ACTL3142

Week 8: Moving Beyond Linearity

Tetian Madfouni

t.madfouni@unsw.edu.au

Announcements

Not much. Welcome to the “fun” part of the course though (depending how you want to think about it)

Unfortunately the tutorial side of this is very much rapid fire different possible models

Recap: Moving Beyond Linearity

Up to now, we've been using GLMs and linear regressions for everything, what is the key drawback of these simplistic models?

- Assumption of linearity. Generally an extremely unrealistic assumption and while we get an interpretable model, it's normally going to be pretty bad

Polynomial regression, how does it work/differ from linear regression? Any other important things to note

- Basically just adding extra powers of the predictor, i.e. $y = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \cdots$. This allows for non-linear relationships that follow polynomial patterns to be modelled. However, generally we should increase these powers over 4-5ish as it tends to overfit and become very erratic

Step functions, same thing as above?

- Pretty much just fitting a step function where we have cuts at $c_0, \dots, c_k$, and $y = \beta_0 + \beta_1 I(c_1 \leq x_i < c_2) + \beta_2 I(c_2 \leq x_i < c_3) + \cdots$. In practice not super useful, but illustrates a more general idea of basis functions where $y = \beta_0 + \beta_1 b_1(x_i) + \beta_2 b_2(x_i) + \cdots$. What would the basis functions be for polynomial regression?

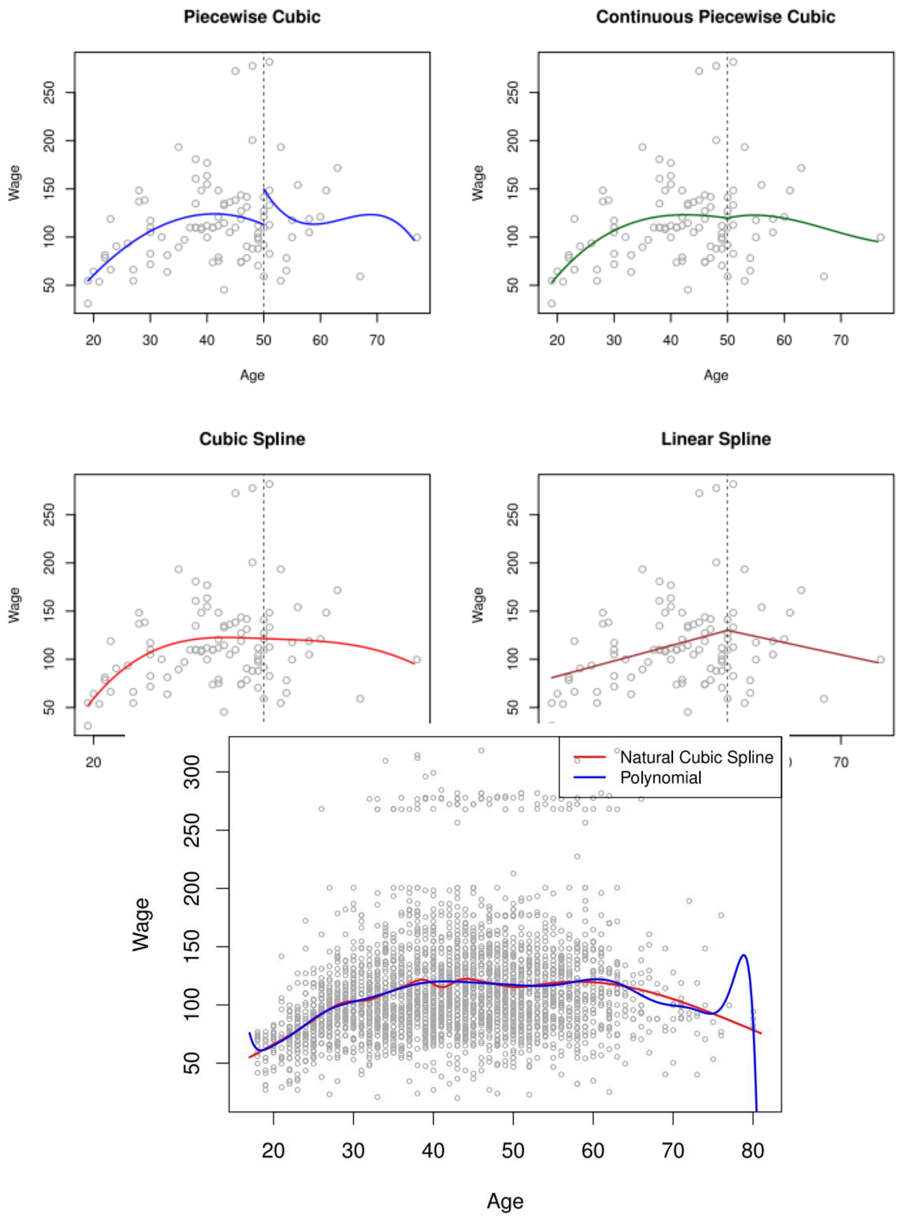
Recap: Moving Beyond Linearity

Regression splines, same thing as above ☺

- Effectively just splitting data at points called "knots", and between each of these knots we fit a different polynomial regression. This is just a piecewise regression, but becomes a spline when we require that for a degree d polynomial piecewise regression, the function is continuous up to the $(d-1)$th derivative.

How do natural splines extend on regression splines and why?

- Using something like a cubic spline, the function before the first and past the last knots tends to look very strange. A natural spline enforces a linear regression in these two regions



Recap: Moving Beyond Linearity

Smoothing splines

- Best thought of as a regression spline with a knot at every point. Obviously this would blatantly overfit, so this is corrected by adding a penalty (much like ridge/lasso) where $L = MSE + \lambda \int_{-\infty}^{\infty} [g''(x)]^2 dx$. Effectively penalising how erratic the curve is.

Local regression

- Quite similar to the logic of KNN. When we want to predict y for some x we fit a weighted linear regression on the neighbouring points of x. Very much an algorithm that suffers from curse of dimensionality.

Generalised Additive Model (GAM)

- Function of the form $y_i = \beta_0 + f_1(x_i, 1) + f_2(x_{i,2}) + \cdots$. Effectively allowing us to apply any function we want to each predictor, this can include things like splines, and all the things we've covered this week. So overall it's an extremely general type of model.