STATISTICAL SIGNAL EXTRACTION AND FILTERING:

by D.S.G. Pollock

Linear Time-Invariant (LTI) Filters

An LTI filter entails a linear combination of successive elements of a discrete-time signal $x(t) = \{x_t; t = \pm 1, \pm 2, \ldots\}$. The filter's output is

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

By associating z^t to each element y_t and by summing the sequence, we get

$$\sum_t y_t z^t = \sum_t \Big\{ \sum_j \psi_j x_{t-j} \Big\} z^t \quad \text{or} \quad y(z) = \psi(z)x(z), \tag{2}$$

where the constituent z-transforms are

$$x(z) = \sum_t x_t z^t, \quad y(z) = \sum_t y_t z^t \quad \text{and} \quad \psi(z) = \sum_j \psi_j z^j. \tag{3}$$

The Impulse Response

The sequence $\{\psi_j\}$ of the filter's coefficients is its impulse reponse. Finite moving averages are called fini impulse-response (FIR) filters. When the impulse response has an indefinite duration, there is an infinite impulse-response (IIR) filter. A filter is causal or backward-looking if none of its coefficients is associated with a negative power of z.

Causal Filters

A practical filter must comprise only a finite number of distinct parameters. Linear IIR filters that are causal entail recursive equations

$$\sum_{j=0}^{g} \phi_j y_{t-j} = \sum_{j=0}^{d} \theta_j x_{t-j}, \quad \text{with} \quad \phi_0 = 1, \tag{4}$$

of which the z-transform is

$$\phi(z)y(z) = \theta(z)x(z), \tag{5}$$

wherein $\phi(z) = \phi_0 + \phi_1 z + \cdots + \phi_p z^p$ and $\theta(z) = \theta_0 + \theta z + \cdots + \theta_q z^q$. Setting $\phi_0 = 1$ identifies $y(t)$ as the output.

The lagged values of the output constitute feedback.

The recursive equation may be assimilated to the equation under (2) by writing it in rational form:

$$y(z) = \frac{\theta(z)}{\phi(z)}x(z) = \psi(z)x(z). \tag{6}$$

On the condition that the filter is stable, the expression $\psi(z)$ stands for the series expansion of the ratio of the polynomials.

The Series Expansion of a Rational Transfer Function

The method of finding the coefficients of the series expansion can be illustrated by the second-order case:

$$\frac{\theta_0 + \theta_1 z}{\phi_0 + \phi_1 z + \phi_2 z^2} = \left\{\psi_0 + \psi_1 z + \psi_2 z^2 + \cdots\right\}. \tag{7}$$

We rewrite this equation as

$$\theta_0 + \theta_1 z = \left\{\phi_0 + \phi_1 z + \phi_2 z^2\right\}\left\{\psi_0 + \psi_1 z + \psi_2 z^2 + \cdots\right\}. \tag{8}$$

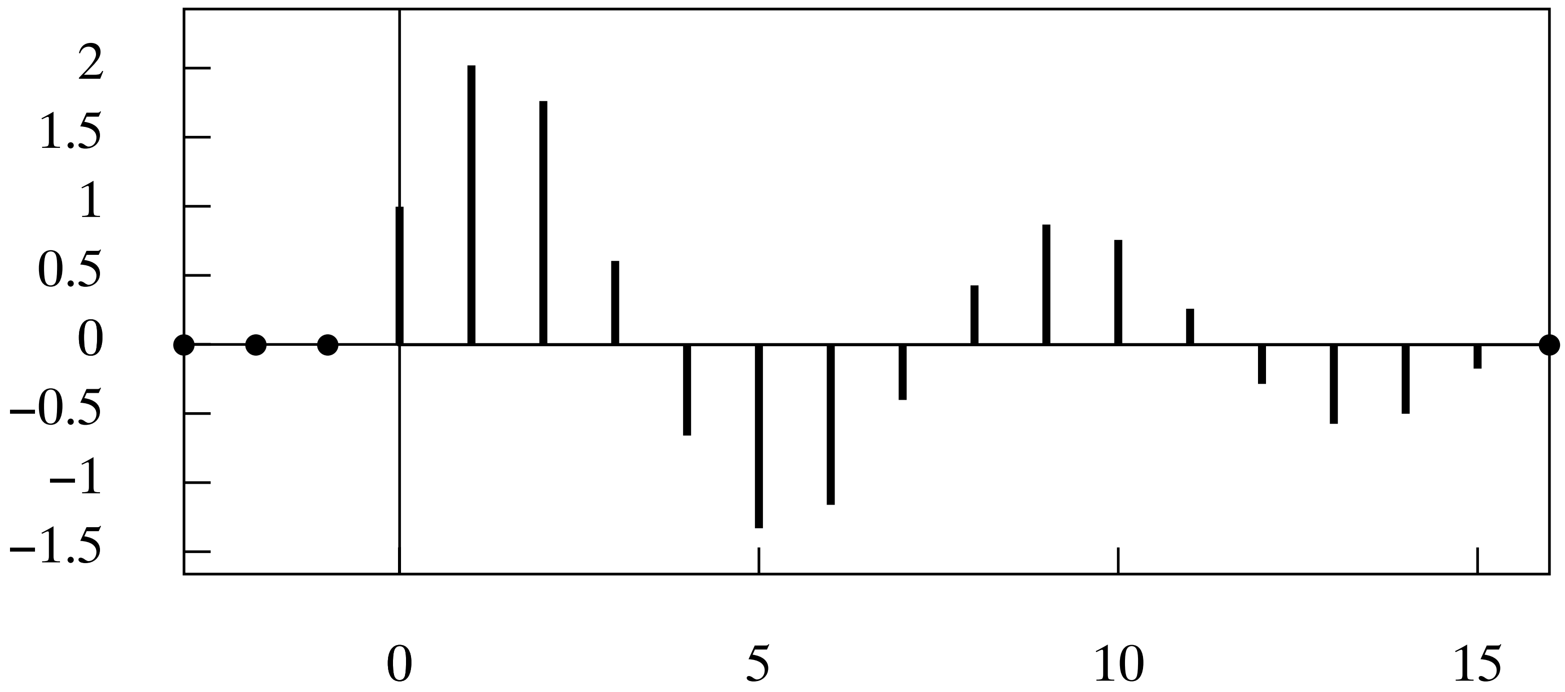

Figure 1. The impulse response of the transfer function $\theta(z)/\phi(z)$ with $\phi(z) = 1.0 - 1.2728z + 0.81z^2$ and $\theta(z)(z) = 1.0 + 0.75z$.

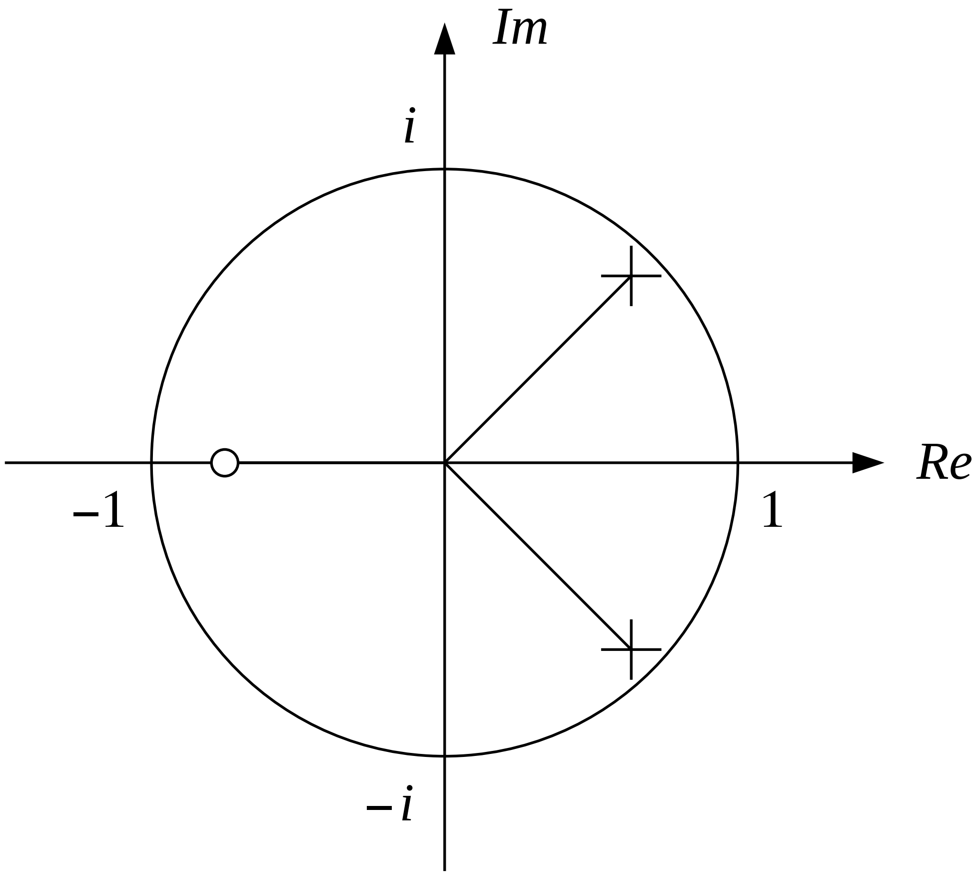

Figure 2. The pole–zero diagram corresponding to the transfer function of Figure 1. The conjugate complex poles have arguments of $\pm\pi/4$ and a modulus of 0.9. The single real-valued zero has the value of -0.75.

The following table assists us in multiplying together the two polynomials:

$$
(9) \qquad \begin{array}{c|cccc}
 & \psi_0 & \psi_1 z & \psi_2 z^2 & \cdots \\
\hline
\phi_0 & \phi_0\psi_0 & \phi_0\psi_1 z & \phi_0\psi_2 z^2 & \cdots \\
\phi_1 z & \phi_1\psi_0 z & \phi_1\psi_1 z^2 & \phi_1\psi_2 z^3 & \cdots \\
\phi_2 z^2 & \phi_2\psi_0 z^2 & \phi_2\psi_1 z^3 & \phi_2\psi_2 z^4 & \cdots
\end{array}
$$

Performing the multiplication on the RHS of the equation, and by equating the coefficients of the same powers of z on the two sides, we find that

$$
\begin{aligned}
\theta_0 &= \phi_0\psi_0, & \psi_0 &= \theta_0/\phi_0, \\
\theta_1 &= \phi_0\psi_1 + \phi_1\psi_0, & \psi_1 &= (\theta_1 - \phi_1\psi_0)/\phi_0, \\
0 &= \phi_0\psi_2 + \phi_1\psi_1 + \phi_2\psi_0, & \psi_2 &= -(\phi_1\psi_1 + \phi_2\psi_0)/\phi_0, \\
&\vdots & &\vdots \\
0 &= \phi_0\psi_n + \phi_1\psi_{n-1} + \phi_2\psi_{n-2}, & \psi_n &= -(\phi_1\psi_{n-1} + \phi_2\psi_{n-2})/\phi_0.
\end{aligned}
$$

Bi-directional (Non causal) Filters

A two-sided symmetric filter, which has the form of

$$(11) \qquad \psi(z) = \theta(z^{-1})\theta(z) = \psi_0 + \psi_1(z^{-1} + z) + \cdots + \psi_m(z^{-m} + z^m),$$

imposes no delays on any of the components of the signal. This is explained by the Cramér–Wold factorisation $\psi(z) = \theta(z^{-1})\theta(z)$, which gives rise to two equations

$$\text{(i)} \quad q(z) = \theta(z)y(z) \quad \text{and} \quad \text{(ii)} \quad x(z) = \theta(z^{-1})q(z).$$

The corresponding operations are

$$(12) \qquad \text{(i)} \quad q_t = \sum_j \theta_j y_{t-j} \qquad \text{and} \qquad \text{(ii)} \quad x_t = \sum_j \theta_j q_{t+j}.$$

The first operation, which runs in real time, imposes a time delay 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.

If $\psi(z)$ corresponds to an FIR filter, then $x(t)$ may be generated via a single application of the two-sided filter $\psi(z)$ to the signal $y(t)$, or it may be generated in two operations via the successive applications of $\theta(z)$ to $y(z)$ and $\theta(z^{-1})$ to $q(z) = \theta(z)y(z)$. It is a matter of indifference of which of these techniques is used to generate $x(t)$.

Two-Sided Rational Filters

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

$$\psi(z) = \frac{\theta(z^{-1})\theta(z)}{\phi(z^{-1})\phi(z)}. \tag{13}$$

This must be applied in two separate passes running forwards and backwards in time and described, respectively, by the equations

$$\text{(i)} \quad \phi(z)q(z) = \theta(z)y(z) \quad \text{and} \quad \text{(ii)} \quad \phi(z^{-1})x(z) = \theta(z^{-1})q(z). \tag{14}$$

The Response to a Sinusoidal Input

Conside mapping the sequence $\{x_t = \cos(\omega t)\}$ through the transfer function with the coefficients $\{\psi_0, \psi_1, \ldots\}$. The output is

$$(15) \qquad y(t) = \sum_j \psi_j \cos\big(\omega[t-j]\big).$$

Using the identity $\cos(A-B) = \cos A \cos B + \sin A \sin B$, this becomes

$$(16) \qquad \begin{aligned} y(t) &= \Big\{\sum_j \psi_j \cos(\omega j)\Big\}\cos(\omega t) + \Big\{\sum_j \psi_j \sin(\omega j)\Big\}\sin(\omega t) \\ &= \alpha\cos(\omega t) + \beta\sin(\omega t) = \rho\cos(\omega t - \theta), \end{aligned}$$

Expanding the final expression gives $\alpha = \rho\cos(\theta)$ and $\beta = \rho\sin(\theta)$, whence,

$$(17) \qquad \rho^2 = \alpha^2 + \beta^2 \qquad \text{and} \qquad \theta = \tan^{-1}\left(\frac{\beta}{\alpha}\right).$$

Also, if $\lambda = \alpha + \mathrm{i}\beta$ and $\lambda^* = \alpha - \mathrm{i}\beta$ are conjugate complex numbers, then ρ would be their modulus. This is illustrated in Figure 3.

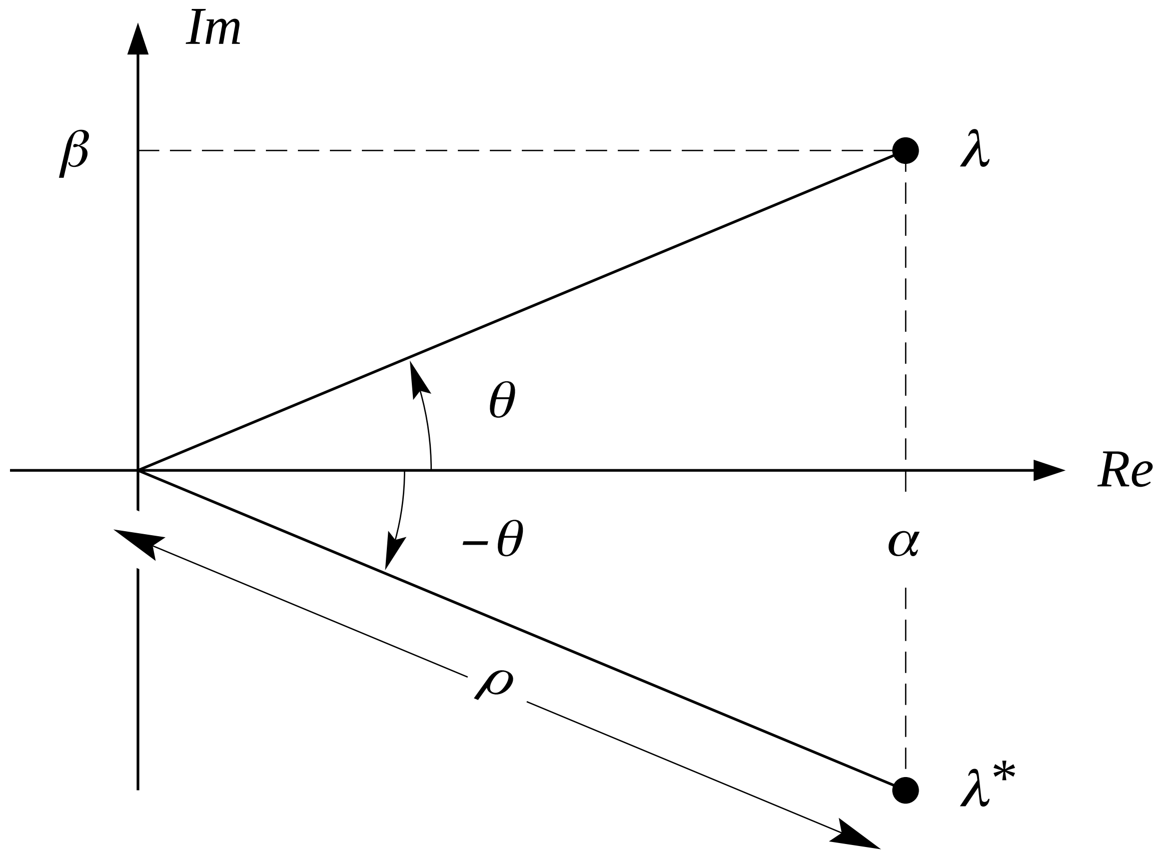

Figure 3. The Argand Diagram showing a complex number $\lambda = \alpha + i\beta$ and its conjugate $\lambda^* = \alpha - i\beta$.

The transfer function has a *gain effect,* whereby the amplitude of the sinusoid is increased or diminished by the factor ρ. It has a *phase effect,* whereby the peak of the sinusoid is displaced by a time delay of θ/ω periods. The frequency of the output is the same as the frequency of the input, which is a fundamental feature of all linear dynamic systems.

To obtain full information on the system, it is necessary to excite it over a full range of frequencies.

Aliasing and the Shannon–Nyquist Sampling Theorem

In a discrete-time system, signal frequencies in excess of π radians per sampling interval are confounded with frequencies within the interval $[0, \pi]$. Consider a cosine wave of a frequency ω in the interval $\pi < \omega < 2\pi$ that is sampled at unit intervals. Let $\omega^* = 2\pi - \omega$. Then,

$$
\begin{aligned}
\cos(\omega t) &= \cos\bigl\{(2\pi - \omega^*)t\bigr\} \\
&= \cos(2\pi)\cos(\omega^* t) + \sin(2\pi)\sin(\omega^* t) \\
&= \cos(\omega^* t);
\end{aligned}
\tag{18}
$$

which indicates that ω and ω^* are observationally indistinguishable. Here, $\omega^* \in [0, \pi]$ is described as the alias of $\omega > \pi$.

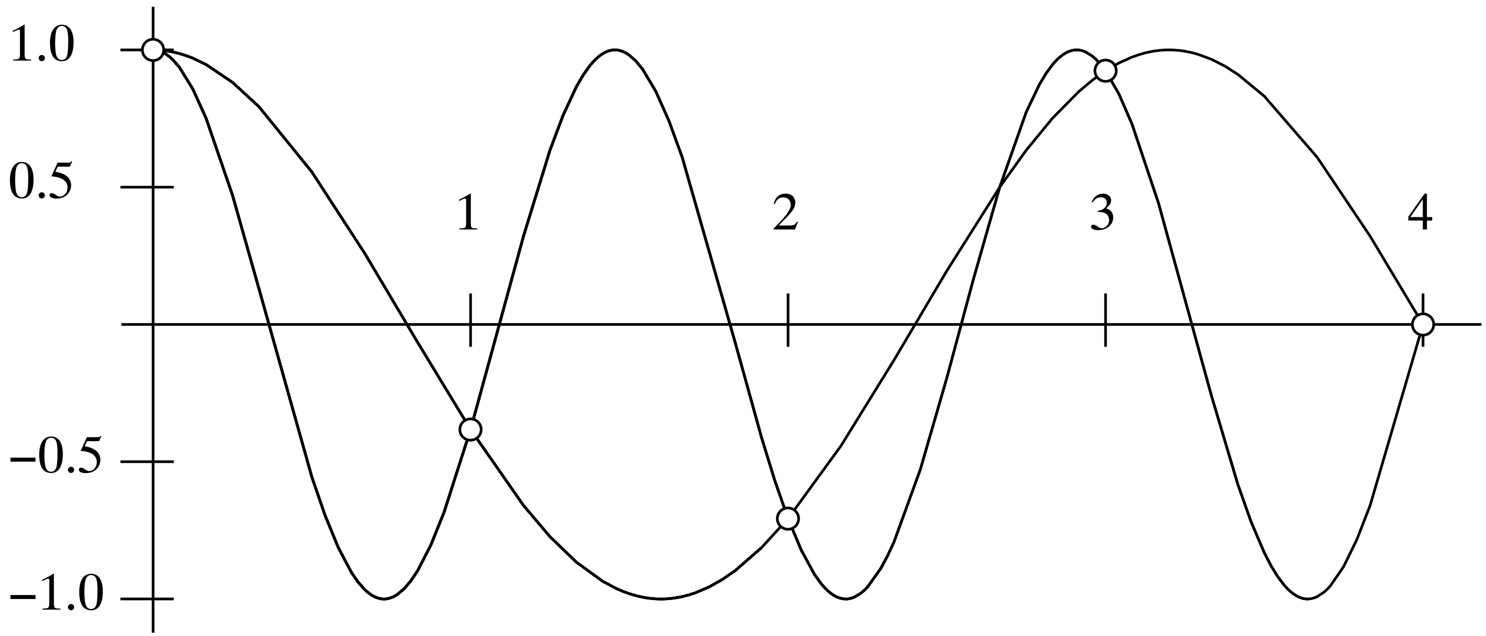

Figure 4. The values of the function $\cos\{(11/8)\pi t\}$ coincide with those of its alias $\cos\{(5/8)\pi t\}$ at the integer points $\{t = 0, \pm 1, \pm 2, \ldots\}$.

The Frequency Response of a Linear Filter

The frequency response of a linear filter $\psi(z)$ is its response to the set of sinusoidal inputs of all frequencies ω within the Nyquist interval $[0, \pi]$. This entails the squared gain of the filter, defined by

$$\rho^2(\omega) = \psi_\alpha^2(\omega) + \psi_\beta^2(\omega), \tag{19}$$

where

$$\psi_\alpha(\omega) = \sum_j \psi_j \cos(\omega j) \quad \text{and} \quad \psi_\beta(\omega) = \sum_j \psi_j \sin(\omega j), \tag{20}$$

and the phase displacement, defined by

$$\theta(\omega) = \text{Arg}\{\psi(\omega)\} = \tan^{-1}\{\psi_\beta(\omega)/\psi_\alpha(\omega)\}. \tag{21}$$

It is convenient to replace the trigonometrical functions of (20) by the complex exponential functions

$$e^{\mathrm{i}\omega j} = \frac{1}{2}\{\cos(\omega j) + \sin(\omega j)\} \quad \text{and} \quad e^{-\mathrm{i}\omega j} = \frac{1}{2}\{\cos(\omega j) - \sin(\omega j)\}. \tag{22}$$

These enable the trigonometrical functions to be expressed as

$$\text{(23)} \qquad \cos(\omega t) = \frac{1}{2}\{e^{\mathrm{i}\omega j} + e^{-\mathrm{i}\omega j}\} \quad \text{and} \quad \sin(\omega j) = \frac{\mathrm{i}}{2}\{e^{-\mathrm{i}\omega j} - e^{\mathrm{i}\omega j}\}.$$

Setting $z = \exp\{-\mathrm{i}\omega j\}$ in $\psi(z)$ gives

$$\text{(24)} \qquad \psi(e^{-\mathrm{i}\omega j}) = \psi_\alpha(\omega) - \mathrm{i}\psi_\beta(\omega),$$

which we shall write hereafter as $\psi(\omega) = \psi(e^{-\mathrm{i}\omega j})$.

The squared gain of the filter, previously denoted by $\rho^2(\omega)$, is the square of the complex modulus

$$\text{(25)} \qquad |\psi(\omega)|^2 = \psi_\alpha^2(\omega) + \psi_\beta^2(\omega),$$

which is obtained by setting $z = \exp\{-\mathrm{i}\omega j\}$ in $\psi(z^{-1})\psi(z)$.

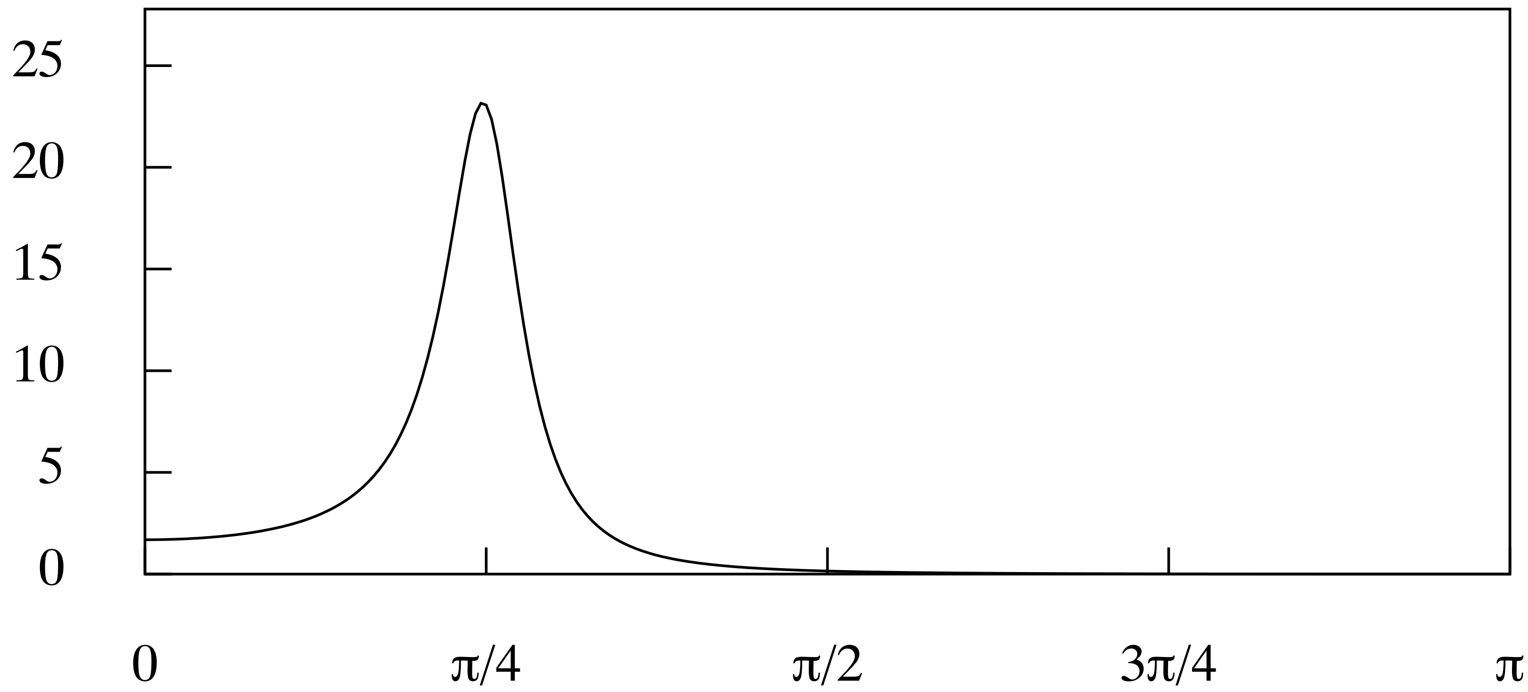

Figure 5. The spectral density function of the ARMA(2, 1) process $y(t) = 1.2728y(t-1) - 0.81y(t-1) + \varepsilon(t) + 0.0.75\varepsilon(t-1)$ with $V\{\varepsilon(t)\} = 1$.

The Spectrum of a Stationary Stochastic Process

Consider a stationary stochastic process $y(t) = \{y_t; t = 0, \pm 1, \pm 2, \ldots\}$ The generic element of the process can be expressed as $y_t = \sum_j \psi_j \varepsilon_{t-j}$, where ε_t is an element of a sequence $\varepsilon(t)$ of independently and identically distributed random variables with $E(\varepsilon_t) = 0$ and $V(\varepsilon_t) = \sigma^2$ for all t.

The autocovariance generating function of the process is

$$\sigma^2\psi(z^{-1})\psi(z) = \gamma(z) = \{\gamma_0 + \gamma_1(z^{-1} + z) + \gamma_2(z^{-2} + z^2) + \cdots\}. \tag{26}$$

The following table assists in forming the product $\gamma(z) = \sigma^2\psi(z^{-1})\psi(z)$:

$$\begin{array}{c|cccc} & \psi_0 & \psi_1 z & \psi_2 z^2 & \cdots \\ \hline \psi_0 & \psi_0^2 & \psi_0\psi_1 z & \psi_0\psi_2 z^2 & \cdots \\ \psi_1 z^{-1} & \psi_1\psi_0 z^{-1} & \psi_1^2 & \psi_1\psi_2 z & \cdots \\ \psi_2 z^{-2} & \psi_2\psi_0 z^{-2} & \psi_2\psi_1 z^{-1} & \psi_2^2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{array} \tag{27}$$

The autocovariances are obtained by summing along the NW–SE diagonals:

$$(28)\qquad \begin{aligned} \gamma_0 &= \sigma^2\{\psi_0^2+\psi_1^2+\psi_2^2+\psi_3^2+\cdots\},\\ \gamma_1 &= \sigma^2\{\psi_0\psi_1+\psi_1\psi_2+\psi_2\psi_3+\cdots\},\\ \gamma_2 &= \sigma^2\{\psi_0\psi_2+\psi_1\psi_3+\psi_2\psi_4+\cdots\},\\ &\ \ \vdots \end{aligned}$$

By setting $z=\exp\{-\mathrm{i}\omega j\}$ and dividing by 2π, we get the spectral density function of the process:

$$(29)\qquad f(\omega)=\frac{1}{2\pi}\left\{\gamma_0+2\sum_{\tau=1}^{\infty}\gamma_\tau\cos(\omega\tau)\right\}.$$

This entails the cosine Fourier transform of the sequence of autocovariances.

Wiener–Kolmogorov Filtering of Stationary Sequences

The purpose of a Wiener–Kolmogorov (W–K) filter is to extract an estimate of a signal sequence $\xi(t)$ from an observable data sequence (t) afflicted by the noise $\eta(t)$.

$$y(t) = \xi(t) + \eta(t). \tag{30}$$

The signal and the noise are zero-mean stationary stochastic processes that are mutually independent. For the present, we assume that the data constitute a doubly-infinite sequence.

The autocovariance generating function of the data is

$$\gamma^{yy}(z) = \gamma^{\xi\xi}(z) + \gamma^{\eta\eta}(z) \quad \text{and} \quad \gamma^{\xi\xi}(z) = \gamma^{y\xi}(z). \tag{31}$$

These functions have Cramér–Wold factorisations:

$$\gamma^{yy}(z) = \phi(z^{-1})\phi(z), \quad \gamma^{\xi\xi}(z) = \theta(z^{-1})\theta(z), \quad \gamma^{\eta\eta}(z) = \theta_\eta(z^{-1})\theta_\eta(z).$$

The estimate x_t of the signal element ξ_t, is a linear combination of the elements of the data sequence:

$$x_t = \sum_j \psi_j y_{t-j}. \tag{33}$$

The principle of minimum-mean-square-error estimation indicates that the estimation errors must be statistically uncorrelated with the elements of the information set. Thus, the following condition applies for all k:

$$\begin{aligned} 0 &= E\Big\{y_{t-k}(\xi_t - x_t)\Big\} \\ &= E(y_{t-k}\xi_t) - \sum_j \psi_j E(y_{t-k}y_{t-j}) \\ &= \gamma_k^{y\xi} - \sum_j \psi_j \gamma_{k-j}^{yy}. \end{aligned} \tag{34}$$

The equation may be expressed, in terms of the z-transforms, as

$$\gamma^{y\xi}(z) = \psi(z)\gamma^{yy}(z). \tag{35}$$

It follows that

$$\begin{aligned} \psi(z) &= \frac{\gamma^{y\xi}(z)}{\gamma^{yy}(z)} \\ &= \frac{\gamma^{\xi\xi}(z)}{\gamma^{\xi\xi}(z) + \gamma^{\eta\eta}(z)} = \frac{\theta(z^{-1})\theta(z)}{\phi(z^{-1})\phi(z)}. \end{aligned} \tag{36}$$

The same principle applies to the estimation of the residual component. This is obtained using the complementary filter

$$\psi^c(z) = 1 - \psi(z) = \frac{\gamma^{\eta\eta}(z)}{\gamma^{\xi\xi}(z) + \gamma^{\eta\eta}(z)}. \tag{37}$$

The estimated signal component may be obtained by filtering the data in two passes according to the following equations:

$$\phi(z)q(z) = \theta(z)y(z), \qquad \phi(z^{-1})x(z^{-1}) = \theta(z^{-1})q(z^{-1}). \tag{38}$$

The first equation relates to a process that runs forwards in time to generate the elements of an intermediate sequence, represented by the coefficients of $q(z)$. The second equation represents a process that runs backwards to deliver the estimates of the signal, represented by the coefficients of $x(z)$.

By setting $z = \exp\{-\mathrm{i}\omega\}$, one can derive the frequency-response function of the filter that is used in estimating the signal $\xi(t)$. The effect of the filter is to multiply each of the frequency elements of $y(t)$ by the fraction of its variance that is attributable to the signal.

The Hodrick–Prescott (Leser) Filter and the Butterworth Filter

The Wiener–Kolmogorov methodology can be applied to non stationary data with minor adaptations. A model of the processes underlying the data can be adopted that has the form of

$$
\begin{aligned}
\nabla^d(z)y(z) &= \nabla^d(z)\{\xi(z)+\eta(z)\} = \delta(z)+\kappa(z) \\
&= (1+z)^n\zeta(z)+(1-z)^m\varepsilon(z),
\end{aligned} \tag{39}
$$

where $\zeta(z)$ and $\varepsilon(z)$ are the z-transforms of two independent white-noise sequences $\zeta(t)$ and $\varepsilon(t)$.

The model of $y(t)=\xi(t)+\eta(t)$ entails a pair of stochastic processes, of which the z-transform are.

$$
\xi(z)=\frac{(1+z)^n}{\nabla^d(z)}\zeta(z) \quad \text{and} \quad \eta(z)=\frac{(1-z)^m}{\nabla^d(z)}\varepsilon(z). \tag{40}
$$

The condition $m \geq d$ is necessary to ensure the stationarity of $\eta(t)$, which is obtained from $\varepsilon(t)$ by differencing $m-d$ times.

The filter that is applied to $y(t)$ to estimate $\xi(t)$, which is the d-fold integral of $\delta(t)$, takes the form of

$$\psi(z) = \frac{\sigma_\zeta^2(1+z^{-1})^n(1+z)^n}{\sigma_\zeta^2(1+z^{-1})^n(1+z)^n + \sigma_\varepsilon^2(1-z^{-1})^m(1-z)^m}, \tag{41}$$

regardless of the degree d of differencing that would be necessary to reduce $y(t)$ to stationarity.

By setting $d = m = 2$ and $n = 0$ in (39), a model is obtained of a second-order random walk $\xi(t)$ affected by white-noise errors of observation $\eta(t) = \varepsilon(t)$. The resulting lowpass W–K filter, in the form of

$$\psi(z) = \frac{1}{1 + \lambda(1-z^{-1})^2(1-z)^2} \quad \text{with} \quad \lambda = \frac{\sigma_\eta^2}{\sigma_\delta^2}, \tag{42}$$

is the **Hodrick–Prescott** filter. The complementary highpass filter is

$$\psi^c(z) = \frac{(1-z^{-1})^2(1-z)^2}{\lambda^{-1} + (1-z^{-1})^2(1-z)^2}. \tag{43}$$

Here, λ is the adjustable smoothing parameter.

By setting $m = n$, a filter for estimating $\xi(t)$ is obtained that takes the form of

$$\begin{aligned}\psi(z) &= \frac{\sigma_\zeta^2(1+z^{-1})^n(1+z)^n}{\sigma_\zeta^2(1+z^{-1})^n(1+z)^n + \sigma_\varepsilon^2(1-z^{-1})^n(1-z)^n} \\ &= \frac{1}{1+\lambda\left(\mathrm{i}\dfrac{1-z}{1+z}\right)^{2n}} \quad \text{with} \quad \lambda = \frac{\sigma_\varepsilon^2}{\sigma_\zeta^2}.\end{aligned} \tag{44}$$

This is the **Butterworth** lowpass digital filter. The filter has two adjustable parameters, and, it is a more flexible than the H–P filter.

First, there is the parameter λ. This can be expressed as

$$\lambda = \{1/\tan(\omega_d)\}^{2n}, \tag{45}$$

where ω_d is the nominal cut-off point of the filter, which is the mid point in the transition of the filter's frequency response from its pass band to its stop band.

The second of the adjustable parameters is n, which denotes the order of the filter. As n increases, the transition between the pass band and the stop band becomes more abrupt.

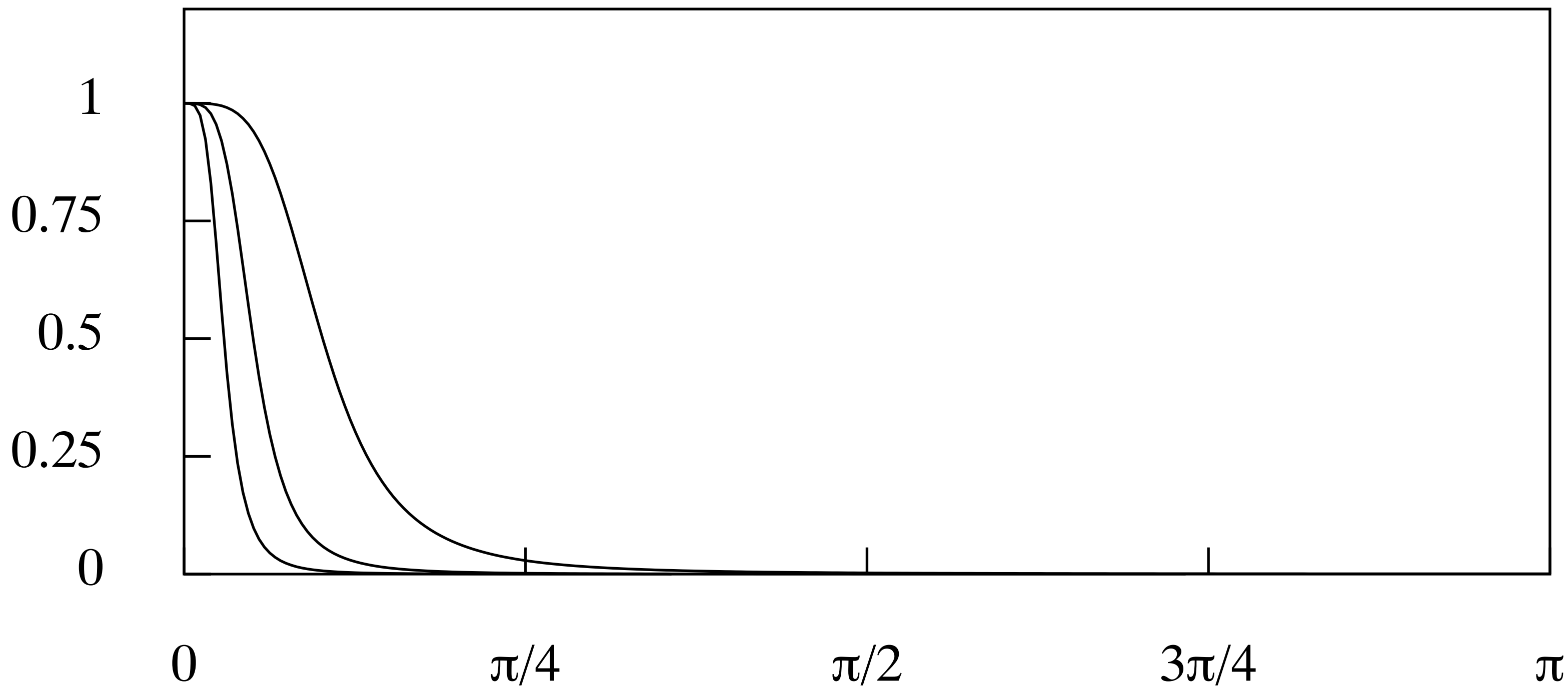

Figure 6. The gain of the Hodrick–Prescott lowpass filter with a smoothing parameter set to 100, 1600 and 14400.

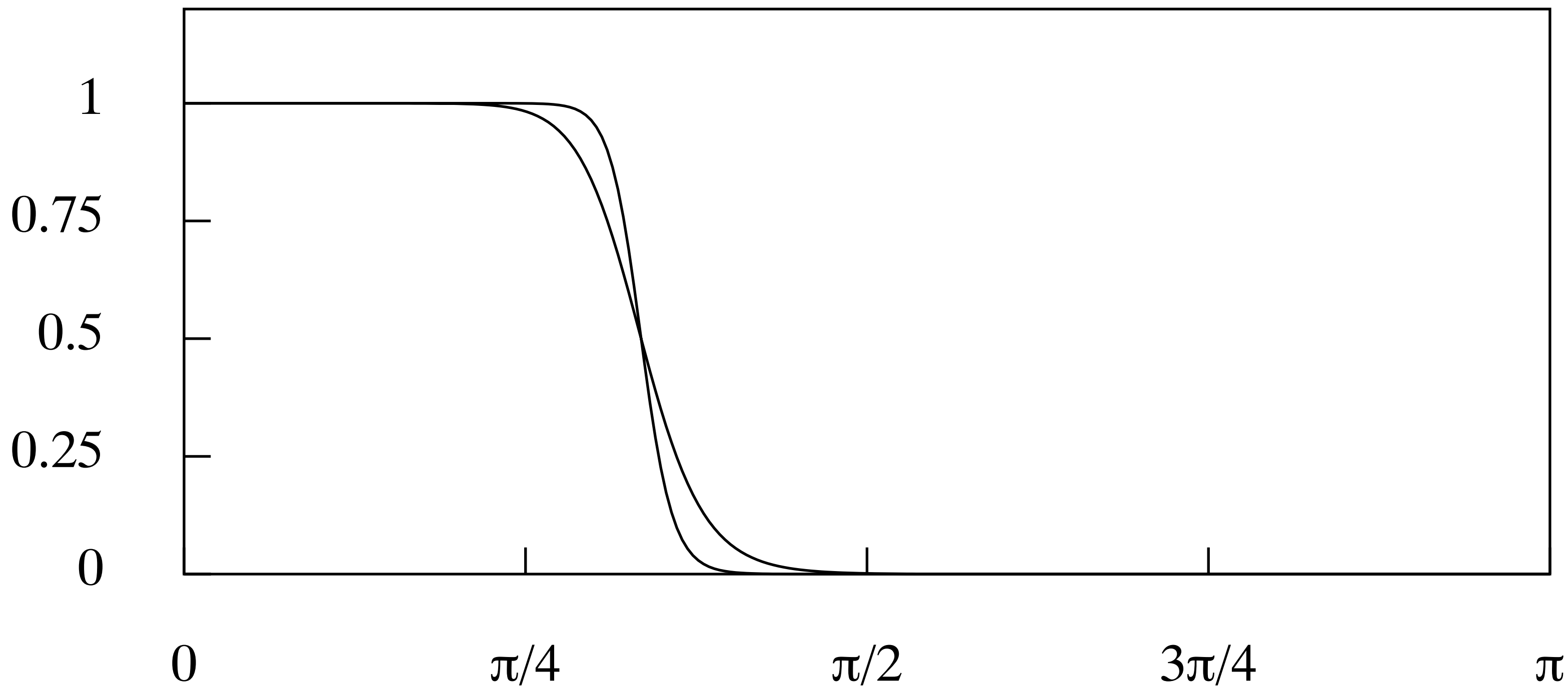

Figure 7. The squared gain of the lowpass Butterworth filters of orders $n = 6$ and $n = 12$ with a nominal cut-off point of $2\pi/3$ radians.

Wiener–Kolmogorov Filters for Finite Sequences

The W–K theory can be adapted to finite data sequences. Consider a data vector $y = [y_0, y_1, \ldots, y_{T-1},]'$ with components ξ and η:

$$y = \xi + \eta. \tag{46}$$

The components are assumed to be independently normally distributed with zero means and with positive-definite dispersion matrices. Then,

$$\begin{aligned} E(\xi) &= 0, \qquad D(\xi) = \Omega_\xi, \\ E(\eta) &= 0, \qquad D(\eta) = \Omega_\eta, \\ \text{and} \quad & C(\xi, \eta) = 0. \end{aligned} \tag{47}$$

Here, Ω_ξ and Ω_η may be obtained, from $\gamma_\xi(z)$ and $\gamma_\eta(z)$, respectively, by replacing z by the matrix lag operator $L_T = [e_1, e_2, \ldots, e_{T-1}, 0]$ obtained from $I_T = [e_0, e_1, e_2, \ldots, e_{T-1}]$ by deleting the leading column and by appending a zero vector to the end. Negative powers of z are replaced by powers of $F_T = L_T^{-1}$.

The independence of ξ and η implies that $D(y) = \Omega_\xi + \Omega_\eta$.

Time-Varying Filter Coefficients

We may begin by considering the determination of the vector of the T filter coefficients $\psi_{t.} = [\psi_{t,0}, \psi_{t,1}, \ldots, \psi_{t,T-1}]$ that determine x_t, which is the tth element of the filtered vector $x = [x_0, x_1, \ldots, x_{T-1}]'$. The estimate of ξ_t based on the sample $y_0, y_1, \ldots, y_{T-1}$ is

$$x_t = \sum_{j=-t}^{t-T+1} \psi_{t,j} y_{t-j},$$

and the principle of minimum-mean-square-error estimation indicates that

$$\begin{aligned} 0 &= E\Big\{y_{t-k}(\xi_t - x_t)\Big\} \\ &= E(y_{t-k}\xi_t) - \sum_{j=-t}^{t-T+1} \psi_{t,j} E(y_{t-k} y_{t-j}) \\ &= \gamma_k^{y\xi} - \sum_{j=-t}^{t-T+1} \psi_{t,j} \gamma_{j-k}^{yy}. \end{aligned} \tag{48}$$

Equation (48) can be rendered also in a matrix format. By running from $k = -t$ to $k = T - t - 1$, we get

$$(49) \qquad \begin{bmatrix} \gamma_t^{\xi\xi} \\ \gamma_{t+1}^{\xi\xi} \\ \vdots \\ \gamma_{T-1-t}^{\xi\xi} \end{bmatrix} = \begin{bmatrix} \gamma_0^{yy} & \gamma_1^{yy} & \cdots & \gamma_{T-1}^{yy} \\ \gamma_1^{yy} & \gamma_0^{yy} & \cdots & \gamma_{T-2}^{yy} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{T-1}^{yy} & \gamma_{T-2}^{yy} & \cdots & \gamma_0^{yy} \end{bmatrix} \begin{bmatrix} \psi_{t,0} \\ \psi_{t,1} \\ \vdots \\ \psi_{t,T-1} \end{bmatrix}.$$

Here, on the LHS, we have set $\gamma_j^{y\xi} = \gamma_j^{\xi\xi}$ in accordance with (31).

This equation above can be written as $\Omega_\xi e_t = \Omega_y \psi'_{t.}$, where e_t is a vector of order T containing a single unit preceded by t zeros and followed by $T - 1 - t$ zeros. The coefficient vector $\psi_{t.}$ is given by

$$(50) \qquad \psi_{t.} = e'_t \Omega_\xi \Omega_y^{-1} = e'_t \Omega_\xi (\Omega_\xi + \Omega_\eta)^{-1}.$$

and the estimate of ξ_t is $x_t = \psi_{t.} y$.

Given the data vector y, the estimate of the signal vector ξ is

$$(51) \qquad x = \Omega_\xi \Omega_y^{-1} y = \Omega_\xi (\Omega_\xi + \Omega_\eta)^{-1} y.$$

The Estimates as Conditional Expectations

The optimal predictors of the signal and the noise components are the conditional expectations:

$$
\begin{aligned}
\text{(52)} \qquad E(\xi|y) &= E(\xi) + C(\xi, y)D^{-1}(y)\{y - E(y)\} \\
&= \Omega_\xi(\Omega_\xi + \Omega_\eta)^{-1} = x,
\end{aligned}
$$

$$
\begin{aligned}
\text{(53)} \qquad E(\eta|y) &= E(\eta) + C(\eta, y)D^{-1}(y)\{y - E(y)\} \\
&= \Omega_\eta(\Omega_\xi + \Omega_\eta)^{-1}y = h.
\end{aligned}
$$

which are their minimum-mean-square-error estimates.

The error dispersion matrices, from which confidence intervals for the estimated components may be derived, are

$$
\begin{aligned}
\text{(54)} \qquad D(\xi|y) &= D(\xi) - C(\xi, y)D^{-1}(y)C(y, \xi) \\
&= \Omega_\xi - \Omega_\xi(\Omega_\xi + \Omega_\eta)^{-1}\Omega_\xi,
\end{aligned}
$$

$$
\begin{aligned}
\text{(55)} \qquad D(\eta|y) &= D(\eta) - C(\eta, y)D^{-1}(y)C(y, \eta), \\
&= \Omega_\eta - \Omega_\eta(\Omega_\xi + \Omega_\eta)^{-1}\Omega_\eta.
\end{aligned}
$$

The Least-Squares Derivation of the Estimates

The estimates x and h of ξ and η can also be derived according to the following criterion:

$$(56) \quad \text{Minimise} \quad S(\xi,\eta) = \xi'\Omega_\xi^{-1}\xi + \eta'\Omega_\eta^{-1}\eta \quad \text{subject to} \quad \xi + \eta = y.$$

The resulting estimates may be described, also, as the minimum chi-square estimates or the maximum-likelihood estimates.

Substituting for $\eta = y - \xi$ gives the concentrated criterion function $S(\xi) = \xi'\Omega_\xi^{-1}\xi + (y-\xi)'\Omega^{-1}(y-\xi)$. Differentiating in respect of ξ and setting the result to zero gives the condition

$$\Omega^{-1}(y - x) = \Omega_\xi^{-1}x.$$

Pre multiplication by Ω_η, gives $y = x + \Omega\Omega_\xi^{-1}x = (\Omega_\xi + \Omega_\eta)\Omega_\xi^{-1}x$. Therefore, the solution for x is

$$(57) \qquad x = \Omega_\xi(\Omega_\xi + \Omega_\eta)^{-1}y.$$

Since ξ and η are interchangeable, and since $h + x = y$, there are also

$$(58) \qquad h = \Omega_\eta(\Omega_\xi + \Omega_\eta)^{-1}y \qquad \text{and} \qquad x = y - \Omega_\eta(\Omega_\xi + \Omega_\eta)^{-1}y.$$

Computing the Estimates

The filter matrices $\Psi_\xi = \Omega_\xi(\Omega_\xi + \Omega_\eta)^{-1}$ and $\Psi_\eta = \Omega_\eta(\Omega_\xi + \Omega_\eta)^{-1}$ of (57) and (58) are the matrix analogues of the z-transforms displayed in equations (36) and (37).

To calculate the estimates x and h, first solve the equation

$$(\Omega_\xi + \Omega_\eta)b = y \tag{59}$$

for the value of b. Then, one can generate

$$x = \Omega_\xi b \qquad \text{and} \qquad h = \Omega_\eta b. \tag{60}$$

If Ω_ξ and Ω_η are narrow-band moving-average dispersion matrices, then the solution to (58) is via a Cholesky factorisation $\Omega_\xi + \Omega_\eta = GG'$, where G is a lower-triangular matrix with a limited number of nonzero bands.

The system $GG'b = y$ may be cast in the form of $Gp = y$ and solved for p. Then, $G'b = p$ can be solved for b.

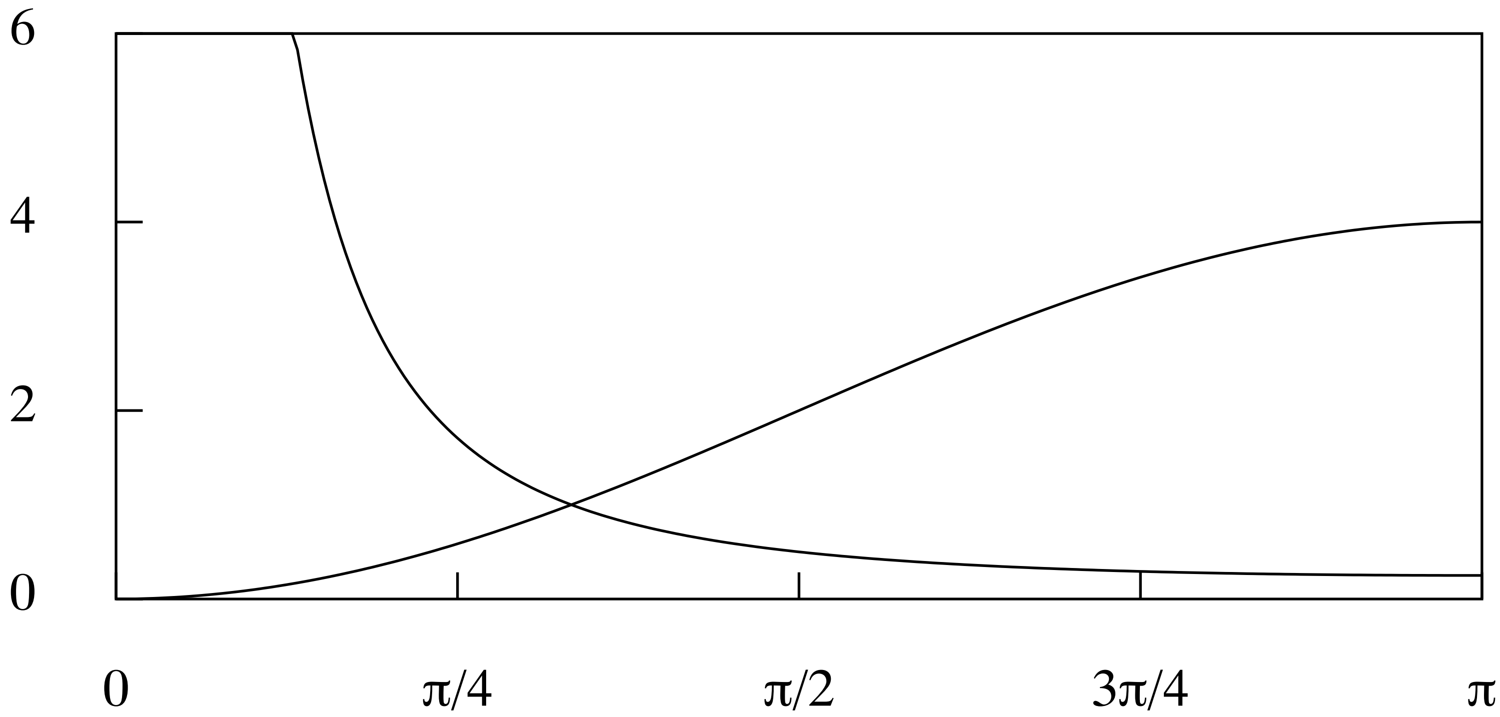

Figure 8. The squared gain of the difference operator, which has a zero at zero frequency, and the squared gain of the summation operator, which is unbounded at zero frequency.

The Difference and Summation Operators

The trend can be eliminated from the data $y(t) = \{y_t; t = 0, \pm 1, \pm 2, \ldots\}$ by taking the first differences $y(t) - y(t-1)$ or the second differences $y(t) - 2y(t-1) + y(t-2)$. Differences of higher orders are rare.

The difference operator is $\nabla(z) = 1 - z$. The z-transform of the first difference is $\nabla(z)y(z) = y(z) - zy(z)$. For the second difference, it is $\nabla^2(z)y(t) = (1 - 2z + z^2)y(z)$.

The inverse of the difference operator is the summation operator

$$\Sigma(z) = (1 - z)^{-1} = \{1 + z + z^2 + \cdots\}. \tag{61}$$

The z-transform of the d-fold summation operator is

$$\Sigma^d(z) = \frac{1}{(1 - z)^d} = 1 + dz + \frac{d(d+1)}{2!}z^2 + \frac{d(d+1)(d+2)}{3!}z^3 + \cdots. \tag{62}$$

The difference operator nullifies the trend and it severely attenuates the elements of the data that are adjacent in frequency to the zero frequency of the trend. It also amplifies the high frequency elements of the data.

The Matrix Difference Operators

For a sample of T elements in the vector $y = [y_0, y_1, \ldots, y_{T-1}]'$, one must use the matrix difference operator $\nabla(L_T) = I_T - L_T$.

Examples of the first-order and second-order matrix difference operators are

$$(63) \qquad \nabla_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}, \quad \nabla_4^2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{bmatrix}.$$

The correspoding inverse matrices are

$$(64) \qquad \Sigma_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \quad \Sigma_4^2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 \\ 4 & 3 & 2 & 1 \end{bmatrix}.$$

The elements of the leading vector are the coefficients associated with the expansion of $\Sigma^d(z)$ of (62).

Polynomial Interpolation

The first p columns of the matrix Σ_T^p provide a basis of the set of polynomials of degree $p-1$ defined on the integers $t = 0, 1, 2, \ldots, T-1$. Consider, for example, the first three columns of the matrix Σ_4^3, which may be transformed as follows:

$$
\begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 6 & 3 & 1 \\ 10 & 6 & 3 \end{bmatrix}
\begin{bmatrix} 1 & 1 & 1 \\ -2 & -1 & 1 \\ 1 & 0 & 0 \end{bmatrix}
=
\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{bmatrix}. \tag{65}
$$

The first column of the matrix on the LHS contains the ordinates of the quadratic function $(t^2+t)/2$. The columns of the transformed matrix contain the ordinates t^0, t^1 and t^2 for the integers $t = 1, 2, 3, 4$. The extension of the matrix to T rows provides a basis for the quadratic functions $q(t) = at^2 + bt + c$ defined on T consecutive integers.

The matrix of the powers of the integers is notoriously ill-conditioned.

Consider, the matrix that takes the p-th differences:

$$\nabla_T^p = (I - L_T)^p. \tag{66}$$

Let $\nabla_T^p = [Q_*, Q]'$, where Q_*' has p rows. Then

$$\nabla_T^p y = \begin{bmatrix} Q_*' \\ Q' \end{bmatrix} y = \begin{bmatrix} g_* \\ g \end{bmatrix}; \tag{67}$$

and g_* is liable to be discarded, whereas g will be regarded as the vector of the p-th differences of the data.

If $\nabla_T^{-p} = [S_*, S]$ is partitioned conformably then

$$\begin{bmatrix} S_* & S \end{bmatrix} \begin{bmatrix} Q_*' \\ Q' \end{bmatrix} = S_* Q_*' + SQ' = I_T, \tag{68}$$

and

$$\begin{bmatrix} Q_*' \\ Q' \end{bmatrix} \begin{bmatrix} S_* & S \end{bmatrix} = \begin{bmatrix} Q_*' S_* & Q_*' S \\ Q' S_* & Q' S \end{bmatrix} = \begin{bmatrix} I_p & 0 \\ 0 & I_{T-p} \end{bmatrix}. \tag{69}$$

If g_* is available, then y can be recovered from g via

$$y = S_* g_* + Sg. \tag{70}$$

Polynomial Interpolation by Least Squares

Since S_*, provides a basis for all polynomials of degree $p-1$ that are defined on the integer points $t = 0, 1, \ldots, T-1$, it follows that $S_* g_* = S_* Q_*' y$ contains the ordinates of a polynomial of degree $p-1$, which is interpolated through the first p elements of y, indexed by $t = 0, 1, \ldots, p-1$.

Let $y = \xi + \eta$, where ξ contains the ordinates of a polynomial of degree $p-1$ and η is a disturbance term with $E(\eta) = 0$ and $D(\eta) = \sigma_\eta^2 I_T$. To estimate ξ via $x = S_* r_*$ we should minimise the residual sum of squares

$$(y - x)'(y - x) = (y - S_* r_*)'(y - S_* r_*) \tag{71}$$

with respect to r_*. The resulting values are

$$r_* = (S_*' S_*)^{-1} S_*' y \quad \text{and} \quad x = S_* (S_*' S_*)^{-1} S_*' y. \tag{72}$$

An Alternative Polynomial Representation

An alternative representation of the estimated polynomial is available. This is provided by the identity

$$S_*(S_*'S_*)^{-1}S_*' = I - Q(Q'Q)^{-1}Q'. \tag{73}$$

To prove this identity, consider the fact that $Z = [Q, S_*]$ is square matrix of full rank and that Q and S_* are mutually orthogonal such that $Q'S_* = 0$. Then

$$\begin{aligned} Z(Z'Z)^{-1}Z' &= \begin{bmatrix} Q & S_* \end{bmatrix} \begin{bmatrix} (Q'Q)^{-1} & 0 \\ 0 & (S_*'S)^{-1} \end{bmatrix} \begin{bmatrix} Q' \\ S_*' \end{bmatrix} \\ &= Q(Q'Q)^{-1}Q' + S_*(S_*'S_*)^{-1}S_*'. \end{aligned} \tag{74}$$

Since $Z(Z'Z)^{-1}Z' = Z(Z^{-1}Z'^{-1})Z' = I$, the result of (73) follows; and the vector the ordinates of the polynomial regression is also given by

$$x = y - Q(Q'Q)^{-1}Q'y. \tag{75}$$

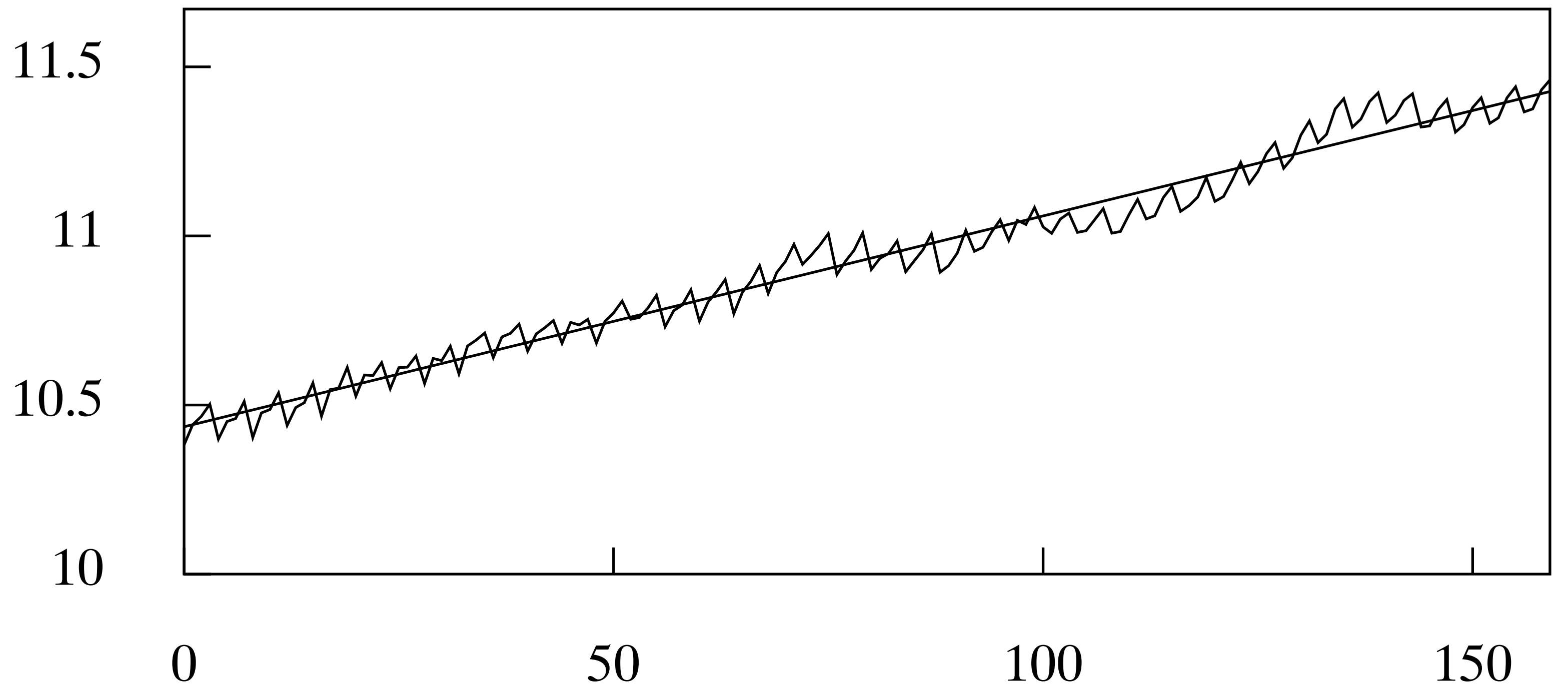

Figure 9. The quarterly series of the logarithms of consumption in the U.K., for the years 1955 to 1994, together with a linear trend interpolated by least-squares regression.

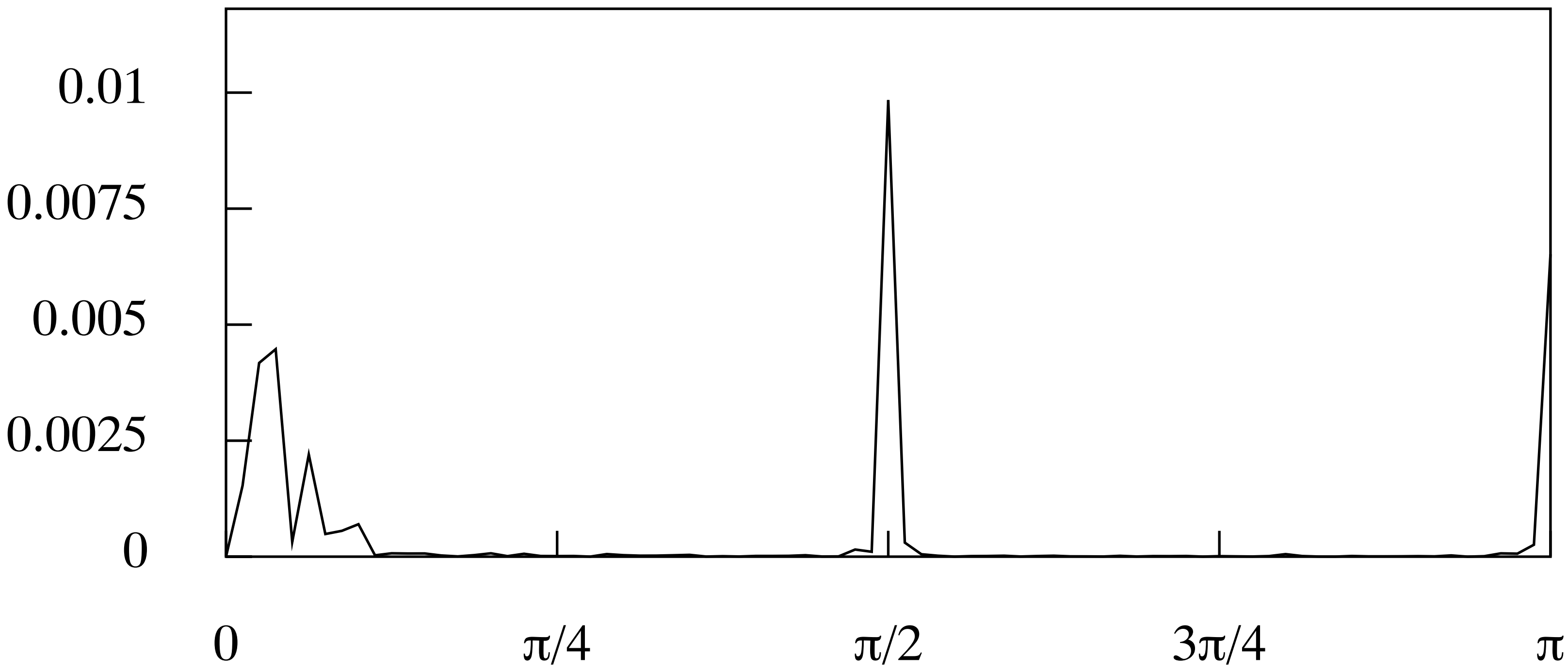

Figure 10. The periodogram of the residual sequence obtained from the linear detrending of the logarithmic consumption data.

Polynomial Regression and Trend Extraction

Polynomial regression can be used in a preliminary detrending of the data. Once the trend has been eliminated from the data, one can proceed to assess their spectral structure by examining the periodogram of the residual sequence.

Often the periodogram will reveal the existence of a cut-off frequency that bounds a low frequency trend/cycle component and separates it from the remaining elements of the spectrum.

An example is given in Figures 5 and 6. Figure 5 represents the logarithms of the quarterly data on aggregate consumption in the United Kingdom for the years 1955 to 1994.

The linear trend that has been interpolated by least-squares regression establishes a benchmark of constant exponential growth, against which the fluctuations of consumption can be measured.

The periodogram of the residual sequence is plotted in Figure 6. This shows that the low-frequency structure is bounded by a frequency value of $\pi/8$. This value can used in specifying the appropriate filter for extracting the low-frequency trajectory of the data

Filters for Short Trended Sequences

One way of eliminating the trend is to take differences of the data. Usually, twofold differencing is appropriate. The matrix analogue of the second-order backwards difference operator in the case of $T = 5$ is given by

$$\text{(76)} \qquad \nabla_5^2 = \begin{bmatrix} Q'_* \\ Q' \end{bmatrix} = \left[\begin{array}{ccccc} 1 & 0 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 & 0 \\ \hline 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \end{array}\right].$$

Applying Q' to the equation $y = \xi + \eta$, representing the trended data, gives

$$\text{(77)} \qquad \begin{aligned} Q'y &= Q'\xi + Q'\eta \\ &= \delta + \kappa = g. \end{aligned}$$

The moments of the differenced vectors are

$$(78)\qquad \begin{aligned} E(\delta) &= 0, \qquad & D(\delta) &= \Omega_\delta = Q'D(\xi)Q,\\ E(\kappa) &= 0, \qquad & D(\kappa) &= \Omega_\kappa = Q'D(\eta)Q. \end{aligned}$$

The starting-value problem can be circumvented by concentrating on the estimation of η. The conditional expectation of η given $g = Q'y$ is

$$(79)\qquad \begin{aligned} h = E(\eta|g) &= E(\eta) + C(\eta, g)D^{-1}(g)\{g - E(g)\}\\ &= C(\eta, g)D^{-1}(g)g, \end{aligned}$$

Within this expression, there are

$$(80)\qquad D(g) = \Omega_\delta + Q'\Omega_\eta Q \quad \text{and} \quad C(\eta, g) = \Omega_\eta Q.$$

Putting these details into (79) gives the following estimate of η:

$$(81)\qquad h = \Omega_\eta Q(\Omega_\delta + Q'\Omega_\eta Q)^{-1}Q'y.$$

Putting this into the equation

$$(82)\qquad x = E(\xi|g) = y - E(\eta|g) = y - h$$

gives

$$(83)\qquad x = y - \Omega_\eta Q(\Omega_\delta + Q'\Omega_\eta Q)^{-1}Q'y.$$

The Least-Squares Derivation of the Filter

The least-squares criterion is

$$\text{Minimise}\quad (y-\xi)'\Omega_\eta^{-1}(y-\xi)+\xi' Q\Omega_\delta^{-1}Q'\xi. \tag{84}$$

The first term penalises departures of the curve from the data, and the second term imposes a penalty for a lack of smoothness. Differentiating the function with respect to ξ and setting the result to zero gives

$$\Omega_\eta^{-1}(y-x)=-Q\Omega_\delta^{-1}Q'x=Q\Omega_\delta^{-1}d, \tag{85}$$

where x stands for the estimated value of ξ and $d=Q'x$. Premultiplying by $Q'\Omega_\eta$ gives

$$Q'(y-x)=Q'y-d=Q'\Omega_\eta Q\Omega_\delta^{-1}d, \tag{86}$$

whence

$$\begin{aligned} Q'y&=d+Q'\Omega_\eta Q\Omega_\delta^{-1}d\\ &=(\Omega_\delta+Q'\Omega_\eta Q)\Omega_\delta^{-1}d, \end{aligned} \tag{87}$$

which gives

$$\Omega_\delta^{-1} d = (\Omega_\delta + Q'\Omega_\eta Q)^{-1} Q'y. \tag{88}$$

Putting this into

$$x = y - \Omega_\eta Q \Omega_\delta^{-1} d, \tag{89}$$

which comes from premultiplying (64) by Ω_η, gives

$$x = y - \Omega_\eta Q(\Omega_\delta + Q'\Omega_\eta Q)^{-1} Q'y. \tag{90}$$

One should observe that

$$\Omega_\eta Q(\Omega_\delta + Q'\Omega_\eta Q)^{-1} Q'y = \Omega_\eta Q(\Omega_\delta + Q'\Omega_\eta Q)^{-1} Q'e, \tag{91}$$

where $e = Q(Q'Q)^{-1}Q'y$ is the vector of residuals obtained by interpolating a straight line through the data by a least-squares regression.

It makes no difference to the estimate of the component that is complementary to the trend whether the filter is applied to the data vector y or the residual vector e.

The Leser (HP) Filter

The specific cases that have been considered in the context of the classical form of the Wiener–Kolmogorov filter can now be adapted to the circumstances of short trended sequences. First there is the Leser filter. This is derived by setting

$$(92) \qquad D(\eta) = \Omega_\eta = \sigma_\eta^2 I, \quad D(\delta) = \Omega_\delta = \sigma_\delta^2 I \quad \text{and} \quad \lambda = \frac{\sigma_\eta^2}{\sigma_\delta^2}$$

within (90) to give

$$(93) \qquad x = y - Q(\lambda^{-1}I + Q'Q)^{-1}Q'y$$

Here, λ is the so-called smoothing parameter. As $\lambda \to \infty$, the vector x tends to that of a linear function interpolated into the data by least-squares regression, which is

$$(75) \qquad x = y - Q(Q'Q)^{-1}Q'y.$$

The Butterworth Filter

The Butterworth filter that is appropriate to short trended sequences can be represented by the equation

$$(94) \qquad x = y - \lambda\Sigma Q(M + \lambda Q'\Sigma Q)^{-1}Q'y.$$

Here, the matrices

$$(95) \qquad \Sigma = \{2I_T - (L_T + L_T^{-1})\}^{n-2} \quad \text{and} \quad M = \{2I_T + (L_T + L_T^{-1})\}^{n}$$

are obtained from the RHS of the equations $\{(1-z)(1-z^{-1})\}^{n-2} = \{2-(z+z^{-1})\}^{n-2}$ and $\{(1+z)(1+z^{-1})\}^{n} = \{2+(z+z^{-1})\}^{n}$, respectively, by replacing z by L_T and z^{-1} by L_T^{-1}. Observe that the equalities no longer hold after the replacements. However, it can be verified that

$$(96) \qquad Q'\Sigma Q = \{2I_T - (L_T + L_T^{-1})\}^{n}.$$

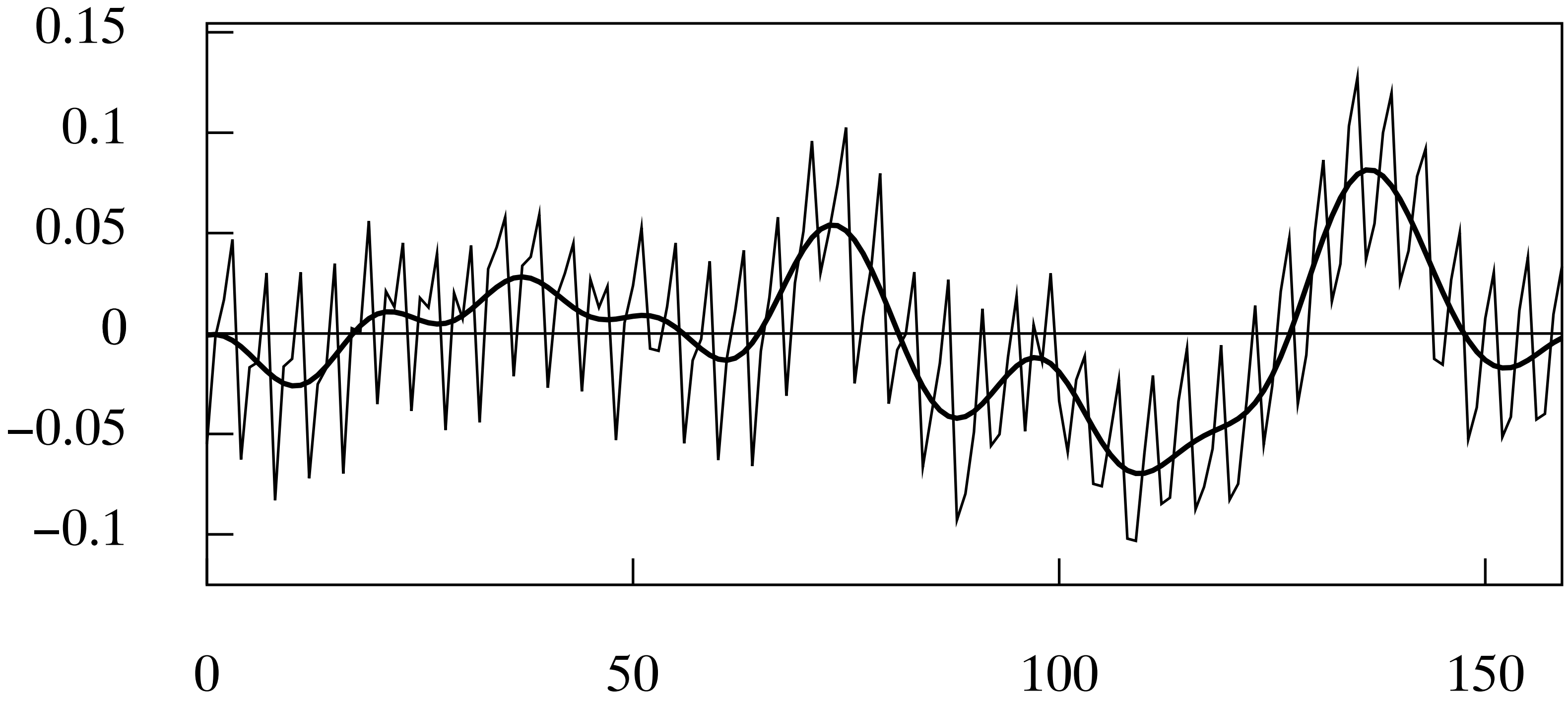

Figure 11. The residual sequence from fitting a linear trend to the logarithmic consumption data with an interpolated line representing the business cycle, obtained by the frequency-domain method.

Filtering in the Frequency Domain

The method of Wiener–Kolmogorov filtering can also be implemented using the circulant dispersion matrices that are given by

$$\begin{aligned}
&\Omega^{\circ}_{\xi} = \bar{U}\gamma_{\xi}(D)U, \quad \Omega^{\circ}_{\eta} = \bar{U}\gamma_{\eta}(D)U \quad \text{and} \\
&\Omega^{\circ} = \Omega^{\circ}_{\xi} + \Omega^{\circ}_{\eta} = \bar{U}\{\gamma_{\xi}(D) + \gamma_{\eta}(D)\}U,
\end{aligned} \tag{97}$$

wherein the diagonal matrices $\gamma_{\xi}(D)$ and $\gamma_{\eta}(D)$ contain the ordinates of the spectral density functions of the component processes.

Here, $U = T^{-1/2}[W^{jt}]$, with $t, j = 0, \ldots, T-1$, is the matrix of the Fourier transform, in which $W^{jt} = \exp(-\mathrm{i}2\pi tj/T)$, and $\bar{U}$ is its conjugate transpose. Also, $D = \mathrm{diag}\{1, W, W^2, \ldots, W^{T-1}\}$ is a diagonal matrix comprising the T roots of unity.

By replacing the dispersion matrices within (52) and (53) by their circulant counterparts, we derive the following formulae:

$$x = \bar{U}\gamma_{\xi}(D)\{\gamma_{\xi}(D) + \gamma_{\eta}(D)\}^{-1}Uy = P_{\xi}y, \tag{98}$$

$$h = \bar{U}\gamma_{\eta}(D)\{\gamma_{\xi}(D) + \gamma_{\eta}(D)\}^{-1}Uy = P_{\eta}y. \tag{99}$$

Implementing the Frequency-Domain Filter

A Fourier transform may applied to the data vector y to give Uy, which resides in the frequency domain. Then, Uy is multiplied by

$$J_\xi = \gamma_\xi(D)\{\gamma_\xi(D) + \gamma_\eta(D)\}^{-1} \quad \text{and} \quad J_\eta = \gamma_\eta(D)\{\gamma_\xi(D) + \gamma_\eta(D)\}^{-1}.$$

Finally, the products are carried back into the time domain by the inverse Fourier transform $\bar{U}$.

The Frequency-Domain Method and Nonstationary Data

Data may be reduced to stationarity by twofold differencing before filtering. The filtered sequence may be reinflated by summation.

Let the original data be $y = \xi + \eta$ and let the differenced data be $g = Q'y = \delta + \kappa$. If the estimates of $\delta = Q'\xi$ and $\kappa = Q'\eta$ are denoted by d and k respectively, then the estimates of ξ and η will be

$$x = S_*d_* + Sd \quad \text{where} \quad d_* = (S_*'S_*)^{-1}S_*'(y - Sd) \tag{100}$$

and

$$h = S_*k_* + Sk \quad \text{where} \quad k_* = -(S_*'S_*)^{-1}S_*'Sk. \tag{101}$$

Here, d_* an k_* are the initial conditions that are obtained via the minimisation of the function

$$\begin{aligned}(y-x)'(y-x) &= (y - S_*d_* - Sd)'(y - S_*d_* - Sd)\\ &= (S_*k_* + Sk)'(S_*k_* + Sk) = h'h.\end{aligned} \tag{102}$$

The minimisation ensures that the estimated trend x adheres as closely as possible to the data y.

An Alternative Frequency Domain Method

An alternative method concentrates on estimating the stationary high frequency component. First, the data are reduced to stationarity by twofold differencing.

Then the re-inflation of the high frequency component occurs in the frequency domain after it has been extracted. The resulting vector of Fourier coefficients is transformed to the time domain.

The centralised difference operator reduces the trended data sequence to stationary This is

$$
\begin{aligned}
N(z) &= z^{-1} - 2 + z = z^{-1}(1-z)^2 \\
&= z^{-1}\nabla^2(z).
\end{aligned} \tag{108}
$$

The matrix version of the operator is obtained by setting $z = L_T$ and $z^{-1} = L_T'$, which gives

$$
N(L_T) = N_T = L_T - 2I_T + L_T'. \tag{109}
$$

$$N_5 = \begin{bmatrix} Q'_{-1} \\ Q' \\ Q_{+1} \end{bmatrix} = \left[\begin{array}{ccccc} -2 & 1 & 0 & 0 & 0 \\ \hline 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ \hline 0 & 0 & 0 & 1 & -2 \end{array}\right]. \tag{110}$$

On deleting the first and last elements of the vector $N_T y$, which are $Q'_{-1}y = e'_1 \nabla^2_T y$ and $Q_{+1}y$, respectively, we get $Q'y = [q_1, \ldots, q_{T-2}]'$. The loss of these elements can be overcome by supplementing the original data vector y with two extrapolated end points y_{-1} and y_T.or by attributing appropriate values to q_0 and q_{T-1}.

Let Λ be the matrix which selects the appropriate ordinates of the Fourier transform $\gamma = Uq$ of the twice differenced data. These ordinates must be reinflated to compensate for the differencing operation, which has the frequency response

$$f(\omega) = 2 - 2\cos(\omega). \tag{111}$$

Reinflation in the Frequency Domain

The response of the anti-differencing operation is $1/f(\omega)$; and γ is reinflated by pre-multiplying by the diagonal matrix

$$V = \mathrm{diag}\{v_0, v_1, \ldots, v_{T-1}\}, \tag{112}$$

comprising the values $v_j = 1/f(\omega_j); j = 0, \ldots, T-1$, where $\omega_j = 2\pi j/T$.

Let $H = V\Lambda$ be the matrix that is is applied to $\gamma = Uq$ to generate the Fourier ordinates of the filtered vector. The resulting vector is transformed to the time domain to give

$$h = \bar{U}H\gamma = \bar{U}HUq. \tag{113}$$

Since $f(\omega)$ is zero-valued when $\omega = 0$ and that $1/f(\omega)$ is unbounded in the neighbourhood of $\omega = 0$.

The low-frequency trend component that is complementary to h is

$$x = y - h = y - \bar{U}HUq. \tag{114}$$

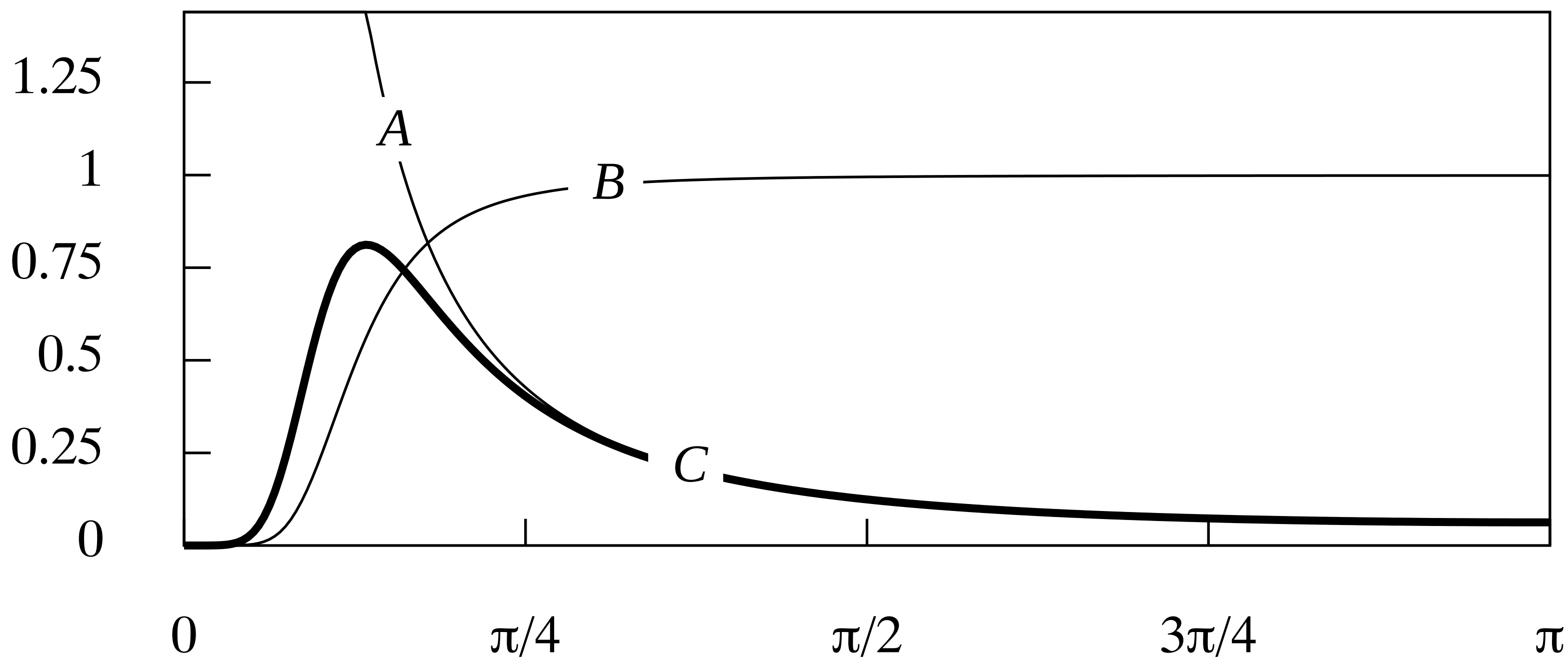

Figure 12. The pseudo-spectrum of a random walk, labelled A, together with the squared gain of the highpass Hodrick–Prescott filter with a smoothing parameter of $\lambda = 100$, labelled B. The curve labelled C represents the spectrum of the filtered process.

Business Cycles and Spurious Cycles

For the original data, the decomposition is usually a multiplicative one and, for the logarithmic data, the corresponding decomposition is an additive one. The filters are usually applied to the logarithmic data, and the sum of the estimated components should equal the logarithmic data.

The manner in which any component is defined and and extracted is liable to all of the other components. In particular, variations in the definition of the trend will have substantial effects upon the representation of the business cycle.

It has been proposed by several authors, including Harvey and Jaeger (1993) and Cogley and Nason (1995), that the effect of using the Hodrick–Prescott filter to extract a trend from the data is to create or induce spurious cycles in the complementary component, which includes the cyclical component. Others have strongly disputed this idea.

We analyse the effects of the H–P filter in the following sequence of graphics:

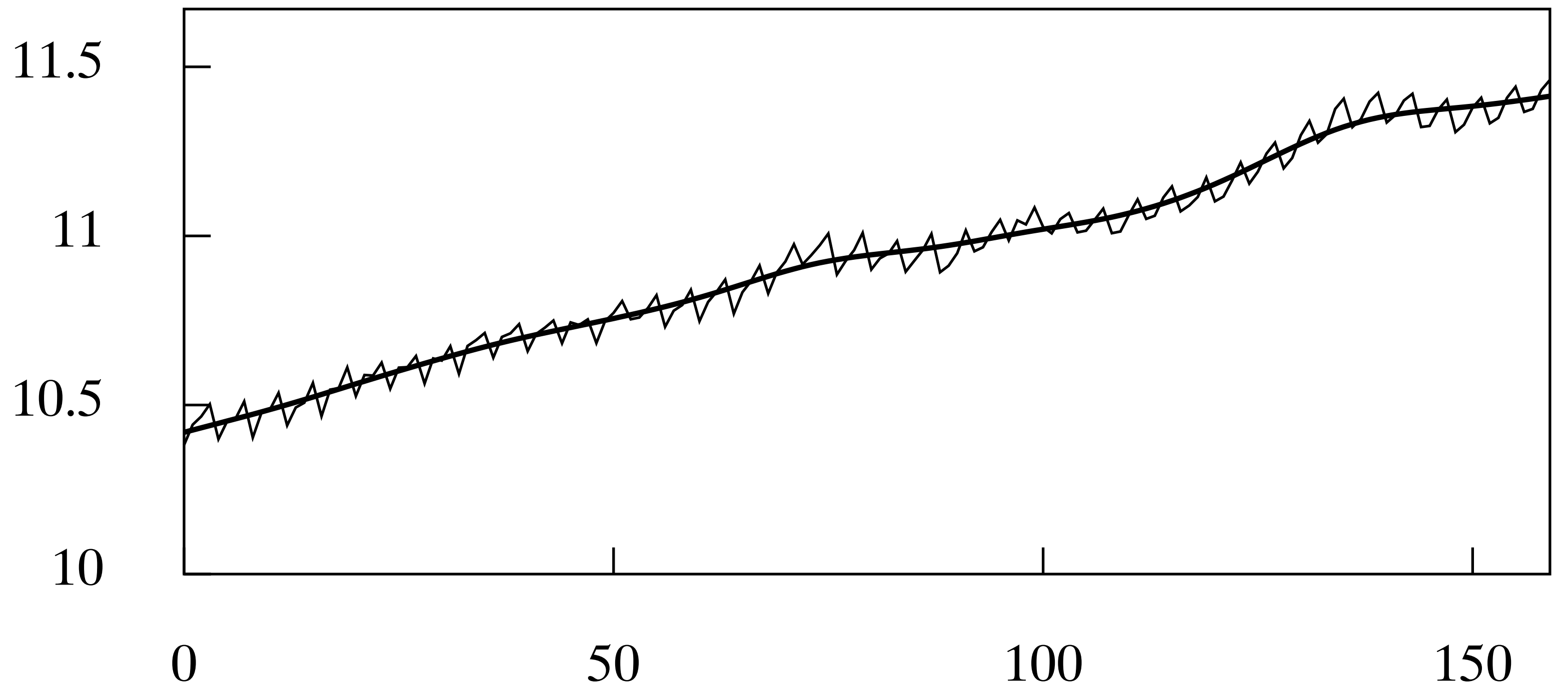

Figure 13. the quarterly logarithmic consumption data together with a trend interpolated by the lowpass Hodrick–Prescott filter with the smoothing parameter set to $\lambda = 1600$.

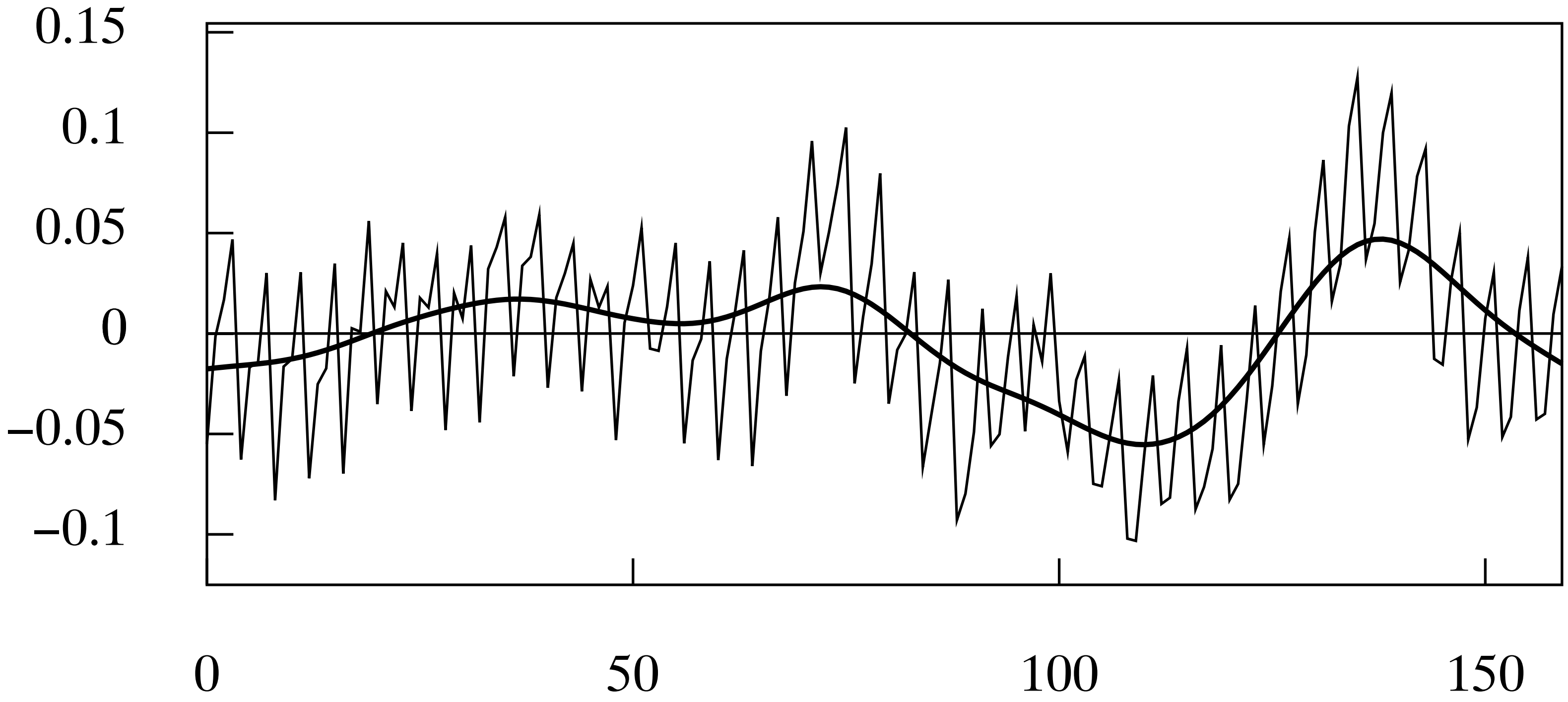

Figure 14. The residual sequence obtained by extracting a linear trend from the logarithmic consumption data, together with a low-frequency trajectory that has been obtained via the lowpass Hodrick–Prescott filter.

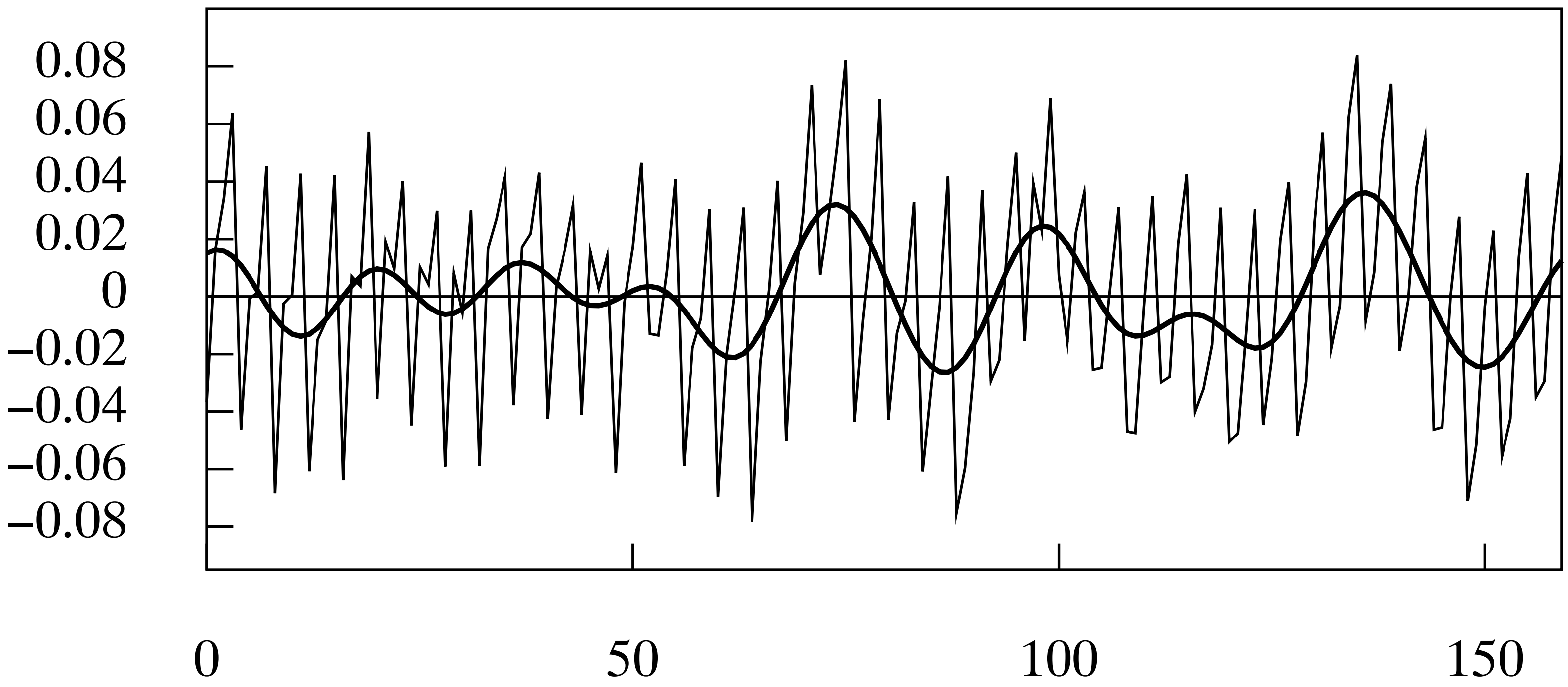

Figure 15. The sequence obtained by using the Hodrick–Prescott filter to extract the trend, together with a fluctuating component obtained by a lowpass frequency-domain filter with a cut-off point at $\pi/8$ radians.

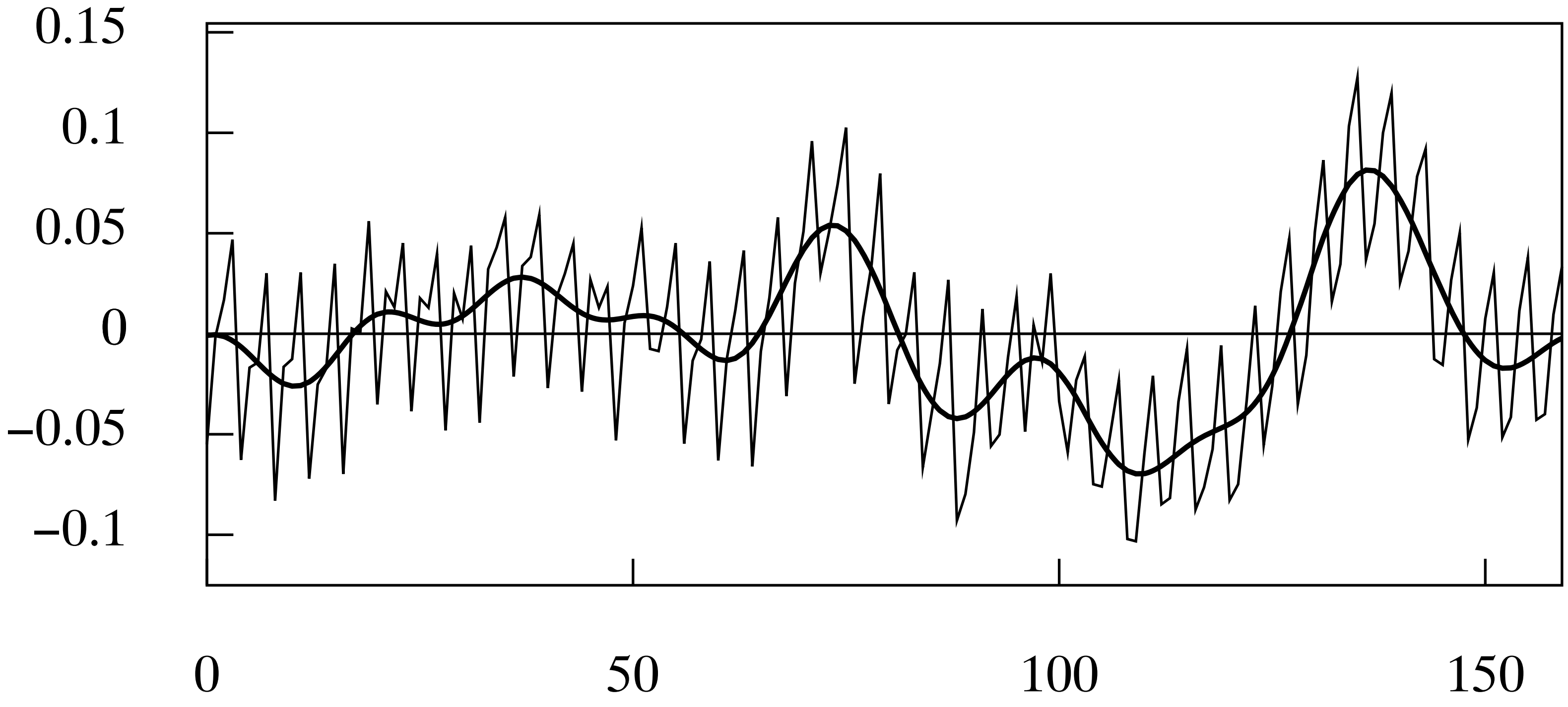

Figure 11. The residual sequence from fitting a linear trend to the logarithmic consumption data with an interpolated line representing the business cycle, obtained by the frequency-domain method.