

TIME SERIES HW 1L

ROBERT RICHARDSON

2. Consider the AR(1) model $y_t = \phi y_{t-1} + \epsilon_t$
- (a) Find the MLE of (θ, ν) for the conditional likelihood.
The likelihood is

$$p(y_{1:n}|\theta) = \prod_{i=2}^n p(y_i|y_{i-1}, \theta)$$

where $y_t|y_{t-1} \sim N(\phi y_{t-1}, v)$
The log likelihood is

$$\ell = -\frac{n-1}{2} \log 2\pi v - \frac{1}{2v} \sum_{t=2}^n (y_t - \phi y_{t-1})^2$$

Find the maximum likelihood values by taking derivatives with respect to each parameter individually and set equal to 0.

$$\begin{aligned} \frac{\partial \ell}{\partial \phi} &= -\frac{1}{v} \sum_{t=2}^n (y_t - \phi y_{t-1})(-y_{t-1}) \\ &= -\frac{1}{v} \left(-\sum_{t=2}^n y_t y_{t-1} + \phi \sum_{t=2}^n y_{t-1} y_{t-1} \right) \\ 0 &= \sum_{t=2}^n y_t y_{t-1} - \hat{\phi} \sum_{t=2}^n y_{t-1} y_{t-1} \\ \hat{\phi} &= \frac{\sum_{t=2}^n y_t y_{t-1}}{\sum_{t=2}^n y_{t-1}^2} \\ \frac{\partial \ell}{\partial v} &= -\frac{n-1}{2v} + \frac{1}{2v^2} \sum_{t=2}^n (y_t - \phi y_{t-1})^2 \\ \hat{v} &= \frac{1}{n-1} \sum_{t=2}^n (y_t - \hat{\phi} y_{t-1})^2 \end{aligned}$$

Check second order conditions by taking the second derivative and showing that the MLE will give the second derivative a negative value.

- (b) Find the MLE of (θ, ν) for the unconditional likelihood.
The unconditional likelihood is given in the text as

$$p(y_{1:n}|\theta) = \frac{(1 - \phi^2)^{1/2}}{(2\pi v)^{n/2}} \exp \left\{ -\frac{Q^*(\phi)}{2v} \right\}$$

where

$$Q^*(\phi) = y_1^2(1 - \phi^2) + \sum_{t=2}^n (y_t - \phi y_{t-1})^2$$

The log likelihood is then

$$\ell = \frac{1}{2} \log(1 - \phi^2) - \frac{n}{2} \log(2\pi v) - \frac{1}{2v} y_1^2 (1 - \phi^2) - \frac{1}{2v} \sum_{t=2}^n (y_t - \phi y_{t-1})^2$$

The maximum likelihood values are found by setting the respective derivatives equal to 0.

$$\begin{aligned} \frac{\partial \ell}{\partial \phi} &= -\frac{2\phi}{2(1 - \phi^2)} + \frac{1}{v} y_1^2 \phi - \frac{1}{v} \sum_{t=2}^n (y_t - \phi y_{t-1})(-y_{t-1}) \\ 0 &= -\phi v + y_1^2 \phi (1 - \phi^2) + (1 - \phi^2) \sum_{t=2}^n y_t y_{t-1} - \phi (1 - \phi^2) \sum_{t=2}^n y_{t-1} y_{t-1} \\ &= \hat{\phi}^3 \left(\sum_{t=2}^n y_{t-1} y_{t-1} - y_1^2 \right) - \hat{\phi}^2 \sum_{t=2}^n y_t y_{t-1} + \hat{\phi} (y_1^2 - \sum_{t=2}^n y_{t-1} y_{t-1} - \hat{v}) + \sum_{t=2}^n y_t y_{t-1} \\ \frac{\partial \ell}{\partial v} &= -\frac{n}{2v} + \frac{1}{v^2} \left(y_1^2 (1 - \phi^2) + \sum_{t=2}^n (y_t - \phi y_{t-1})^2 \right) \\ \hat{v} &= \frac{1}{n} \left(y_1^2 (1 - \hat{\phi}^2) + \sum_{t=2}^n (y_t - \hat{\phi} y_{t-1})^2 \right) \end{aligned}$$

This gives a set of 2 equations which can be solved using E.M. or Newton-Rhapson methods to solve. Second order conditions will hold.

- (c) Assuming v is known, find the MAP for ϕ using a uniform prior for both the conditional and unconditional likelihood.

(1) Conditional: The posterior is

$$\begin{aligned} p(\phi|v, y_{1:n}) &\propto p(\phi) p(y_{1:n}|\phi, v) \\ &\propto p(y_{1:n}|\phi, v) \end{aligned}$$

Thus the MAP estimate is the same as the MLE estimate.

$$\hat{\phi}_{MAP} = \frac{\sum_{t=2}^n y_t y_{t-1}}{\sum_{t=2}^n y_{t-1}^2}$$

- (2) Using the same exact logic, the MAP estimate will be equal to the MLE estimate for the unconditional likelihood. It will be the solution of

$$\hat{\phi}_{MAP}^3 \left(\sum_{t=2}^n y_{t-1} y_{t-1} - y_1^2 \right) - \hat{\phi}_{MAP}^2 \sum_{t=2}^n y_t y_{t-1} + \hat{\phi}_{MAP} (y_1^2 - \sum_{t=2}^n y_{t-1} y_{t-1} - \hat{v}) + \sum_{t=2}^n y_t y_{t-1} = 0$$

There is possibly an issue to see if the MAP is in the range $(-1, 1)$. If it is not, the MAP is the endpoint which yields the maximum value of the posterior.

3. Find the posterior for ϕ and v in the reference analysis using the conditional likelihood. The posterior distribution is

$$\begin{aligned} p(\phi, v|y_{1:n}) &\propto p(y_{1:n}|\phi, v) p(\phi, v) \\ &\propto \frac{1}{v} \prod_{t=2}^n p(y_t|y_{t-1}, \theta) \\ &\propto v^{-(n-1)/2-1} \exp \left\{ -\frac{1}{2v} \sum_{t=2}^n (y_t - \phi y_{t-1})^2 \right\} \end{aligned}$$

To find the marginals, I integrate out the parameters individually. Note that in terms of v , the joint is an inverse gamma, so

$$\begin{aligned}
p(\phi|y_{1:n}) &= \int p(\phi, v|y_{1:n}) dv \\
&\propto \frac{\Gamma((n-1)/2)}{(\frac{1}{2} \sum_{t=2}^n (y_t - \phi y_{t-1})^2)^{(n-1)/2}} \int \frac{(\frac{1}{2} \sum_{t=2}^n (y_t - \phi y_{t-1})^2)^{(n-1)/2}}{\Gamma((n-1)/2)} v^{(n-1)/2-1} \exp \left\{ \frac{1}{2v} \sum_{t=2}^n (y_t - \phi y_{t-1})^2 \right\} \\
&= \frac{\Gamma((n-1)/2)}{(\frac{1}{2} \sum_{t=2}^n (y_t - \phi y_{t-1})^2)^{(n-1)/2}} \\
&\propto \left(\sum_{t=2}^n y_t^2 - \phi \sum_{t=2}^n y_t y_{t-1} + \phi^2 \sum_{t=2}^n y_{t-1}^2 \right)^{-(n-1)/2} \\
&\propto \left(\sum_{t=2}^n y_t^2 - \frac{(\sum_{t=2}^n y_t y_{t-1})^2}{\sum_{t=2}^n y_{t-1}^2} + \sum_{t=2}^n y_{t-1}^2 \left(\phi - \frac{\sum_{t=2}^n y_t y_{t-1}}{\sum_{t=2}^n y_{t-1}^2} \right)^2 \right)^{-(n-1)/2} \\
&\propto \left(1 + \frac{(n-2) \sum_{t=2}^n y_{t-1}^2}{\sum_{t=2}^n y_t^2 - \frac{(\sum_{t=2}^n y_t y_{t-1})^2}{\sum_{t=2}^n y_{t-1}^2}} \frac{1}{n-2} \left(\phi - \frac{\sum_{t=2}^n y_t y_{t-1}}{\sum_{t=2}^n y_{t-1}^2} \right)^2 \right)^{-(n-1)/2}
\end{aligned}$$

This is a kernel for a t distribution with $n-2$ degrees of freedom, a mean of $\frac{\sum_{t=2}^n y_t y_{t-1}}{\sum_{t=2}^n y_{t-1}^2}$ and a variance

$$\text{of } \frac{\sum_{t=2}^n y_t^2 - \frac{(\sum_{t=2}^n y_t y_{t-1})^2}{\sum_{t=2}^n y_{t-1}^2}}{(n-2) \sum_{t=2}^n y_{t-1}^2}$$

The posterior for v is found similarly

$$\begin{aligned}
p(v|y_{1:n}) &= \int p(\phi, v|y_{1:n}) d\phi \\
&\propto \int v^{-(n-1)/2-1} \exp \left\{ -\frac{1}{2v} \sum_{t=2}^n (y_t - \phi y_{t-1})^2 \right\} d\phi \\
&\propto \int v^{-(n-1)/2-1} \exp \left\{ -\frac{1}{2v} \left(\sum_{t=2}^n y_t^2 - \phi \sum_{t=2}^n y_t y_{t-1} + \phi^2 \sum_{t=2}^n y_{t-1}^2 \right) \right\} d\phi \\
&\propto \int v^{-(n-1)/2-1} \exp \left\{ -\frac{1}{2v} \left(\sum_{t=2}^n y_t^2 - \frac{(\sum_{t=2}^n y_t y_{t-1})^2}{\sum_{t=2}^n y_{t-1}^2} + \sum_{t=2}^n y_{t-1}^2 \left(\phi - \frac{\sum_{t=2}^n y_t y_{t-1}}{\sum_{t=2}^n y_{t-1}^2} \right)^2 \right) \right\} d\phi \\
&\propto v^{-(n-2)/2-1} \exp \left\{ -\frac{1}{2v} \left(\sum_{t=2}^n y_t^2 - \frac{(\sum_{t=2}^n y_t y_{t-1})^2}{\sum_{t=2}^n y_{t-1}^2} \right) \right\} \int v^{-1/2} \exp \left\{ -\frac{1}{2v} \sum_{t=2}^n y_{t-1}^2 \left(\phi - \frac{\sum_{t=2}^n y_t y_{t-1}}{\sum_{t=2}^n y_{t-1}^2} \right)^2 \right\} d\phi \\
&\propto v^{-(n-2)/2-1} \exp \left\{ -\frac{1}{2v} \left(\sum_{t=2}^n y_t^2 - \frac{(\sum_{t=2}^n y_t y_{t-1})^2}{\sum_{t=2}^n y_{t-1}^2} \right) \right\}
\end{aligned}$$

This is an inverse gamma distribution with parameters $\frac{n-2}{2}$ and $\frac{1}{2} \left(\sum_{t=2}^n y_t^2 - \frac{(\sum_{t=2}^n y_t y_{t-1})^2}{\sum_{t=2}^n y_{t-1}^2} \right)$

4. Confirm the posterior for the conditional likelihood assuming conjugate priors. Use priors $\phi \sim N(0, v)$ and $v \sim IG(n_0/2, d_0/2)$

The posterior distribution is

$$\begin{aligned}
p(\phi, v|y_{1:n}) &\propto p(y_{1:n}|\phi, v)p(\phi, v) \\
&\propto v^{-(n_0-2)/2} \exp\left\{-\frac{d_0}{2v}\right\} v^{-1/2} \exp\left\{-\frac{\phi^2}{2v}\right\} \prod_{t=2}^n p(y_t|y_{t-1}, \theta) \\
&\propto v^{-(n+n_0)/2-1} \exp\left\{-\frac{d_0}{2v} - \frac{\phi^2}{2v} - \frac{1}{2v} \sum_{t=2}^n (y_t - \phi y_{t-1})^2\right\}
\end{aligned}$$

We integrate out the parameters to find the marginals

$$\begin{aligned}
p(\phi|y_{1:n}) &= \int p(\phi, v|y_{1:n}) dv \\
&\propto \frac{\Gamma((n+n_0)/2)}{\frac{1}{2} (d_0 + \phi^2 + \sum_{t=2}^n (y_t - \phi y_{t-1})^2)^{(n+n_0)/2}} \\
&\times \int \frac{\frac{1}{2v} (d_0 + \phi^2 + \sum_{t=2}^n (y_t - \phi y_{t-1})^2)^{(n+n_0)/2}}{\Gamma((n+n_0)/2)} v^{-(n+n_0)/2-1} \exp\left\{-\frac{1}{2v} \left(d_0 + \phi^2 + \sum_{t=2}^n (y_t - \phi y_{t-1})^2\right)\right\} \\
&= \frac{\Gamma((n+n_0)/2)}{\left(\frac{1}{2} (d_0 + \phi^2 + \sum_{t=2}^n (y_t - \phi y_{t-1})^2)\right)^{(n+n_0)/2}} \\
&\propto \left(d_0 + \sum_{t=2}^n y_t^2 - \phi \sum_{t=2}^n y_t y_{t-1} + \phi^2 \left(\sum_{t=2}^n y_{t-1}^2 + 1\right)\right)^{-(n+n_0)/2} \\
&\propto \left(d_0 + \sum_{t=2}^n y_t^2 - \frac{(\sum_{t=2}^n y_t y_{t-1})^2}{\sum_{t=2}^n y_{t-1}^2 + 1} + \left(\sum_{t=2}^n y_{t-1}^2 + 1\right) \left(\phi - \frac{\sum_{t=2}^n y_t y_{t-1}}{\sum_{t=2}^n y_{t-1}^2 + 1}\right)^2\right)^{-(n+n_0)/2} \\
&\propto \left(1 + \frac{(n+n_0-1) \left(1 + \sum_{t=2}^n y_{t-1}^2\right)}{d_0 + \sum_{t=2}^n y_t^2 - \frac{(\sum_{t=2}^n y_t y_{t-1})^2}{1 + \sum_{t=2}^n y_{t-1}^2}} \frac{1}{n+n_0-1} \left(\phi - \frac{\sum_{t=2}^n y_t y_{t-1}}{1 + \sum_{t=2}^n y_{t-1}^2}\right)^2\right)^{-(n+n_0)/2}
\end{aligned}$$

This is a t-distribution with $n + n_0 - 1$ degrees of freedom. The mean is $\frac{\sum_{t=2}^n y_t y_{t-1}}{1 + \sum_{t=2}^n y_{t-1}^2}$ and the variance term is $\frac{d_0 + \sum_{t=2}^n y_t^2 - \frac{(\sum_{t=2}^n y_t y_{t-1})^2}{1 + \sum_{t=2}^n y_{t-1}^2}}{(n+n_0-1)(1 + \sum_{t=2}^n y_{t-1}^2)}$. Note that this disagrees slightly with the general formula in the book, but only because in the book, n is the number of rows of the design matrix F whereas here, n is the length of $y_{1:n}$, but we delete one from the response because we are using the conditional likelihood.

$$\begin{aligned}
p(v|y_{1:n}) &= \int p(\phi, v|y_{1:n}) d\phi \\
&\propto \int v^{-(n+n_0)/2-1} \exp \left\{ -\frac{d_0}{2v} - \frac{\phi^2}{2v} - \frac{1}{2v} \sum_{t=2}^n (y_t - \phi y_{t-1})^2 \right\} d\phi \\
&\propto \int v^{-(n+n_0)/2-1} \exp \left\{ -\frac{1}{2v} \left(d_0 + \phi^2 + \sum_{t=2}^n y_t^2 - \phi \sum_{t=2}^n y_t y_{t-1} + \phi^2 \sum_{t=2}^n y_{t-1}^2 \right) \right\} d\phi \\
&\propto \int v^{-(n+n_0)/2-1} \exp \left\{ -\frac{1}{2v} \left(d_0 + \sum_{t=2}^n y_t^2 - \frac{(\sum_{t=2}^n y_t y_{t-1})^2}{1 + \sum_{t=2}^n y_{t-1}^2} + \left(1 + \sum_{t=2}^n y_{t-1}^2 \right) \left(\phi - \frac{\sum_{t=2}^n y_t y_{t-1}}{1 + \sum_{t=2}^n y_{t-1}^2} \right)^2 \right) \right\} d\phi \\
&\propto v^{-(n+n_0-1)/2-1} \exp \left\{ -\frac{1}{2v} \left(d_0 + \sum_{t=2}^n y_t^2 - \frac{(\sum_{t=2}^n y_t y_{t-1})^2}{1 + \sum_{t=2}^n y_{t-1}^2} \right) \right\} \int v^{-1/2} \exp \left\{ -\frac{1}{2v} \left(1 + \sum_{t=2}^n y_{t-1}^2 \right) \left(\phi - \frac{\sum_{t=2}^n y_t y_{t-1}}{1 + \sum_{t=2}^n y_{t-1}^2} \right)^2 \right\} d\phi \\
&\propto v^{-(n+n_0-1)/2-1} \exp \left\{ -\frac{1}{2v} \left(d_0 + \sum_{t=2}^n y_t^2 - \frac{(\sum_{t=2}^n y_t y_{t-1})^2}{1 + \sum_{t=2}^n y_{t-1}^2} \right) \right\}
\end{aligned}$$

This is the kernel for an inverse gamma with parameters $(n+n_0-1)/2$ and $\frac{1}{2} \left(d_0 + \sum_{t=2}^n y_t^2 - \frac{(\sum_{t=2}^n y_t y_{t-1})^2}{1 + \sum_{t=2}^n y_{t-1}^2} \right)$

5. Comment that for each of these parts, it is much easier to use a linear model design matrix to find MLE and MAP estimates. For the first equation, this looks like

$$\begin{aligned}
\mathbf{Y} &= \mathbf{F}'\beta + \epsilon \\
\begin{pmatrix} y_3 \\ y_4 \\ \vdots \\ y_n \end{pmatrix} &= \begin{pmatrix} y_2 & y_1 \\ y_3 & y_2 \\ \vdots & \vdots \\ y_{n-1} & y_{n-2} \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} + \begin{pmatrix} \epsilon_3 \\ \epsilon_4 \\ \vdots \\ \epsilon_n \end{pmatrix}
\end{aligned}$$

For the second, this is

$$\begin{aligned}
\mathbf{Y} &= \mathbf{F}'\beta + \epsilon \\
\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} &= \begin{pmatrix} \sin(2\pi\omega_0 \cdot 1) & \cos(2\pi\omega_0 \cdot 1) \\ \sin(2\pi\omega_0 \cdot 2) & \cos(2\pi\omega_0 \cdot 2) \\ \vdots & \vdots \\ \sin(2\pi\omega_0 \cdot n) & \cos(2\pi\omega_0 \cdot n) \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} \epsilon_3 \\ \epsilon_4 \\ \vdots \\ \epsilon_n \end{pmatrix}
\end{aligned}$$

Using these design matrices, the MLE and MAP estimates for the coefficients is the OLS estimate $\hat{\beta} = (\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'\mathbf{Y}$. To show this, write the likelihood as

$$p(y_{1:n}|\mathbf{F}, \beta, v) \propto v^{-(n-2)/2} \exp \left\{ -\frac{1}{2v} \left((\beta - \hat{\beta})'\mathbf{F}\mathbf{F}'(\beta - \hat{\beta}) + (\mathbf{y} - \mathbf{F}'\beta)'(\mathbf{y} - \mathbf{F}'\beta) \right) \right\}$$

This likelihood is maximized with respect to β when $\beta = \hat{\beta}$. The posterior with a reference prior is

$$p(\beta, v|y_{1:n}\mathbf{F}) \propto v^{-(n-2)/2-1} \exp \left\{ -\frac{1}{2v} \left((\beta - \hat{\beta})'\mathbf{F}\mathbf{F}'(\beta - \hat{\beta}) + (\mathbf{y} - \mathbf{F}'\beta)'(\mathbf{y} - \mathbf{F}'\beta) \right) \right\}$$

Again, with respect to β , the OLS estimate is the MAP. To find the MLE for v , the full story with logs and derivatives is a little more necessary.

$$\begin{aligned}
\ell &= \frac{-(n-2)}{2} \log(v) - \frac{1}{2v} \left((\beta - \hat{\beta})' \mathbf{F} \mathbf{F}' (\beta - \hat{\beta}) + (\mathbf{y} - \mathbf{F}' \beta)' (\mathbf{y} - \mathbf{F}' \beta) \right) \\
\frac{\partial \ell}{\partial v} &= \frac{-(n-2)}{2v} + \frac{1}{2v^2} \left((\beta - \hat{\beta})' \mathbf{F} \mathbf{F}' (\beta - \hat{\beta}) + (\mathbf{y} - \mathbf{F}' \beta)' (\mathbf{y} - \mathbf{F}' \beta) \right) \\
\hat{v} &= \frac{1}{n-2} \left((\beta - \hat{\beta})' \mathbf{F} \mathbf{F}' (\beta - \hat{\beta}) + (\mathbf{y} - \mathbf{F}' \beta)' (\mathbf{y} - \mathbf{F}' \beta) \right)
\end{aligned}$$

The MAP of the posterior of a reference prior is found similarly to be

$$\hat{v}_{MAP} = \frac{1}{n} \left((\beta - \hat{\beta})' \mathbf{F} \mathbf{F}' (\beta - \hat{\beta}) + (\mathbf{y} - \mathbf{F}' \beta)' (\mathbf{y} - \mathbf{F}' \beta) \right)$$

Exact answers for this will vary greatly, due to simulation variance.

7. The posterior distribution for $p(\phi_1, \phi_2, v, \theta | y_{1:n})$ is

$$p(\phi_1, \phi_2, v, \theta | y_{1:n}) \propto \prod_{t: y_{t-1} > \theta} p(y_t | y_{t-1}, \phi_1, v) \prod_{t: y_{t-1} < \theta} p(y_t | y_{t-1}, \phi_2, v) p(\phi_1 | c) p(\phi_2 | c) p(v | \alpha_0, \beta_0) p(\theta | a)$$

Inference for the posterior can be done using Metropolis-Hastings for θ , but conditionals for the other parameters can be found given θ , so a gibbs step may be used.

$$\begin{aligned}
p(\phi_1 | y_{1:n}, \theta, v) &\propto \prod_{t: y_{t-1} > \theta} p(y_t | y_{t-1}, \phi_1, v) p(\phi_1 | c) \\
&\propto \exp \left\{ -\frac{1}{2v} \sum_{t: y_{t-1} > \theta} (y_t - \phi_1 y_{t-1})^2 \right\} \exp \left\{ \frac{1}{2c} \phi^2 \right\} \\
&\propto \exp \left\{ -\frac{1}{2} \left(\frac{1}{v} \sum_{t: y_{t-1} > \theta} y_t^2 - \frac{1}{v} \phi \sum_{t: y_{t-1} > \theta} y_t y_{t-1} + \frac{1}{v} \phi^2 \sum_{t: y_{t-1} > \theta} y_{t-1}^2 + \frac{1}{c} \phi^2 \right) \right\}
\end{aligned}$$

This simplifies to a normal with mean $\frac{c \sum_{t: y_{t-1} > \theta} y_t y_{t-1}}{c \sum_{t: y_{t-1} > \theta} y_{t-1}^2 + v}$ and variance $\frac{cv}{c \sum_{t: y_{t-1} > \theta} y_{t-1}^2 + v}$. Switch the index to find the posterior for ϕ_2 . The posterior for v will be

$$p(v | y_{1:n}, \theta, \phi_1, \phi_2) \propto \prod_{t: y_{t-1} > \theta} p(y_t | y_{t-1}, \phi_1, v) \prod_{t: y_{t-1} < \theta} p(y_t | y_{t-1}, \phi_2, v) p(v | \alpha_0, \beta_0)$$

This simplifies to an inverse gamma kernel with parameters $\frac{\alpha_0 + n - 1}{2}$ and

$$\frac{1}{2} \left(\sum_{t: y_{t-1} > \theta} (y_t - \phi_1 y_{t-1})^2 + \sum_{t: y_{t-1} < \theta} (y_t - \phi_2 y_{t-1})^2 + \beta_0 \right)$$

The posterior for θ is proportional to $\prod_{t: y_{t-1} > \theta} p(y_t | y_{t-1}, \phi_1, v) \prod_{t: y_{t-1} < \theta} p(y_t | y_{t-1}, \phi_2, v) p(\theta | a)$. Using Metropolis hastings, samples from the posterior can be found.