SIMPLE PANEL GROWTH REGRESSION EXERCISE

1. The model

In this exercise you fit a second-order autoregressive model to the log of per capita real GDP from 156 countries. The model is

$$Ay_i(t) = A\rho_1 y_i(t-1) + A\rho_2 y_i(t-2) + Ac_i + \lambda_i \varepsilon_i(t) , \tag{1}$$

where i indexes countries and t indexes time, $\varepsilon_i(t)$ is i.i.d. $N(0,1)$ and uncorrelated with $y_i(s)$ for $s < t$. This implies, for $A \neq 0$,

$$y_i(t) = \rho_1 y_i(t-1) + \rho_2 y_i(t-2) + c_i + \eta_i()t) \tag{2}$$

with $\eta_i(t) = \frac{\lambda_i \varepsilon_i(t)}{A}$ distributed as $N(0, \lambda_i^2/A^2)$.

The model imposes the same dynamics (ρ_1 and ρ_2) on all countries, but allows the constant terms c_i to vary across countries. If $\rho_1 + \rho_2$ is within $1/T$ (with T sample size) of 1 and the absolute value of $c_i/(1-\rho_1-\rho_2)$ is large relative to the values of y in the sample, the model implies near-linear growth through the sample. One can calculate the "local steady state growth rate" near y as the value of δ satisfying:

$$y + \delta = \rho_1 y + \rho_2(y - \delta) + c \tag{3}$$

$$\therefore \delta = \frac{(\rho_1 + \rho_2 - 1)y + c}{1 + \rho_2} . \tag{4}$$

This is close to $c/(1+\rho_2)$ when $c/(1-\rho_1-\rho_2) >> y$, but decreases as y increases.

As you know from class discussion, estimating this model requires specifying a joint distribution for the initial conditions (here $y_i(1970)$ $y_i(1971)$) and the constant term c_i, or at least (as in Liu's paper) the conditional distribution of c_i given the initial conditions. The R software available for this exercise specifies that conditional distribution as

$$c_i \mid (y_i(1970), y_i(1971)) \sim N(\bar{y}(1-\rho_1-\rho_2), \nu^2) , \tag{5}$$

where $\bar{y} = \big(y_i(1970) + y_i(1971)\big)/2$ and $\nu^2 = \lambda_i^2/(A^2\tau^2)$ and τ is the prior parameter called `cointight` in the software. While this specification, like Liu's, makes c_i a normal linear regression on initial conditions, it has only one free hyperparameter (τ) instead of the four free in Liu's specification, and uses just one regression model, while Liu allows for a mixture of several reagression models. Liu's specification can be shown to deliver consistency as the number of countries $N \to \infty$ with fixed T, while the specification in this exercise probably does not. Nonetheless, with 43 observations, the bias in estimates of the ρ values is likely to be small.

Date: January 9, 2019.

2. How to do the estimation

The course web site contains the raw Penn World Table data in `.csv` format, as an R data frame in `.RData` format, and as a list of R data frames in `.RData` format. The first two stack all the data and indicate the country via an R "factor" variable called countrycode. You probably want to work with the data already logged, truncated to 1970-2014, as a named R list of time series objects, which is available as `ylist` in `ylist.RData`. That file also contains the full country names corresponding to the three-letter country codes in the `country` vector. The `PanelVAR.zip` file contains an R source package that you should be able to install as a local package after unzippping it on your computer. The programs invoked below are in that package. Though I've done the exercise with the package, it's possible that there are bugs in it or that it has dependencies that I didn't notice. Let me know promptly if you run into problems with it. The `csminwelNew()` program is an optimization program most of the class has used before, but you can of course use other optimization programs.

Here are the commands to estimate the model conditional on `A, lambda`.

```
pvpout <- pvarprior(ylist, lags=2, sig=.05, tight=1, decay=.3,
   urtight=1, cointight=1)
## Constructs the MN prior dummy observations
stout <- StackData(ylist, dumObs=pvpout, lags=2)
## Stacks the real data Y and X matrices on top of the dummy observations.
psvout <- pSVARmdd(stout, A, lambda)
## Evaluates the likelihood maximum and the integrated posterior density
## (mdd) conditional on A and lambda
```

To get estimates that ignore heteroskedasticity across countries, set

```
lambda <- matrix(rep(1, length(ylist)), nrow=1)
A <- matrix(20, 1, 1)
```

The `A` value above is just a good initial guess, implying a standard error of forecast of about .05 in a typical (`lambd[i]=1`) country. Note that `A` and `lambda` must be dimensioned matrices (since this code is set up to handle multivariate systems). To get estimates that use no prior dummy observations, leave out the `dumObs` argument for `StackData` or set it to its default value of `NULL`.

Obtain optimized estimates of `A` and `lambda` as follows:

```
pvobj <- pvarObj(stout)
## Returns a function that invokes pSVARmdd, but saves the big Y and X
## matrices locally, so they aren't passed as arguments on each function
## evaluation
optout <- csminwelNew(pvobj, x0=Alambda2vec(A, lambda),
     H0=diag(156) * 1e-3, nit=200)
## Finds optimal A, lambda pair, maximizing mdd conditional on A and lambda
## (not exactly the same as maximizing posterior density over all parameters)
AL <- alvec2mat(optout$xh, nv=1)
## Sorts the optimized parameter vector into an A matrix and lambda matrix.
psvopt <- pSVARmdd(stout, AL$A, AL$lambda)
```

```
## The full output of pSVARmdd at the optimized A,lambda.
## (pvobj returns only the log mdd)
```

On my laptop, this takes about 10 seconds per iteration (almost all the time is used in taking the 156-dimensional numerical derivative) and converges in around 60 iterations with the default 1e-7 convergence criterion. However, adequate accuracy could be obtained with a looser convergence criterion.

The `A` and `lambda` obtained this way maximize the marginal posterior density for those parameters, with the autoregressive parameters and constants integrated out. It would be feasible to start from there and use `pvobj` as the posterior density in an MCMC chain, taking account of the uncertainty in `A`, `lambda`, but this exercise is already long enough. A program that could be used to do the MCMC, `pvarMCMC`, is in the Panel. (It is actually a generic Metropolis random walk program.)

The `psvout` or `psvopt` objects above contain information about the fitted model in somewhat inconvenient form. The function `vcov.psv` calculates the covariance matrix of all estimated coefficients, **`coef`**`.psv` calculates the reduced form autoregressive parameters and the reduced form constant terms c_i. **`residuals`**`.psv` extracts the estimated values of $\varepsilon_i(t)$ (which the model assumes zero-mean and unit variance) as a list indexed by country. Bear in mind that for each country the last four residuals in this model are for dummy observations. Since `psvopt` is of class `"psv"`, these functions can be invoked without the `.psv`, e.g. as `vmat <-vcov(psvopt)`. Note that **`residuals`**`.psv` returns a list indexed by countries, so you will want to do something like

```
ul <- matrix(unlist(residuals(psvout)$u, ncol=length(ylist)|
dimnames(ul)[[2]] <- names(ylist)
```

These same programs and steps would work if the model were a structural VAR rather than a univariate AR.

3. Questions to address

(1) The estimates emerge with $\rho_1 + \rho_2$ near, but less than, 1 and c_i's all positive. This implies positive trend-like growth, with the δ in (4) falling as y increases. Is this implied growth slowdown realistic for most economies? Are there examples of countries where it clearly doesn't fit? You can look at residuals or at the time path of $\delta(y)$ from (4) for some individual countries to check this.

(2) That growth slows as y increases is consistent with the idea of "convergence", that poor countries are on their way to catching up with richer ones. Is there evidence for convergence? One version of this question is whether $c_i/(1-\rho_1-\rho_2)$ is more or less dispersed than $y_i(2014)$.

(3) How important is the non-normality in the residuals? And does accounting for heteroskedasticity across countries reduce it? You can compare estimates that hold variances fixed across countries to estimates with optimized `A, lambda`. You can look at histograms of the residuals across all countries and **`qqnorm`** plots and/or **`qqplot`** plots against, say, t distributions with low degrees of freedom.

(4) Liu allows for a mixture of normals in the distribution of c_i conditional on initial conditions. Our specfication more tightly restricts the distribution of $c_i \mid (y_i(1970), y_i(1971)$. If our restrictions are too tight, this might show up in the residuals from the `cointight` dummy observations not looking $N(0,1)$. These are the 44th residuals of each country. Check whether they look $N(0,1)$.
(5) The model assumes no correlation of residuals across countries. Assess whether the estimates suggest this assumption is incorrect.
(6) What are the most important directions in which you might consider making the model more flexible, assuming we are staying with the same data set?
(7) Extra Credit: Try varying some of the Minnesota prior parameters you give to `pvarprior` to see if reasonable changes in them change conclusions. You can also try varying A and λ around the posterior-maximizing values. Of course in principle or in a serious research paper, you would explore this with an MCMC chain, but trying a few *ad hoc* variations is a check on whether the posterior sampling might alter conclusions.