TREND ESTIMATION AND DE-TRENDING

D.S.G. POLLOCK
Department of Economics, Queen Mary College,
University of London,
Mile End Road, London E1 4NS, UK

Abstract. An account is given of a variety of linear filters which can be used for extracting trends from economic time series and for generating de-trended series. A family of rational square-wave filters is described which enable designated frequency ranges to be selected or rejected. Their use is advocated in preference to other filters which are commonly used in quantitative economic analysis.

1. Introduction: The Variety of Linear Filters

Whenever we form a linear combination of successive elements of a discrete-time signal $y(t) = \{y_t; t = 0, \pm 1, \pm 2, \ldots\}$, we are performing an operation which is described as linear filtering. Such an operation can be represented by the equation

$$x(t) = \psi(L)y(t) = \sum_j \psi_j y(t - j), \tag{1}$$

wherein

$$\psi(L) = \{ \cdots + \psi_{-1}L^{-1} + \psi_0 I + \psi_1 L + \cdots \}, \tag{2}$$

is described as the filter.

The effect of the operation it to modify the signal $y(t)$ by altering the amplitudes of its cyclical components and by advancing or delaying them in time. These modifications are described, respectively, as the gain effect and the phase effect of the filter. The gain effect is familiar through the example of the frequency-specific amplification of sound recordings which can be achieved with ordinary domestic sound systems. A phase effect in the form of a time delay is bound to accompany any signal processing that takes place in real time.

In quantitative economic analysis, filters are used for smoothing data series, which is a matter of attenuating or even discarding the high-frequency components of the series and preserving the low-frequency components. The converse operation, which is also common, is to extract and discard the low-frequency trend components so as to leave a stationary sequence of residuals, from which the dynamics of short-term economic relationships can be estimated more easily.

As it stands, the expression under (2) represents a Laurent series comprising an indefinite number of terms in powers of the lag operator L and its inverse $L^{-1} = F$ whose effects on the sequence $y(t)$ are described by the equations $Ly(t) = y(t-1)$ and $L^{-1}y(t) = Fy(t) = y(t+1)$.

In practice, $\psi(L)$ often represents a finite polynomial in positive powers of L which is described as a one-sided moving-average operator. Such a filter can only impose delays upon the components of $y(t)$.

Alternatively, the expression $\psi(L)$, might stand for the series expansion of a rational function $\delta(L)/\gamma(L)$; in which case the series is liable to comprise an indefinite number of ascending powers of L, beginning with $L^0 = I$. Such a filter is realised via a process of feedback, which may be represented by the equation

$$\gamma(L)x(t) = \delta(L)y(t), \tag{3}$$

or, more explicitly, by

$$\begin{aligned} x(t) \;=\; & \delta_0 y(t) + \delta_1 y(t-1) + \cdots + \delta_d y(t-d) \\ & -\gamma_1 x(t-1) - \cdots - \gamma_g x(t-g). \end{aligned} \tag{4}$$

Once more, the filter can only impose time delays upon the components of $x(t)$; and, because the filter takes a rational form, there are bound to be different delays at the various frequencies.

Occasionally, a two-sided symmetric filter in the form of

$$\psi(L) = \delta(F)\delta(L) = \psi_0 I + \psi_1(F + L) + \cdots + \psi_d(F^d + L^d) \tag{5}$$

is employed in smoothing the data or in eliminating its seasonal components. The advantage of such a filter is the absence of a phase effect. That is to say, no delay is imposed on any of the components of the signal. The so-called Cramér–Wold factorisation which sets $\psi(L) = \delta(F)\delta(L)$, and which must be available for any properly-designed filter, provides a straightforward way of explaining the absence of a phase effect. For the factorisation enables the transformation of (1) to be broken down into two operations:

$$\text{(i)}\quad z(t) = \delta(L)y(t) \qquad \text{and} \qquad \text{(ii)}\quad x(t) = \delta(F)z(t). \tag{6}$$

The first operation, which runs in real time, imposes time delays on every component of $x(t)$. The second operation, which works in reversed time, imposes an equivalent reverse-time delay on each component. The reverse-time delays, which are advances in other words, serve to eliminate the corresponding real-time delays.

The processed sequence $x(t)$ may be generated via a single application of the two-sided filter $\psi(L)$ to the signal $y(t)$, or it may be generated in two operations via the successive applications of $\delta(L)$ to $y(t)$ and $\delta(F)$ to $z(t) = \delta(L)y(t)$. The question of which of these techniques has been used to generate $y(t)$ in a particular instance should be a matter of indifference.

The final species of linear filter that may be used in the processing of economic time series is a symmetric two-sided rational filter of the form

$$\psi(L) = \frac{\delta(F)\delta(L)}{\gamma(F)\gamma(L)}. \tag{7}$$

Such a filter must, of necessity, be applied in two separate passes running forwards and backwards in time and described, respectively, by the equations

$$\text{(i)} \quad \gamma(L)z(t) = \delta(L)y(t) \qquad \text{and} \qquad \text{(ii)} \quad \gamma(F)x(t) = \delta(F)z(t). \tag{8}$$

Such filters represent a most effective way of processing economic data in pursuance of a wide range of objectives.

The essential aspects of linear filtering are recounted in numerous texts devoted to signal processing. Two that are worthy of mention are by Haykin (1989) and by Oppenheim and Shafer (1989). The text of Pollock (1999) bridges the gap between signal processing and time-series analysis.

In this paper, we shall concentrate on the dual objectives of estimating economic trends and of de-trending data series. However, before we present the methods that we wish to advocate, it seems appropriate to provide a critical account of some of the methods that are in common use.

2. Differencing Filters

The means of reducing time series to stationarity, which has been employed traditionally in quantitative economics, has been to take as many differences of the series as are necessary to eliminate the trend and to generate a series that has a convergent autocovariance function. A sequence of d such operations can be represented by the equation

$$x(t) = (I - L)^d y(t). \tag{9}$$

This approach to trend-elimination has a number of disadvantages, which can prejudice the chances of using the processed data successfully in estimating economic relationships.

The first of the deleterious effects of the difference operator, which is easily emended, is that it induces a phase lag. Thus, when it is applied to data observed quarterly, the operator induces a time lag of one-and-a-half months. To compensate for the effect, the differenced data may be

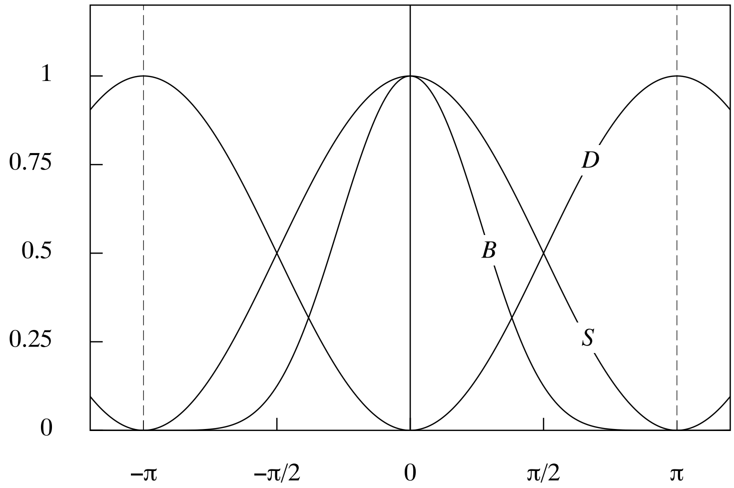

Figure 1. The frequency-response functions of the lowpass filter $\psi_S(z) = \frac{1}{4}(z+2+z^{-1})$, the highpass filter $\psi_D(z) = \frac{1}{4}(-z+2-z^{-1})$ and the binomial filter $\psi_B(z) = \frac{1}{64}(1+z)^3(1+z^{-1})^3$.

shifted forwards in time to the points that lie midway between the observations. When applied twice, the operator induces a lag of three months. In that case, the appropriate recourse in avoiding a phase lag is to apply the operator both in real time and in reversed time. The resulting filter is

$$(I-F)(I-L) = -F + 2I - L, \tag{10}$$

which is a symmetric two-sided filter with no phase effect.

As Figure 1 shows, this filter serves to attenuate the amplitude of the components of $y(t)$ over a wide range of frequencies. It also serves to increase the amplitude of the high-frequency components. If the intention is only to remove the trend from the data, then the amplitude of these components should not be altered. In order not to affect the high-frequency components, the filter coefficients must be scaled by a factor of 0.25.

To understand this result, one should consider the transfer-function of the resulting filter, which is obtained by replacing the lag operator L by the complex argument z^{-1} to give

$$\psi_D(z) = \frac{1}{4}(-z+2-z^{-1}). \tag{11}$$

The effect of the filter upon the component of the highest observable frequency—which is the so-called Nyquist frequency of $\omega = \pi$—is revealed by setting $z = \exp\{\mathrm{i}\pi\}$, which creates the filter's frequency-response function. This is

$$\psi_D(e^{i\pi}) \;=\; \frac{1}{4}\Big\{2 - (e^{\mathrm{i}\pi} + e^{-\mathrm{i}\pi})\Big\} \tag{12}$$

$$= \frac{1}{4}\Big\{2 - 2\cos(\pi)\Big\} = 1.$$

Thus, the gain of the filter, which is the factor by which the amplitude of a cyclical component is altered, is unity at the frequency $\omega = \pi$, which is what is required.

The condition that has been fulfilled by the filter may be expressed most succinctly by writing $|\psi_D(-1)| = 1$, where the vertical lines denote the operation defined by

$$|\psi(z)| = \sqrt{\psi(z)\psi(z^{-1})}, \tag{13}$$

which, in the case where $z = \exp\{i\omega\}$, amounts to taking the complex modulus. In that case, z is located on the unit circle; and, when it is expressed as a function of ω, $|\psi(\exp\{\mathrm{i}\omega\})|$ becomes the so-called amplitude-response function, which indicates the absolute value of the filter gain at each frequency.

In the case of the phase-neutral differencing filter of (10), as in the case of any other phase-free filter, the condition $\psi(z) = \psi(z^{-1})$ is fulfilled. This condition implies that the transfer function $\psi(z) = |\psi(z)|$ is a non-negative real-valued function. Therefore, the operation of finding the modulus is redundant. In general, however, the transfer function is a complex-valued function $\psi(z) = |\psi(z)|\exp\{\mathrm{i}\theta(\omega)\}$ whose argument $\theta(\omega)$, evaluated at a particular frequency, corresponds to the phase shift at that frequency.

Observe that the differencing filter also obeys the condition $|\psi_D(1)| = 0$. This indicates that the gain of the filter is zero at zero frequency, which corresponds to the fact that it annihilates a linear trend, which may be construed as a zero-frequency component.

The adjunct of the highpass trend-removing filter $\psi_D(z)$ is a complementary lowpass trend-estimation or smoothing filter defined by

$$\psi_S(z) = 1 - \psi_D(z) = \frac{1}{4}(z + 2 + z^{-1}). \tag{14}$$

As can be seen from Figure 1, the two filters $\psi_S(z)$ and $\psi_D(z)$ bear a relation of symmetry, with is to say that, when they are considered as functions on the interval $[0, \pi]$, they represent reflections of each other about a vertical axis drawn through the frequency value of $\omega = \pi/2$. The symmetry condition can be expressed succinctly via the equations $\psi_S(-z) = \psi_D(z)$ and $\psi_D(-z) = \psi_S(z)$.

The differencing filter $\psi_D(z) = \frac{1}{4}(1 - z)(1 - z^{-1})$ and its complement $\psi_S(z) = \frac{1}{4}(1 + z)(1 + z^{-1})$ can be generalised in a straightforward manner to generate higher-order filters. Thus, we may define a binomial lowpass filter via the equation

$$\psi_B(z) = \frac{1}{4^n}(1 + z)^n(1 + z^{-1})^n. \tag{15}$$

This represents a symmetric two-sided filter whose coefficients are equal to the ordinates of the binomial probability function $b(2n; p = \frac{1}{2}, q = \frac{1}{2})$. The gain or frequency response of this filter is depicted in Figure 1 for the case where $2n = 6$.

An n increases, the profile of the coefficients of the binomial filter tends increasingly to resemble that of a Gaussian normal probability density function. The same is true of the profile of the frequency-response function defined over the interval $[-\pi, \pi]$, which is the Fourier transform of the sequence of coefficients. In this connection, one might recall that the Fourier (integral) transform of a Gaussian distribution is itself a Gaussian distribution. As n increases, the span of the filter coefficients widens. At the same time, the dispersion of the frequency-response function diminishes, with the effect that the filter passes an ever-diminishing range of low-frequency components.

It is clear that, for the family of binomial filters, the symmetry of the relationship between the highpass and lowpass filters prevails only in the case of $n = 1$. Thus, if $\psi_C(z) = 1 - \psi_B(z)$, then, in general, $\psi_C(z) \neq \psi_B(-z)$. This is to be expected from the characterisation that we have given above.

It remains to conclude this section by demonstrating the effect that the simple differencing filter of equation (10) is liable to have on a typical economic time series. An example is provided by a series of monthly measurements on the U.S. money stock from January 1960 to December 1970. Over the period in question, the stock appears to grow at an accelerating rate.

Figure 2 shows the effect of fitting a polynomial of degree five in the temporal index t to the logarithms of the data. This constitutes a rough-and-ready means of estimating the trend.

The periodogram of the residuals from the polynomial regression is displayed in Figure 3. Here, there is evidence of a strong seasonal component at the frequency of $\omega = \pi/6$. Components of a lesser amplitude are also evident at the harmonic frequencies of $\omega = \pi/3$, $\pi/2$, $2\pi/3$, and there is a barely perceptible component at the frequency of $\omega = 5\pi/6$.

Apart from these components, which are evidently related to an annual cycle in the money stock, there is a substantial low-frequency component, which spreads over a range of adjacent frequencies and which attains its maximum amplitude at a frequency that corresponds to a period of roughly four years. This component belongs to the trend; and the fact that it is evident in the periodogram of the residuals is an indication of the inadequacy of the polynomial as a means of estimating the trend.

Figure 4 shows the periodogram of the logarithmic money-stock sequence after is has been subjected to the differencing filter of (10). As might be expected, the effect of the filter has been to remove the low-

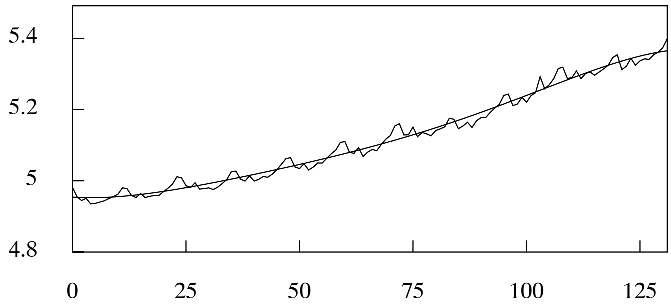

Figure 2. The logarithms of 132 monthly observations on the U.S. money stock with an interpolated polynomial time trend of degree 5.

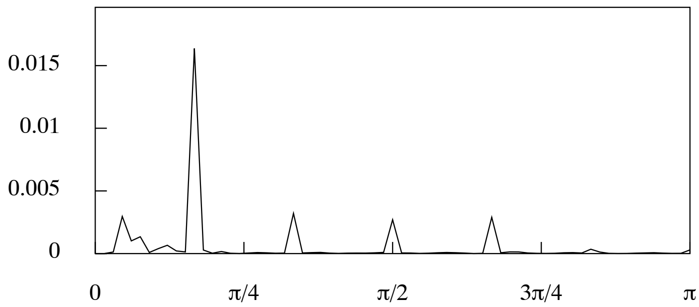

Figure 3. The periodogram of the residuals from fitting a 5th degree polynomial time trend to the logarithms of the U.S. money stock.

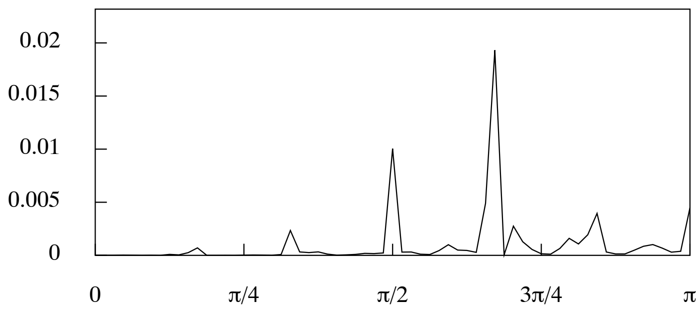

Figure 4. The periodogram of a sequence obtained by applying the second-order differencing filter to the logarithms of the U.S. money stock.

frequency trend components. However, it also has an effect which spreads into the mid and high-frequency ranges. In summary, we might say that the differencing filter has destroyed or distorted much of the information that would be of economic interest. In particular, the pattern of the seasonal effect has been corrupted. This distortion is liable to prejudice our ability to build effective forecasting models that are designed to take account of the seasonal fluctuations.

One might be tempted to use the lowpass binomial filter, defined under (15), as a means of extracting the trend. However, as Figure 1 indicates, even with a filter order of 6, there would be substantial leakage from the seasonal components into the estimated trend; and we should need to deseasonalise the data before applying the filter.

In the ensuing sections, we shall describe alternative procedures for trend extraction and trend estimation. The first of these procedures, which is the subject of the next section, is greatly superior to the differencing procedure. Nevertheless, it is still subject to a variety of criticisms. The procedure of the ultimate section is the one which we shall recommend.

3. Notch Filters

The binomial filter $\psi_B(z)$, which we have described in the previous section, might be proposed as a means of extracting the low-frequency components of an economic time series, thereby estimating the trend. The complementary filter, which would then serve to generate the de-trended series, would take the form of $\psi_C(z) = 1 - \psi_B(z)$.

Such filters, however, would be of limited use. In order to ensure that a sufficiently restricted range of low-frequency components are passed by the binomial filter, a large value of n would be required. This would entail a filter with numerous coefficients and a wide time span. When a two-sided filter of $2n + 1$ coefficients reaches the end of the data sample, there is a problem of overhang. Either the final n sample elements must remain unprocessed, or else n forecast values must be generated in order to allow the most recent data to be processed. The forecasts, which could be provided by an ARIMA model, for example, might be of doubtful accuracy.

In applied economics, attention is liable to be focussed on the most recent values of a data series; and therefore a wide-span symmetric filter, such as the binomial filter, is at a severe disadvantage. It transpires that methods are available for constructing lowpass filters which require far fewer parameters.

To describe such methods, let us review the original highpass differencing filter of equation (10). Such a filter achieves the effect of annihilating a trend component by placing a zero of the function $\psi(z)$ on the unit circle at

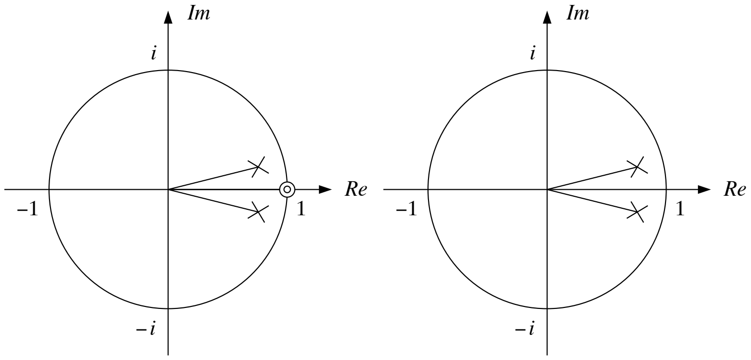

Figure 5. The pole–zero diagram of the real-time components of the notch filter ψ_N (left) and of the Hodrick–Prescott filter $\psi_P = 1 - \psi_N$ (right) in the case where $\lambda = 64$. The poles are marked by crosses and, in the case of the notch filter, the double zero at $z = 1$ is marked by concentric circles.

the point $z = 1$, which corresponds to a frequency value of $\omega = 0$. Higher-order differencing filters are obtained by placing more than one zero at this location. However, the effect of the zeros is likely to be felt over the entire frequency range with the deleterious consequences that we have already illustrated with a practical example.

In order to limit the effects of a zero of the filter, the natural recourse is to place a pole in the denominator of the filter's transfer function located at a point in the complex plane near to the zero. The pole should have a modulus that is slightly less than unity. The effect will be that, at any frequencies remote from the target frequency of $\omega = 0$, the pole and the zero will virtually cancel, leaving the frequency response close to unity. However, at frequencies close to $\omega = 0$, the effect of the zero, which is on the unit circle, will greatly outweigh the effect of the pole, which is inside it, and a narrow notch will be cut in the frequency response of the transfer function.

The device that we have described is called a notch filter. It is commonly used in electrical engineering to eliminate unwanted components, which are sometimes found in the recordings of sensitive electrical transducers and which are caused by the inductance of the alternating current of the mains electrical supply. In that case, the zero of the transfer function is placed, not at $z = 1$, but at some point on the unit circle whose argument corresponds to the mains frequency Also, the pole and the zero must be accompanied by their complex conjugates.

The poles in the denominator of the electrical notch filter are commonly placed in alignment with the corresponding zeros. However, the notch can be widened by placing the pole in a slightly different alignment. Such a recourse is appropriate when the mains frequency is unstable. Considerations of symmetry may then dictate that there should be a double zero on the unit circle flanked by two poles. If μ denotes a zero and κ denotes a pole, then this prescription would be met by setting

$$\mu_1, \mu_2 = e^{i\omega} \qquad \text{and} \qquad \kappa_1, \kappa_2 = \rho e^{i\omega \pm \epsilon} \quad \text{with} \quad 0 < \rho < 1, \tag{16}$$

where ω denotes the target frequency and ϵ denotes a small offset. The accompanying conjugate values are obtained by reversing the sign of the imaginary number i.

The concept of a notch filter with offset poles leads directly to the idea of a rational trend-removal filter of the form

$$\frac{\delta(z^{-1})}{\gamma(z^{-1})} = \frac{(1 - z^{-1})^2}{(1 - \kappa z^{-1})(1 - \kappa^* z^{-1})}, \tag{17}$$

where $\kappa = \rho \exp\{i\epsilon\}$ is a pole which may be specified in terms of its modulus ρ and its argument ϵ, and where $\kappa^* = \rho \exp\{-i\epsilon\}$ is its conjugate. To generate a phase-neutral filter, this function must be compounded with the function $\delta(z)/\gamma(z)$, which corresponds to the same filter applied in reversed time. Although only two parameters ρ and ϵ are involved, the search for an appropriate specification for the filter is liable to be difficult and time-consuming in the absence of a guiding design formula.

A notch filter, which has acquired considerable popularity amongst economists, and which depends on only one parameter, is given by the formula

$$\psi_N(z) = \frac{\delta(z)\delta(z^{-1})}{\gamma(z)\gamma(z^{-1})} = \frac{(1 - z)^2(1 - z^{-1})^2}{(1 - z)^2(1 - z^{-1})^2 + \lambda^{-1}}. \tag{18}$$

The placement of its poles and zeros within the complex plane is illustrated in Figure 5. The complement of the filter, which is specified by

$$\psi_P(z) = 1 - \psi_N(z) = \frac{\lambda^{-1}}{(1 - z)^2(1 - z^{-1})^2 + \lambda^{-1}} \tag{19}$$

is know to economists as the Hodrick–Prescott smoothing filter.

The filter was presented originally by Hodrick and Prescott (1980) in a widely circulated discussion paper. The paper was published as recently as (1997). Examples of the use of this filter have been provided by Kydland and Prescott (1990), King and Rebelo (1993) and by Cogley and Nason (1995).

The Hodrick–Prescott filter has an interesting heuristic. It transpires that it is the optimal estimator of the trajectory of a second-order random walk observed with error. Its single adjustable parameter λ^{-1} corresponds to the signal-to-noise ratio, which is the ratio of the variance of the white-noise process that drives the random walk and the variance of the error

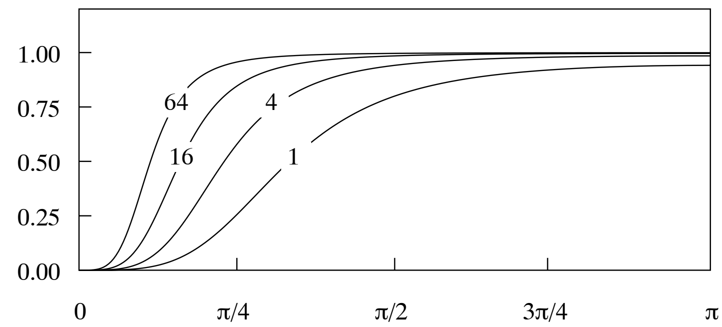

Figure 6. The frequency-response function of the notch filter ψ_N for various values of the smoothing parameter λ.

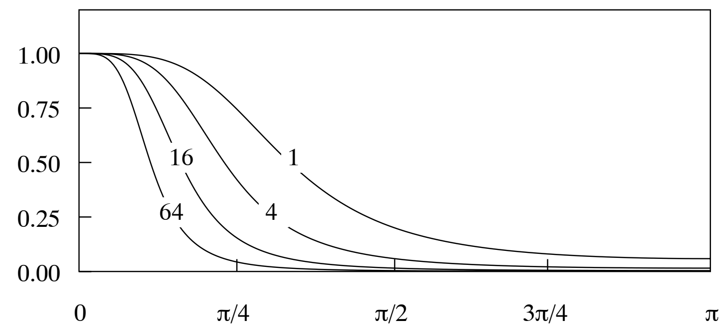

Figure 7. The frequency-response function of the Hodrick–Prescott smoothing filter ψ_P for various values of the smoothing parameter λ.

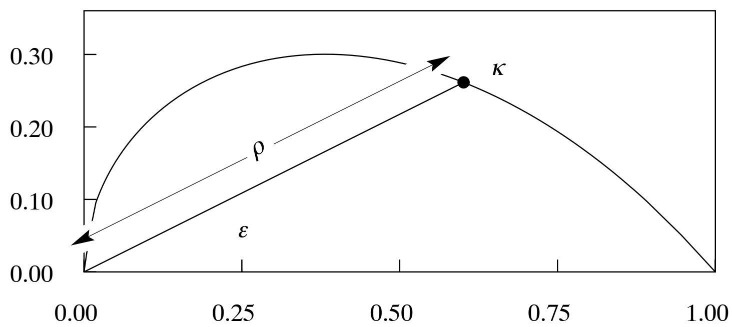

Figure 8. The trajectory in the complex plane of a pole of the norch filter ψ_N. The pole approaches $z = 1$ as $\lambda^{-1} \to 0$.

that obscures its observations. It is usual to describe λ as the smoothing parameter.

The filter is also closely related to the Reinsch (1976) smoothing spline, which is used extensively in industrial design. With the appropriate choice of the smoothing parameter, the latter represents the optimal estimator of the underlying trajectory of an integrated Wiener process observed with error.

The effect of increasing the value of λ in the formula for the smoothing filter is to reduce the range of the low-frequency components that are passed by the filter. The converse effect upon the notch filter is to reduce the width of the notch that impedes the passage of these components. These two effects are illustrated in Figures 6 and 7, which depict the frequency-response functions of the two filters. Figure 8 shows the trajectory of the poles of the filter as a function of the value of λ.

In order to implement either the smoothing filter or the notch filter, it is necessary to factorise their common denominator to obtain an expression for $\gamma(z)$. Since $z^2\gamma(z)\gamma(z^{-1})$ is a polynomial of degree four, one can, in principle, find analytic expressions for the poles which are in terms of the smoothing parameter λ. Alternatively, one may apply the iterative procedures which are used in the obtaining the Cramér–World factorisation of a Laurent polynomial. This is, in fact, how Figure 8 has been constructed.

The Hodrick–Prescott smoothing filter has been subjected to criticisms from several sources. In particular, it has been claimed—by Harvey and Jaeger (1993) amongst others—that thoughtless de-trending using the filter can lead investigators to detect spurious cyclical behaviour in economic data. The claim can only be interpreted to mean that, sometimes, the notch filter will pass cyclical components which ought to be impeded and attributed to the trend. One might say, in other words, that in such circumstances, the trend has been given a form which is too inflexible. This problem, which cannot be regarded as a general characteristic of the filter, arises from a mismatch of the chosen value of the smoothing parameter with the characteristics of the data series. However, it must be admitted that it is often difficult to find an appropriate value for the parameter.

A more serious shortcoming of the filter concerns the gradation between the stopband, which is the frequency range which is impeded by the filter, and the passband which is the frequency range where the components of a series are unaffected by the filter. This gradation may be too gentle for some purposes, in which case there can be no appropriate choice of value for the smoothing parameter.

In order to construct a frequency-selective filter which is accurately attuned to the characteristics of the data, and which can discriminate adequately between the trend and the residue, a more sophisticated methodology may be called for. We shall attempt to provide this in the ensuing sections of the paper.

4. Rational Square-Wave Filters

In the terminology of digital signal processing, an ideal frequency-selective filter is one for which the frequency response is unity over a certain range of frequencies, described as the passband, and zero over the remaining frequencies, which constitute the stopband. In a lowpass filter ψ_L, the passband covers a frequency interval $[0, \omega_c]$ ranging from zero to a cut-off point. In the complementary highpass filter ψ_H, it is the stopband which stands on this interval. Thus

$$|\psi_L(e^{i\omega})| = \begin{cases} 1, & \text{if } \omega < \omega_c \\ 0, & \text{if } \omega > \omega_c \end{cases} \quad \text{and} \quad |\psi_H(e^{i\omega})| = \begin{cases} 0, & \text{if } \omega < \omega_c \\ 1, & \text{if } \omega > \omega_c. \end{cases} \tag{20}$$

In this section, we shall derive a pair of complementary filters that fulfil this specification approximately for a cut-off frequency of $\omega_c = \pi/2$. Once we have designed these prototype filters, we shall be able to apply a transformation that shifts the cut-off point from $\omega = \pi/2$ to any other point $\omega_c \in [0, \pi]$.

The idealised conditions of (20), which define a periodic square wave, are impossible to fulfil in practice. In fact, the Fourier transform of the square wave is an indefinite sequence of coefficients defined over the positive and negative integers; and, in constructing a practical moving-average filter, only a limited number of central coefficient can be taken. In such a filter, the sharp disjunction between the passband and the stopband, which characterises the ideal filter, is replaced by a gradual transition. The cost of a more rapid transition is bound to be an increased number of coefficients.

A preliminary step in designing a pair of complementary filters is to draw up a list of specifications that can be fulfilled in practice. We shall be guided by the following conditions:

$$\begin{array}{lll} \text{(i)} & \psi_L(z) + \psi_H(z) = 1, & \textit{Complementarity} \\ \text{(ii)} & \psi_L(-z) = \psi_H(z), \quad \psi_H(-z) = \psi_L(z), & \textit{Symmetry} \\ \text{(iii)} & \psi_L(z^{-1}) = \psi_L(z), \quad \psi_H(z^{-1}) = \psi_H(z), & \textit{Phase-Neutrality} \\ \text{(iv)} & |\psi_L(1)| = 1, \quad |\psi_L(-1)| = 0, & \textit{Lowpass Conditions} \\ \text{(v)} & |\psi_H(1)| = 0, \quad |\psi_H(-1)| = 1. & \textit{Highpass Conditions} \end{array} \tag{21}$$

There is no reference here to the rate of the transition from the passband to the stopband. In fact, the condition under (iv) and (v) refer only to the end points of the frequency range $[0, \pi]$, which are the furthest points from the cut-off.

Observe that the symmetry condition $\psi_L(-z) = \psi_H(z)$ under (ii) necessitates placing the cut-off frequency at $\omega_c = \pi/2$. The condition implies that, when it is reflected about the axis of $\omega_c = \pi/2$, the frequency response of the lowpass filter becomes the frequency response of the highpass filter. This feature is illustrated by Figure 10.

It will be found that all of the conditions of (21) are fulfilled by the highpass differencing filter ψ_D defined under (11) in conjunction with the complementary lowpass smoothing filter $\psi_S = 1 - \psi_D$ defined under (14). However, we have already rejected ψ_D and ψ_S on the grounds that their transitions between the passband to the stopband are too gradual.

In order to minimise the problem of spectral leakage whilst maintaining a transition that is as rapid as possible, we now propose to fulfil the conditions of (21) via a pair of rational functions that take the forms of

$$\psi_L(z) = \frac{\delta_L(z)\delta_L(z^{-1})}{\gamma(z)\gamma(z^{-1})} \qquad \text{and} \qquad \psi_H(z) = \frac{\delta_H(z)\delta_H(z^{-1})}{\gamma(z)\gamma(z^{-1})}. \tag{22}$$

The condition of phase neutrality under (iii) is automatically satisfied by these forms. We propose to satisfy the lowpass and highpass conditions under (iv) and (v) by specifying that

$$\delta_L(z) = (1+z)^n \qquad \text{and} \qquad \delta_H(z) = (1-z)^n. \tag{23}$$

Similar specifications are also to be found in the binomial filter ψ_B of (15) and in the notch filter ψ_N of (18).

Given the specifications under (22), it follows that the symmetry condition of (ii) will be satisfied if and only if every root of $\gamma(z) = 0$ is a purely imaginary number. It follows from (i) that the polynomial $\gamma(z)$ must fulfil the condition that

$$\gamma(z)\gamma(z^{-1}) = \delta_L(z)\delta_L(z^{-1}) + \delta_H(z)\delta_H(z^{-1}). \tag{24}$$

On putting the specifications of (23) and (24) into (22), we find that

$$\begin{aligned}\psi_L(z) &= \frac{(1+z)^n(1+z^{-1})^n}{(1+z)^n(1+z^{-1})^n + (1-z)^n(1-z^{-1})^n} \\ &= \frac{1}{1 + \left(i\dfrac{1-z}{1+z}\right)^{2n}}\end{aligned} \tag{25}$$

and that

$$\begin{aligned}\psi_H(z) &= \frac{(1-z)^n(1-z^{-1})^n}{(1+z)^n(1+z^{-1})^n + (1-z)^n(1-z^{-1})^n} \\ &= \frac{1}{1+\left(i\dfrac{1+z}{1-z}\right)^{2n}}.\end{aligned} \tag{26}$$

These will be recognised as instances of the Butterworth filter, which is familiar in electrical engineering—see, for example, Roberts and Mullis (1987).

The Butterworth filter, in common with the Hodrick–Prescott filter can also be derived by applying the Wiener–Kolmogorov theory of signal extraction to an appropriate statistical model. In that context, the filter represents a device for obtaining the minimum-mean-square-error estimate of the component in question. See Kolmogorov (1941) and Wiener (1950) for the original expositions of the theory and Whittle (1983) for a modern account.

A defining characteristic of the Wiener–Kolmogorov filters is the condition of complementarity of (21) (i). On that basis, we might also regard the complementary binomial filters $\psi_D(z)$ and $\psi_S(z)$ of (11) and (15), respectively, as Wiener–Kolmogorov filters; but they are unusual in being represented by polynomials of finite degree, whereas filters of this class are more commonly represented by rational functions.

Since $\delta_L(z)$ and $\delta_H(z)$ are now completely specified, it follows that $\gamma(z)$ can be determined via the Cramér–Wold factorisation of the polynomial of the RHS of (24). However, it is relatively straightforward to obtain analytic expressions for the roots of the equation $\gamma(z)\gamma(z^{-1}) = 0$. The roots come in reciprocal pairs; and, once they are available, they may be assigned unequivocally to the factors $\gamma(z)$ and $\gamma(z^{-1})$. Those roots which lie outside the unit circle belong to $\gamma(z)$ whilst their reciprocals, which lie inside the unit circle, belong to $\gamma(z^{-1})$. Therefore, consider the equation

$$(1+z)^n(1+z^{-1})^n + (1-z)^n(1-z^{-1})^n = 0, \tag{27}$$

which is equivalent to the equation

$$1+\left(i\frac{1-z}{1+z}\right)^{2n} = 0. \tag{28}$$

Solving the latter for

$$s = i\frac{1-z}{1+z} \tag{29}$$

is a matter of finding the $2n$ roots of -1. These are given by

$$s = \exp\left\{\frac{i\pi j}{2n}\right\}, \quad \text{where} \quad j = 1, 3, 5, \ldots, 4n-1, \quad \text{or} \quad j = 2k-1;\ k = 1, \ldots, 2n. \tag{30}$$

The roots correspond to a set of $2n$ points which are equally spaced around the circumference of the unit circle. The radii, which join the points to the centre, are separated by angles of π/n; and the first of the radii makes an angle of $\pi/(2n)$ with the horizontal real axis.

The inverse of the function $s = s(z)$ is the function

$$z = \frac{i-s}{i+s} = \frac{i(s+s^*)}{2-i(s-s^*)}. \tag{31}$$

Here, the final expression comes from multiplying top and bottom of the second expression by $s^* - i = (i+s)^*$, where s^* denotes the conjugate of the complex number s, and from noting that $ss^* = 1$. On substituting the expression for s from (29), it is found that the solutions of (28) are given, in terms of z, by

$$z_k = i\frac{\cos\{\pi(2k-1)/2n\}}{1+\sin\{\pi(2k-1)/2n\}}, \quad \text{where} \quad k = 1, \ldots, 2n. \tag{32}$$

The roots of $\gamma(z^{-1}) = 0$ are generated when $k = 1, \ldots, n$. Those of $\gamma(z) = 0$ are generated when $k = n+1, \ldots, 2n$.

Figure 9 shows the disposition in the complex plane of the poles and zeros of the prototype lowpass filter $\psi(z)_L$ for the case where $n = 6$, whilst Figure 10 shows the gain of this filter together with that of the complementary filter $\psi(z)_H$.

5. Frequency Transformations

The object of the filter $\psi_L(z)$ is to remove from a time series a set of trend components whose frequencies range from $\omega = 0$ to a cut-off value of $\omega = \omega_c$. The prototype version of the filter has a cut-off at the frequency $\omega = \pi/2$. In order to convert the prototype filter to one that will serve the purpose, a means must be found for mapping the frequency interval $[0, \pi/2]$ into the interval $[0, \omega_c]$. This can be achieved by replacing the argument z, wherever it occurs in the filter formula, by the argument

$$g(z) = \frac{z-\alpha}{1-\alpha z}, \tag{33}$$

where $\alpha = \alpha(\omega_c)$ is an appropriately specified parameter.

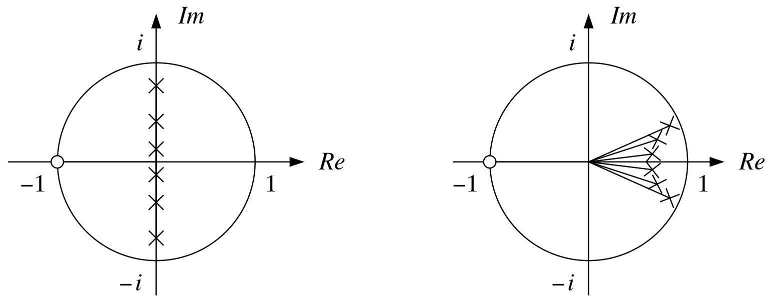

Figure 9. The pole–zero diagrams of the lowpass square-wave filters for $n = 6$ when the cut-off is at $\omega = \pi/2$ (left) and at $\omega = \pi/8$.

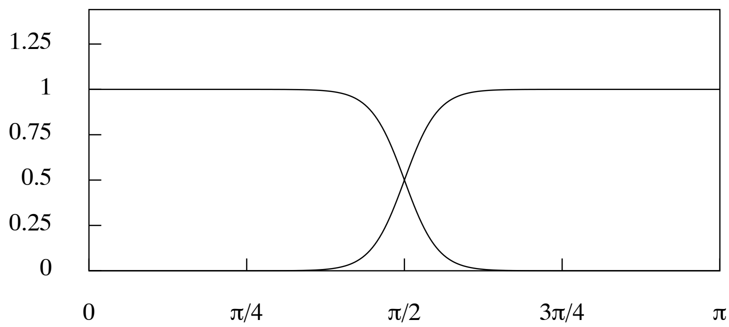

Figure 10. The frequency-responses of the prototype square-wave filters with $n = 6$ and with a cut-off at $\omega = \pi/2$.

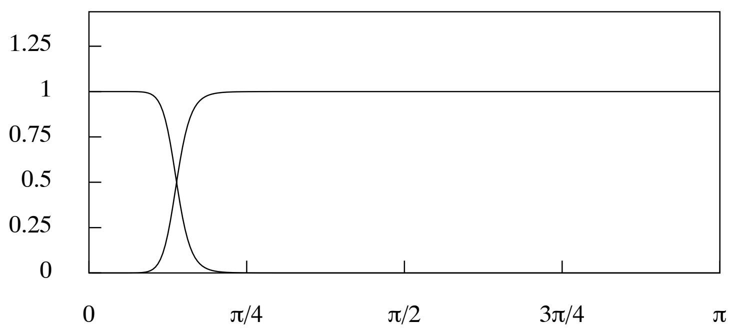

Figure 11. The frequency-responses of the square-wave filters with $n = 6$ and with a cut-off at $\omega = \pi/8$.

The function $g(z)$ fulfils the following conditions:

$$\begin{aligned}
&\text{(i)} && g(z)g(z^{-1}) = 1, \\
&\text{(ii)} && g(z) = z \quad \text{if} \quad \alpha = 0, \\
&\text{(iii)} && g(1) = 1 \quad \text{and} \quad g(-1) = -1, \\
&\text{(iv)} && \text{Arg}\{g(z)\} \geq \text{Arg}\{z\} \quad \text{if} \quad \alpha > 1, \\
&\text{(v)} && \text{Arg}\{g(z)\} \leq \text{Arg}\{z\} \quad \text{if} \quad \alpha < 1.
\end{aligned} \tag{34}$$

The conditions (i) and (ii) indicate that, if $g(z) \neq z$, then the modulus of the function is invariably unity. Thus, as z encircles the origin, $g = g(z)$ travels around the unit circle. The conditions of (iii) indicate that, if $z = e^{i\omega}$ travels around the unit circle, then g and z will coincide when $\omega = 0$ and when $\omega = \pi$—which are the values that bound the positive frequency range over which the transfer function of the filter is defined. Finally, conditions (iv) and (v) indicate that, if $g \neq z$, then g either leads z uniformly or lags behind it as the two travel around the unit circle from $z = 1$ to $z = -1$.

The value of α is completely determined by any pair of corresponding values for g and z. Thus, from (33), it follows that

$$\begin{aligned}
\alpha &= \frac{z - g}{1 - gz} \\
&= \frac{g^{1/2}z^{-1/2} - g^{-1/2}z^{1/2}}{g^{1/2}z^{1/2} - g^{-1/2}z^{-1/2}}.
\end{aligned} \tag{35}$$

Imagine that the cut-off of a prototype filter is at $\omega = \theta$ and that it is desired to shift it to $\omega = \kappa$. Then $z = e^{i\kappa}$ and $g = e^{i\theta}$ will be corresponding values; and the appropriate way of shifting the frequency would be to replace the argument z within the filter formula by the function $g(z)$ wherein the parameter α is specified by

$$\begin{aligned}
\alpha &= \frac{e^{i(\theta-\kappa)/2} - e^{-i(\theta-\kappa)/2}}{e^{i(\theta+\kappa)/2} - e^{-i(\theta+\kappa)/2}} \\
&= \frac{\sin\{(\theta-\kappa)/2\}}{\sin\{(\theta+\kappa)/2\}}.
\end{aligned} \tag{36}$$

To find an explicit form for the transformed filter, we may begin by observing that, when $g(z)$ is defined by equation (33), we have

$$\frac{1 - g(z)}{1 + g(z)} = \left\{\frac{1+\alpha}{1-\alpha}\right\}\left\{\frac{1-z}{1+z}\right\}. \tag{37}$$

Here there is

$$\frac{1+\alpha}{1-\alpha} = \frac{\sin\{(\theta+\kappa)/2\} + \sin\{(\theta-\kappa)/2\}}{\sin\{(\theta+\kappa)/2\} - \sin\{(\theta-\kappa)/2\}} \tag{38}$$

$$= \frac{\sin(\theta/2)\cos(\kappa/2)}{\cos(\theta/2)\sin(\kappa/2)}.$$

In the prototype filter, we are setting $\theta = \pi/2$ and, in the transformed filter, we are setting $\kappa = \omega_c$, which is the cut-off frequency. The result of these choices is that

$$\frac{1+\alpha}{1-\alpha} = \frac{1}{\tan(\omega_c/2)}. \tag{39}$$

It follows that the lowpass filter with a cut-off at ω_c takes the form of

$$\begin{aligned} \psi_L(z) &= \frac{1}{1+\lambda\left(i\dfrac{1-z}{1+z}\right)^{2n}} \\ &= \frac{(1+z)^n(1+z^{-1})^n}{(1+z)^n(1+z^{-1})^n + \lambda(1-z)^n(1-z^{-1})^n}, \end{aligned} \tag{40}$$

where $\lambda = \{1/\tan(\omega_c)\}^{2n}$. The same reasoning shows that the highpass filter with a cut-off at ω_c takes the form of

$$\begin{aligned} \psi_H(z) &= \frac{1}{1+\dfrac{1}{\lambda}\left(i\dfrac{1+z}{1-z}\right)^{2n}} \\ &= \frac{\lambda(1-z)^n(1-z^{-1})^n}{(1+z)^n(1+z^{-1})^n + \lambda(1-z)^n(1-z^{-1})^n}. \end{aligned} \tag{41}$$

In applying the frequency transformation to the prototype filter, we are also concerned with finding revised values for the poles. The conditions under (iii) indicate that the locations of the zeros will not be affected by the transformation. Only the poles will be altered. Consider, therefore, the generic factor within the denominator of the prototype. This is $z - i\rho$, where $i\rho$ is one of the poles specified under (30). Replacing z by $g(z)$ and setting the result to zero gives the following condition:

$$\frac{z-\alpha}{1-\alpha z} - i\rho = 0. \tag{42}$$

This indicates that the pole at $z = \rho$ will be replaced by a pole at

$$z = \frac{\alpha + i\rho}{1 + i\rho\alpha} = \frac{\alpha(1-\rho^2) + i\rho(1-\alpha^2)}{1-\rho^2\alpha^2}, \tag{43}$$

where the final expression comes from multiplying top and bottom of its predecessor by $1 - i\rho\alpha$.

Figure 11, displays the pole-zero diagram of the prototype filter and of a filter with a cut-off frequency of $\pi/8$. It also suggests that one of the effects of a frequency transformation may be to bring some of poles closer to the perimeter of the unit circle. This can lead to stability problems in implementing the filter, and it is liable to prolong the transient effects of ill-chosen start-up conditions.

6. Implementing the Filters

The classical signal-extraction filters are intended to be applied to lengthy data sets. The task of adapting them to limited samples often causes difficulties and perplexity. The problems arise from not knowing how to supply the initial conditions with which to start a recursive filtering process. By choosing inappropriate starting values for the forwards or the backwards pass, one can generate a so-called transient effect, which is liable, in fact, to affect all of the processed values.

Of course, when the values of interest are remote from either end of a long sample, one can trust that they will be barely affected by the start-up conditions. However, in many applications, such as in the processing of economic data, the sample is short and the interest is concentrated at the upper end where the most recent observations are to be found.

One approach to the problem of the start-up conditions relies upon the ability to extend the sample by forecasting and backcasting. The additional exta-sample values can be used in a run-up to the filtering process wherein the filter is stabilised by providing it with a plausible history, if it is working in the direction of time, of with a plausible future, if it is working in reversed time. Sometimes, very lengthy extrapolations are called for—see Burman (1980), for example.

The approach that we shall adopt in this paper is to avoid the start-up problem altogether by deriving specialised finite-sample versions of the filters on the basis of the statistical theory of conditional expectations.

Some of the more successful methods for treating the problem of the start-up conditions that have been proposed have arisen within the context of the Kalman filter and the associated smoothing algorithms—see Ansley and Kohn (1985), De Jong (1991), and Durbin and Koopman (2001), for example. The context of the Kalman filter is a wide one; and it seems that the necessary results can be obtained more easily by restricting the context.

Let us begin, therefore, by considering a specific model for which the square-wave filter would represent the optimal device for extracting the signal, given a sample of infinite length. The model is represented by the equation

$$\begin{aligned} y(t) &= \xi(t) + \eta(t) \\ &= \frac{(1+L)^n}{(1-L)^2}\nu(t) + (1-L)^{n-2}\varepsilon(t), \end{aligned} \tag{44}$$

where $\nu(t)$ and $\varepsilon(t)$ are statistically independent sequences generated by normal white-noise processes. This can be rewritten as

$$\begin{aligned}(1-L)^2 y(t) &= (1+L)^n \nu(t) + (1-L)^n \varepsilon(t) \\ &= \zeta(t) + \kappa(t),\end{aligned} \tag{45}$$

where $\zeta(t) = (1-L)^2\xi(t) = (1+L)^n\nu(t)$ and $\kappa(t) = (1-L)^2\eta(t) = (1-L)^n\varepsilon(t)$ both follow noninvertible moving-average processes.

The statistical theory of signal extraction, as expounded by Whittle (1983), for example, indicates that the lowpass filter $\psi_L(z)$ of equation (40) will generate the minimum mean-square-error estimate of the sequence $\xi(t)$, provided that the smoothing parameter has the value of $\lambda = \sigma_\varepsilon^2/\sigma_\nu^2$. The theory also indicates that the Hodrick–Prescott filter will generate the optimal estimate in the case where $\xi(t)$ is a second-order random walk and $\eta(t)$ is a white-noise process:

$$\begin{aligned}y(t) &= \xi(t) + \eta(t) \\ &= \frac{1}{(1-L)^2}\nu(t) + \eta(t).\end{aligned} \tag{46}$$

Now imagine that there are T observations of the process $y(t)$ of equation (44), which run from $t = 0$, to $t = T-1$. These are gathered in a vector

$$y = \xi + \eta. \tag{47}$$

To find the finite-sample the counterpart of equation (45), we need to represent the second-order difference operator $(1-L)^2$ in the form of a matrix. The matrix that finds the differences $d_2, \ldots, d_{T-1}$ of the data points $y_0, y_1, y_2, \ldots, y_{T-1}$ is in the form of

$$Q' = \begin{bmatrix} 1 & -2 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & -2 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & 0 & 0 & \ldots & -2 & 1 \end{bmatrix}. \tag{48}$$

Premultiplying equation (47) by this matrix gives

$$\begin{aligned}d = Q'y &= Q'\xi + Q'\eta \\ &= \zeta + \kappa,\end{aligned} \tag{49}$$

where $\zeta = Q'\xi$ and $\kappa = Q'\eta$. The first and second moments of the vector ζ may be denoted by

$$E(\zeta) = 0 \qquad \text{and} \qquad D(\zeta) = \sigma_\nu^2 M, \tag{50}$$

and those of κ by

$$E(\kappa) = 0 \qquad \text{and} \qquad \begin{aligned} D(\kappa) &= Q'D(\eta)Q \\ &= \sigma_\varepsilon^2 Q'\Sigma Q, \end{aligned} \tag{51}$$

where both M and $Q'\Sigma Q$ are symmetric Toeplitz matrices with $2n+1$ nonzero diagonal bands. The generating functions for the coefficients of these matrices are, respectively, $\delta_L(z)\delta_L(z^{-1})$ and $\delta_H(z)\delta_H(z^{-1})$, where $\delta_L(z)$ and $\delta_H(z)$ are the polynomials defined in (23).

The optimal predictor z of the twice-differenced signal vector $\zeta = Q'\xi$ is given by the following conditional expectation:

$$\begin{aligned} E(\zeta|d) &= E(\zeta) + C(\zeta, d)D^{-1}(d)\{d - E(d)\} \\ &= M(M + \lambda Q'\Sigma Q)^{-1}d = z, \end{aligned} \tag{52}$$

where $\lambda = \sigma_\varepsilon^2/\sigma_\nu^2$. The optimal predictor k of the twice-differenced noise vector $\kappa = Q'\eta$ is given, likewise, by

$$\begin{aligned} E(\kappa|d) &= E(\kappa) + C(\kappa, d)D^{-1}(d)\{d - E(d)\} \\ &= \lambda Q'\Sigma Q(M + \lambda Q'\Sigma Q)^{-1}d = k. \end{aligned} \tag{53}$$

It may be confirmed that $z + k = d$.

The estimates are calculated, first, by solving the equation

$$(M + \lambda Q'\Sigma Q)g = d \tag{54}$$

for the value of g and, thereafter, by finding

$$z = Mg \qquad \text{and} \qquad k = \lambda Q'\Sigma Qg. \tag{55}$$

The solution of equation (54) is found via a Cholesky factorisation which sets $M + \lambda Q'\Sigma Q = GG'$, where G is a lower-triangular matrix. The system $GG'g = d$ may be cast in the form of $Gh = d$ and solved for h. Then $G'g = h$ can be solved for g.

There is a straightforward correspondence between the finite-sample implementations of the filter and the formulations that assume an infinite sample. In terms of the lag-operator polynomials, equation (54) would be rendered as

$$\begin{aligned} \gamma(F)\gamma(L)g(t) &= d(t), \qquad \text{where} \\ \gamma(F)\gamma(L) &= \delta_L(F)\delta_L(L) + \lambda\delta_H(F)\delta_H(L). \end{aligned} \tag{56}$$

The process of solving equation (54) via a Cholesky decomposition corresponds to the application of the filter in separate passes running forwards and backwards in time respectively:

$$\text{(i)} \quad \gamma(L)f(t) = d(t) \qquad \text{(ii)} \quad \gamma(F)g(t) = f(t). \tag{57}$$

The coefficients of successive rows of the Cholesky factor G converge upon the values of the coefficients of $\gamma(z)$; and, at some point, it may become appropriate to use the latter instead. This will save computer time and computer memory.

The two equations under (55) correspond respectively to

$$z(t) = \delta_L(F)\delta_L(L)g(t) \qquad \text{and} \qquad k(t) = \delta_H(F)\delta_H(L)q(t). \tag{58}$$

Our object is to recover from z an estimate x of the trend vector ξ. This would be conceived, ordinarily, as a matter of integrating the vector z twice via a simple recursion which depends upon two initial conditions. The difficulty is in discovering the appropriate initial conditions with which to begin the recursion.

We can circumvent the problem of the initial conditions by seeking the solution to the following problem:

$$\text{Minimise} \qquad (y-x)'\Sigma^{-1}(y-x) \qquad \text{Subject to} \qquad Q'x = z. \tag{59}$$

The problem is addressed by evaluating the Lagrangean function

$$L(x,\mu) = (y-x)'\Sigma^{-1}(y-x) + 2\mu'(Q'x - z). \tag{60}$$

By differentiating the function with respect to x and setting the result to zero, we obtain the condition

$$\Sigma^{-1}(y-x) - Q\mu = 0. \tag{61}$$

Premultiplying by $Q'\Sigma$ gives

$$Q'(y-x) = Q'\Sigma Q\mu. \tag{62}$$

But, from (54) and (55), it follows that

$$\begin{aligned} Q'(y-x) &= d - z \\ &= \lambda Q'\Sigma Q g, \end{aligned} \tag{63}$$

whence we get

$$\begin{aligned} \mu &= (Q'\Sigma Q)^{-1}Q'(y-x) \\ &= \lambda g. \end{aligned} \tag{64}$$

Putting the final expression for μ into (61) gives

$$x = y - \lambda\Sigma Q g. \tag{65}$$

This is our solution to the problem of estimating the trend vector ξ. Notice that there is no need to find the value of z explicitly, since the value of x can be expressed more directly in terms of $g = \Sigma^{-1}z$.

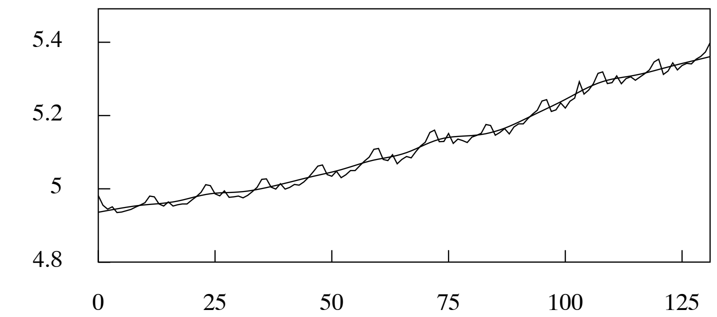

Figure 12. The data on the U.S. money stock with an interpolated trend estimated by a lowpass square-wave filter with $n = 6$ and a cut off at $\omega = \pi/8$.

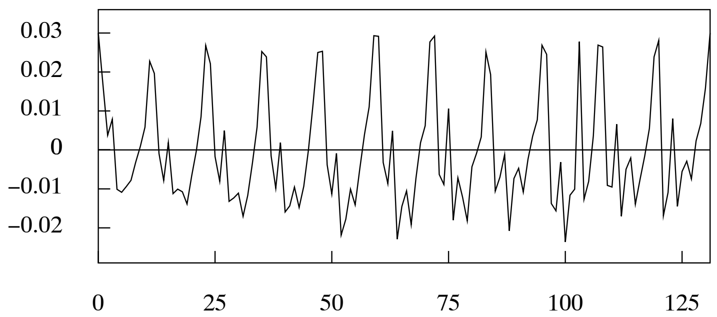

Figure 13. The residual sequence obtained by detrending the logarithm of the money stock data with a square-wave filter.

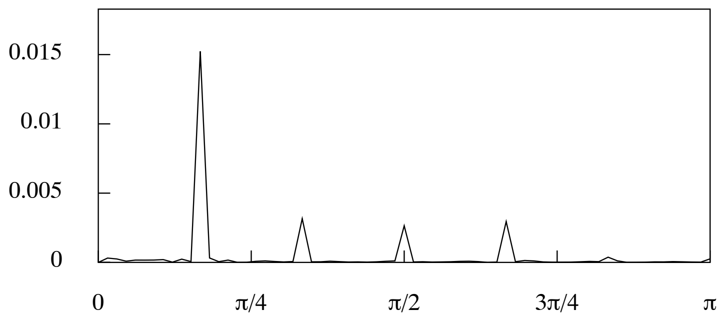

Figure 14. The periodogram of the residuals from detrending the logarithm of the U.S. money stock data.

It is notable that there is a criterion function which will enable us to derive the equation of the trend estimation filter in a single step. The function is

$$L(x) = (y - x)'\Sigma^{-1}(y - x) + \lambda x'QM^{-1}Q'x, \tag{66}$$

wherein $\lambda = \sigma_\varepsilon^2/\sigma_\nu^2$ as before. This is minimised by the value specified in (65). The criterion function becomes intelligible when we allude to the assumptions that $y \sim N(\xi, \sigma_\varepsilon^2\Sigma)$ and that $Q'\xi = \zeta \sim N(0, \sigma_\nu^2 M)$; for then it plainly resembles a combination of two independent chi-square variates.

The effect of the square-wave filter is illustrated in Figures 12–14 which depict the detrending of the logarithmic series of the U.S. money stock. It is notable that, in contrast to periodogram of Figure 3, which relates to the the residuals from fitting a polynomial trend, the periodogram of Figure 14 shows virtually no power in the range of frequencies below that of the principal seasonal frequency.

We should point out that our derivation and the main features of our algorithm are equally applicable to the task of implementing the Hodrick–Prescott (H–P) filter and the Reinsch smoothing spline. In the case of the H–P filter, we need only replace the matrices Σ and M in the equations above by the matrices I and $Q'Q$ respectively. Then equation (52) becomes

$$(I + \lambda Q'Q)^{-1}d = z, \tag{67}$$

whilst equation (65), which provides the estimate of the signal or trend, becomes

$$x = y - \lambda Qz. \tag{68}$$

References

Ansley, C.F., and R. Kohn, (1985), Estimation, Filtering and Smoothing in State Space Models with Incompletely Specified Initial Conditions, *The Annals of Statistics,* **13,** 1286–1316.

Burman, J.P., (1980), Seasonal Adjustment by Signal Extraction, *Journal of the Royal Statistical Society, Series A,* **143,** 321–337.

Cogley, T., and J.M. Nason, (1995), Effects of the Hodrick–Prescott Filter on Trend and Difference Stationary Time Series, Implications for Business Cycle Research, *Journal of Economic Dynamics and Control,* **19,** 253–278.

De Jong, P., (1991), The Diffuse Kalman Filter, *The Annals of Statistics,* **19,** 1073–1083.

Durbin, J., and S.J. Koopman, (2001), *Time Series Analysis by State Space Methods,* Oxford University Press.

Harvey, A.C., and A. Jaeger, (1993), Detrending, Stylised Facts and the Business Cycle, *Journal of Applied Econometrics,* **8,** 231–247.

Haykin, S., (1989), *Modern Filters, Macmillan Publishing Company,* New York.

Hodrick, R.J., and E.C Prescott, (1980), *Postwar U.S. Business Cycles: An Empirical Investigation,* Working Paper, Carnegie–Mellon University, Pittsburgh, Pennsylvania.

Hodrick R.J., and Prescott, E.C., (1997), Postwar U.S. business bycles: An Empirical Investigation, Journal of Money, Credit and Banking, 29, 1–16.

King, R.G., and S.G. Rebelo, (1993), Low Frequency Filtering and Real Business Cycles, *Journal of Economic Dynamics and Control,* **17,** 207–231.

Kolmogorov, A.N., (1941), Interpolation and Extrapolation. *Bulletin de l'academie des sciences de U.S.S.R.,* Ser. Math., 5, 3–14.

Kydland, F.E., and C. Prescott, (1990), Business Cycles: Real Facts and a Monetary Myth, *Federal Reserve Bank of Minneapolis Quarterly Review,* **14,** 3–18.

Oppenheim A.V. and R.W. Schafer, (1989), *Discrete-Time Signal Processing,* Prentice-Hall, Englewood Cliffs, New Jersey.

Pollock, D.S.G., (1999), *Time-Series Analysis, Signal Processing and Dynamics,* The Academic Press, London.

Reinsch, C.H., (1976), Smoothing by Spline Functions, *Numerische Mathematik,* **10,** 177–183.

Roberts, R.A., and C.T. Mullis, (1987), *Digital Signal Processing,* Addison Wesley, Reading, Massachusetts.

Whittle, P., (1983), *Prediction and Regulation by Linear Least-Square Methods, Second Revised Edition,* Basil Blackwell, Oxford.

Wiener, N., (1950), *Extrapolation, Interpolation and Smoothing of Stationary Time Series,* MIT Technology Press, John Wiley and Sons, New York.