Non-linear mixed effect models

Philip Dixon

10 July 2020

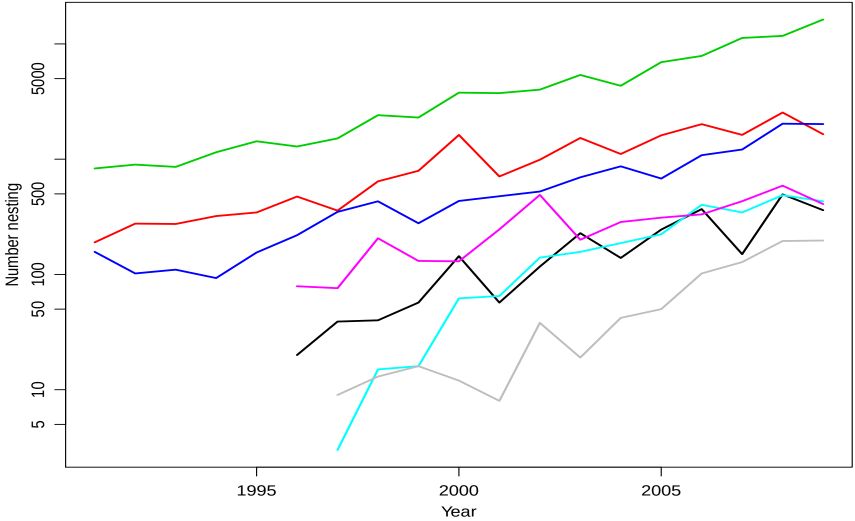
Non-linear mixed effect models

- Combine ideas from non-linear models and linear mixed effect models
- Very powerful, can be very difficult (theoretically and practically)
- Software is making more and more possible
- Outline:
  - Linear mixed effect models
  - Non-linear models
  - Non-linear mixed effect models
  - Alternatives: meta analysis
  - Alternatives: Bayesian

Linear mixed effect models

- Use random effects to describe clusters of data
  - Random effects also describe correlation between observations
- Number of sea turtles nesting on 7 beaches over time
- Cluster = beach
- Goals:
  - Estimate trend over time
  - Is trend consistent on all beaches

Sea turtle data



Sea turtle data

- Building up a model, *i* indexes Year; *j* indexes beach
  - linear regression, log count vs time

$$Y_{ij} = \beta_0 + \beta_1 X_i + \varepsilon_{ij}$$

  - Allow intercept to vary by beach

$$Y_{ij} = \beta_{0j} + \beta_1 X_i + \varepsilon_{ij}$$

  - Make intercept a random effect

$$\begin{aligned} Y_{ij} &= \beta_0 + u_{0j} + \beta_1 X_i + \varepsilon_{ij} \\ u_{0j} &\sim N(0,\, \sigma^2_{beach}) \\ Y_{ij} &= u_{0j} + \beta_1 X_i + \varepsilon_{ij} \\ u_{0j} &\sim N(\beta_0,\, \sigma^2_0) \end{aligned}$$

Sea turtle data

- Add a slope random effects, assume independent of intercept RE

$$
\begin{aligned}
Y_{ij} &= u_{0j} + u_{1j} X_i + \varepsilon_{ij} \\
u_{0j} &\sim N(\beta_0,\, \sigma_0^2) \\
u_{1j} &\sim N(\beta_1,\, \sigma_1^2)
\end{aligned}
$$

- Make intercept and slope RE correlated

$$
\begin{aligned}
Y_{ij} &= u_{0j} + u_{1j} X_i + \varepsilon_{ij} \\
\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} &\sim N\left(\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{bmatrix} \right)
\end{aligned}
$$

- Practical issue:
  - Intercept is predicted log count in year 0, far from the data
  - Redefine intercept as year 1990, or 2000

Linear mixed effect models

- The linear mixed effect model:

$$\begin{aligned} Y &= X\beta + Zu + \varepsilon \\ Y \mid u &\sim N(\,X\beta + Zu, \sigma_e^2 \end{aligned}$$

- Need marginal distribution of Y to construct a likelihood
- $Y \sim ??$
- In general, need to integrate: $f(Y) = \int f(Y \mid u) f(u) du$
- When u and ε are normally distributed, statistical magic happens:

$$\begin{aligned} Y &\sim N(\,X\beta,\; ZGZ' + R) \\ \text{Var } u &= G \\ \text{Var } \varepsilon &= R \end{aligned}$$

Linear mixed effect models

- Random effects result in correlations between observations

$$\text{Var } Y = ZGZ' + R$$

- Intercept RE only: Var Y is compound symmetric
- both RE: Var Y more complicated
- “Easy” to fit: Y is a large multivariate normal distribution
- Iterative algorithm:
  - Use REML to estimate random effect parameters given fixed effect parameters
  - Condition on those estimated RE parameters $\rightarrow$ Var Y
  - Use generalized least squares to estimate fixed effect parameters

Non-linear models

- Relax the assumption of linear effects:

$$Y_i = f(\theta, X_i) + \epsilon_i$$

- Exponential growth model

$$Y_i = Y_0 e^{rX_i} + \epsilon_i$$

- Differences between two models for sea turtles

	$\log Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$	$Y_i = Y_0 e^{rX_i} + \epsilon_i$
Fixed model	Same, $\beta_0 = \log Y_0$	$\beta_1 = r$
error variance	constant for $\log Y_i$	constant for Y_i
error distrib.	normal for $\log Y_i$	normal for Y_i

Non-linear models

- Can rewrite models using more interpretable parameters
- Quadratic response model

$$\begin{aligned} Y_i &= \beta_0 + \beta_1 X_i + \beta_2 X^2 + \varepsilon_i \\ Y_i &= \beta_0 + \beta_2 \ (X_i - X_m)^2 + \varepsilon_i \end{aligned}$$

- X_m is location of maximum / minimum, β_0 is height at that max/min
- Two parameter logistic curve

$$\begin{aligned} Y_i &= \frac{1}{1 + \exp(\beta_0 + \beta_1 X_i)} + \varepsilon_i \\ Y_i &= \frac{1}{1 + \exp\left[\beta_1 (X_i - LC_{50}\right]} + \varepsilon_i \end{aligned}$$

- LC_{50} is X at which response is 0.5

Non-linear mixed effect models

- Clusters of observations
  - individuals followed over time
  - plots with multiple individuals with different X values
  - blocks with multiple treatments
- Could:
  - Fit one NL model to all clusters - probably violating independence
  - Fit separate NL models to each cluster
  - Combine NL and LME ideas: some or all NL parameters get distributions
- NLME allow inference to new individuals in the population

Non-linear mixed effect models

- Easy to write down the model
- Hard to fit it to data, estimate parameters
- Conditional distribution given by the NL model

$$Y \mid u = f(X, \beta + u) + \varepsilon$$

- Want $f(Y)$, need to integrate out the random effects
- Almost never an analytic solution
    - No nice multivariate normal distribution for Y

Non-linear mixed effect models

- Three computational approaches
  - Linearize the model: pseudolikelihood methods
  - Approximate the integral:
    - Laplace approximation or (adaptive) Gaussian quadrature
  - Add prior distributions, simulate samples from the posterior: Bayesian MCMC
- As far as I can tell:
  - nlme() linearizes the model to estimate fixed effects
  - nlmer() uses a Laplace approximation
- And a fourth idea: focus on the experimental unit (eu)
  - Fit NL model to each eu, extract estimate and se
  - Use meta-analysis to estimate overall estimate and precision
  - And estimate heterogeneity among eu's
  - Use meta-regression to evaluate fixed effect models

Comparison of approaches

- All methods:
  - Provides estimates, se's and inference for fixed effects
  - Provides estimates of variability
  - Predictions with se's, e.g. to interpolate for a subject
- Frequentist mixed effects
  - Can predict for new subjects
  - Can pool information for some parameters, separately est. others
  - Inference based on normal approximations for estimates
  - Use AIC, BIC or LR tests to do model selection on random effects
- What do I mean by "partial pooling"?
  - Quadratic response model, 3 parameters: β_0, β_2, and X_m
  - All groups have same β_2, but differ in β_0 and X_m
  - Model has random effects for β_0 and X_m, not for β_2

nlme or nlmer?

	nlme	nlmer
Fitting fixed effects	easy	cumbersome
Starting values	SS function	have to specify
Predict *Y* for new subjects	easy	
Nonconstant variance	easy-ish	not possible
Se/CI for predictions	??	bootMer()
Graphics	nice	do yourself
Active development	no	yes

Comparisons

- Bayesian mixed effects
  - Predictions with se's for old and new subjects
  - Can pool information for some parameters, separately est. others
  - se's and ci's for RE variances
  - Inference based directly on posterior distribution
  - Can use Bayesian model selection tools
  - Have to specify prior distributions - requires thought
  - "Non-linear models are incredibly flexible and powerful, but require much more care with respect to model specification and priors than typical generalized linear models." (from brms vignette)
  - Lots of traps for the unaware

Comparisons

- Meta analysis
    - Probably the easiest to fit complicated FE models
    - But can not partially pool information
    - Start with separate fits to each group
        - Ignores a group when model doesn't fit

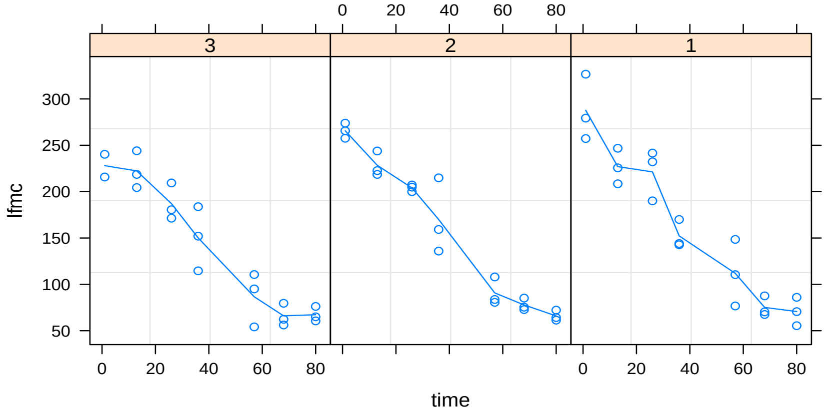
Examples: Leaf Fire Moisture Content

- Data in nlraa package (lfmc)
- Paper in Ecology and Evolution, 2019, Oddi et al.
  - Available at `https://doi.org/10.1002/ece3.5543`
- Four species, measured multiple times during fire season
- 2 sites, each with 6 plots
- Destructively sampled 3 plants per plot each time
- Oddi et al. fit 4 parameter logistic curves to each plot.
- Oddi et al. code in their supplemental material.

Examples: Leaf Fire Moisture Content

- My example: one species *S. bracteolactus*, in 1 site, 3 plots
- Goals
  - Which parameters appear to vary among plots?
  - What is the typical curve?
- My code in nlme.r and with comments in nlme.Rmd
  - Using `nlme()`: non-linear ME model
  - Using `nlmer()`: non-linear ME model
  - Using `stan_nlmer()`: Bayesian NL ME
    - Fernando prefers `brms`
    - More difficult to set up
    - But requiring more thinking is a good thing here
  - Using `metafor()`: meta analysis

Examples: Leaf Fire Moisture Content



Examples: N leaching, from Gina

- N leaching (kg/ha) as a function of N fertilizer applied (kg/ha)
    - 9 sites, 20 years, 2 crop rotations at each site/year
    - 7 N rates, same N for all site years.
    - leaching derived from a crop model
- Goals
    - Choose a non-linear model
    - Estimate difference between cropping systems for each parameter
    - Correlation among parameters?
    - Variation between sites? between years?
    - Which sites have highest year-year variation in leaching?

Examples: N leaching, 2010 data