Problem Set 4
Due 13. October, 14:15

Exercise 4.1: Business cycle facts

In this problem set we establish business cycle facts about Norway and compare them to the ones established in Kydland and Prescott (1990, Minneapolis Fed, Quarterly Review 14(2)) for the United States. The data analysis and the questions of the problem set are cooked up to solve it with Matlab (I also believe it is good practice for future problem sets). You are required to program the data filters in part (d) yourself.

(a) Visit the website of Statistics Norway (SBB) and download aggregate quarterly data for Norway on:

- Statbank, Table 09190 (current prices, seasonally adjusted):
  Household final consumption expenditure (total ($x_{:,1}$), goods ($x_{:,2}$), services ($x_{:,3}$)); Gross fixed capital formation (total ($x_{:,4}$), mainland ($x_{:,5}$)); Gross capital formation ($x_{:,6}$); Total exports ($x_{:,7}$); Total imports ($x_{:,8}$); Gross domestic product, market values (total ($x_{:,9}$), mainland $x_{:,10}$).

- Statbank, Table 09175 (mainland Norway):
  Compensation of employees ($x_{:,11}$); Total hours worked for employees and self-employed, seasonally adjusted ($x_{:,12}$);

We will refer to date t in time-series n as $x_{t,n}$, and the notion $x_{:,n}$ corresponds to the whole range of dates for the time-series n (a column vector). For $n = 1, ..., 10$ the date runs from 1978K1 to 2015K2, for $n = 11, 12$ the date runs from 1995K1 to 2015K2. After extracting the data from the SSB Statbank, manually rotate the tables such that each time series $x_{:,n}$ appears as a column vector, this will make the import into Matlab more convenient. Save the tables in the Semi-colon delimited format and merge the two data files (with Excel for example) such that you end up with the structure ($x_{:,11}$ and $x_{:,12}$ will have missing entries for dates before 1995K1):

date	x1	x2	$\cdots$	x12
1978K1	27900	17503	$\cdots$	.
1978K2	28861	18160	$\cdots$	.
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
1995K1	109344	60729	$\cdots$	758.1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
2015K2	320934	155991	$\cdots$	959.1

(b) Import the merged file into Matlab (in the Home tab, choose the option Import Data, delimiter: Semicolon). Import the dates as a cell array, and the collection of time-series $x_{:,1}, \ldots, x_{:,12}$ as one numeric matrix. Rename the cell array of dates to

date (if you haven't done it already) and the data matrix to x and save your transformed data using the command `save('data.mat','date','x')`. If you want to fetch the data again after closing the program you can load them with the command `load('data.mat','date','x')` given that you set the Matlab path to the same directory as in the previous session.

(c) Normalize the gathered macro data by dividing each time series, $x_{:,n}$, by its empirical mean

$$\hat{x}_{:,n} \equiv \frac{x_{:,n}}{T^{-1}\sum_{t=1}^{T} x_{t,n}}.$$

Plot the data on a logarithmic scale. Do the normalized time-series look stationary or do they have a trend?

(d) Detrend the normalized data, $\hat{x}_{:,n}$, using three different filters:

  (i) Linear trend:

$$\log(\hat{x}_{t,n}^{cycle}) = \log(\hat{x}_{t,n}) - \log(\hat{x}_{0,n}) - tg_n.$$

Estimate the two parameters $\log(\hat{x}_{0,n})$ and g_n with Ordinary Least Squares (OLS):

$$\langle\log(\hat{x}_{0,n}), g_n\rangle = \arg\min \sum_{t=1}^{T}\left[\log(\hat{x}_{t,n}^{cycle})\right]^2.$$

To do so, rewrite the linear trend filter as the generic equation

$$y = X\beta + \varepsilon$$

where $y_{t,n} \equiv \log(\hat{x}_{t,n})$, $\varepsilon_{t,n} \equiv \log(\hat{x}_{t,n}^{cycle})$, and

$$y \equiv \begin{bmatrix} y_{1,1} & \cdots & y_{T,1} \\ \vdots & \ddots & \vdots \\ y_{T,1} & \cdots & y_{T,N} \end{bmatrix}, \quad \varepsilon \equiv \begin{bmatrix} \varepsilon_{1,1} & \cdots & \varepsilon_{T,1} \\ \vdots & \ddots & \vdots \\ \varepsilon_{T,1} & \cdots & \varepsilon_{T,N} \end{bmatrix}.$$

Then write a Matlab routine called myols.m that returns the associated OLS estimator

$$\hat{\beta} = \left(X'X\right)^{-1}X'y.$$

  (ii) First-difference:

$$\log(\hat{x}_{t,n}^{cycle}) = \log(\hat{x}_{t,n}) - \log(\hat{x}_{t-1,n}) - g_n.$$

Estimate the parameter g_n again with the above described OLS procedure:

$$g_n = \arg\min \sum_{t=2}^{T}\left[\log(\hat{x}_{t,n}^{cycle})\right]^2.$$

(iii) Hodrick-Prescott with smoothing parameter $\lambda = 1600$:

$$\log(\hat{x}_{:,n}^{trend}) = \arg\min \sum_{t=1}^{T} \left[\log(\hat{x}_{t,n}) - \log(\hat{x}_{t,n}^{trend}) \right]^2$$

$$+ \lambda \sum_{t=2}^{T-1} \left[(\log(\hat{x}_{t+1,n}^{trend}) - \log(\hat{x}_{t,n}^{trend})) - (\log(\hat{x}_{t,n}^{trend}) - \log(\hat{x}_{t-1,n}^{trend})) \right]^2.$$

Note that this is a much more demanding optimization problem compared to (i) and (ii) as we are seeking for the minimum over T variables for each time series. Let's slightly redefine variables in this expression such that

$$\tau_{:,n} = \arg\min \sum_{t=1}^{T} (y_{t,n} - \tau_{t,n})^2 + \lambda \sum_{t=2}^{T-1} \left[(\tau_{t+1,n} - \tau_{t,n}) - (\tau_{t,n} - \tau_{t-1,n}) \right]^2,$$

where the trend component is defined as $\tau_{t,n} \equiv \log(\hat{x}_{t,n}^{trend})$. Derive the first-order conditions of this minimization problem and show that the trend component can be written as

$$\underbrace{\begin{bmatrix} \tau_{1,1} & \cdots & \tau_{1,N} \\ \vdots & \ddots & \vdots \\ \tau_{T,1} & \cdots & \tau_{T,N} \end{bmatrix}}_{\equiv \tau} = (\lambda A + I_T)^{-1} \underbrace{\begin{bmatrix} y_{1,1} & \cdots & y_{1,N} \\ \vdots & \ddots & \vdots \\ y_{T,1} & \cdots & y_{T,N} \end{bmatrix}}_{\equiv y}, \qquad (1)$$

where I_T is a matrix of size $T \times T$ with all ones on the main diagonal and zeros otherwise, and

$$A = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 & 0 & 0 & & \cdots & & 0 \\ -2 & 5 & -4 & 1 & 0 & 0 & 0 & & \cdots & & 0 \\ 1 & -4 & 6 & -4 & 1 & 0 & 0 & & \cdots & & 0 \\ 0 & 1 & -4 & 6 & -4 & 1 & 0 & & \cdots & & 0 \\ & & & & \cdots & & & & & & \\ 0 & & & \cdots & & & 0 & 1 & -4 & 6 & -4 & 1 & 0 \\ 0 & & & \cdots & & & 0 & 0 & 1 & -4 & 6 & -4 & 1 \\ 0 & & & \cdots & & & 0 & 0 & 0 & 1 & -4 & 5 & -2 \\ 0 & & & \cdots & & & 0 & 0 & 0 & 0 & 1 & -2 & 1 \end{bmatrix}.$$

(e) Compute the standard deviation

$$std\left(\log\left(\hat{x}_{:,n}^{cycle} \right) \right)$$

for all the series n and all the three different filters. Are the standard deviations consistent across the three different filters?

(f) Compute the correlation coefficient of the Gross domestic product on the mainland with all variables (including itself) for 5 leading and 5 lagging quarters

$$corrcoef\left(\log\left(\hat{x}_{t,10}^{cycle} \right), \log\left(\hat{x}_{t+s,n}^{cycle} \right) \right), \, n = 1, ..., N, \, s = -5, ..., 5,$$

for the HP-filtered series. How do your results for the Norwegian economy compare with Kydland and Prescott (1990, Tables 1-3)? Pick your favorite time-series (or several of them) and plot the cyclical components against GDP as in Kydland and Prescott (1990, Charts 3-12).