Nonlinear Regression

Why, what and how?

Why nonlinear regression?

- Linear regression models are simple and can reveal many interesting relations between variables.
- When the linearity assumption does not hold, one can apply nonlinear transformations to the variables.

Examples of nonlinear transformations

- Logarithmic

$$y = \beta_1 \log(x_1) + \beta_0 + \epsilon$$

```
> lm(y ~ log(x))
```

Examples of nonlinear transformations

- Polynomial

$$y = \beta_n x^n + \ldots + \beta_1 x + \beta_0 + \epsilon$$

```
> library(Design)
> lm(y ~ pol(x, n))
```

Examples of nonlinear transformations

- Power law

$$\log(y) = \beta \log(x) + \beta_0 + \epsilon$$

```
> lm(log(y) ~ log(x))
```

Generalized Linear Models

- Even better: the outcomes of the dependent variable could be generated from normal, poisson, or any other exponential-family distribution.

Why don't we just call it a day now?

Well, sometimes we would like to define an arbitrary nonlinear relation

$$y = \frac{1}{\beta \log(x) + e^x} + \epsilon$$

Well, sometimes we would like to define an arbitrary nonlinear relation

$$y = \frac{1}{\beta \log(x) + e^x} + \epsilon$$

Highly nonlinear in X!

We may also have theoretical considerations for the distribution of errors

$$y = x^{\beta} + \epsilon \quad \textit{Vs.} \quad \log(y) = \beta \log(x) + \epsilon$$

Here come the benefits of using a nonlinear model

Nonlinear models are often *mechanistic*, i.e., based on a model for the mechanism producing the response.

(Pinheiro & Bates, 2000)

The form of a nonlinear model

$$y = f(x) + \epsilon$$

f(x) is a nonlinear function of x, which is based on theoretical considerations for the mechanism that links xs to the observed ys

Mechanism: one or two?

- Exponential Decay

$$y = \alpha e^{\lambda x} + \epsilon$$

where α and λ are free parameters

- Biexponential

$$y = \alpha_1 e^{\lambda_1 x} + \alpha_2 e^{\lambda_2 x} + \epsilon$$

More free parameters to fit

Improved Interpretability

- Pinheiro & Bates, 2000:
  - A fifth-degree polynomial can approximate a logistic growth model quite well.

$$h = -2.2911 + 16.591t - 44.411t^2 + 56.822t^3 - 31.514t^4 + 6.3028t^5$$

$$\Downarrow$$

$$h = \frac{\phi_1}{1 + exp[-(t - \phi_2)]/\phi_3}$$

Parsimony

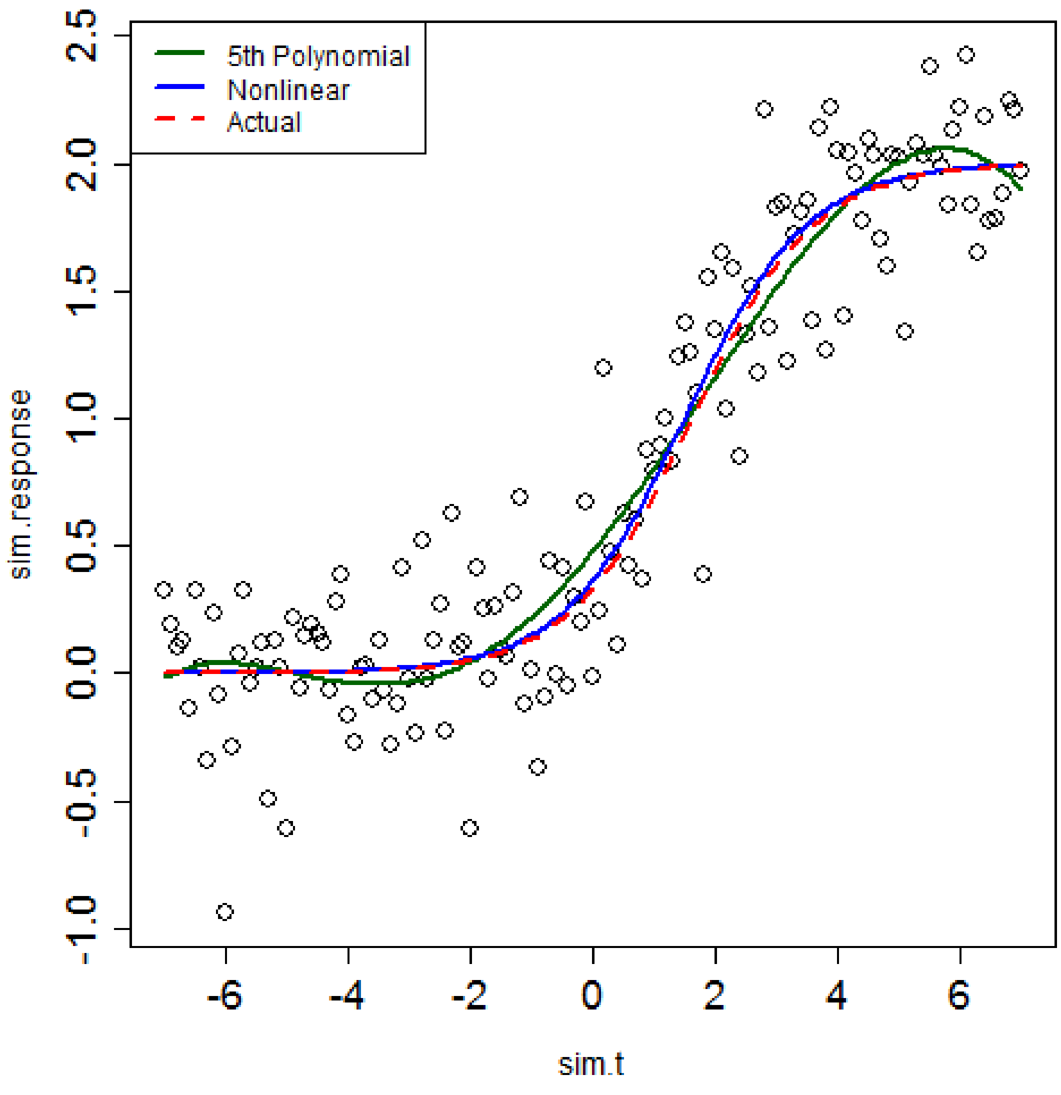
- Pinheiro & Bates, 2000:
  - A fifth-degree polynomial can approximate a logistic growth model quite well.

$$h = -2.2911 + 16.591t - 44.411t^2 + 56.822t^3 - 31.514t^4 + 6.3028t^5$$

$$\Downarrow$$

$$h = \frac{\phi_1}{1+exp[-(t-\phi_2)]/\phi_3}$$

Validity beyond observed range



Residual standard error:

Polynomial: 0.31
Logistic: 0.29

Let’s make a nonlinear model from scratch

The Research Question

Do discourse cues decay? If so, how does it affect the distribution of sentence information?