Lecture 16: Nonlinear and Polynomial Models

ECON 480 - Econometrics - Fall 2018

Ryan Safner

November 26, 2018

Polynomial Functions

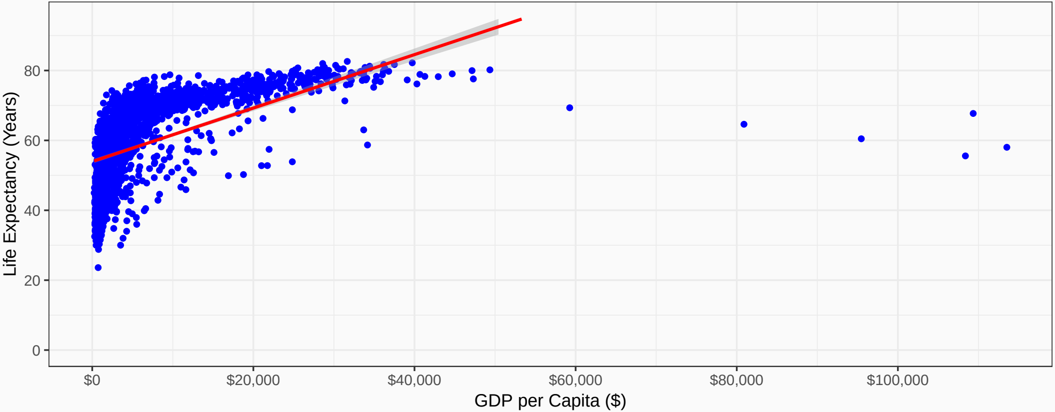
Quadratic Model

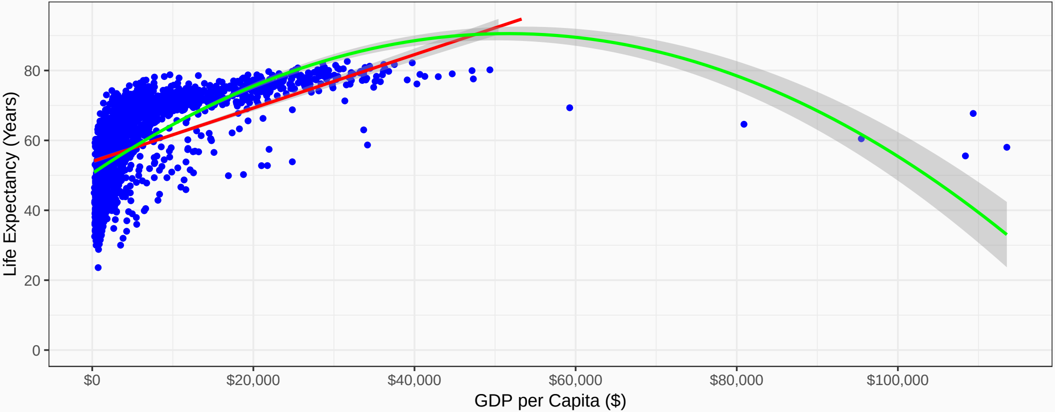
Determining if (Larger) Polynomials are Necessary

- OLS is commonly known as "***linear* regression**" as it fits a *straight line* to data points

- OLS is commonly known as "***linear* regression**" as it fits a *straight line* to data points
- Often, data and relationships between variables may *not* be linear

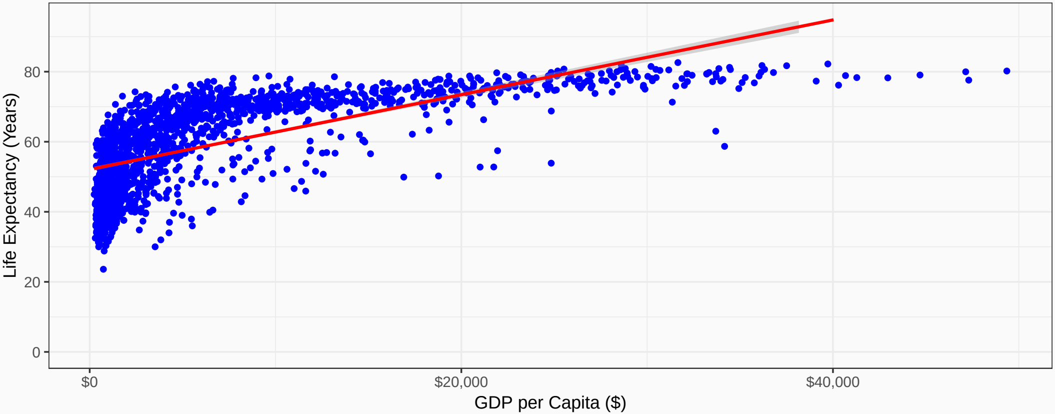
Nonlinearities? Example



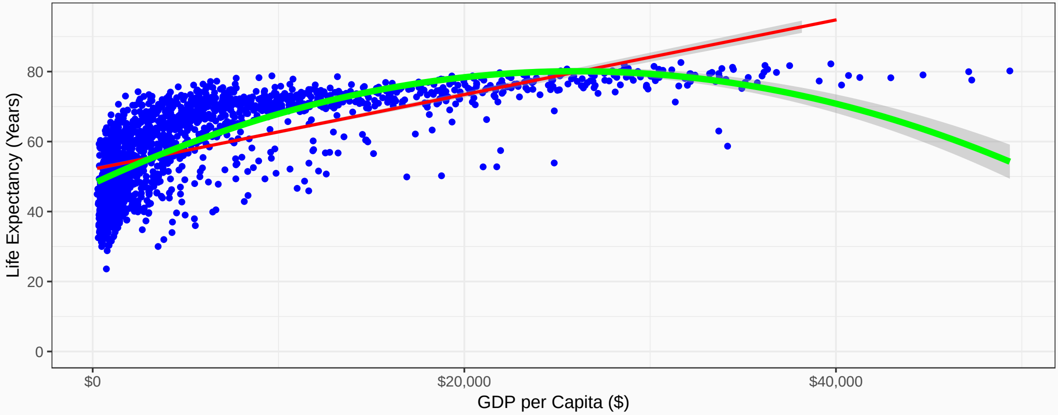


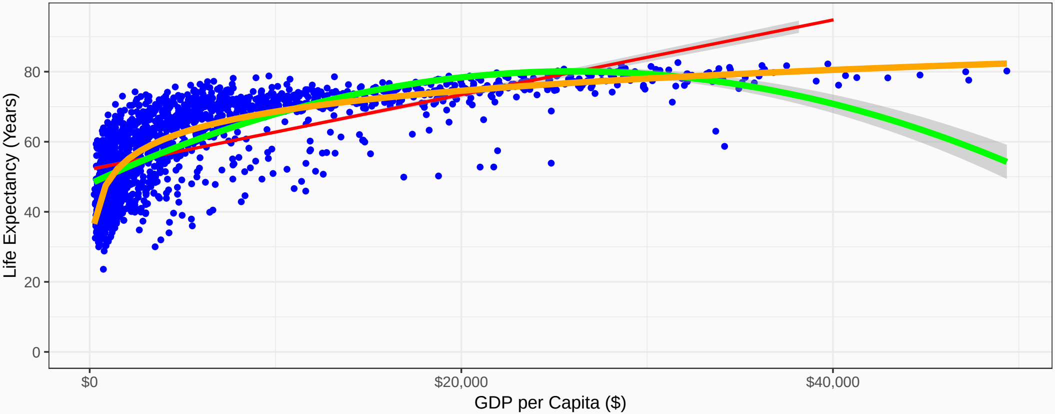
HOOD COLLEGE

Nonlinearities? Example: No Outliers



Nonlinearities? Example: Quadratic Fit




Life Expectancy (Years)
80
60
40
20
0
$0
$20,000
$40,000
GDP per Capita ($)

- Despite being "linear regression", OLS can handle this with an easy fix

- Despite being “linear regression”, OLS can handle this with an easy fix
- OLS requires all *parameters* (i.e. the β's) to be linear, the *regressors* (X's) can be nonlinear:

- Despite being "linear regression", OLS can handle this with an easy fix
- OLS requires all *parameters* (i.e. the β's) to be linear, the *regressors* (X's) can be nonlinear:

$$Y = \beta_0 + \beta_1 X^2 \qquad \checkmark$$

- Despite being "linear regression", OLS can handle this with an easy fix
- OLS requires all *parameters* (i.e. the β's) to be linear, the *regressors* (X's) can be nonlinear:

$$Y = \beta_0 + \beta_1 X^2 \qquad \checkmark$$

$$Y = \beta_0 + \beta_1^2 X \qquad x$$

- Despite being “linear regression”, OLS can handle this with an easy fix
- OLS requires all *parameters* (i.e. the β’s) to be linear, the *regressors* (X’s) can be nonlinear:

$$Y = \beta_0 + \beta_1 X^2 \qquad \checkmark$$

$$Y = \beta_0 + \beta_1^2 X \qquad x$$

$$Y = \beta_0 + \beta_1 \sqrt{X} \qquad \checkmark$$

- Despite being “linear regression”, OLS can handle this with an easy fix
- OLS requires all *parameters* (i.e. the β’s) to be linear, the *regressors* (X’s) can be nonlinear:

$$Y = \beta_0 + \beta_1 X^2 \qquad \checkmark$$

$$Y = \beta_0 + \beta_1^2 X \qquad x$$

$$Y = \beta_0 + \beta_1 \sqrt{X} \qquad \checkmark$$

$$Y = \beta_0 + \sqrt{\beta_1} X \qquad x$$

- Despite being “linear regression”, OLS can handle this with an easy fix
- OLS requires all *parameters* (i.e. the β's) to be linear, the *regressors* (X's) can be nonlinear:

$$Y = \beta_0 + \beta_1 X^2 \qquad \checkmark$$

$$Y = \beta_0 + \beta_1^2 X \qquad x$$

$$Y = \beta_0 + \beta_1 \sqrt{X} \qquad \checkmark$$

$$Y = \beta_0 + \sqrt{\beta_1} X \qquad x$$

$$Y = \beta_0 + \beta_1 X_1 * X_2 \qquad \checkmark$$

HOOD COLLEGE

- Despite being "linear regression", OLS can handle this with an easy fix
- OLS requires all *parameters* (i.e. the β's) to be linear, the *regressors* (X's) can be nonlinear:

$$
\begin{array}{ll}
Y = \beta_0 + \beta_1 X^2 & \checkmark \\
Y = \beta_0 + \beta_1^2 X & x \\
Y = \beta_0 + \beta_1 \sqrt{X} & \checkmark \\
Y = \beta_0 + \sqrt{\beta_1} X & x \\
Y = \beta_0 + \beta_1 X_1 * X_2 & \checkmark \\
Y = \beta_0 + \beta_1 ln(X) & \checkmark
\end{array}
$$

- Despite being "linear regression", OLS can handle this with an easy fix
- OLS requires all *parameters* (i.e. the β's) to be linear, the *regressors* (X's) can be nonlinear:

$$
\begin{array}{ll}
Y = \beta_0 + \beta_1 X^2 & \checkmark \\
Y = \beta_0 + \beta_1^2 X & x \\
Y = \beta_0 + \beta_1 \sqrt{X} & \checkmark \\
Y = \beta_0 + \sqrt{\beta_1} X & x \\
Y = \beta_0 + \beta_1 X_1 * X_2 & \checkmark \\
Y = \beta_0 + \beta_1 ln(X) & \checkmark
\end{array}
$$

- In the end, X will always be just a number, OLS can always estimate parameters for numbers
- *Plotting* the points $(X, \hat{Y})$ can result in a curve for nonlinear X's

- The effect of $X \longrightarrow Y$ may be nonlinear if:

- The effect of $X \longrightarrow Y$ may be nonlinear if:

1. $X \longrightarrow Y$ is different for different levels of X

- The effect of $X \longrightarrow Y$ may be nonlinear if:

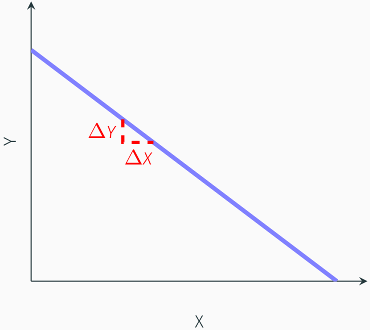
1. $X \longrightarrow Y$ is different for different levels of X
    - e.g. Diminishing returns: $\uparrow X$ increases Y at a decreasing rate

- The effect of $X \longrightarrow Y$ may be nonlinear if:

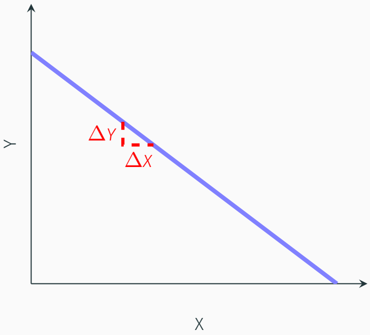
1. $X \longrightarrow Y$ is different for different levels of X
    - e.g. Diminishing returns: $\uparrow X$ increases Y at a decreasing rate
    - e.g. Increasing returns: $\uparrow X$ increases Y at an increasing rate

- The effect of $X \longrightarrow Y$ may be nonlinear if:

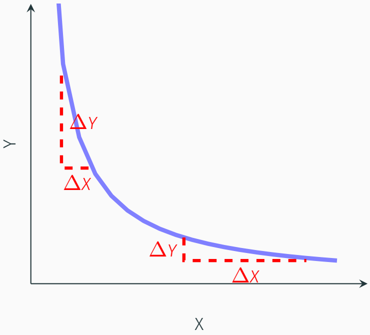
1. $X \longrightarrow Y$ is different for different levels of X
    - e.g. Diminishing returns: $\uparrow X$ increases Y at a decreasing rate
    - e.g. Increasing returns: $\uparrow X$ increases Y at an increasing rate
2. $X \longrightarrow Y$ depends on the value of X_2 - e.g. interaction terms (last lesson)



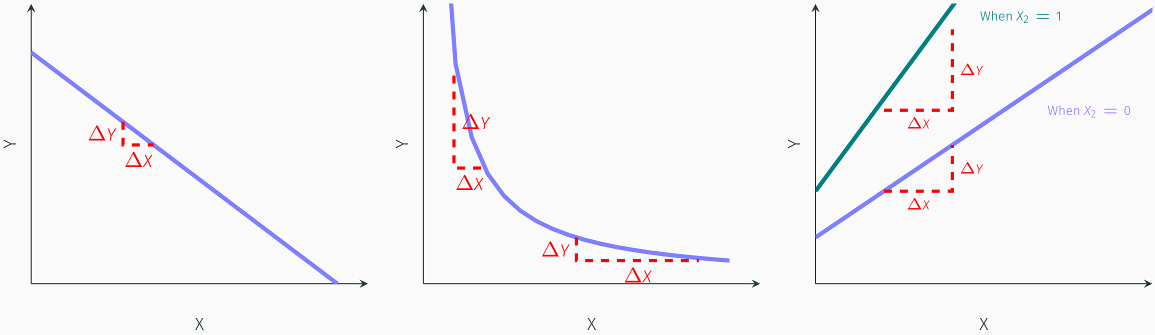
Linear: slope $\beta_1 = \frac{\Delta Y}{\Delta X}$ is constant



Linear: slope $\beta_1 = \frac{\Delta Y}{\Delta X}$ is constant



Quadratic: slope β_1 depends on value of X



Linear: slope $\beta_1 = \frac{\Delta Y}{\Delta X}$ is constant

Quadratic: slope β_1 depends on value of X

Interaction: slope β_1 depends on value of X_2

Polynomial Functions

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{\beta}_2 X_i^2 + \ldots + \hat{\beta}_r X_i^r + \epsilon_i$$

- r is the highest power X_i is raised to (e.g. quadratic $r = 2$, cubic $r = 3$, etc)

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{\beta}_2 X_i^2 + \ldots + \hat{\beta}_r X_i^r + \epsilon_i$$

- r is the highest power X_i is raised to (e.g. quadratic $r = 2$, cubic $r = 3$, etc)
  - The graph of an r^{th}-degree polynomial function has $r - 1$ bends

Polynomial Functions of X

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{\beta}_2 X_i^2 + \ldots + \hat{\beta}_r X_i^r + \epsilon_i$$

- r is the highest power X_i is raised to (e.g. quadratic $r = 2$, cubic $r = 3$, etc)
    - The graph of an r^{th}-degree polynomial function has $r - 1$ bends
- Just another multivariate OLS regression model

Quadratic Model

Quadratic Model

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{\beta}_2 X_i^2 + \epsilon_i$$

- Quadratic model has X and X^2 variables in it (yes, need both!)

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{\beta}_2 X_i^2 + \epsilon_i$$

- Quadratic model has X and X^2 variables in it (yes, need both!)
- How to interpret coefficients?

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{\beta}_2 X_i^2 + \epsilon_i$$

- Quadratic model has X and X^2 variables in it (yes, need both!)
- How to interpret coefficients?
  - β_0 as "intercept" and β_1 as "slope" makes no sense

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{\beta}_2 X_i^2 + \epsilon_i$$

- Quadratic model has X and X^2 variables in it (yes, need both!)
- How to interpret coefficients?
  - β_0 as "intercept" and β_1 as "slope" makes no sense
  - β_1 as effect $X_i \rightarrow Y_i$ holding X_i^2 constant makes no sense

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{\beta}_2 X_i^2 + \epsilon_i$$

- Quadratic model has X and X^2 variables in it (yes, need both!)
- How to interpret coefficients?
    - β_0 as "intercept" and β_1 as "slope" makes no sense
    - β_1 as effect $X_i \rightarrow Y_i$ holding X_i^2 constant makes no sense
    - Note: this is *not* a multicollinearity problem! Correlation only measures *linear* relationships!

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{\beta}_2 X_i^2 + \epsilon_i$$

- Quadratic model has X and X^2 variables in it (yes, need both!)
- How to interpret coefficients?
  - β_0 as "intercept" and β_1 as "slope" makes no sense
  - β_1 as effect $X_i \rightarrow Y_i$ holding X_i^2 constant makes no sense
  - Note: this is *not* a multicollinearity problem! Correlation only measures *linear* relationships!
  - **Calculate marginal effects** by calculating predicted $\hat{Y}_i$ for different X_i

- What is the effect of $\Delta X_i \rightarrow \Delta Y_i$?

- What is the effect of $\Delta X_i \rightarrow \Delta Y_i$?
- Take the **derivative** of Y_i with respect to X_i:

$$\frac{dY_i}{dX_i} = \hat{\beta}_1 + 2\hat{\beta}_2 X_i$$

- What is the effect of $\Delta X_i \rightarrow \Delta Y_i$?
- Take the **derivative** of Y_i with respect to X_i:

$$\frac{dY_i}{dX_i} = \hat{\beta}_1 + 2\hat{\beta}_2 X_i$$

- Marginal effect of a 1 unit change in X_i is a $\hat{\beta}_1 + 2\hat{\beta}_2 X_i$ unit change in Y

Example

$$\widehat{\text{Life Expectancy}}_i = \hat{\beta}_0 + \hat{\beta}_1 GDP_i + \hat{\beta}_2 GDP_i^2$$

- Life Expectancy
- GDP per Capita (GDP for short)

- These coefficients will be very large, let's first transform `gdpPercap` into $1,000s, call it `gdp.t` [1]

[1] Note I am using `dplyr` and `%>%` here for efficiency, I loaded them before without showing it

- These coefficients will be very large, let's first transform `gdpPercap` into $1,000s, call it `gdp.t` [1]

```
gapminder <- gapminder %>%
  mutate(gdp.t=gdpPercap/1000)
```

[1] Note I am using `dplyr` and `%>%` here for efficiency, I loaded them before without showing it

- These coefficients will be very large, let's first transform `gdpPercap` into $1,000s, call it `gdp.t` [1]

```
gapminder <- gapminder %>%
  mutate(gdp.t=gdpPercap/1000)
```

- Let's also make the quadratic term by squaring `gdp.t` and calling it `gdp.sq`

[1] Note I am using `dplyr` and `%>%` here for efficiency, I loaded them before without showing it

- These coefficients will be very large, let's first transform `gdpPercap` into $1,000s, call it `gdp.t` [1]

```
gapminder <- gapminder %>%
  mutate(gdp.t=gdpPercap/1000)
```

- Let's also make the quadratic term by squaring `gdp.t` and calling it `gdp.sq`

```
gapminder <- gapminder %>%
  mutate(gdp.sq=gdp.t^2)
```

[1] Note I am using `dplyr` and `%>%` here for efficiency, I loaded them before without showing it

Quadratic Model: Example Regression in R

- Can "manually" run regression with `gdp.t` and squared term `gdp.sq`

- Can "manually" run regression with `gdp.t` and squared term `gdp.sq`

```
reg1<-lm(lifeExp~gdp.t+gdp.sq, data=gapminder)
summary(reg1)
```

```
##
## Call:
## lm(formula = lifeExp ~ gdp.t + gdp.sq, data = gapminder)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -28.0600  -6.4253   0.2611   7.0889  27.1752
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) 50.5240058  0.2978135  169.65   <2e-16 ***
## gdp.t        1.5509911  0.0373735   41.50   <2e-16 ***
## gdp.sq      -0.0150193  0.0005794  -25.92   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Or can use one of R's shortcuts for polynomial models:

Quadratic Model: Example Regression in R II

- Or can use one of R's shortcuts for polynomial models:
- first term is normal, `x`, second term use `I(x^2)` to add squared term

- Or can use one of R's shortcuts for polynomial models:
- first term is normal, x, second term use `I(x^2)` to add squared term

```
reg1<-lm(lifeExp~gdp.t+I(gdp.t^2), data=gapminder)
summary(reg1)
```

```
##
## Call:
## lm(formula = lifeExp ~ gdp.t + I(gdp.t^2), data = gapminder)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -28.0600  -6.4253   0.2611   7.0889  27.1752
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 50.5240058  0.2978135  169.65   <2e-16 ***
## gdp.t        1.5509911  0.0373735   41.50   <2e-16 ***
## I(gdp.t^2)  -0.0150193  0.0005794  -25.92   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.885 on 1701 degrees of freedom
## Multiple R-squared:  0.5274, Adjusted R-squared:  0.5268
## F-statistic: 949.1 on 2 and 1701 DF,  p-value: < 2.2e-16
```

Quadratic Model: Example Regression in R III

- Another shortcut is just to use the `poly()` command

- Another shortcut is just to use the `poly()` command
- Instead of any `x` variables, just add `poly(x,2)` where 2 is the highest power desired[2]

[2] R gives different coefficient estimates for this version unless you add `, raw=TRUE` inside the `poly()` function, to ensure the polynomials are not computed orthogonally.

Quadratic Model: Example Regression in R III

- Another shortcut is just to use the `poly()` command
- Instead of any `x` variables, just add `poly(x,2)` where 2 is the highest power desired[2]

```
reg1<-lm(lifeExp~poly(gdp.t,2, raw=TRUE), data=gapminder)
reg1
```

```
##
## Call:
## lm(formula = lifeExp ~ poly(gdp.t, 2, raw = TRUE), data = gapminder)
##
## Coefficients:
##                 (Intercept)  poly(gdp.t, 2, raw = TRUE)1
##                    50.52401                      1.55099
## poly(gdp.t, 2, raw = TRUE)2
##                    -0.01502
```

[2] R gives different coefficient estimates for this version unless you add `, raw=TRUE` inside the `poly()` function, to ensure the polynomials are not computed orthogonally.

Quadratic Model: Marginal Effects

$$\widehat{\text{Life Expectancy}}_i = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.02}\, GDP^2$$

- Marginal effect of GDP on Life Expectancy:

$$\widehat{\text{Life Expectancy}}_i = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.02}\, GDP^2$$

- Marginal effect of GDP on Life Expectancy:

$$\frac{d\,Y}{d\,X} = \hat{\beta}_1 + 2\hat{\beta}_2 X_i$$

Quadratic Model: Marginal Effects

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.02}\, GDP^2$$

- Marginal effect of GDP on Life Expectancy:

$$\frac{d\,Y}{d\,X} = \hat{\beta}_1 + 2\hat{\beta}_2 X_i$$

$$\frac{d\,\text{Life Expectancy}}{d\,\text{GDP}} = 1.55 + 2(-0.02)\text{GDP}$$

Quadratic Model: Marginal Effects

$$\widehat{\text{Life Expectancy}}_i = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.02}\, GDP^2$$

- Marginal effect of GDP on Life Expectancy:

$$\begin{aligned}
\frac{d\,Y}{d\,X} &= \hat{\beta}_1 + 2\hat{\beta}_2 X_i \\
\frac{d\,\text{Life Expectancy}}{d\,\text{GDP}} &= 1.55 + 2(-0.02)\text{GDP} \\
&= 1.55 - 0.04\text{GDP}
\end{aligned}$$

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\,GDP - \underset{(0.00)}{0.02}\,GDP^2$$

- Marginal effect of GDP on Life Expectancy:

$$\begin{aligned} \frac{dY}{dX} &= \hat{\beta}_1 + 2\hat{\beta}_2 X_i \\ \frac{d\,\text{Life Expectancy}}{d\,\text{GDP}} &= 1.55 + 2(-0.02)\text{GDP} \\ &= 1.55 - 0.04\text{GDP} \end{aligned}$$

- Positive, with diminishing returns
- Effect on Life Expectancy of increasing GDP depends on initial value of GDP!

Quadratic Model: Marginal Effects II

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\,GDP - \underset{(0.00)}{0.02}\,GDP^2$$

- Marginal effect of GDP if GDP= \$5 (thousand):

Quadratic Model: Marginal Effects II

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.02}\, GDP^2$$

- Marginal effect of GDP if GDP$=$ \$5 (thousand):

$$\frac{d\,\text{Life Expectancy}}{d\,\text{GDP}} = 1.55 - 0.04\text{GDP}$$

Quadratic Model: Marginal Effects II

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.02}\, GDP^2$$

- Marginal effect of GDP if GDP$=$ $5 (thousand):

$$\begin{aligned}\frac{d\,\text{Life Expectancy}}{d\,\text{GDP}} &= 1.55 - 0.04\text{GDP} \\ &= 1.55 - 0.04(5)\end{aligned}$$

Quadratic Model: Marginal Effects II

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.02}\, GDP^2$$

- Marginal effect of GDP if GDP$=$ \$5 (thousand):

$$\begin{aligned} \frac{d \text{ Life Expectancy}}{d \text{ GDP}} &= 1.55 - 0.04\text{GDP} \\ &= 1.55 - 0.04(5) \\ &= 1.55 - 0.20 \end{aligned}$$

Quadratic Model: Marginal Effects II

$$\widehat{\text{Life Expectancy}}_i = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.02}\, GDP^2$$

- Marginal effect of GDP if GDP$=$ \$5 (thousand):

$$\begin{aligned}\frac{d\,\text{Life Expectancy}}{d\,\text{GDP}} &= 1.55 - 0.04\text{GDP}\\ &= 1.55 - 0.04(5)\\ &= 1.55 - 0.20\\ &= 1.35\end{aligned}$$

Quadratic Model: Marginal Effects II

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.02}\, GDP^2$$

- Marginal effect of GDP if GDP$=$ \$5 (thousand):

$$\begin{aligned}\frac{d \text{ Life Expectancy}}{d \text{ GDP}} &= 1.55 - 0.04\text{GDP} \\ &= 1.55 - 0.04(5) \\ &= 1.55 - 0.20 \\ &= 1.35\end{aligned}$$

- i.e. for every addition \$1 (thousand) in GDP per capita, average life expectancy increases by 1.35 years

Quadratic Model: Marginal Effects III

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.02}\, GDP^2$$

- Marginal effect of GDP if GDP $= \$25$ (thousand):

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.02}\, GDP^2$$

- Marginal effect of GDP if GDP$= \$25$ (thousand):

$$\frac{d\,\text{Life Expectancy}}{d\,\text{GDP}} = 1.55 - 0.04\text{GDP}$$

Quadratic Model: Marginal Effects III

$$\widehat{\text{Life Expectancy}}_i = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.02}\, GDP^2$$

- Marginal effect of GDP if GDP$= \$25$ (thousand):

$$\begin{aligned}\frac{d \text{ Life Expectancy}}{d \text{ GDP}} &= 1.55 - 0.04\text{GDP} \\ &= 1.55 - 0.04(25)\end{aligned}$$

Quadratic Model: Marginal Effects III

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.02}\, GDP^2$$

- Marginal effect of GDP if GDP$=\$25$ (thousand):

$$\begin{aligned}\frac{d \text{ Life Expectancy}}{d \text{ GDP}} &= 1.55 - 0.04\text{GDP}\\ &= 1.55 - 0.04(25)\\ &= 1.55 - 1.00\end{aligned}$$

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.02}\, GDP^2$$

- Marginal effect of GDP if GDP $= \$25$ (thousand):

$$\begin{aligned}
\frac{d\,\text{Life Expectancy}}{d\,\text{GDP}} &= 1.55 - 0.04\text{GDP} \\
&= 1.55 - 0.04(25) \\
&= 1.55 - 1.00 \\
&= 0.55
\end{aligned}$$

Quadratic Model: Marginal Effects III

$$\widehat{\text{Life Expectancy}}_i = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.02}\, GDP^2$$

- Marginal effect of GDP if GDP$=$ \$25 (thousand):

$$\begin{aligned}\frac{d\,\text{Life Expectancy}}{d\,\text{GDP}} &= 1.55 - 0.04\text{GDP} \\ &= 1.55 - 0.04(25) \\ &= 1.55 - 1.00 \\ &= 0.55\end{aligned}$$

- i.e. for every addition \$1 (thousand) in GDP per capita, average life expectancy increases by 0.55 years

Quadratic Model: Marginal Effects IV

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.02}\, GDP^2$$

- Marginal effect of GDP if GDP= \$50 (thousand):

Quadratic Model: Marginal Effects IV

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.02}\, GDP^2$$

- Marginal effect of GDP if GDP$=\$50$ (thousand):

$$\frac{d\,\text{Life Expectancy}}{d\,\text{GDP}} = 1.55 - 0.04\text{GDP}$$

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.02}\, GDP^2$$

- Marginal effect of GDP if GDP$=$ \$50 (thousand):

$$\begin{aligned}\frac{d\,\text{Life Expectancy}}{d\,\text{GDP}} &= 1.55 - 0.04\text{GDP} \\ &= 1.55 - 0.04(50)\end{aligned}$$

Quadratic Model: Marginal Effects IV

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.02}\, GDP^2$$

- Marginal effect of GDP if GDP$=$ $50 (thousand):

$$\begin{aligned}
\frac{d\,\text{Life Expectancy}}{d\,\text{GDP}} &= 1.55 - 0.04\text{GDP} \\
&= 1.55 - 0.04(50) \\
&= 1.55 - 2
\end{aligned}$$

HOOD COLLEGE

Quadratic Model: Marginal Effects IV

$$\underset{(0.30)}{\widehat{\text{Life Expectancy}}_i = 50.52} + \underset{(0.04)}{1.55\, GDP} - \underset{(0.00)}{0.02\, GDP^2}$$

- Marginal effect of GDP if GDP $= \$50$ (thousand):

$$\begin{aligned}
\frac{d \text{ Life Expectancy}}{d \text{ GDP}} &= 1.55 - 0.04\text{GDP} \\
&= 1.55 - 0.04(50) \\
&= 1.55 - 2 \\
&= -0.45
\end{aligned}$$

Quadratic Model: Marginal Effects IV

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\,GDP - \underset{(0.00)}{0.02}\,GDP^2$$

- Marginal effect of GDP if GDP$=$ \$50 (thousand):

$$\begin{aligned}\frac{d\,\text{Life Expectancy}}{d\,\text{GDP}} &= 1.55 - 0.04\text{GDP}\\ &= 1.55 - 0.04(50)\\ &= 1.55 - 2\\ &= -0.45\end{aligned}$$

- i.e. for every addition \$1 (thousand) in GDP per capita, average life expectancy *decreases* by 0.45 years

$$\widehat{\text{Life Expectancy}}_i = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.02}\, GDP^2$$

Initial GDP	Marginal Effect of +$1,000 GDP

$$\widehat{\text{Life Expectancy}}_i = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.02}\, GDP^2$$

Initial GDP	Marginal Effect of +$1,000 GDP
$5,000	1.35 years

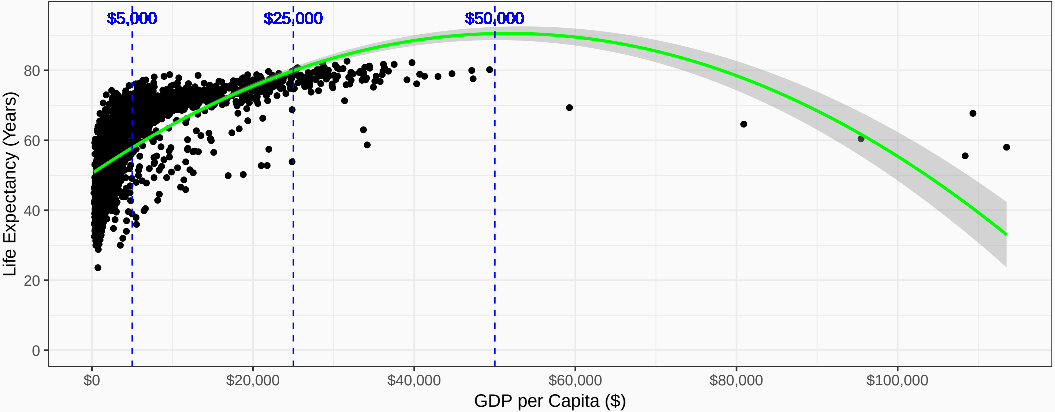
$$\widehat{\text{Life Expectancy}}_i = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.02}\, GDP^2$$

Initial GDP	Marginal Effect of +$1,000 GDP
$5,000	1.35 years
$25,000	0.55 years

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.02}\, GDP^2$$

Initial GDP	Marginal Effect of +$1,000 GDP
$5,000	1.35 years
$25,000	0.55 years
$50,000	−0.45 years

Quadratic Model: Example



- For a quadratic model, we can also find the predicted **maximum** or **minimum** of $\hat{Y}_i$

Quadratic Model: Maxima and Minima

- For a quadratic model, we can also find the predicted **maximum** or **minimum** of $\hat{Y}_i$
- By calculus, a minimum or maximum occurs where:

Quadratic Model: Maxima and Minima

- For a quadratic model, we can also find the predicted **maximum** or **minimum** of $\hat{Y}_i$
- By calculus, a minimum or maximum occurs where:

$$\frac{dY}{dX} = 0$$

Quadratic Model: Maxima and Minima

- For a quadratic model, we can also find the predicted **maximum** or **minimum** of $\hat{Y}_i$
- By calculus, a minimum or maximum occurs where:

$$\frac{dY}{dX} = 0$$
$$\beta_1 + 2\beta_2 X = 0$$

- For a quadratic model, we can also find the predicted **maximum** or **minimum** of $\hat{Y}_i$
- By calculus, a minimum or maximum occurs where:

$$\begin{aligned}\frac{dY}{dX} &= 0\\ \beta_1 + 2\beta_2 X &= 0\\ 2\beta_2 X &= -\beta_1\end{aligned}$$

Quadratic Model: Maxima and Minima

- For a quadratic model, we can also find the predicted **maximum** or **minimum** of $\hat{Y}_i$
- By calculus, a minimum or maximum occurs where:

$$
\begin{aligned}
\frac{dY}{dX} &= 0 \\
\beta_1 + 2\beta_2 X &= 0 \\
2\beta_2 X &= -\beta_1 \\
X^* &= -\frac{1}{2}\frac{\beta_1}{\beta_2}
\end{aligned}
$$

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\,GDP - \underset{(0.00)}{0.015}GDP^2$$

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\,GDP - \underset{(0.00)}{0.015}\,GDP^2$$

$$X^* = -\frac{1}{2}\frac{\beta_1}{\beta_2}$$

HOOD COLLEGE

$$\widehat{\text{Life Expectancy}}_i = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.015} GDP^2$$

$$X^* = -\frac{1}{2}\frac{\beta_1}{\beta_2}$$

$$GDP^* = -\frac{1}{2}\frac{(1.55)}{(-0.015)}$$

Quadratic Model: Maxima and Minima Example

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\,GDP - \underset{(0.00)}{0.015}GDP^2$$

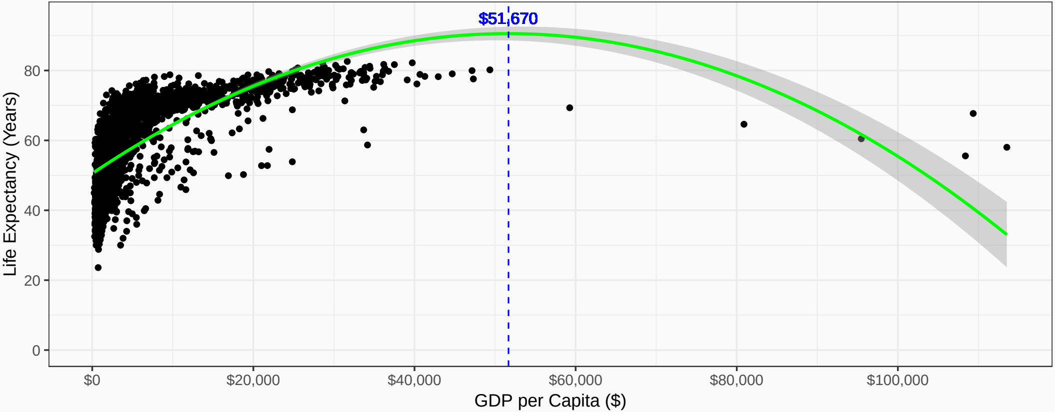
$$
\begin{aligned}
X^* &= -\frac{1}{2}\frac{\beta_1}{\beta_2} \\
GDP^* &= -\frac{1}{2}\frac{(1.55)}{(-0.015)} \\
X^* &= -\frac{1}{2}(-103.333)
\end{aligned}
$$

Quadratic Model: Maxima and Minima Example

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.00)}{0.015} GDP^2$$

$$
\begin{aligned}
X^* &= -\frac{1}{2}\frac{\beta_1}{\beta_2} \\
GDP^* &= -\frac{1}{2}\frac{(1.55)}{(-0.015)} \\
X^* &= -\frac{1}{2}(-103.333) \\
X^* &= 51.67
\end{aligned}
$$

Quadratic Model: Maxima and Minima Example II



Determining if (Larger) Polynomials are Necessary

Quadratic Model: Necessary?

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.0005)}{0.015}\, GDP^2$$

- Do we *need* a quadratic model?

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.0005)}{0.015}\, GDP^2$$

- Do we *need* a quadratic model?
- We can determine if $\hat{\beta}_2$ is statistically significant:

Quadratic Model: Necessary?

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.0005)}{0.015}\, GDP^2$$

- Do we *need* a quadratic model?
- We can determine if $\hat{\beta}_2$ is statistically significant:
  - H_0: $\beta_2 = 0$

$$\widehat{\text{Life Expectancy}}_i = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.0005)}{0.015}\, GDP^2$$

- Do we *need* a quadratic model?
- We can determine if $\hat{\beta}_2$ is statistically significant:
    - H_0: $\beta_2 = 0$
    - H_1: $\beta_2 \neq 0$

Quadratic Model: Necessary?

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.0005)}{0.015}\, GDP^2$$

- Do we *need* a quadratic model?
- We can determine if $\hat{\beta}_2$ is statistically significant:
  - H_0: $\beta_2 = 0$
  - H_1: $\beta_2 \neq 0$

$$t = \frac{\hat{\beta}_2 - 0}{SE(\hat{\beta}_2)} = \frac{0.015}{0.0005} = 30$$

Quadratic Model: Necessary?

$$\widehat{\text{Life Expectancy}_i} = \underset{(0.30)}{50.52} + \underset{(0.04)}{1.55}\, GDP - \underset{(0.0005)}{0.015}\, GDP^2$$

- Do we *need* a quadratic model?
- We can determine if $\hat{\beta}_2$ is statistically significant:
  - H_0: $\beta_2 = 0$
  - H_1: $\beta_2 \neq 0$

$$t = \frac{\hat{\beta}_2 - 0}{SE(\hat{\beta}_2)} = \frac{0.015}{0.0005} = 30$$

- Statistically significant $\implies$ we should keep the quadratic model
  - If we only ran a linear model, it would be biased!

Example

- Should we keep going up in polynomials?

Example

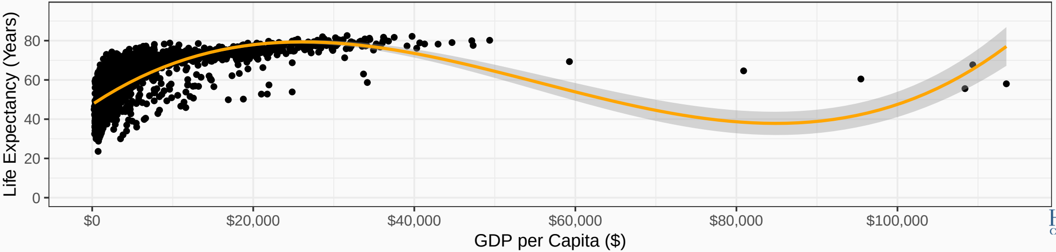
- Should we keep going up in polynomials?

$$\widehat{\text{Life Expectancy}}_i = \hat{\beta}_0 + \hat{\beta}_1 GDP_i + \hat{\beta}_2 GDP_i^2 + \hat{\beta}_3 GDP_i^3$$

Example

- Should we keep going up in polynomials?

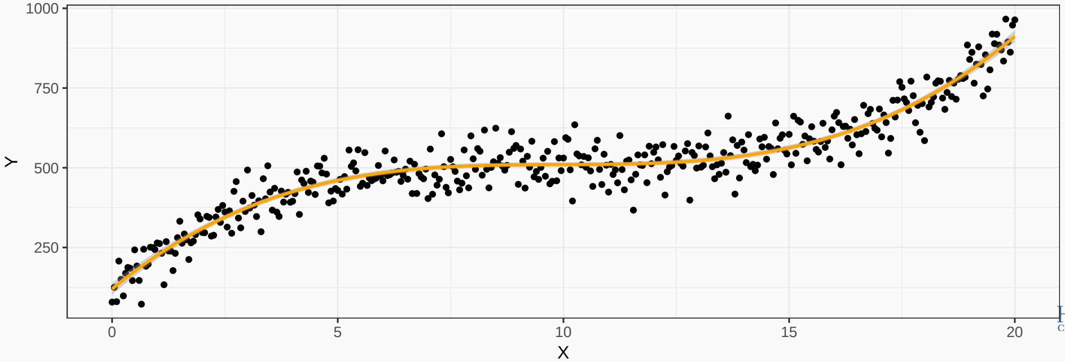
$$\widehat{\text{Life Expectancy}}_i = \hat{\beta}_0 + \hat{\beta}_1 GDP_i + \hat{\beta}_2 GDP_i^2 + \hat{\beta}_3 GDP_i^3$$



- In general, should have a compelling theoretical reason why data or relationships should “change direction” multiple times

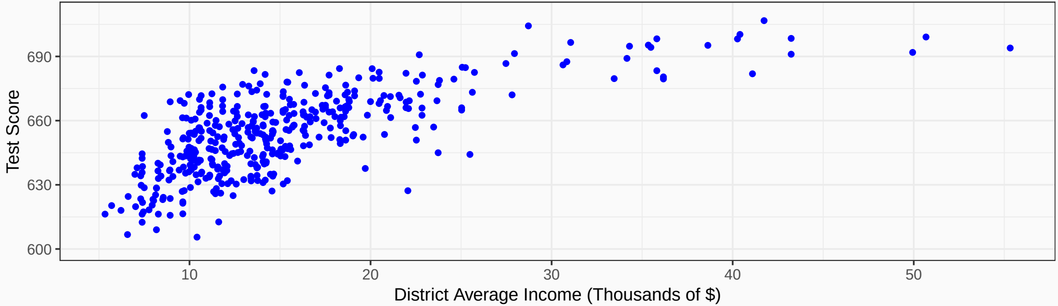
Higher-Order Polynomials

- In general, should have a compelling theoretical reason why data or relationships should "change direction" multiple times
- Or clear data patterns that have multiple "bends"



Example

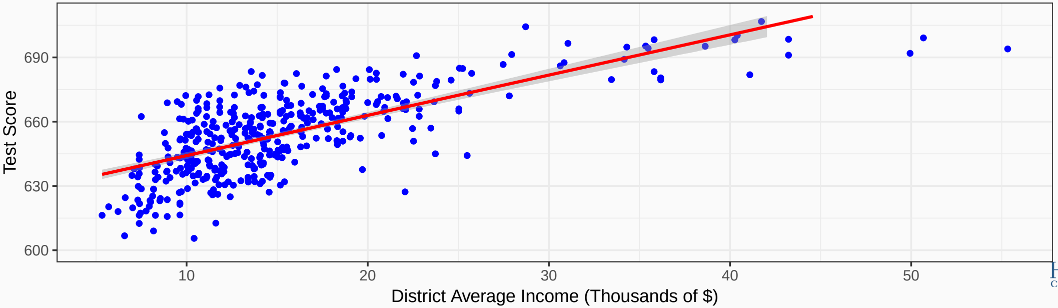
Test Scores: does school district's average income matter?



Example

Test Scores: does school district's average income matter?

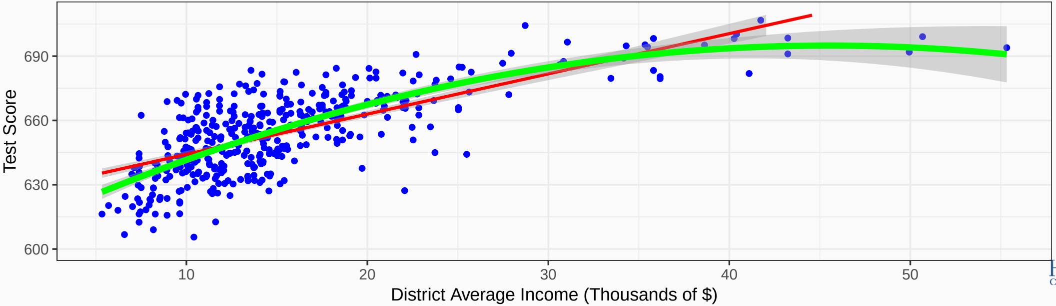
$$TestScore_i = \hat{\beta}_0 + \hat{\beta}_1 Income_i$$



Example

Test Scores: does school district's average income matter?

$$\widehat{TestScore_i} = \hat{\beta}_0 + \hat{\beta}_1 Income_i + \hat{\beta}_2 Income_i^2$$



A Second Example IV

- Let's manually generate a squared term, `avgincsq`

- Let's manually generate a squared term, `avgincsq`

```
CASchool <- CASchool %>%
  mutate(avgincsq=avginc^2)
```

A Second Example V

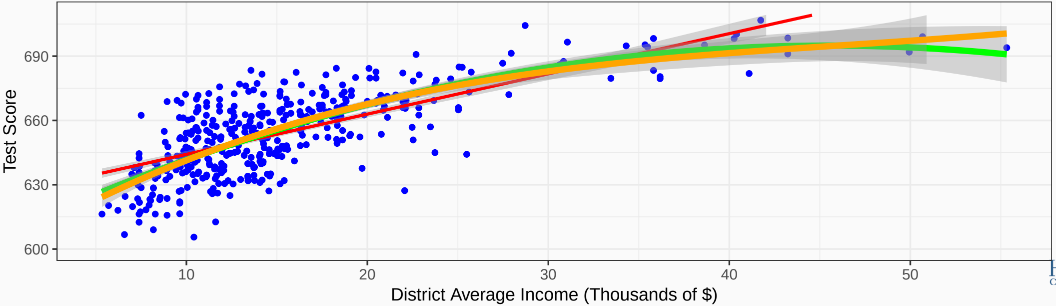
```r
regsc<-lm(testscr~avginc+avgincsq, data=CASchool)
summary(regsc)
```

```
##
## Call:
## lm(formula = testscr ~ avginc + avgincsq, data = CASchool)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -44.416  -9.048   0.440   8.348  31.639
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 607.30174    3.04622 199.362  < 2e-16 ***
## avginc        3.85100    0.30426  12.657  < 2e-16 ***
## avgincsq     -0.04231    0.00626  -6.758 4.71e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

$$\widehat{\text{Test Score}} = \underset{(3.05)}{607.30} + \underset{(0.30)}{3.85}\, Income - \underset{(0.01)}{0.04}\, Income^2$$

Example
Test Scores: does school district's average income matter?

$$TestScore_i = \hat{\beta}_0 + \hat{\beta}_1 Income_i + \hat{\beta}_2 Income_i^2 + \hat{\beta}_3 Income_i^3$$



- Let's manually generate a cubic term, `avginc3`:

- Let's manually generate a cubic term, `avginc3`:

```
CASchool <- CASchool %>%
  mutate(avginc3=avginc^3)
```

A Second Example VIII

```r
regcu<-lm(testscr~avginc+avgincsq+avginc3, data=CASchool)
summary(regcu)
```

```
##
## Call:
## lm(formula = testscr ~ avginc + avgincsq + avginc3, data = CASchool)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -44.28  -9.21   0.20   8.32  31.16
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  6.001e+02  5.830e+00 102.937  < 2e-16 ***
## avginc       5.019e+00  8.595e-01   5.839 1.06e-08 ***
## avgincsq    -9.581e-02  3.736e-02  -2.564   0.0107 *
## avginc3      6.855e-04  4.720e-04   1.452   0.1471
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

1. Are there good theoretical reasons for relationships changing (e.g. increasing/decreasing returns)?

1. Are there good theoretical reasons for relationships changing (e.g. increasing/decreasing returns)?
2. Plot your data: does a straight line fit well enough?

1. Are there good theoretical reasons for relationships changing (e.g. increasing/decreasing returns)?
2. Plot your data: does a straight line fit well enough?
3. Specify a polynomial function of a higher power (start with 2) and estimate OLS regression

Strategy for Polynomial Model Specification

1. Are there good theoretical reasons for relationships changing (e.g. increasing/decreasing returns)?
2. Plot your data: does a straight line fit well enough?
3. Specify a polynomial function of a higher power (start with 2) and estimate OLS regression
4. Use t-test to determine if higher-power term is significant

Strategy for Polynomial Model Specification

1. Are there good theoretical reasons for relationships changing (e.g. increasing/decreasing returns)?
2. Plot your data: does a straight line fit well enough?
3. Specify a polynomial function of a higher power (start with 2) and estimate OLS regression
4. Use t-test to determine if higher-power term is significant
5. Interpret effect of change in X on Y

Strategy for Polynomial Model Specification

1. Are there good theoretical reasons for relationships changing (e.g. increasing/decreasing returns)?
2. Plot your data: does a straight line fit well enough?
3. Specify a polynomial function of a higher power (start with 2) and estimate OLS regression
4. Use t-test to determine if higher-power term is significant
5. Interpret effect of change in X on Y
6. Repeat steps 3-5 as necessary