NON-LINEAR MODELS

Outline

- Polynomial Regression
- Step functions
- Splines
- Generalized Additive Models

Non-Linear Models

- Real relationships are seldom linear
- However, in the absence of strong theory for a specific functional relationship, relationships are assumed to be linear.
- Linear models are parsimonious and linear relationships are easy to interpret.
- When the linearity assumption is not good enough, we can turn to non-linear models such as
  - *polynomials*
  - *splines*
  - *generalized additive models*

  … which offer a lot of flexibility without losing the ease and interpretability of linear models

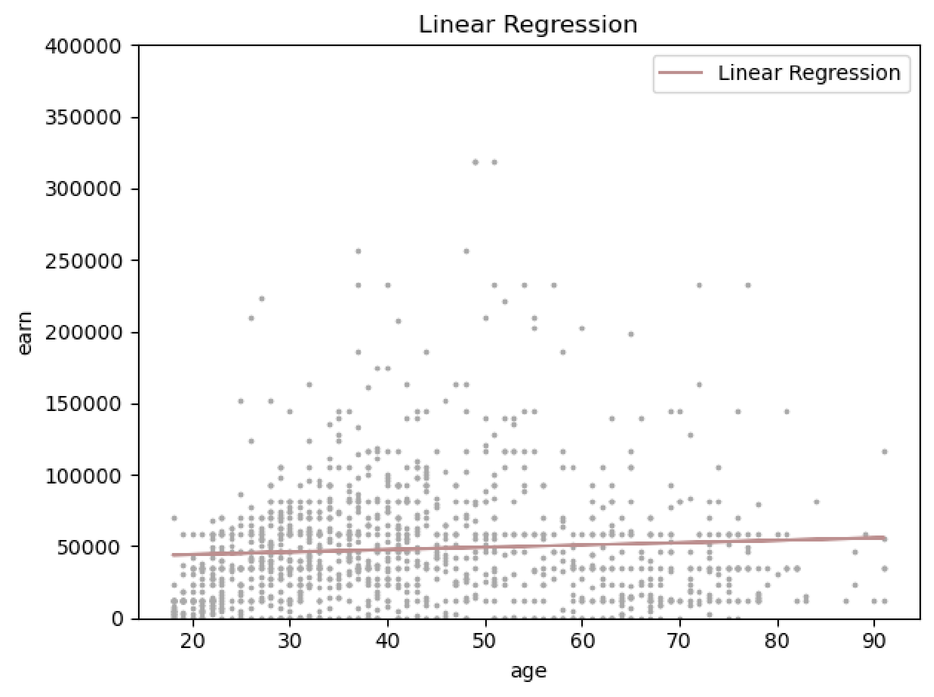
Polynomial Regression

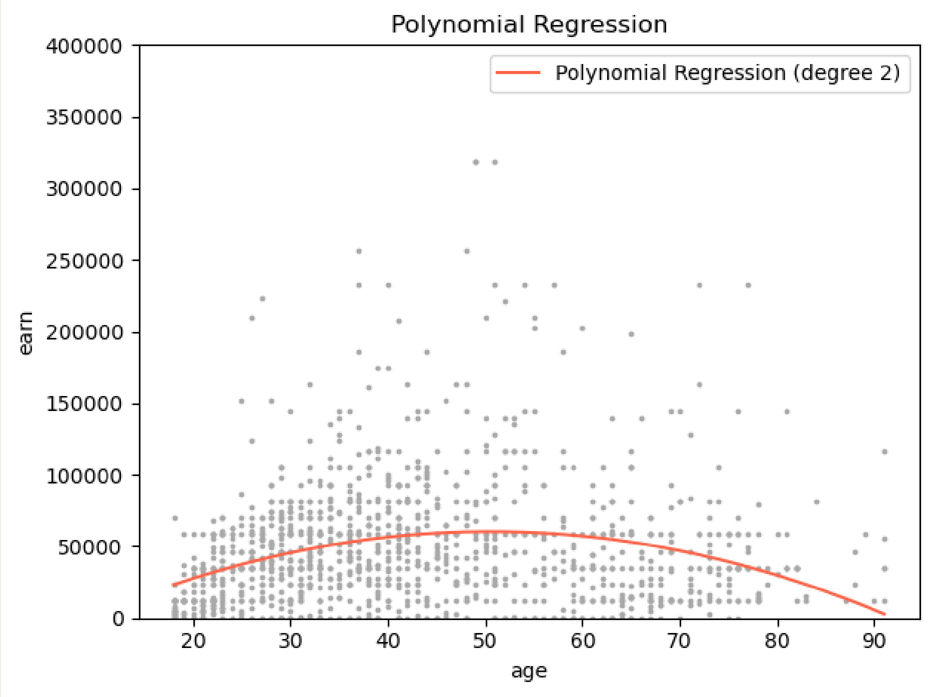
- Generate new features consisting of polynomial terms of the original feature set
- Degree of polynomial refers to the highest power
- This is still a weighted linear combination of features, so still a linear model and can uses OLS estimation procedures.

Linear Regression: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$

Polynomial Regression: $y = \beta_0 + \beta_{11} x_1 + \beta_{12} x_1^2 + \beta_{13} x_1^3 + \beta_{21} x_2 + \beta_{22} x_2^2 + \beta_{23} x_2^3 + \varepsilon$

Polynomial Regression





Polynomial Regression

- Coefficients of polynomial regression may indicate the nature of the relationship
- However, with polynomial regression, one is usually more interested in predictions
- Appropriate degree to use can be determined by examining errors of competing polynomial regressions. For models with equivalent errors, simpler lower degree models are preferred.
- Polynomial terms must be used cautiously as high degree polynomials may overfit the train data.

- Polynomial terms can also be included within logistic regression. The process is similar to that for linear regression.

Polynomial Regression

- Can be applied to more that one numeric predictor variable.
- When conducting polynomial regression with a variable of degree d, all lower degree polynomials must also be included.

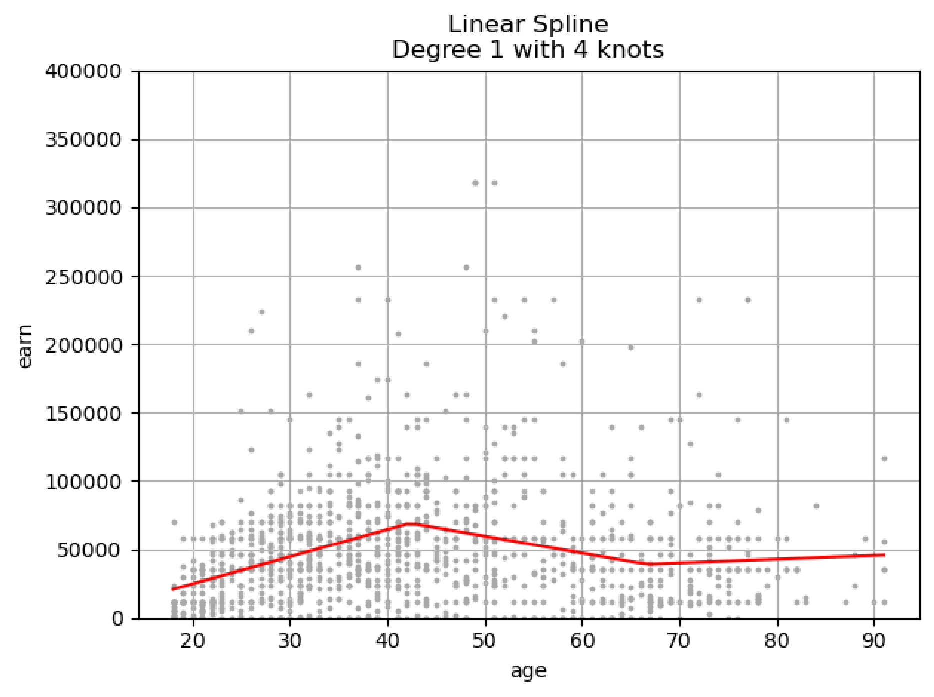
Step Functions

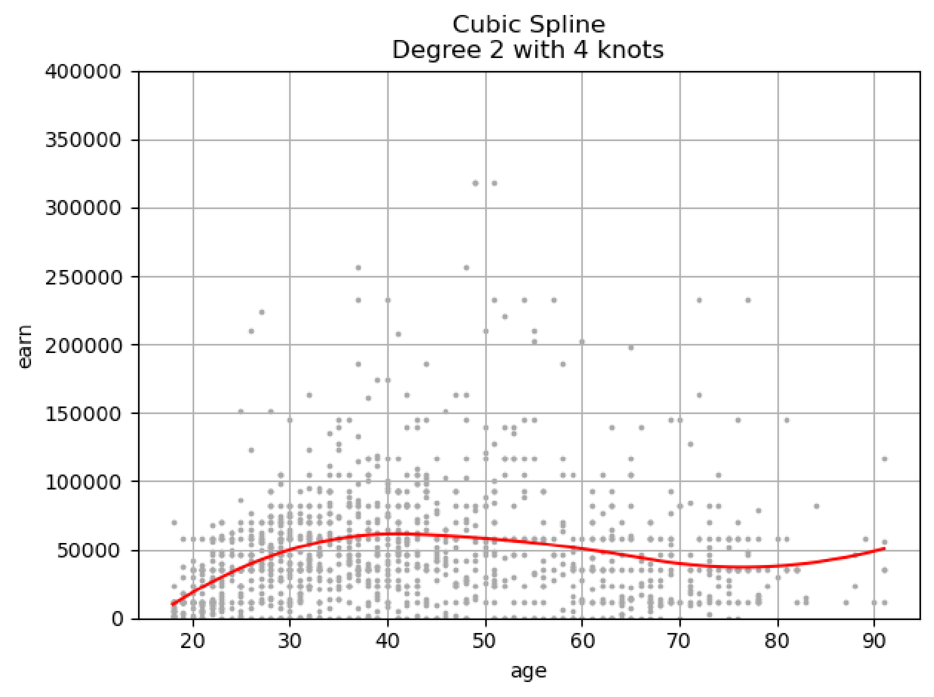
- Predictor variable is split into different sections and predictions computed for each section.
- In practice, this amounts to creating a set of dummy variables representing each group
- Modeled as interactions between dummy variable and predictor

Splines

- Polynomial functions fit the same function over the entire range of predictor variable
- Step functions segment the predictor and allow different functions to be fit to each section. However, such functions tend to be discontinuous at the knots.
- Splines ensure that the functions are continuous at the knots
- Linear Spline is a piecewise linear polynomial continuous at each knot
- A cubic spline is a piecewise cubic polynomial with continuous derivatives up to order 2 at each knot

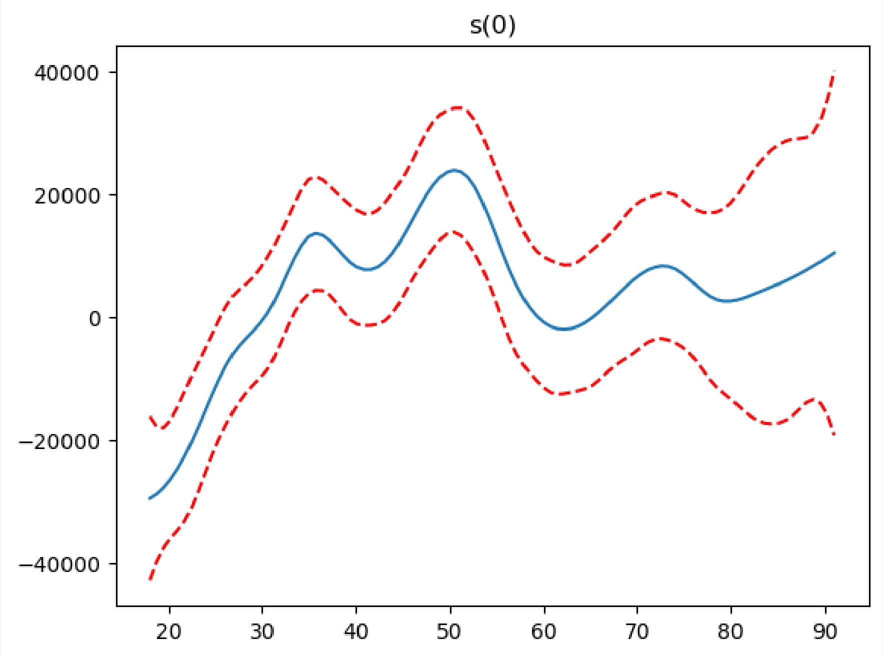
Splines


Linear Spline
Degree 1 with 4 knots
earn
age
400000
350000
300000
250000
200000
150000
100000
50000
0
20
30
40
50
60
70
80
90


Cubic Spline
Degree 2 with 4 knots
earn
age
400000
350000
300000
250000
200000
150000
100000
50000
0
20
30
40
50
60
70
80
90

Generalized Additive Models

- Allows for flexible nonlinearities in several variables, but retains the additive structure of linear models
- $y = \beta_0 + f_1(x_1) + f_2(x_2) + ... + f_p(x_p) + \varepsilon$
- Can include a mix of linear and non-linear terms
- GAMs are additive. However, low-order interactions can be included.
- GAMs can be used for classification in the same way as for linear regression.



Conclusion

- In this module, we examined
  - *polynomial regression*
  - *step functions*
  - *splines*
  - *generalized additive models*