1 The Hodrick-Prescott Filter

1.1 Introduction

The Hodrick-Prescott (HP) filter estimates an unobserved time trend for time series variables. The procedure was first introduced by Hodrick and Prescott in the context of business cycle estimation in 1980 as a working paper, but was published 17 years later in 1997.

Let $y_t$ denote an observable (macroeconomic) time series. The HP filter decomposes $y_t$ into a nonstationary time trend ($g_t$) and a stationary residual component ($c_t$), that is:

$$y_t = g_t + c_t$$

Note that both $g_t$ and $c_t$ are unobservables. Since $c_t$ is a stationary process we can think of $y_t$ as a noisy signal for the nonstationary time trend $g_t$. Hence, the problem is how to extract an estimate for $g_t$ from $y_t$.

The HP filter solves this problem by allocating some weight for the signal against a linear time trend. Let $\lambda$ represent that weight. For $\lambda = 0$ we assume that there is no noise, i.e. $g_t = y_t$, and as $\lambda$ gets larger more and more weight is allocated for the linear trend. That is, as $\lambda \to \infty$, $g_t$ approaches the ordinary least squares estimate of $y_t$'s linear time trend. Hodrick and Prescott show that under some conditions the best choice (in the MSE sense) of $\lambda$ is driven by the relative variances of $c_t$ and the second difference of $g_t$. In Eviews, for example, the default values of $\lambda$ are 100 for annual data, 1600 for quarterly data, and 14400 for monthly data.

1.2 Derivation

The HP filter solves the following problem:

$$\min_{(g_t)_{t=1}^{T}} \sum_{t=1}^{T} (y_t - g_t)^2 + \lambda \sum_{t=2}^{T-1} [(g_{t+1} - g_t) - (g_t - g_{t-1})]^2$$

The FOCs of this problem are given by:

$$g_1 : c_1 = \lambda (g_1 - 2g_2 + g_3)$$
$$g_2 : c_2 = \lambda (-2g_1 + 5g_2 - 4g_3 + g_4)$$
$$g_t : c_t = \lambda (g_{t-2} - 4g_{t-1} + 6g_t - 4g_{t+1} + g_{t+2}) \quad \text{for } t = 3, 4, \ldots, T - 2$$
$$g_{T-1} : c_{T-1} = \lambda (g_{T-3} - 4g_{T-2} + 5g_{T-1} - 2g_T)$$
$$g_T : c_T = \lambda (g_{T-2} - 2g_{T-1} + g_T)$$
Writing the FOCs in matrix notation:

\[
\hat{c} = \lambda F \hat{g} \tag{4}
\]

Where the \(^\sim\) notation denotes estimates of the unobservables, and \(F\) is a \(T \times T\) coefficients matrix that is given by:

\[
F_{T \times T} = \begin{bmatrix}
1 & -2 & 1 & 0 & \ldots & 0 \\
-2 & 5 & -4 & 1 & 0 & \ldots & 0 \\
1 & -4 & 6 & -4 & 1 & 0 & \ldots & 0 \\
0 & 1 & -4 & 6 & -4 & 1 & 0 & \ldots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & \ldots & 0 & 1 & -4 & 6 & -4 & 1 & 0 \\
0 & \ldots & 0 & 1 & -4 & 6 & -4 & 1 \\
0 & \ldots & 0 & 1 & -4 & 5 & -2 \\
0 & \ldots & 0 & 1 & -2 & 1
\end{bmatrix} \tag{5}
\]

Using (1) and (4) we get \(y - \hat{g} = \lambda F \hat{g}\), and by rearranging we get the filtered time trend:

\[
\hat{g} = (\lambda F + I)^{-1} y \tag{6}
\]

Notice also that by the structure of the matrix \(F\), the sum each column is zero, and therefore by construction the sum of the estimated noises is zero:

\[
\sum_{t=1}^{T} \hat{c}_t = 0 \tag{7}
\]

References

