MATH5805: Advanced Time Series Analysis
Session 1, 2014
Assignment 3

Final Report Due: Wednesday May 21 at 5pm in lectures.
PLEASE NOTE: I intend to discuss progress on this assignment during lectures in Weeks 9 and 10 to ensure there are no 'road blocks' to you completing it. Please start work on the assignment in before Week 9 Lecture so that any issues or concerns can be raised and addressed.

This assignment focuses on analysis of a binary time series representing by a value 1 for a win and a value 0 for a loss by Cambridge University in the annual Thames boat race between Cambridge and Oxford Universities. A covariate to be used in some analyses is the weight difference between the winning and losing boat crews.

Resources:

- Data in file called **boatrace.csv**. In this assignment use the complete data set.
- R program to read in data and perform basic plotting and analysis is in **BoatRace_InputDataBasicAnalysis.R**
- Journal article by Klingenberg(2008) giving background and comprehensive analysis of a conditionally independent binary response with Gaussian autoregressive state equation is in file: **Klingenberg2007_CSDA_RegModelsBinaryTimeSeries.pdf**
- Durbin and Koopman §14.6.
- KFAS R-package.
- R-programs used for Topic 8.

**Note: There are missing years (discussed in both reference sources). Our aim is to develop methods that cope with these but if you encounter problems with doing any of the analysis when there are missing data please discuss with me and, as a fall-back, analyse the data by ignoring the missing gaps (treating the series values as contiguous in time).**

**Part 1: Fitting the model proposed by Durbin and Koopman**
Review Section 14.6 of Durbin and Koopman and do the following:

1. Write out the model used in state space form and carefully specify all components (matrices, vectors and noise terms) in the state equation and the observation equation.

2. Summarise their findings.

3. Use the KFAS package to fit this model and produce a graphical display similar to Figure 14.8 of Durbin and Koopman.

4. Compare your results with those reported in Durbin and Koopman.

**Part 2: Fitting the model proposed by Klingenberg** Review Section 4 of Klingenberg (2008) concerning his analysis of the Boat Race series and do the following:

1. Write out Klingenberg's model in state space form and carefully specify all components (matrices, vectors and noise terms) in the state equation and the observation equation.

2. Summarise the method he used to obtain his estimates of the binary response, Gaussian linear state model.

3. Summarise his key findings including point estimates and their standard errors of all parameters in the model.

4. Using the Topic 8 ideas develop the Laplace approximation methodology for this model. Find the following quantities:

   (a) The mode of the conditional distribution of the state given the observations using parameter values reported by Klingenberg.

   (b) Produce profile plots of the Laplace approximate log likelihood $\log L_{\mathrm{LA}}$ versus each parameter over a range around Klingenbergs estimates (similar to Figure 8.9 or 8.13 in Topic 8).

   (c) Use these plots to determine good starting values for a Newton-Raphson iterative update to find the parameters that optimise $\log L_{\mathrm{LA}}$. Summarise how this procedure converged, the approximate estimates $\hat{\psi}_{\mathrm{LA}}$ and their standard errors.

   (d) Compare your results with those obtained by Klingenberg (your item 3) above). In particular you should produce similar graphical displays to Figure 2 of Klingenberg.

**Part 3: Summary of both models, analyses and synthesis of results**
Write a brief summary comparing the models and results in Part 1 and Part 2 and make any recommendations for additional analysis that might be useful.