

Nonlinear Models

Explore the use of nonlinear models using some tools in R.

```
require(ISLR)
```

```
## Loading required package: ISLR
```

```
attach(Wage)
```

Polynomials

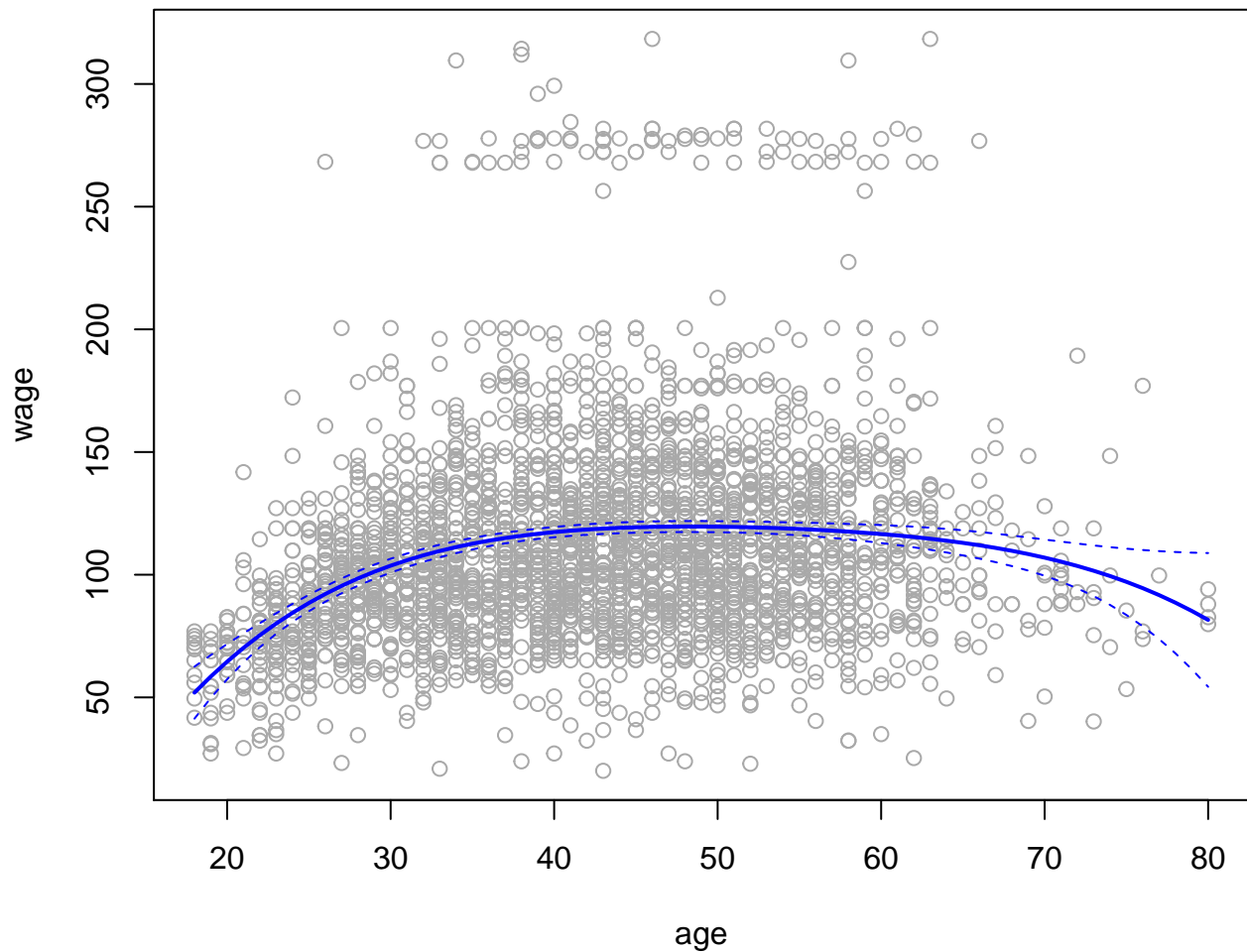
First we will use polynomials and focus on a single predictor age: (They are orthogonal polynomials)

```
fit = lm(wage~poly(age,4), data=Wage)
summary(fit)
```

```
##
## Call:
## lm(formula = wage ~ poly(age, 4), data = Wage)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -98.707 -24.626  -4.993  15.217  203.693
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   111.7036     0.7287  153.283 < 2e-16 ***
## poly(age, 4)1   447.0679     39.9148   11.201 < 2e-16 ***
## poly(age, 4)2  -478.3158     39.9148  -11.983 < 2e-16 ***
## poly(age, 4)3   125.5217     39.9148    3.145  0.00168 **
## poly(age, 4)4  -77.9112     39.9148   -1.952  0.05104 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 39.91 on 2995 degrees of freedom
## Multiple R-squared:  0.08626,    Adjusted R-squared:  0.08504
## F-statistic: 70.69 on 4 and 2995 DF,  p-value: < 2.2e-16
```

Lets make a plot of the fitted function, along with the standard errors of the fit.

```
agelims = range(age)
age.grid = seq(from=agelims[1], to=agelims[2])
preds = predict(fit, newdata=list(age=age.grid),se=TRUE)
se.bands = cbind(preds$fit+2*preds$se, preds$fit-2*preds$se)
plot(age,wage,col='darkgrey')
lines(age.grid, preds$fit, lwd=2,col='blue')
matlines(age.grid,se.bands,col='blue',lty=2)
```



There are other more direct ways of doing this in R. For example,

```
fita = lm(wage~age+I(age^2)+I(age^3)+I(age^4),data=Wage)
summary(fita)
```

```
##
## Call:
## lm(formula = wage ~ age + I(age^2) + I(age^3) + I(age^4), data = Wage)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -98.707 -24.626  -4.993  15.217  203.693
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.842e+02  6.004e+01  -3.067  0.002180 **
## age          2.125e+01  5.887e+00   3.609  0.000312 ***
## I(age^2)     -5.639e-01  2.061e-01  -2.736  0.006261 **
## I(age^3)      6.811e-03  3.066e-03   2.221  0.026398 *
## I(age^4)     -3.204e-05  1.641e-05  -1.952  0.051039 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 39.91 on 2995 degrees of freedom
```

```
## Multiple R-squared:  0.08626,    Adjusted R-squared:  0.08504  
## F-statistic: 70.69 on 4 and 2995 DF,  p-value: < 2.2e-16
```

The coefficients are different to those we got before! However the fits are the same:

```
plot(fitted(fit),fitted(fita))
```

