Introduction to Template Model Builder

for fitting complex models

Overview

- Automatic differentiation
- TMB features
- Mixed-effects example
- Demonstration
  - The C++ template fle
  - Running from within R

Automatic differentiation

AD is simply the software applying the chain rule of differentiation when it performs numerical operations.

By way of example, if the data are iid $N(\mu, \sigma^2)$, then each y_i contributes a term of the form

$$-\log\sigma - \frac{(y_i - \mu)^2}{2\sigma^2} \ , \qquad (1)$$

to $l(\boldsymbol{\theta})$. To match the template file seen later, this model will be parameterized using $\boldsymbol{\theta} = (\theta_1, \theta_2) = (\mu, \log\sigma)$. So, if $\sigma^2 = \exp(2\theta_2)$ is a variable defined in the TMB template file, then this object is stored in TMB as

$$\begin{aligned}
\sigma^2 &\equiv \left(\sigma^2, \left[\frac{\partial\sigma^2}{\partial\theta_1}, \frac{\partial\sigma^2}{\partial\theta_2}\right]\right) \\
&= \left(\exp(2\theta_2), \left[\frac{\partial\exp(2\theta_2)}{\partial\theta_1}, \frac{\partial\exp(2\theta_2)}{\partial\theta_2}\right]\right) \\
&= \big(\exp(2\theta_2), [0, 2\exp(2\theta_2)]\big) \ . \qquad (2)
\end{aligned}$$

Automatic differentiation

Similarly, if the second term in (1) is denoted by ζ, then

$$\begin{aligned}\zeta &= \frac{(y_i-\mu)^2}{2\sigma^2} \\ &= \frac{(y_i-\theta_1)^2}{2\exp(2\theta_2)}\ ,\end{aligned}$$

is stored in TMB as

$$\zeta \equiv \left(\zeta, \left[\frac{\partial\zeta}{\partial\theta_1}, \frac{\partial\zeta}{\partial\theta_2}\right]\right)\ ,$$

where, for example,

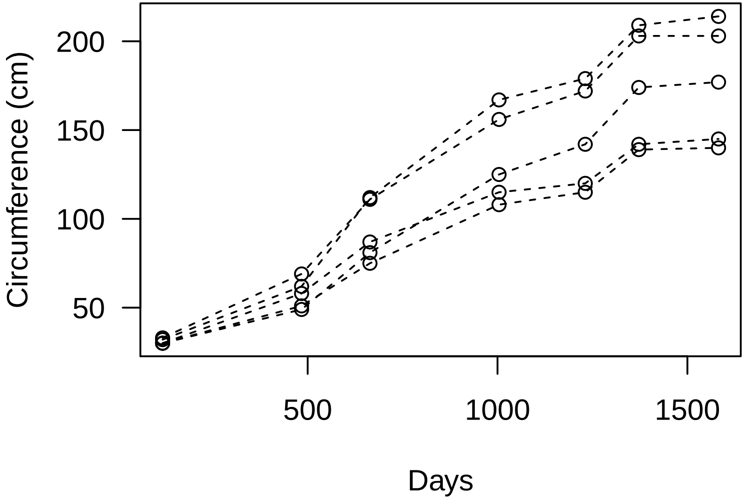
$$\begin{aligned}\frac{\partial\zeta}{\partial\theta_2} &= \frac{\partial\frac{(y_i-\theta_1)^2}{2\sigma^2}}{\partial\sigma^2} \times \frac{\partial\sigma^2}{\partial\theta_2} \qquad (3)\\ &= -\frac{(y_i-\theta_1)^2}{2\sigma^4} \times 2\exp(2\theta_2) \\ &= -\frac{(y_i-\theta_1)^2}{\exp(2\theta_2)}\ . \qquad (4)\end{aligned}$$

Why use TMB?

- In addition to returning the objective function (e.g., negative log-likelihood, or negative log-joint density function) it provides the exact derivatives with respect to *all* parameters.
  - Faster and more reliable optimization of negative log-likelihood.
  - Derivatives can be used by Hamiltonian MCMC.
- Fits mixed effects models
  - Uses Laplace approximation to marginalize over random effects.
  - Doesn't provide quadrature or importance sampling (unlike ADMB).
- Provides powerful functionality for spatio-temporal models.
  - Gaussian Markov Random Fields
  - Revolutionizing quantitative ecology, e.g., spatial factor analysis.
  - Sophisticated concepts (triangulations, Matern covariance function etc) — see Andrea Havron's talk in Sept.
- Well integrated with R

Orange tree example

The data are measurements of the circumference (mm) of five orange trees at seven different sampling occasions (from Draper&Smith 1981)



Orange tree mixed-effects model

We assume growth follows a logistic curve. We fit a mixed-effects model, with random tree effects ($u_i, i = 1, ..., 5$) crossed with random day effects ($v_j, j = 1, ...7$).

$$
\begin{aligned}
u_i &\sim N(0, \sigma_u^2) \\
v_j &\sim N(0, \sigma_v^2) \\
y_{ij}|u_i, v_j &\sim N(\mu_{ij}, \sigma^2) \ ,
\end{aligned}
$$

and with x_j denoting days at the jth sampling time,

$$
\mu_{ij} = \frac{a + u_i + v_j}{1 + \exp(-(x_j - b)/c)} \ .
$$

Parameter a corresponds to the full-grown (as age$\rightarrow$infinity) expected circumference, and b corresponds to the age at which the expected circumference is equal to $a/2$. Parameter $c > 0$ is an inverse growth rate parameter, with smaller values of c corresponding to faster growth.

R Demo

- .cpp template file
- Dynamic loading in R and optimization