean weight increases by 1.1 lbs for each 1 inch increase in height. However, because of the presence of the quadratic term, the amount of mean weight increase varies with the subject's height; without the quadratic term, the effect of height on mean weight would be constant, or the same for all heights.

It helps to look at a plot of the predicted values.

```
xvec <-
seq(min(nutrition.data$hgt.ctr),max(nutrition.data$hgt.ctr),length.out=101)
yhat1 <- predict(mod.hgt1,newdata=list(hgt.ctr=xvec))
```

```
yhat2 <- predict(mod.hgt5,newdata=list(hgt.ctr=xvec))
plot(xvec,yhat2,type='l',ylab='Mean weight',xlab='Centered height')
lines(xvec,yhat1,lty=2)
legend(x='topleft',lty=1:2,legend=c('quadratic','linear'))
```



You can see the slight curvature of mean weight as a function of centered height. The dashed line shows the linear model, which agrees very closely with the quadratic fit. This is also consistent with the fact that the quadratic term is not significant: it doesn't really improve the model fit overall.

Think back to basic calculus. The derivative of the function defining mean weight using height and height squared would be

```
weight.prime <- function(x)
  # linear term
  coef(mod.hgt5)[2]+
  # twice the quadratic term times x
  2*coef(mod.hgt5)[3]*x
# look at the slope at different centered heights
# (so 0 is at the mean height)
weight.prime(-2:2)

## [1] 1.037035 1.068654 1.100273 1.131892 1.163511
```

This is the effect of height on mean weight for the quadratic model: it depends on height. Each additional inch of height increases the slope by

```r
2*coef(mod.hgt5)[3]

## I(hgt.ctr^2)
##   0.03161906
```

lbs.

For the linear model, the effect of height on weight is just

```r
coef(mod.hgt1)[2]

## hgt.ctr
## 1.07223
```

for any value of height.