BSTS Part 1 Script

Slide 1

Hello fellow traveler! I can see by the look on your face that you are ready to know about how Baysean statistics can be applied to time series analysis. Well, great news! That is exactly what you're about to learn on this adventure.

Slide 3

This adventure is divided into 2 stages: theory and practice.

(->) We must first learn the mysteries of structural time series analysis.

(->) Then we will learn about structural time series in the Bayesian framework. Be warned! This will be the most perilous part of your journey. There will be a great deal of math involved, but you cannot truly unravel the secrets of Bayesian structural time series without it.

(->) Once we understand how Bayesian structural time series ticks, we will be ready to roll initiative and take on this beast using the power of R! This will consist of 2 parts. First we see how to implement BSTS models. Then we will harness the full power of BSTS by including regression terms. (-> Slide 4)

Slide 4

But before your adventure begins, we would like to point out some resources. Much of what we will talk about is based on the paper "Predicting the Present with Bayesian Structural Time Series" by Steven L. Scott and Hal Varian.

(->) For more background on structural time series see "An Introduction to State Space Time Series Analysis" by Commandeur and Koopman.

(->) Or, for the truly adventurous, see "Time Series Analysis By State Space Methods" by Durbin and Koopman.

(->) For implementing BSTS, we will be using the *bsts* package in R so the package documentation makes for a good reference.

(->) You may also be interested to know a GitHub repository exists for this tutorial series. There you can find all the presentations, code, and data that we use. A link to this is provided in the description below. (-> Slide 5)

Slide 5

Let's begin our journey with a brief review of time series analysis. We assume that you already have some familiarity with time series. We try to keep things simple but if you have never seen time series before, you might have a difficult time following along. Also note that we assume you are familiar with the basic ideas of Bayesian statistics. If you don't know anything about either of these concepts, you'll have disadvantage and you should probably learn a bit about them before continuing on your quest.

A time series is a stochastic process indexed by time.

(->) That is, a collection of random variables with an index set T that we think of as time. We typically consider discrete time processes with T being either the integers or the natural numbers.

(->) An important concept in time series is stationarity. There are several varieties but we usually mean weakly stationary. This is sometimes called covariance stationary or just stationary.

(->) To be stationary, a time series must have a constant mean.

(->) In addition, the covariance between random variables at each time point only depends on the difference in time.

(->) The main issue in time series is that the data is not i.i.d. like we assume for many statistical methods. Sometimes the identical distribution part holds but the independence part doesn't. Sometimes neither of the i's hold. (-> Slide 6)

Slide 6

The classical time series model is the autoregressive moving averge, or ARMA model. The simplest autoregressive model is the AR(1) which looks like this.

(->) Here e_t is the error term which we assume to be i.i.d. with mean zero and variance σ^2.

(->) And we require the coefficient ϕ to be less than one. (-> Slide 7)

Slide 7

An AR(1) process might look like this. (-> Slide 8)

Slide 8

The simplest moving average model is the MA(1). It looks like this.

(->) Again, e_t is the error term here which we assume to be i.i.d. with mean zero and variance σ^2. (-> Slide 9)

Slide 9

An MA(1) process might look something like this. (-> Slide 10)

Slide 10

The AR(1) and MA(1) models can be extended and combined to form an ARMA model. A general ARMA process is usually written like this.

(->) $\phi(B)$ is the AR part which is a linear combination of previous values of the series X_t.

(->) And $\theta(B)$ is the MA part which is a linear combination of previous error terms.

Well, that's it for the time series refresher. If you haven't encountered time series before and had trouble following along you want to study up a bit and come back. Otherwise...

Slide 11

Glory and reward await those who are brave enough to navigate the trials and tribulations involved in learning BSTS. Come stalwart adventurer! let us explore this dungeon of mysterious treasures. In the following tutorials, we will battle untold horrors and demons together, who guard the secrects of BSTS.