

Phylogeny model updates

10 June 2020

What I will cover...

- ▶ The way phylogeny is generally handled in ecological statistical models
 - ▶ PGLS
 - ▶ PMM
- ▶ The way I want to handle it for OSPREE (but I could be wrong)
- ▶ My progress on that goal

A couple quick notes ...

- ▶ I am skipping over why we care about phylogeny, check out [phylogenyfounding.pdf](#) for some of that.
- ▶ Phylogenetic structure in a value (say, a 'trait') is measured based on a value called λ . When $\lambda = 1$ the trait is perfectly predicted by the phylogeny; when it's 0, the phylogeny does not matter.
- ▶ You can play with the code in `ubermini.R` or `max_sims.R` to get a feel for this.

Common modeling approaches

Part 1: PGLS (adapted from second edition of *Statistical Rethinking*)

Consider...

$$y \sim MVN(\mu, S) \quad (1)$$

$$\mu_i = \alpha + \beta * x_i \quad (2)$$

μ is a usual linear model. P is a vector of phenological dates (one per species), and S is a covariance matrix with as many rows and columns as species. In ordinary regression this takes the form:

$$S = \sigma^2 I \quad (3)$$

where I is just an identity matrix (all 1s) so we can ignore it. In PGLS we replace S with the phylogenetic covariance matrix (Σ).

Common modeling approaches

Part 1: PGLS

You have to make sure of a few things:

- ▶ Phylogeny must go in as correlation matrix (this makes the diagonals 1s and the off-diagonals the correlation across species due to evolutionary history) and make sure the rows and columns are in the same order as the species will be ordered numerically.
- ▶ This model forces the correlation structure you give it—it does not adjust the correlation structure at all.

If you don't want to assume $\lambda = 1$, then you need to estimate a value to multiply the matrix by such that:

$$S = \sigma^2(\Sigma * \lambda) \quad (4)$$

Common modeling approaches

Part 2: PMM (phylogenetic mixed model)

Now PMM...

$$y = \alpha + \beta x + a + e \quad (5)$$

$$a \sim \text{normal}(0, \sigma_P^2 \Sigma) \quad (6)$$

$$e \sim \text{normal}(0, \sigma_R^2 I) \quad (7)$$

$$\text{PGLS: } y \sim \text{normal}(\alpha + \beta x, \sigma_P^2 \Sigma) \quad (8)$$

... where α is the intercept β is the slope for the co-factor x , a is the phylogenetic random effect, and e is the residual error. This model estimated two variances: V_P is the variance of the phylogenetic effect and V_R is the residual error (environment effects, intraspecific variance, measurement error, etc.).

Common modeling approaches

Part 3: PMM vs PGLS

- ▶ People often say PGLS does not allow for non-phylogenetically structured error (but I think it sort of does once you scale the phylogenetic effect by λ , no?)
- ▶ PMM explicitly models other sources of error through e .

In PMM the strength of the phylogenetic effect is measured as:

$$\lambda = \frac{\sigma_P^2}{\sigma_P^2 + \sigma_R^2} \quad (9)$$

which is equivalent to just saying 'what proportion of the variance is due to phylogeny?'

What's wrong with this model?

What I want for OSPREE

What is wrong with this?

$$y = \alpha_0 + \alpha + \beta x + e \quad (10)$$

$$\alpha \sim MVN(0, \sigma_{P\alpha}^2) \quad (11)$$

$$\beta \sim MVN(0, \sigma_{P\beta}^2) \quad (12)$$

$$e \sim normal(0, \sigma_y^2) \quad (13)$$

$$\sigma_{P\alpha}^2 = \alpha_\alpha + \lambda_\alpha * \Sigma \quad (14)$$

$$\sigma_{P\beta}^2 = \alpha_\beta + \lambda_\beta * \Sigma \quad (15)$$

Where α_0 is a grand mean and species-level intercepts are partially pooled by phylogeny, scaled by λ_α and slopes are also are partially pooled by phylogeny, scaled by λ_β , and there is some residual error σ_y .

What's wrong with this model?

What I want for OSPREE

Did we answer 'What's wrong with this model?'

If not, go back, discuss.

What's wrong with this model?

What I want for OSPREE

Also, I have another question....

In linear regression like this, what's the difference between error and an intercept?

Where am I at?

What I have done

- ▶ I have stolen code from Will Pearse, which he tells me should work.
- ▶ I have it running for forcing, chilling, photo slopes... with the real OSPREE data, but it does not want to have phylogeny on the intercepts also (I want to try it with non-partially pooled intercepts ...).
- ▶ I tried to do fake data
- ▶ And now I am stuck ... (see ubermini.R and ubermini.stan)

Where I am at

I am stuck

The new code is this line I stole from Will Pearse's code:

```
bforce ~ MVN(rep.vector(0,n.sp),  
diag.matrix(repector(null.interceptsb, n.sp)) +  
lam.interceptsb*Vphy);
```

Where I am at

I am stuck

```
bforce ~ MVN(rep.vector(0,n.sp),  
diag.matrix(repector(null.interceptsb, n.sp)) +  
lam.interceptsb*Vphy);
```

It says that my vector of slopes is multinormal centered around 0 (why zero? That's how Gaussian processes work, it somehow gets the 'centering' if you will from the $y \sim \text{normal}(\hat{y}, \sigma_y)$ bit of the model) and the variance should be the within-species variance on the diagonal, and the between-species variance on the off-diagonals.

Where I am at

I am stuck

Let α be `null.interceptsb` and λ be `lam.interceptsb`.

`bforce` =

$$\begin{bmatrix} \alpha + \lambda * Vphy & \cdots & \lambda * Vphy \\ & \ddots & \\ \lambda * Vphy & \cdots & \alpha + \lambda * Vphy \end{bmatrix}$$

Where I am at

I am stuck

If V_{phy} is set to have 1s down the diagonal, we can simplify to this:

bforce =

$$\begin{bmatrix} \alpha + \lambda & \cdots & \lambda * V_{phy} \\ & \ddots & \\ \lambda * V_{phy} & \cdots & \alpha + \lambda \end{bmatrix} \quad (16)$$

Where I am at

I am stuck

If V_{phy} is set to have 0s down the diagonal, we can simplify to this:

bforce =

$$\begin{bmatrix} \alpha & \cdots & \lambda * V_{phy} \\ & \ddots & \\ \lambda * V_{phy} & \cdots & \alpha \end{bmatrix} \quad (17)$$

This seems good—now the within-species estimate is α and the between-species is $\lambda * V_{phy}$.

And the crowd goes wild!

No, just kidding

It's not working

- ▶ I can return the slopes
- ▶ λ is wrong and I don't know why.
- ▶ Any ideas? I have a tried a few things ...
 - ▶ Set `nind` to 1 (1 value per species) no dice.
 - ▶ Tried taking it off slopes and putting on intercepts ... no dice.
 - ▶ Your idea here?
- ▶ I asked Will, but it will be a while.