STA 642: Homework # 5 Exercises

1. **Course mini-project proposal and development**. Continue progressing on this. Talk/email to MW and TAs as convenient. The next homework will be based on (only) the summary proposal for your project.

2. **TVAR models.** Next week we will start looking at another class of DLMs– revisiting AR models and then their key practical and widely-used extensions to time-varying autogressive (TVAR) models. Read P&W Section 5.1 (skip over 5.1.3, or at least take it "very lightly" for now) and 5.2; and review the detailed course slides on TVAR models linked to the web page.

   There is code and examples on TVAR models in the course code base. Example code/scripts explore TVAR analyses of the quarterly changes in US inflation series and a subset of the EEG series from class discussion/slides.

3. **A start in stochastic volatility (SV) modelling.** Before we get into TVAR models, we will first look a bit more about the recurrent question of observational variances in DLMs that may/appear to change over time. We will focus in detail on the venerable and widely-used discount *stochastic volatility model* detailed in P&W, Section 4.3.7, and in (very) detailed course slides on this SV model linked to the web page. Read and explore that in preparation, and in helping with Question 4 in this homework.

   Some/most of you will likely want to integrate SV into projects later in semester, and the basic ideas and methodology in the univariate model underlie a main initial class of multivariate volatility models– i.e., time-varying variance matrices– in one of our first multivariate DLM settings coming along (P&W chapter 10). This basic model is also useful as a first model for time-varying rates in time series models for integer counts (e.g., for flows in networks ....); again, we will see examples later on.

4. The basic distribution theory in this question underlies the venerable and widely-used discount volatility model (P&W, Section 4.3.7). We will review the basic model, ideas and results in class; here, you will visit and work through some basic theory in advance to develop conceptual and technical understanding. We will also build on this later, for other kinds of models as noted above.

   The theory in this exercise concerns aspects of a bivariate distribution for two positive scalars ϕ_0 and ϕ_1. Mapping to the SV model in DLMs is made by noting the same setup arises there with, at any times $t-1, t,$ the match $\phi_0 \leftarrow v_{t-1}^{-1}$ and $\phi_1 \leftarrow v_t^{-1}$, and then the bivariate distribution relates to $p(v_{t-1}, v_t|\mathcal{D}_t)$.

   Two positive scalar random quantities ϕ_0 and ϕ_1 have a joint distribution defined by:

   - the margin $p(\phi_0)$ given by $\phi_0 \sim Ga(a, b)$ for some scalars $a > 0, b > 0$; and
   - the conditional $p(\phi_1|\phi_0)$ that is implicitly defined by

   $$\phi_1 = \phi_0\eta/\beta, \quad \text{where} \quad \eta \sim Be(\beta a, (1-\beta)a) \quad \text{and} \quad \eta \perp\!\!\!\perp \phi_0,$$

   and where $\beta \in (0, 1)$ is a known, constant discount factor.

(a) What is $E(\phi_1|\phi_0)$?

Directly, $E(\phi_1|\phi_0) = \phi_0 E(\eta)/\beta = \phi_0$ since $E(\eta) = \beta a/[\beta a + (1-\beta)a] = \beta$.

(b) What are $E(\phi_0)$ and $E(\phi_1)$?

$E(\phi_0) = a/b$ and so $E(\phi_1) = E[E(\phi_1|\phi_0)] = E(\phi_0) = a/b$ as well.

(c) Starting with the joint density $p(\phi_0)p(\eta)$ (a product form since ϕ_0 and η are independent), make the bivariate transformation to (ϕ_0, ϕ_1) and show that

$$p(\phi_0, \phi_1) = c\,\mathrm{e}^{-b\phi_0}\,\phi_1^{\beta a-1}\,(\phi_0 - \beta\phi_1)^{(1-\beta)a-1}\,, \quad \text{on } 0 < \phi_1 < \phi_0/\beta,$$

being zero otherwise. Here c is a normalizing constant that does not depend on the conditioning value of ϕ_0 (and we do not care about the value of c for the derivations here).

Since $\eta = \beta\phi_1/\phi_0$, the transformation from (ϕ_0, η) to (ϕ_0, ϕ_1) has reverse Jacobian β/ϕ_0. Further, we know $0 < \eta < 1$ so the transformed range is immediately $0 < \phi_1 < \phi_0/\beta$. Hence, substituting $\eta = \beta\phi_1/\phi_0$ in the $Be(\beta a, (1-\beta)a)$ p.d.f. and multiplying by the Jacobian– and ignoring some constants but being careful to account for all terms involving ϕ_0, ϕ_1– we have

$$p(\phi_1, \phi_0) \propto \frac{\beta}{\phi_0}\,\{\phi_0^{a-1}\mathrm{e}^{-b\phi_0}\}\left(\frac{\beta\phi_1}{\phi_0}\right)^{\beta a-1}\left(1 - \frac{\beta\phi_1}{\phi_0}\right)^{(1-\beta)a-1}\,, \quad \text{on } 0 < \phi_1 < \phi_0/\beta,$$

which, after simplifying terms, yields

$$p(\phi_1, \phi_0) \propto \mathrm{e}^{-b\phi_0}\,\phi_1^{\beta a-1}\,(\phi_0 - \beta\phi_1)^{(1-\beta)a-1}, \quad \text{on } 0 < \phi_1 < \phi_0/\beta,$$

as stated.

(d) Derive the p.d.f. $p(\phi_1)$ (up to a proportionality constant). Deduce that the marginal distribution of ϕ_1 is $\phi_1 \sim Ga(\beta a, \beta b)$.

Directly,

$$p(\phi_1) = \int p(\phi_0, \phi_1) d\phi_0 \propto \phi_1^{\beta a-1}\int_{\beta\phi_1}^{\infty} \mathrm{e}^{-b\phi_0}(\phi_0 - \beta\phi_1)^{(1-\beta)a-1} d\phi_0. \tag{1}$$

We can now use $\gamma = \phi_0 - \beta\phi_1$ to transform from ϕ_0 to γ in the integrand, giving

$$\begin{aligned} p(\phi_1) &\propto \phi_1^{\beta a-1}\mathrm{e}^{-\beta b\phi_1}\int_0^{\infty} \mathrm{e}^{-b\gamma}\,\gamma^{(1-\beta)a-1} d\gamma \\ &\propto \phi_1^{\beta a-1}\mathrm{e}^{-\beta b\phi_1} \end{aligned}$$

as the integral over γ is that of $\gamma \sim Ga((1-\beta)a, b)$ so depends only on a, b, β.

Hence the margin for ϕ_1 is $\phi_1 \sim Ga(\beta a, \beta b)$.

(e) Using the technical details of your derivations above (and without much more work), show that the reverse conditional $p(\phi_0|\phi_1)$ is implicitly defined by

$$\phi_0 = \beta\phi_1 + \gamma \quad \text{where} \quad \gamma \sim Ga((1-\beta)a, b) \quad \text{with } \gamma \perp\!\!\!\perp \phi_1.$$

This is implicit in the integrand of eqn. (1) above in the derivation of $p(\phi_1)$. Since the integrand is proportional to $p(\phi_1, \phi_0)$, then

$$p(\phi_0|\phi_1) \propto p(\phi_1, \phi_0) \propto (\phi_0 - \beta\phi_1)^{(1-\beta)a-1}\, e^{-b\phi_0}, \quad \text{on } \phi_0 > \beta\phi_1.$$

Transform the random quantity ϕ_0 to $\gamma = \phi_0 - \beta\phi_1$; the Jacobian is 1 and we see that the resulting $p(\gamma|\phi_1)$ is that of $\gamma \sim Ga((1-\beta)a, b)$ independently of ϕ_1. The result follows.

5. P&W Chapter 4, Section 4.6: Problem 3.

Use the results derived in Question 4 above to answer this. (Do not redevelop technical results already shown.)

Matching notations $\phi_0 \leftarrow v_{t-1}^{-1}$, $\phi_1 \leftarrow v_t^{-1}$, $\eta \leftarrow \gamma_t$, and matching parameters $a \leftarrow n_{t-1}/2, b \leftarrow d_{t-1}/2 = n_{t-1}s_{t-1}/2,$ means that the setup of this question is precisely that of the discount volatility model of eqn. (4.17) in P&W. Our result in part 4d above then proves the evolution theory of eqn. (4.18) in P&W.

6. P&W Chapter 4, Section 4.6: Problem 4.

Again, use the results derived in Question 4 above to answer this. (Do not redevelop technical results already shown.)

(a) With the matches of notation made in the previous question, and additionally matching $\gamma \leftarrow \nu_{t-1}^*$, the $1-$step back recursion in the discount volatility model follows from our result in part 4e above.

(b) Use the notation for precisions given by $\phi_t = v_t^{-1}$. The first-order Markov evolution of the v_t, hence the ϕ_t, means that, conditional on ϕ_t, ϕ_{t-1} is conditionally independent of all future ϕ_r as well as data at times t onward. The result follows immediately. (This is precisely the same use of conditional independence in first-order Markov models that we have already used in developing backward recursions for retrospective smoothing and sampling of states in DLMs.)

(c) Starting with $E(\phi_T|\mathcal{D}_T) = s_T^{-1}$, recurse back over time $t = T, T-1, \ldots,$ at each time evaluating expectations in the defining equation $\phi_{t-1} = \beta\phi_t + (1-\beta)\nu_{t-1}^*$. Immediately, this gives the backward smoothing recursion

$$E(\phi_{t-1}|\mathcal{D}_T) = E\{E(\phi_{t-1}|\phi_t, \mathcal{D}_{t-1})|\mathcal{D}_T\} = \beta E(\phi_t|\mathcal{D}_T) + (1-\beta)s_{t-1}^{-1}.$$

(d) Starting with a sample $v_T = \phi_T^{-1}$ from the inverse gamma posterior $p(v_T|\mathcal{D}_T)$, recurse back over time $t = T, T-1, \ldots,$ at each time $t-1$ sampling $v_{t-1} = \phi_{t-1}^{-1}$ via $v_{t-1}^{-1} = \beta v_t^{-1} + \nu_{t-1}^*$ where $\nu_{t-1}^* \sim Ga((1-\beta)n_{t-1}/2, n_{t-1}s_{t-1}/2)$ independently.