ECON 8185 - Homework 3

João Lazzaro *

April 11, 2019

Question 1

In this question we were asked to write a code to compute the MLE estimator for the parameters of HW2 model:

$$\max_{c_t,x_t,l_t} E\sum_{t=0}^{\infty}\beta^t\frac{\left(c_t l_t^{\psi}\right)^{1-\sigma}}{1-\sigma}N_t$$

S.T.

$$c_t+(1+\tau_{xt})x_t=r_tk_t+(1-\tau_{ht})w_th_t+T_t$$
$$N_{t+1}k_{t+1}=[(1-\delta)k_t+x_t]N_t$$
$$h_t+l_t=1$$
$$S_t=PS_{t-1}+Q\epsilon_t,\quad S_t=[\ln z_t,\tau_{ht},\tau_{xt},\ln g_t]$$
$$c_t,x_t\geq 0$$

Where $N_t=(1+\gamma_n)^t$ and firm tecnhology is $Y_t=K_t^{\theta}(Z_tL_t)^{1-\theta}$. γ_z is the rate of labor-augmenting technical progress. The stochastic shocks hitting this economy affect technology, tax rates, and government spending and the stochastic processes are modeled as a VAR(1) process. The resource constraint is $Y_t=N_t(c_t+x_t+g_t)$. We must work with detrended variables, we should use "hat" variables, but typing this is time consuming so from now on all variables should be understood as detrended unless I say something

*santo279@umn.edu

about them. Let's first define some parameters for this model: $\delta = 0.0464$, $\theta = 1/3$, $\beta = 0.9$, $\sigma = 2$, $\psi = 1$, $\gamma_z = 0$ and $\gamma_n = 0$.

In HW2, I describe how to approximate and write the model in the state space form:

$$X_{t+1} = AX_t + B\varepsilon_{t+1}$$

$$Y_t = CX_t + \omega_t$$

Where, $X_t = [k_t, S_t], S_t = [z_t, \tau_{ht}, \tau_{xt}, g_t]$, $Y_t = [y_t, x_t, h_t]$ and as before:

$$S_{t+1} = PS_t + Q\varepsilon_{t+1}$$

We will assume that the true shocks are uncorrelated, but when we proceed to estimation, we will not restrict the parameter space. Hence, P and Q are:

$$P = \begin{bmatrix} \rho_z & \rho_{zh} & \rho_{zx} & \rho_{zg} \\ \rho_{hz} & \rho_h & \rho_{hx} & \rho_{hg} \\ \rho_{xz} & \rho_{xh} & \rho_x & \rho_{xg} \\ \rho_{gz} & \rho_{gh} & \rho_{gx} & \rho_g \end{bmatrix} = \begin{bmatrix} 0.95 & 0 & 0 & 0 \\ 0 & 0.75 & 0 & 0 \\ 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.25 \end{bmatrix}$$

$$Q = \begin{bmatrix} \sigma_z & \sigma_{zh} & \sigma_{zx} & \sigma_{zg} \\ & \sigma_h & \sigma_{hx} & \sigma_{hg} \\ & & \sigma_x & \sigma_{xg} \\ & & & \sigma_g \end{bmatrix} = \begin{bmatrix} 0.05 & 0 & 0 & 0 \\ & 0.03 & 0 & 0 \\ & & 0.02 & 0 \\ & & & 0.04 \end{bmatrix}$$

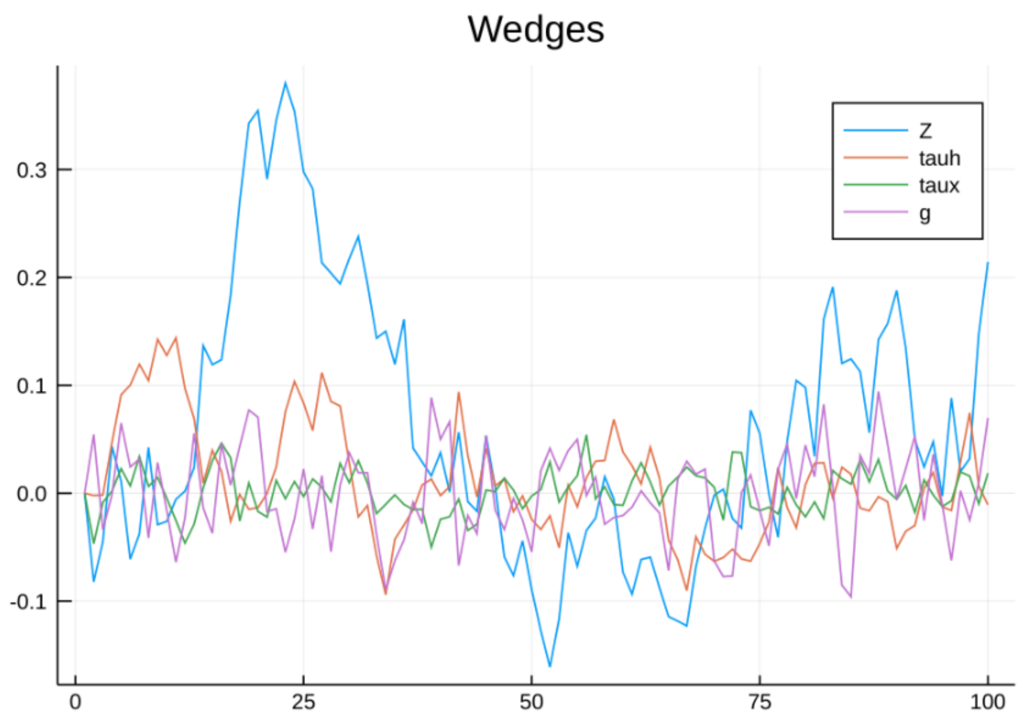
The Steady State values for the stochastic shocks are defined below:

$$SS = \begin{bmatrix} g_{ss} & \tau_{xss} & \tau_{hss} & z_{ss} \end{bmatrix} = \begin{bmatrix} 0.01 & 0.05 & 0.02 & 1 \end{bmatrix}$$

The function State_Space in the file State_Space.jl was created for HW2 and finds Matrices A, B and C, given the parameters. With these matrices, we may simulate some data:

Given the simulated data above, we want to build an algorithm to estimate the matrices P and Q (and the steady state values) such that we would have the same data generating process. To do that, we'll build a likelihood function using the Kalman Filter. We follow Ljunqvist Sargent exposition of the Kalman filter. Assume for a slightly more general case that ω_t is an iid vector sequence of normal random variances with mean zero and covariance matrix R (In our case, $R = 0$).

We observe Y_t but not X_t. We want to estimate $\hat{X}_t = E[X_t|Y^{t-1}]$ and also the covariances $\Sigma_t = E[(X_t - \hat{X}_t)(X_t - \hat{X}_t)']$. The trick is to use the



new information in y_0 relative to what we know $(\hat{X}_0)$: $a_0 := Y_0 - C\hat{X}_0$. $a_t := Y_t - C\hat{X}_t$. Regressing the model:

$$X_0 - \hat{X}_0 = L_0 a_0 + \eta$$

We get that:

$$L_0 = \Sigma_0 C' \left(C\Sigma_0 C' + R\right)^{-1}$$

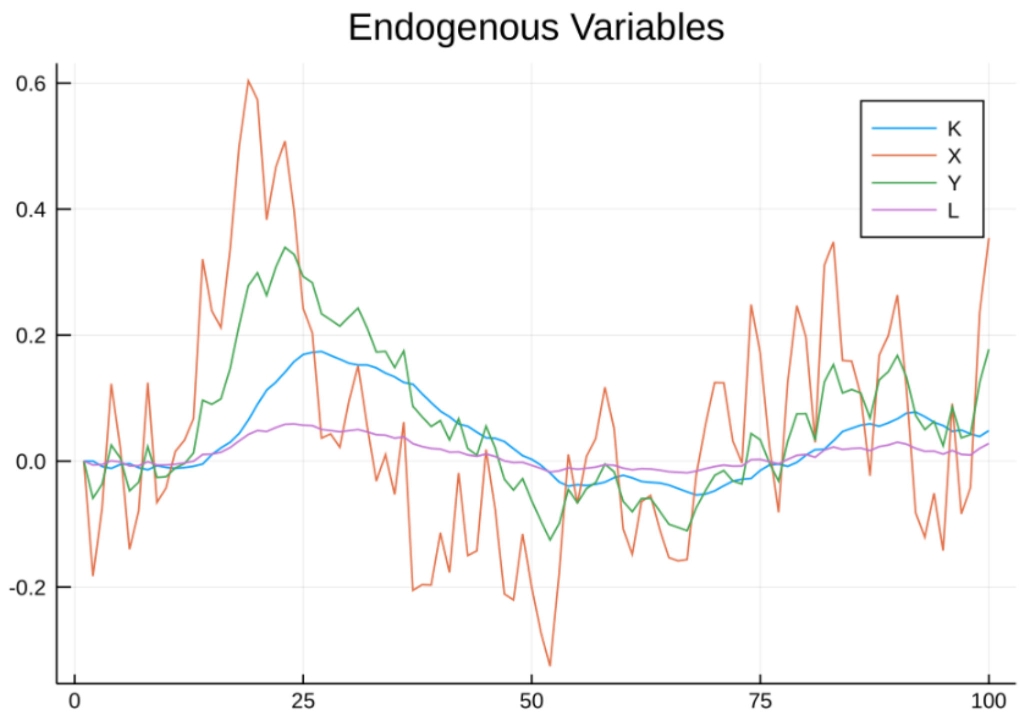
Defining $\hat{X}_1 = E[X_1|Y_0] = AX_0$:

$$X_1 = A\hat{X}_0 + A\left(X_0 - \hat{X}_0\right) + B\varepsilon_1$$

Substituting the linear model and L_0:

$$\hat{X}_1 = A\hat{X}_0 + K_0(Y_0 - C\hat{X}_0)$$

Where $K_0 = AL_0$ is the Kalman gain. Plugging these equations: $X_1 - \hat{X}_1 = A(X_0 - \hat{X}_0) + B\varepsilon_1 - K_0(Y_0 - C\hat{X}_0)$ and we may compute a new variance



matrix $\Sigma_1 = E(X_1 - \hat{X}_1)(X_1 - \hat{X}_1)$:

$$\Sigma_1 = (A - K_0 C)\Sigma_0 (A - K_0 C)' + (BB' + K_0 R K_0')$$

Iterating the above steps we get te Kalman filter recursion:

$$a_t = Y_t - C\hat{X}_t$$

$$K_t = A\Sigma_t C'(C\Sigma_t C' + R)^{-1}$$
$$\hat{X}_{t+1} = A\hat{X}_t + K_t a_t$$
$$\Sigma_{t+1} = (A - K_t C)\Sigma_t (A - K_t C)' + (BB' + K_t R K_t')$$

In the file KalmanFilter.jl a function was defined to implement the recursion above. The innovations representation that emerges from the Kalman filter is:

$$\hat{X}_{t+1} = A\hat{X}_t + K_t a_t$$
$$Y_t = C\hat{X}_t + a_t$$

Define: $\Omega_t := E[a_t a_t'] = C\Sigma_t C' + R$. We have that $E[Y_t|Y^{t-1}] = C\hat{X}_t$ and the conditional distribution of Y_t is $N(C\hat{X}_t, \Omega_t)$. Therefore, $C\hat{X}_t, \Omega_t$ are sufficient statistics for the conditional distribution of Y_t. Note that we can factor the likelihood function as:

$$f(Y_T, ...Y_0) = f(Y_T|Y^{T-1})f(Y_{T-1}|Y^{T-2})...f(Y_1|Y_0)f(Y_0)$$

And the log of the conditional density of the $m\times 1$ vector Y_t, ignoring constant terms, is:

$$\log f(Y_t|Y^{t-1}) = -0.5\log(|\Omega_t|) - 0.5a_t'\Omega_t^{-1}a_t$$

In the file KalmanFilter.jl, I defined a function to compute the log-likelihood.

Question 2

In total, we have 30 parameters to be estimated: 16 correlations, 10 covariances and 4 Steady State values.

We maximize log likelihood function using Julia Optim package. Our optimizer will use the box-constrained[1] LBFGS which is a variant of the Newton method with approximated Hessians and Gradients. Our initial guess will be points normally distributed around the true ones, with standard deviation of 0.5. The sample size is of 100 observations.

Estimating the Diagonal of P

As a warm-up, we start by estimating only the diagonal of matrix P. We get that the estimated matrix is:

$$\hat{P} = \begin{bmatrix} 0.9501 & 0 & 0 & 0 \\ 0 & 0.7936 & 0 & 0 \\ 0 & 0 & 0.3725 & 0 \\ 0 & 0 & 0 & 0.2511 \end{bmatrix}$$

The estimation was satisfactory, except for ρ_x that was far from the true parameter. To fix that, we simulate more data and get a sample size of 600. The result with the increased sample is:

[1] We impose that the parameters must make sense, variances cannot be negative and correlations not greater than 1

$$\hat{Q} = \begin{bmatrix} 0.9499 & 0 & 0 & 0 \\ 0 & 0.7930 & 0 & 0 \\ 0 & 0 & 0.4773 & 0 \\ 0 & 0 & 0 & 0.2723 \end{bmatrix}$$

Estimating the Diagonal of P and Q

We'll keep the new sample size from now on. The next subset of parameters estimated includes the diagonal of P and Q. The results are:

$$\hat{P} = \begin{bmatrix} 0.9499 & 0 & 0 & 0 \\ 0 & 0.7855 & 0 & 0 \\ 0 & 0 & 0.4167 & 0 \\ 0 & 0 & 0 & 0.2723 \end{bmatrix}$$

$$\hat{Q} = \begin{bmatrix} 0.0492 & 0 & 0 & 0 \\ & 0.0294 & 0 & 0 \\ & & 0.0192 & 0 \\ & & & 0.0402 \end{bmatrix}$$

The estimation also works fine.

Estimating everything except the Steady States

Now, we estimate all the Parameters with the exception of the Steady States and things start to get a little trickier. I was forced to make the initial guesses closer to the original parameter, now the standard deviation of the true normal distribution is of 0.1. The results are:

$$\hat{P} = \begin{bmatrix} 0.9506 & -0.017 & 0.0557 & 0.0114 \\ 0.117 & 0.7634 & -0.23 & 0.11 \\ 0.0878 & 0.114 & 0.6119 & 0.012 \\ 0.0858 & 0.0133 & 0.0105 & 0.2500 \end{bmatrix}$$

$$\hat{Q} = \begin{bmatrix} 0.0499 & 0.022 & 0.012 & 0 \\ & 0.003 & 0.0062 & 0.0434 \\ & & 0.0333 & 0 \\ & & & 0.0294 \end{bmatrix}$$

Estimating 30 parameters

With the addition of the Steady State parameters, the model often fails to converge to a reasonable point. For some parameters, the Kalman matrices are non invertible and the solver frequently tries to evaluate the function on those regions, returning an error message and stopping the estimation. The solver also tries combinations of parameters that would make an explosive root, returning more errors.

To get a somewhat reasonable estimation, the LBFGS algorithm was working poorly. I obtained my best result using a Simulated Annealing algorithm.

$$\hat{P} = \begin{bmatrix} 0.9413 & 0.0517 & 0.0417 & 0.0194 \\ 0.117 & 0.8628 & 0.0023 & 0.11 \\ 0.1278 & 0.0354 & 0.5678 & 0.0494 \\ 0.0278 & 0 & 0.0307 & 0.2345 \end{bmatrix}$$

$$\hat{Q} = \begin{bmatrix} 0.0494 & 0.0224 & 0.0042 & 0 \\ & 0.0245 & 0.0062 & 0.0034 \\ & & 0.0294 & 0.0136 \\ & & & 0.0346 \end{bmatrix}$$

$$\hat{SS} = \begin{bmatrix} 0.0231 & 0.0434 & 0.00629 & 1 \end{bmatrix}$$