Lesson 6 - More Random Effects and examples

MLE Software Online Course

Click here to view presentation online

Jim Bence

16 December 2023

Topics

- Overdispersion via random effects
- What about REML?
- Residuals
- So is the Laplace approximation working?
- More examples
  - Age comp data - multinomial and Dirichlet-multinomial
  - Multivariate normal data
  - Correlated random effects for vonB model
  - Other undocumented examples

Overdispersion via random effects

- For distributions where variance cannot be controlled separately from mean (e.g., Poisson, multinomial)
  - Treat parameters of these distributions as random
- New probability distributions have been defined this way.
  - E.g., the NB (Poisson, gamma rate parameter), Dirichlet-multinomial (multinomial, p vector Dirichlet)
  - Compound pdf found by integrating the joint likelihood
- Alternatively could specify observation-specific random effects. E.g., Poisson with log of rate normal (this is GLMM, with log link function) - Easy to generalize using RTMB.

REML

- ML variance estimates are known to be biased
- REML variance estimates are unbiased in linear normal models and generally less biased than ML estimates
- REML estimates can be obtained by declaring all the fixed effects other than the variances as random (don't add anything to the function you minimize)
- Pretty much ignored and not studied/evaluated in stock assessment

Demonstration of REML for "known" bias case

- In standard regression (with normal errors) the maximum likelihood estimate for residual error variance is (residual SS)/n.
- The minimum variance unbiased estimate for linear regression is (residual SS)/(n-2).
- This is approximately unbiased for nonlinear regression.
- The RTMB reml procedure will produce the approximatley unbiased estimates (see musky_vonb_reml.R)
- This is just to demo of what REML is doing for a known answer case!

Residuals

- standard Pearson residuals
- Problems with standard Pearson residuals
- One step ahead (osa) aka recursive quantile residuals

Pearson residuals

- defined as (obs-pred)/sd
- sd is what the standard deviation for (obs-pred) should be given your model and model estimates
- Idea is that if raw residuals are approximately normal and independent then Pearson residuals will be approximately normal, independent, with equal (1) variance.
  - So all the residuals can be looked at together

Problems with Pearson residuals

- Actual residuals typically:
  - Not normal
  - Not independent
- In addition, we really want to look at residuals in some sense integrated over random effects, rather than at the best estimates of random effects

Solutions: OSA = Recursive quantile residuals:

- The capability is built into RTMB and in theory can be applied almost automatically: oneStepPredict(obj) - with data set up using OBS
- Numerically intensive and can be numerically tricky.
- The theory underlying this is pretty intense. See: Thygesen et al. Environ Ecol Stat 24(2): 317–339.
- Near automatic approach relies on TMB/RTMB understanding the density functions you are using.

Checking on the Laplace approximation

- Approximation depends on approximate normality of the combined vector or parameter estimates and random estimates.
- This is why we generally don't specify non-normal distributions for random effects.
- RTMB includes a helper function that checks the Laplace approximation

RTMB function checkConsistency to check on Laplace approximation

- Call as checkConsistency(obj) or as checkConsistency(obj,estimate=TRUE). Run summary on result.
- Requires you have set up your function for simulation (using OBS) (and the simulations work!).
- Usually want estimate=TRUE optional argument. This conducts a full simulation and evaluates parameter bias and whether simulated data are consistent with assumed distributions.of simulation.
  - Without this it evaluates the approximation in an approximate way (but faster).

More examples

- Age comp data
    - multinomial and Dirichlet-multinomial
- Multivariate normal data
- Correlated random effects for vonB model
- Other undocumented examples

Age comp example

- Constant recruitment and survival (necessary in example because the only type of data being used is age comps).
- multiple samples of age composition. Initially assumed to be multinomial, then Dirichlet multinomial
- Alternatives
  - Assume selectivity known (correctly)
  - Estimate selectivity for each age forced to increase monotonically
  - Estimate selectivity via a logistic function

Model

$$N_a = R\exp(-Z(a-r)), r \le a \le \max$$

$$C_{i,a} = q_i S_a N_a$$

$$p_a = \frac{S_a N_a}{\sum_{j=r}^{\max} S_j N_j}$$

$$\underline{n}_i \sim \text{multinom}\left(n_i, \underline{p}_a\right), n_i = \sum_{a=r}^{a=\max} n_{i,a}$$

Setting effective sample size

- Multinomial apps often use data as proportions and “effective sample size” (ESS)
- In such applications the negative log likelihood was written in terms of proportions and ESS
- In RTMB emulate by providing dmultinom product of proportions and ESS as the data (don’t set size!)
  - The NLL returned will depend on the ESS but ESS cannot be estimated (nothing new for multinomial here)
  - Works because the internal dmultinom calcs allow non-integers.

Dirichlet-Multinomial

- Compound distribution, p vector comes from dirichlet then used as parameter of multinomial.
- Used to introduce overdispersion relative to multinomial
- Using linear form where ESS proportional to sample size

$$\underline{n}_i \sim \mathrm{DirMult}(n_i, \underline{\alpha}_i), \alpha_{i,a} > 0$$

$$E\left(n_{i,a}\right) = n_i \frac{\alpha_{i,a}}{\sum_j \alpha_{i,j}}, ESS_i = \sum_a \alpha_{j,a}$$

$$\alpha_{i,a} = \theta n_i p_a, ESS_i = \theta n_i$$

Multivariate Normal Distribution with unstructured var-cov matrix

- Parameterize so that standard deviations and correlations can be calculated and converted into a variance-covariance matrix
- Estimate the standard deviations on log-scale
- Not sufficient to restrict correlations to -1 to 1 as some combinations of correlations are not consistent with feasible correlation matrices
    - e.g., if *a* and *b* are highly negatively correlated, you can't have *c* highly positively correlated with *a* and highly negatively correlated with *b*

Converting SDs and correlation matrix to a var-cov matrix

- Typical element of correlation matrix P is $\rho_{i,j} = \sigma^2_{i,j}/(\sigma_{i,i}\sigma_{j,j})$
  - $\sigma^2_{i,j}$ is a typical element of the variance-covariance matrix, with $\sigma^2_{i,j} = \rho_{i,j}\sigma_{i,i}\sigma_{j,j}$
- Matrix algebra to convert correlation matrix and SDs to var-cov matrix

$$\Sigma = \mathrm{diag}(\sigma_{1,1}, \cdots, \sigma_{k,k})\mathrm{P}\,\mathrm{diag}(\sigma_{1,1}, \cdots, \sigma_{k,k})$$

Motivation for being able to deal with unstructured MVN

- You might actually want to assume some data are multivariate normal (e.g., some catch-at-age assessments assume log of catch-at-age MVN). But often we impose structure on correlations.
- You might want to allow different random effects to be correlated. E.g., we might expect the vonB function parameters for a pond will be either positively or negatively correlated with one another.
- Illustrate with another very simple example

Simple application of unstructured multivariate normal

- We have a set of multivariate observations assumed to come from a multivariate normal distribution
- We estimate the mean vector and parameters that determine the variance-covariance matrix
  - Illustrate how to estimate parameters on real number line that determine a "legal" correlation matrix
  - Illustrate how to get the variance-covariance matrix from correlation matrix and vectors of standard deviations
  - See "unstructured" example R script

Correlated random effects for vonB model

- We have already seen how to make a vonB parameter a random effect. We could have made more than one parameter random at the same time using dnorm
    - This assumes the random effects are independent but often that is not plausible [when fish ultimately get big (Linf) they might approach asymptotic size more slowly (K)].
- We can implement this by having the vonB parameters follow a multivariate normal distribution
- Larger and more realistic example data set, reparameterized the vonB (use L2 rather than t0)

New growth data set

- 20 different ponds
- Each observation gives the pond ID and the length and age for an individual fish

Reparameterized vonB

- This is not changing the underlying growth function, just how it is parameterized
- Standard parameterization has been criticized for t0 being hard to interpret and for correlation with other parameters
- Use L2 (length at age 2) instead. Chose age within range of observed data

$$L_a = \widetilde{L}_2 + \left(L_\infty - \widetilde{L}_2\right) \exp(-K(a-2))$$

The multivariate random effects model

$$\underline{\omega}_i = \left[\log L\infty_i, \log K_i, \log \widetilde{L2}_i\right]^T \sim$$

$$MVN\left(\left[\overline{\log L\infty_i}, \overline{\log K_i}, \overline{\log \widetilde{L2}}\right]^T, \Sigma\right)$$

$$L_{i,j} \sim N\left(\mathrm{vonb}(\underline{\omega}_i, a_{i,j}), \sigma^2_{L,i,j}\right)$$

$$\sigma_{L,i,j} = \exp[\mathrm{int} + \mathrm{slp} * \mathrm{vonb}(\underline{\omega}_i, a_{i,j})]\,\mathrm{vonb}(\underline{\omega}_i, a_{i,j})$$