Most macroeconomic time series are not stationary. The stationary ones are usually normalized variables such as trade balance to GDP ratio, etc. Suppose $y_t$ is a non-stationary macro series. We think of $y_t$ as an aggregate of many components differentiated by their frequencies. We borrow the concept of "frequency" and "period" from the engineering field. Easily speaking, frequency is the number of cycles per period. Hence, high frequency components of the data are realized in a short time period such as daily or weekly. On the other hand, low frequency components take years or decades to be realized. In general, we can decompose $y_t$ into 3 components.

1. Low-frequency components, or trend

2. Medium-frequency components, or business cycles or cyclical components

3. High-frequency components, or seasonal and irregular components

The goal of a filter design is to extract only the components with desired frequencies, and to make them stationary. In studies of (international) business cycles, we focus on medium frequencies and aim to remove trend and seasonal or irregular components. The basis of filtering with stochastic trends is the frequency domain.

1 Frequency domain

Studies in the time domain expressing $y_t$ as a function of time $y(t)$. For example,

$$y_t = \mu + \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j},$$

and $\{\epsilon_t\}_{t=-\infty}^{\infty}$ follows an i.i.d. white noise process with variance $\sigma^2$. Business cycles studies are concerned with the covariance or correlation of $y(t)$ and $y(\tau)$.
The frequency domain defines series as a function of frequency \( \omega \) in radian. (See Hamilton (1994) for details.)

\[
y_t = \mu + \int_0^\pi \xi(\omega)d\omega, \quad \xi(\omega) = \gamma(\omega)\cos(\omega t) + \delta(\omega)\sin(\omega t). \tag{2}
\]

\( \xi(\omega) \) represents the periodic component. Filtering a time series hence is nothing more than multiplying frequency response function \( \alpha(\omega) \) to \( \xi(\omega) \), where \( \alpha(\omega) \) has non-zero elements for the desired range of \( \omega \), and zero otherwise. Let \( y_t^\ast \) denoted the filtered series.

\[
y_t^\ast = \mu + \int_0^\pi \alpha(\omega)\xi(\omega)d\omega. \tag{3}
\]

The variance of the filtered series depends on the frequency response function and the variance of the original series.

\[
\text{var}(y_t^\ast) = \int_0^\pi |\alpha(\omega)|^2f_y(\omega)d\omega \tag{4}
\]

\( f_y(\omega) = \text{var}(\xi(\omega)) \) and it is called the spectral density of \( y \) at the frequency \( \omega \). \( |\alpha(\omega)|^2 \) is called squared gain. It relates the effect of the filter on the variance of the filtered series \( y_t^\ast \) relative to that of the original series \( y_t \). We can also compute autocovariance of \( y_t^\ast \) in the frequency domain. (Chapter 11 in Sargent (1987) also gives detailed discussion about this.)

The other way to define cycles is to use "periodity." We define periodity \( p \) as the length of period of one cycle. Hence, frequency \( \omega = 2\pi/p \). Mitchell and Burns (1946) define business cycles as the components in the data that last at least 6 quarters and at most 32 quarters. Based on this concept of business cycles, we are interested in the range of frequencies \( \omega \in [2\pi/32, 2\pi/6] \).

2 Frequently used filters

2.1 Deterministic trends

These detrending method assumes a certain functional form of time trend.

\[
y_t = \text{trend} + y_t^\ast
\]

Hence, \( y_t^\ast \) is simply an estimate of the residual from regressing \( y_t \) onto the time trend.

- Linear trend: \( \text{trend} = \alpha_0 + \alpha_1 t \)
- Quadratic trend: \( \text{trend} = \alpha_0 + \alpha_1 t + \alpha_2 t^2 \)
The weakness of these detrending methods are two folds. First, they make studying business cycles meaningless since these methods assume that we know the actual functional form of trend. However, we study business cycles fundamentally because we believe there are possibly stochastic components in trend in the first place. Second, the detrended series $y_t^*$ still contains the high-frequency components.

2.2 Stochastic trends

We can express $y_t$ as a function of several stochastic components.

$$y_t = \mu_t + \psi_t + u_t \quad (5)$$
$$\mu_t = \mu_{t-1} + \beta_{t-1} + v_t \quad (6)$$
$$\beta_t = \beta_{t-1} + h_t \quad (7)$$

$\mu_t$ is the stochastic trend and its shifter $v_t$ follows $NID(0, \sigma_v^2)$. $\psi_t$ is the cyclical component we are interested in. $\beta_t$ is the slope of the trend. Its shifter $h_t$ follows $NID(0, \sigma_h^2)$. $u_t$ is the irregular component and follows $NID(0, \sigma_u^2)$. The most popular filters assuming stochastic trends are the first-differenced filter, the Hodrick-Prescott (HP) and Band-pass (BP) filters.

2.2.1 First-differenced filter

The first-differenced filter assumes following the stochastic process.

$$y_t = y_{t-1} + \epsilon_t \quad (8)$$

This is equivalent to assuming that $\psi_t = 0$ and $\beta_t = 0$. Then $\epsilon_t = v_t + u_t - u_{t-1}$ and it follows a white noise process. The stochastic trend is removed by taking the first difference.

$$\Delta y_t = y_t - y_{t-1} = \epsilon_t \quad (9)$$

By construction, the filtered series $\Delta y_t$ is not cyclical component but high-frequency irregular component. The other obvious weakness of this filter is that it does not generate a stationary series when $y_t$ is integrated of higher order than 1.

2.2.2 Hodrick-Prescott filter

Hodrick and Prescott (1980) finds the optimal filter by assuming $\psi_t = 0$, and $\sigma_v = 0$.

$$y_t = \mu_t + u_t \quad (10)$$
In other words, the cyclical components are not directly measured and the stochastic component in trend comes from its slope-shifter only. The optimal filtered series $y_{t}^{HP}$ minimizes the following loss function:

$$\Sigma_{t=1}^{T}\{(y_{t} - \mu_{t})^2 + \lambda ((\mu_{t+1} - \mu_{t}) - (\mu_{t} - \mu_{t-1}))^2\}$$

Hence, conceptually $y_{t}^{HP}$ is not the cyclical component but the smoothed irregular components. When we work with slow-moving series, this may serve as a good approximation since the weight of the high-frequency components are small. The solution for $y_{t}^{HP}$ is as follows.

$$y_{t}^{HP} = \left[ \frac{\lambda(1 - L)^2(1 - L^{-1})^2}{1 + \lambda(1 - L)^2(1 - L^{-1})^2} \right] y_{t}$$

$$\lambda = \sigma_{u}^2 / \sigma_{h}^2 \quad (11)$$

$L$ is the lag operator where $L_{k} y_{t} = y_{t-k}$. The HP filter gives a stationary series $y_{t}^{HP}$ when $y_{t}$ is integrated of order up to 4.

Hodrick and Prescott (1980) use U.S. quarterly GNP data to estimate the value of $\lambda$. They set $\lambda = 1,600$. Since then, the HP filter has been widely used with this value of $\lambda$. However, there is a problem with doing this mechanically. Since $\lambda$ is supposed to be an estimate of variance of the underlying stochastic components, it is likely to be different across series and data frequencies. (See the detailed discussion in Harvey and Jaeger, 1993). So, beware the studies that universally apply the HP filter with $\lambda = 1,600$ across the board. There are a number of studies attempting to find the right $\lambda$ for the annual frequency, but zero attempts are made in finding variable-specific $\lambda$.

### 2.2.3 Band-Pass or Baxter-King Filter

Baxter and King (1999) use Mitchell and Burns’ concept of business cycles to construct a filter as a moving average (MA) of lags and leads for $K$ periods. The goal is to derive an ideal band-pass linear filter that extracts the frequencies $\omega \in \left[\omega_{L}, \omega_{U}\right]$. Assume zero mean.

$$y_{t}^{BP} = \Sigma_{h=-K}^{K} a_{h} y_{t-h}, \quad a_{h} = a_{-h} \quad (12)$$

Or,

$$y_{t}^{BP} = a(L) y_{t}, \quad a(L) = \Sigma_{h=-K}^{K} a_{h} L^{h} \quad (13)$$

This is equivalent to transform $y_{t}$ to $y_{t}^{BP}$ in the following ways.

$$y_{t} = \int_{-\pi}^{\pi} \xi(\omega) d\omega \quad (14)$$

$$y_{t}^{BP} = \int_{\omega_{L}}^{\omega_{U}} a(\omega) \xi(\omega) d\omega \quad (15)$$
Baxter and King derive an ideal approximation of $\alpha(\omega)$ in the frequency domain, and then transform its back to the time domain. Here are the steps to construct a band-pass filter that passes thorough the frequency range $\omega \in [\underline{\omega}, \overline{\omega}]$ where $\underline{\omega} = 2\pi/32$ and $\overline{\omega} = 2\pi/6$.

1. Construct a low-pass filter that passes thorough the frequency range $\omega \in [0, \overline{\omega}]$

   The ideal frequency response function is,
   
   \[ \beta(\omega) = \begin{cases} 
   1 & \text{if } |\omega| \leq \overline{\omega}, \\
   0 & \text{if } |\omega| > \overline{\omega}. 
   \end{cases} \]

   Then convert this back to the time domain using the inverse Fourier transform.\(^1\)
   
   \[ b_h = \frac{1}{2\pi} \int_{-\pi}^{\pi} \beta(\omega)e^{i\omega h} d\omega \]

   Note that $e^{i\omega h} = \cos(\omega h) + i \sin(\omega h)$. Then, the ideal filter is as follows.
   
   \[ b(L) = \sum_{h=-\infty}^{\infty} b_h L^h \]

   Hence, the ideal low-pass filter is the infinite MA process. We approximate this with $K$ lags and leads.
   
   \[ \hat{b}(L) = \sum_{h=-K}^{K} b_h L^h \]

   The inverse fourier transform transform the frequency domain coefficient to the time domain coefficient by exploiting the periodic functions of cosine and sine. For $h = 0$,
   
   \[ \hat{b}_h = \frac{\overline{\omega}}{\pi}, \]

   and for $h > 0$,
   
   \[ \hat{b}_h = \frac{\sin(h\overline{\omega})}{h\pi}. \]

   There are no clear criteria for choosing $K$. The higher $K$ is, the better approximation. However, there is a tradeoff from losing more observations of data to compute variance of the filtered series.

2. Construct a low-pass filter that passes through the frequency range $\omega \in [0, \overline{\omega}]$.

   We repeat the first step and get the approximation of the ideal filter.
   
   \[ \hat{b}(L) = \sum_{h=-K}^{K} \hat{b}_h L^h, \]

   \(^1\)The corresponding fourier transform transforming the time domain coefficient to the frequency domain coefficient is $\beta(\omega) = \sum_{h=-\infty}^{\infty} b_h e^{-i\omega h}$
where for $h = 0$,
\[ b_h = \frac{\omega}{\pi}, \]
and for $h = 1, ..., K$,
\[ b_h = \frac{\sin(h\omega)}{h\pi}. \]

3. Subtract the first filter from the second filter. This gives a filter that passes through the frequency range $\omega \in [\omega_1, \omega_2]$. The ideal frequency response function becomes,
\[ \beta_{bp}(L) = \hat{b}(L) - \hat{b}(L). \]

Finally,
\[ y_{tBP} = \beta_{bp}(L)y_t. \quad (16) \]

In other words, for $h = 0$,
\[ b_h = \frac{\overline{\omega} - \omega}{\pi}, \]
and for $h = 1, ..., K$,
\[ b_h = \frac{\sin(h\overline{\omega}) - \sin(h\omega)}{h\pi}. \]

The filter series $y_{tBP}$ is stationary when $y_t$ is integrated of order up to 2. The BP filter is much more flexible than the HP filter for two reasons. First, it has a clear definition of the periodicity of the cycles to be extracted. Second, it does not depend on macro variables or data frequencies. See the detailed comparison of the two filters in Baxter and King (1999).

References


