Math

EE Holmes

9/5/2018

R Markdown can typeset math beautifully. It uses the LaTeX language. You can learn LaTeX with some googling. It can be complex but the basics are fairly easy. Inline math is put between `$ $`. Equations that your want on separate lines can be put between `\begin{equation} \end{equation}` or between `$$ $$`.

This shows part of a R Markdown file for this technical report:

Citation: Holmes, E. E. 2016. Notes on computing the Fisher Information matrix for MARSS models. Part I Background. Technical Report. https://doi.org/10.13140/RG.2.2.27306.11204/1

(Expected) Fisher Information

The Fisher Information is defined as

$$I(\theta) = E_{Y|\theta}\{[\partial \log L(\theta|Y)/\partial\theta]^2\} = \int_x [\partial \log L(\theta|y)/\partial\theta]^2 f(y|\theta)dy \tag{1}$$

In words, it is the expected value (taken over all possible data) of the square of the gradient (first derivative) of the log likelihood surface at θ. It is a measure of how much information data (from our experiment or monitoring) have about θ. The log-likelihood surface is for a fixed set of data and the θ vary. The peak is at the MLE, which is not θ, so the surface has some gradient (slope) at θ since the peak is at the MLE not θ. The Fisher Information is the expected value (over possible data) of those gradients (squared).

It can be shown that the Fisher Information can also be written as

$$I(\theta) = -E_{Y|\theta}\{\partial^2 \log L(\theta|Y)/\partial\theta^2\} = -\int_y [\partial^2 \log L(\theta|y)/\partial\theta^2 f(y|\theta)dy \tag{2}$$

. . .

You can write equations with `$$ $$` also. Take a look at the Rmd file.

$$I(\hat{\theta}) \xrightarrow{P} I(\theta)$$

This is called the *expected* Fisher Information and is computed at the MLE:

$$I(\hat{\theta}) = -E_{Y|\hat{\theta}}\{\partial^2 \log L(\theta|Y)/\partial\theta^2\}|_{\theta=\hat{\theta}}$$

That $|_{\theta=\hat{\theta}}$ at the end means that after doing the derivative with respect to θ, we replace θ with $\hat{\theta}$. It would not make sense to do the substitution before since $\hat{\theta}$ is a fixed value and so you cannot take the derivative with respect to it.

This is a viable approach if you can take the derivative of the log-likelihood with respect to θ and can take the expectation over the data. You could always do that expectation using simulation of course. You just need to be able to simulate data from your model with $\hat{\theta}$.