Modeling interaction effects

ISTA 410 / INFO 510: Bayesian Modeling and Inference

U. of Arizona School of Information

November 2, 2020

Outline

Last time:

- GLMs

This week:

- Interactions to linear models
- Connection to multilevel models
- Intro to Gaussian processes

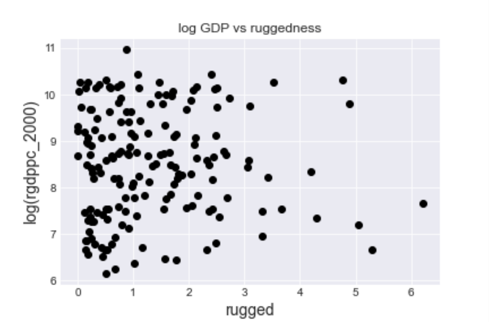
Example: Terrain and GDP

Terrain ”ruggedness” and economy

(Example from Statistical Rethinking Ch7.) What is the relationship between the geographic terrain in a nation and its economy?

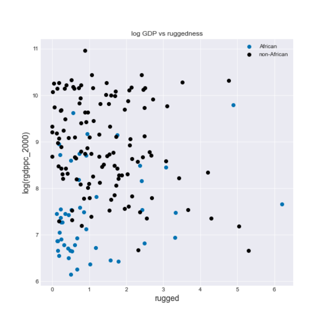
Data: observations on many countries

- Outcome: log GDP (as of 2000, when data was collected)
- Predictor: terrain “ruggedness” index



Terrain ”ruggedness” and economy

Closer examination of the data reveals an interesting phenomenon: the relationship is different for countries in Africa.



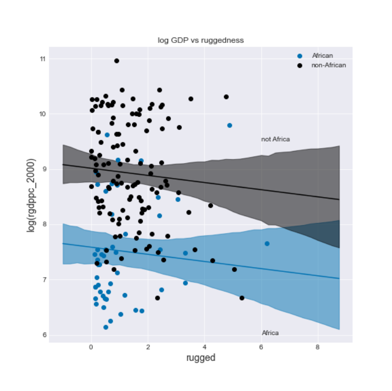
A simple approach that won't work

A simple approach that's not good enough: add an indicator variable for African countries, and do a bivariate regression:

$$
\begin{aligned}
\log GDP &\sim \text{Normal}(\mu_i, \sigma) \\
\mu_i &= \alpha + \beta_R R + \beta_A A \\
\beta_R &\sim \text{Normal}(0, 1) \\
\beta_A &\sim \text{Normal}(0, 1) \\
\sigma &\sim \text{HalfCauchy}(5)
\end{aligned}
$$

A simple approach that won't work

Problem:



Allows for a shift, not a change in slopes.

Allowing interactions

To add interactions:

$$
\begin{aligned}
\log GDP &\sim \mathrm{Normal}(\mu_i, \sigma) \\
\mu_i &= \alpha + \beta_R R + \beta_A A + \beta_{AR} AR \\
\beta_R &\sim \mathrm{Normal}(0, 1) \\
\beta_A &\sim \mathrm{Normal}(0, 1) \\
\beta_{AR} &\sim \mathrm{Normal}(0, 1) \\
\sigma &\sim \mathrm{HalfCauchy}(5)
\end{aligned}
$$

So we have a third slope, for the *product* of R and AR.

Why is this the approach?

Where this comes from: just model the slope β_R as being itself a linear function of A:

$$
\begin{aligned}
\log GDP &\sim \mathrm{Normal}(\mu_i, \sigma) \\
\mu_i &= \alpha + \gamma_i R + \beta_A A \\
\gamma_i &= \beta_R + \beta_{AR} A \\
\beta_R &\sim \mathrm{Normal}(0, 1) \\
\beta_A &\sim \mathrm{Normal}(0, 1) \\
\beta_{AR} &\sim \mathrm{Normal}(0, 1) \\
\sigma &\sim \mathrm{HalfCauchy}(5)
\end{aligned}
$$

The result

Result from the interaction model:

Here, we can see the different slope. Ruggedness has a positive association with GDP for African nations, negative for others.

Example: wine tasting

The "Judgement of Princeton"

The Judgement of Princeton

- 9 judges, 20 wines
- Wines split between red and white, NJ or France
- Judges split between American or French/Belgium

Predictors:

- Wine color: red or white
- Wine origin: NJ or France
- Judge nationality: US or EU

Interactions

Potential for interactions between all predictors:

- Interaction between origin and judge: judge bias.
  Judge bias might depend upon color.
- Interaction between color and judge: taste preference.
  Taste preference might depend upon origin.
- Interaction between origin and color: relative advantage.
  Advantage might depend upon judge.

Let's do some computations and break things down along all axes.

Continuous interactions

So far, the interactions have been conditional on a categorical predictor?

Can we have interactions conditional on a continuous predictor?

Continuous interactions

So far, the interactions have been conditional on a categorical predictor?

Can we have interactions conditional on a continuous predictor?

Why not?

$$y = \alpha_0 + \gamma x_1 + \beta_2 x_2$$
$$\gamma = \beta_1 + \beta_{1,2} x_2$$
$$y = \beta_1 x_1 + \beta 2 x_2 + \beta_{1,2} x_1 x_2$$

Connection to multilevel models

The Africa interaction as a multilevel model

When we did multilevel/hierarchical regression:

- Varying intercepts
- Varying slopes

Ruggedness model: different slope depending on category membership

- In a literal sense, this is just varying slopes
- We could set up the model in a similar way by including two slopes explicitly, drawing from a hyperprior

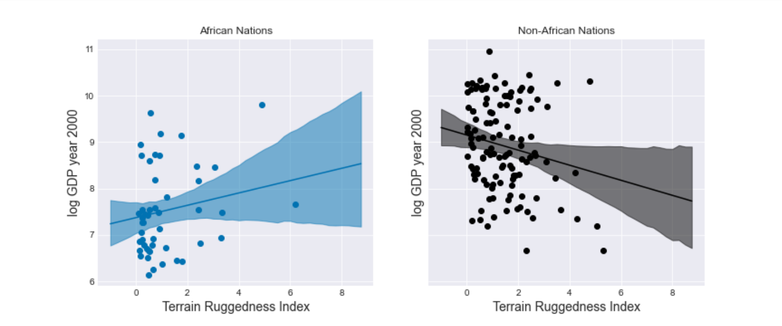
The Africa interaction as a multilevel model

Setting it up as a multilevel model:

$$
\begin{aligned}
\log GDP &\sim \mathrm{Normal}(\mu_i, \sigma) \\
\mu_i &= \alpha + \beta_{R[j]} R + \beta_A A \\
\beta_{R[j]} &\sim \mathrm{Normal}(0, \tau) \\
\beta_A &\sim \mathrm{Normal}(0, 1) \\
\tau &\sim \mathrm{HalfCauchy}(5) \\
\sigma &\sim \mathrm{HalfCauchy}(5)
\end{aligned}
$$

The Africa interaction as a multilevel model

Result from the multilevel version:



Essentially the same result as the conventional interaction version.

Let's inspect the parameters.

Multilevel models as interaction machines

Multilevel models:

- Group observations into clusters
- Parameters for each cluster drawn from a common

Multilevel models as interaction machines

Multilevel models:

- Group observations into clusters
- Parameters for each cluster drawn from a common
- This is just a way of expressing that the parameter values are conditional on group membership
- Multilevel models give a way to encode interactions *and* regularize automatically across groups

Not surprising: remember we constructed the interaction by putting a linear model inside a linear model

What about continuous interactions?

So:

- Multilevel model: slopes can vary across categories, with some information shared between categories
- Can be seen as a form of interaction between category and slope
- But: interactions from before also allow interactions between continuous variables

Can we get the best of both? Use a hyperprior-like structure, but using continuous "categories"?

What about continuous interactions?

So:

- Multilevel model: slopes can vary across categories, with some information shared between categories
- Can be seen as a form of interaction between category and slope
- But: interactions from before also allow interactions between continuous variables

Can we get the best of both? Use a hyperprior-like structure, but using continuous "categories"?

Yes: Gaussian processes