STA 642: Spring 2020 – Homeworks & Assessments

I confirm that all my submissions are my work alone. I have not copied nor adapted the work of others, nor colluded with others on development of my work submitted for assessment. In all respects my work on this course respects and fully abides by the Duke Community Standards.

Name:

Signature:

Date:

STA 642: Spring 2020 – Homework #1 Exercises

Exercises run through basics of AR(1) models and some extensions, to ensure familiarity with key concepts and models. Models/contexts appearing in exercises will be recurrent: they are chosen as a mix of drill, extending class notes and discussion, anticipating material coming along, as building blocks for more elaborate models, and as vehicles for new concepts and methods.

Hand-in solutions to *all* questions, generating understanding and mastery of basic material (as well as good preparation for the mid-term exam), but understand that *not all will be assessed*.

Unless explicitly requested (some questions will explicitly request, most do not ..) do not work derivations of theoretical results from scratch if they relates to known theory. For example, (i) a linear combination of normal random quantities is normal: just quote and define the implied mean and variance, and (ii) a normal prior for a normal, linear regression parameter leads to a normal posterior with mean and variance given by standard formulæ– quote the formulæ, you do not need to reinvent the (quadratic form completion etc etc) wheels! If a question asks "what is the distribution of ... "and the result is a distribution in a known family (e.g., normal, T, etc) then simply state the result; you do not need to derive or write-out density functions.

Hand-in LaTeX drafted solutions to the assessed questions; include whatever numerical summaries and graphs you regard as relevant to support your stories and exploratory conclusions in data analyses, as well as detailed mathematical derivations/solutions.

1. Read and work through the supplementary notes as well as indicated sections of P&W, and explore course code examples. Beyond learning about AR models, this includes revision of basic Bayesian analysis in linear regression– Section 1.2 of Chapter 1 and reference Bayesian posterior distributions arising as laid out there; the Matlab code supports this, and the script for AR(1) model explorations shows use of this in fitting this little model to data (including the SOI time series). Makes sure you are completely on top of the notation– the basic Bayesian regression ideas were covered extensively in prerequisite linear models (and predictive modelling) courses, but notation is always a variable!

   Read ahead into course notes and slides for the coming week(s), and get intimate with relevant sections of the P&W text. Read, digest and anticipate.

2. Become proficient in Matlab. Rerun class examples using class code, explore code scripts and support functions, modify as you like, etc.

3. Consider the stationary model $x_t \leftarrow AR(1|(\phi, v))$ with $s = 1$. Simulate series of length $n = 100$ from the distribution of $x_{1:n}$ *conditional* on an pre-initial value x_0 for the two cases $x_0 = 0$ and $x_0 = 10$. Repeat this a few times with different values of ϕ (large, positive and negative in particular). Describe and interpret the resulting realizations, in comparison to realizations from the stationary joint distribution $p(x_{1:n})$.

4. Figure 1 plots the monthly changes in the US S&P stock market index over 1965 to 2016. The course Schedule page has a little bit of Matlab to read in this (and related) data and create this plot (see link under this HW#1 on the web page). Consider an AR(1) model as a very simple exploratory model– for understanding local dependencies but not for forecasting more

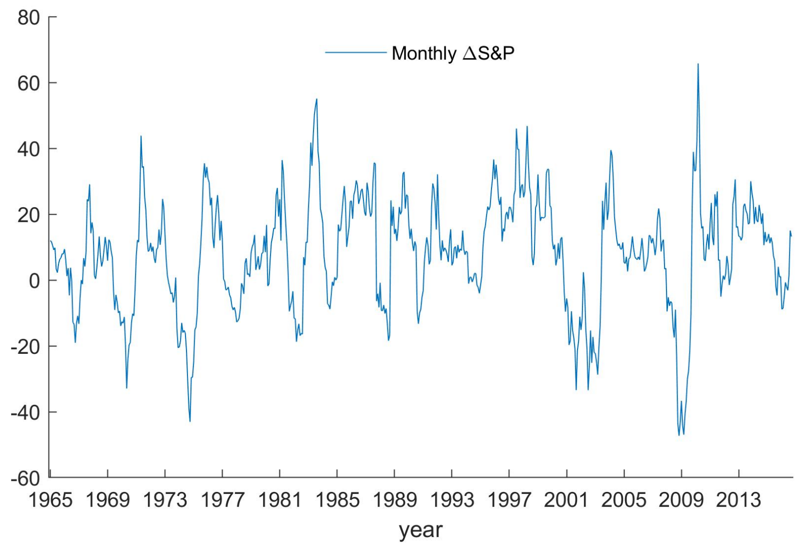

Figure 1: Monthly changes in the US S&P stock market index over 1965–2015

than a month or two ahead. Then, we know there is a great deal of variation across the years in the market economy and that we might expect "change" that an AR(1) model does not capture. To get started on a first, very basic investigation of this, we can simply fit the AR(1) model to shorter sections of the data and examine the resulting inferences on parameters to see if they seem to vary across time. Do this as follows. The full series has $T = 621$ months of data; look at many separate time series by selecting a month m and taking some number k months either side; for example, you might take $k = 84$ and for any month m analyse the data over the "windowed period" from $m - k$ to $m + k$ inclusive. Repeat this for each month m running from $m = k + 1$ to $m = T - k$. These repeat analyses will define a "trajectory" of AR(1) analyses over time, one for each sub-series.

For each sub-series, subtract the sub-series mean and then compute the summaries of the reference posterior for an AR(1) model to just that $2k + 1$ time points– just treating each selected sub-series separately. Using the theoretical posterior T distribution for the ϕ parameter, compute and compare (graphically) the exact posterior 90% credible intervals

(a) Comment on what you see in the plot and comparison, and what you might conclude in terms of changes over time.
(b) Do you believe that short-term changes in S&P have shown real changes in month-month dependencies since 1965?
(c) How would you suggest also addressing the question of whether or not the underlying mean of the series is stable over time?
(d) What about the innovations variance?
(e) What does this suggest for more general models that might do a better job of imitating this data?

5. Sequential Bayesian learning and $1-$step ahead forecasting in AR(1) models. This question starts an investigation of the details of how posterior distributions for (ϕ, v) are sequentially updated as new data arises, forming a prelude to filtering in state-space models generally as well as for extension to time-varying parameter AR models.

   Suppose $x_t \leftarrow AR(1|(\phi, v))$ with (ϕ, v) uncertain. At any time t write $\mathcal{D}_t$ for the past data and information, including all past observations. If no additional information arises over the time interval $(t-1, t]$, then $\mathcal{D}_t$ sequentially updates as the new observation is made via– simply– $\mathcal{D}_t = \{\mathcal{D}_{t-1}, x_t\}$.

   Now suppose you are standing at the end of time interval $t-1$ so that you have current information set $\mathcal{D}_{t-1}$. Suppose the current posterior for $\theta = (\phi, v)$ based on this information has a conjugate normal-inverse gamma form written as

$$(\phi|v, \mathcal{D}_{t-1}) \sim N(m_{t-1}, C_{t-1}(v/s_{t-1})), \tag{1}$$

$$(v^{-1}|\mathcal{D}_{t-1}) \sim Ga(n_{t-1}/2, n_{t-1}s_{t-1}/2) \tag{2}$$

   with known defining parameters. This would be the case, for example, of a reference posterior based on the first $t-1$ observations. Here m_{t-1} and $s_{t-1} > 0$ are natural point estimates of ϕ and v respectively, while $C_{t-1} > 0$ and $n_{t-1} > 0$ relate to uncertainty.

   (a) What is the current marginal posterior for ϕ, namely $p(\phi|\mathcal{D}_{t-1})$?

   (b) Show that, conditional on v and marginalizing over ϕ, the implied $1-$*step ahead forecast distribution* for x_t given v is

$$(x_t|v, \mathcal{D}_{t-1}) \sim N(f_t, q_t v/s_{t-1})$$

   with $f_t = m_{t-1}x_{t-1}$ and $q_t = s_{t-1} + C_{t-1}x_{t-1}^2$.

   (c) Now marginalize also over v to find the implied $1-$*step ahead forecast distribution* for x_t, namely $p(x_t|\mathcal{D}_{t-1})$, i.e., the distribution you will use in practice to predict x_t $1-$step ahead. What is this distribution?

   (d) Now move to time t and observe the outcome x_t. Show that the *time t posterior* $p(\phi, v|\mathcal{D}_t)$ is also normal-inverse gamma, having the same form as in at time $t-1$ above but now with $t-1$ updated to t and updated defining parameters $\{m_t, C_t, n_t, s_t\}$ that can be written in the following forms:

   - $m_t = m_{t-1} + A_t e_t$,
   - $C_t = r_t(C_{t-1} - A_t^2 q_t)$,
   - $n_t = n_{t-1} + 1$,
   - $s_t = r_t s_{t-1}$ with $r_t = (n_{t-1} + e_t^2/q_t)/n_t$,

   where

   - $e_t = x_t - f_t$ is the realized $1-$*step ahead (point) forecast error*, and
   - $A_t = C_{t-1}x_{t-1}/q_t$ is called the *adaptive coefficient*.

   (e) Comment on these expressions, giving particular attention to the following:

      i. How (m_t, C_t) depend on the new data x_t relative to the prior values (m_{t-1}, C_{t-1}).

ii. The role of the adaptive coefficient in the update $(m_{t-1}, C_{t-1}) \longrightarrow (m_t, C_t)$.

iii. The updates for the degrees of freedom n_t and point estimate s_t and how they depend on x_t.

(f) Consider an example in which the forecast error is very large relative to expectation, resulting in a value of e_t^2/q_t much greater than 1. Comment on how the posterior for (ϕ, v) responds.