Series: Economic Forecasting; Time Series: General; Time Series: Nonstationary Distributions and Unit Roots; Time Series: Seasonal Adjustment

Bibliography

Banerjee A, Dolado J J, Galbraith J W, Hendry D F 1993 *Co-Integration, Error Correction, and the Econometric Analysis of Non-stationary Data*. Oxford University Press, Oxford, UK

Box G E P, Tiao G C 1977 A canonical analysis of multiple time series. *Biometrika* **64**: 355–65

Chan N H, Tsay R S 1996 On the use of canonical correlation analysis in testing common trends. In: Lee J C, Johnson W O, Zellner A (eds.) *Modelling and Prediction: Honoring S. Geisser*. Springer-Verlag, New York, pp. 364–77

Chan N H, Wei C Z 1988 Limiting distributions of least squares estimates of unstable autoregressive processes. *Annals of Statistics* **16**: 367–401

Engle R F, Granger C W J 1987 Cointegration and error correction: Representation, estimation, and testing. *Econometrica* **55**: 251–76

Engle R F, Granger C W J 1991 *Long-run Economic Relations: Readings in Co-integration*. Oxford University Press, Oxford, UK

Hiraki T, Shiraishi N, Takezwa N 1996 Cointegration, common factors, and term structure of Yen offshore interest rates. *Journal of Fixed Income* **December**: 69–75

Johansen S 1988 Statistical analysis of co-integration vectors. *Journal of Economic Dynamics and Control* **12**: 231–54

Johansen S 1995 *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press, Oxford, UK

Johansen S 1996 Likelihood-based inference for co-integration of some nonstationary time series. In: Cox D R, Hinkley D V, Barndorff-Nielsen O E (eds.) *Time Series Models: In Econometrics, Finance and Other Fields*. Chapman and Hall, New York

Lee C M C, Myers J, Swaminathan B 1999 What is the intrinsic value of the Dow? *Journal of Finance* **54**: 1693–741

Lien D, Tse Y K 1999 Forecasting the Nikkei spot index with fractional co-integration. *Journal of Forecasting* **18**: 259–73

Lütkepohl H 1993 *Introduction to Multiple Time Series Analysis*. 2nd edn. Springer-Verlag, New York

Maddala G S, Kim I M 1998 *Unit Roots, Cointegration, and Structural Change*. Cambridge University Press, Cambridge, UK

Park J Y 1992 Canonical co-integrating regressions. *Econometrica* **60**: 119–43

Phillips P C B 1991 Optimal inference in co-integrated systems. *Econometrica* **59**: 283–306

Reinsel G C, Ahn S K 1992 Vector AR models with unit roots and reduced rank structure: estimation, likelihood ratio test, and forecasting. *Journal of Time Series Analysis* **13**: 353–75

Stock J H 1994 Unit roots, structural breaks and trends. In: Engle R F, McFadden D L (eds.) *Handbook of Econometrics*, Vol. IV. Elsevier, Netherlands

Stock J H, Watson M W 1988 Testing for common trends. *Journal of the American Statistical Association* **83**: 1097–107

Tanaka K 1996 *Time Series Analysis: Nonstationary and Noninvertibel Distribution Theory*. Wiley, New York

Watson M W 1994 Vector autoregressions and co-integration. In: Engle R F, McFadden D L (eds.) *Handbook of Econometrics* Vol. IV. Elsevier, The Netherlands

N. H. Chan

Time Series: Cycles

Time series data in economics and other fields of social science often exhibit cyclical behavior. For example, aggregate retail sales are high in November and December and follow a seasonal cycle; voter registrations are high before each presidential election and follow an election cycle; and aggregate macroeconomic activity falls into recession every six to eight years and follows a business cycle. In spite of this cyclicality, these series are not perfectly predictable, and the cycles are not strictly periodic. That is, each cycle is unique. Quite generally, it is useful to think of time series as realizations of stochastic processes, and this raises the question of how to represent cyclicality in stochastic processes. The major descriptive tool for this is the 'spectral density function' (or 'spectrum'), which is the subject of this article.

1. Cycles in a Typical Time Series

Figure 1 shows a time series plot of new housing authorizations ('building permits') issued by communities in the USA, monthly, from 1960 through 1999. This plot has characteristics that are typical of many economic time series. First, the plot shows a clear seasonal pattern: permits are low in the late fall and early winter, and rise markedly in the spring and summer. This seasonal pattern is persistent throughout the sample period, but does not repeat itself exactly year-after-year. In addition to a seasonal pattern, there is a slow moving change in the level of the series associated with the business cycle. For example, local minima of the series are evident in 1967, the mid-1970s, the early 1980s, and in 1990, which correspond to periods of macroeconomic slowdown or recession in the US economy.

How can the cyclical variability in the series be represented? One approach is to use a periodic function like a sine or cosine function. But this is inadequate in at least two ways. First, a deterministic periodic function doesn't capture the randomness in the series. Second, since several different periodicities (seasonal, business cycle) are evident in the plot, several periodic functions will be required. The next section presents a representation of a stochastic process that has these ingredients.

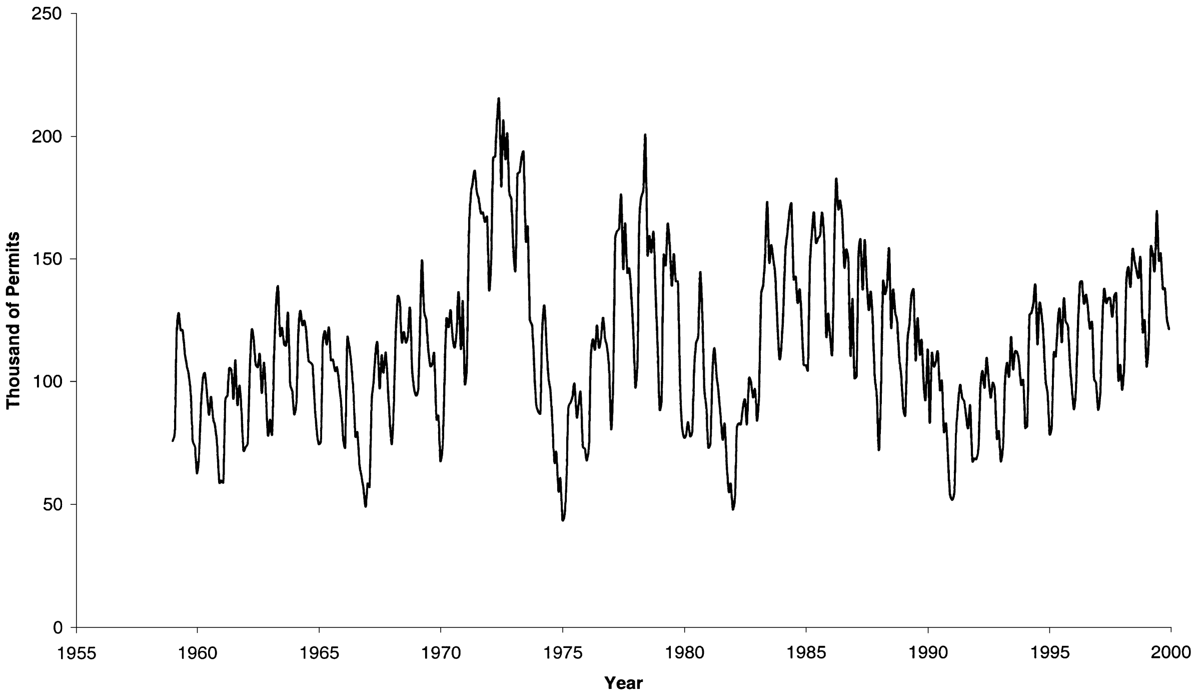

Figure 1
Building permits in the USA

2. *Representing Cyclical Behavior in Time Series: The Spectrum*

2.1 *Spectral Representation of a Covariance Stationary Process*

Consider a sequence of scalar random variables $\{Y_t\}$ where $t = 0, \pm 1, \pm 2, \ldots$, and $E(Y_t) = \mu$, $Cov(Y_t Y_{t-k}) = \lambda_k$. That is, assume that Y_t is observed at regular intervals, like a week, a month, a year, etc. Assume that the first and second moments of the process do not depend on time, that is μ and λ_k are time invariant. To simplify notation, assume that $\mu = 0$.

To motivate the spectral representation of Y_t, it is useful to begin with a very special and simple stochastic process:

$$Y_t = \alpha \cos(\omega t) + \delta \sin(\omega t) \tag{1}$$

where α and δ are random variables with $E(\alpha) = E(\delta) = 0$; $E(\alpha\delta) = 0$; and $E(\alpha^2) = E(\delta^2) = \sigma^2$. While simple, this process has three attractive characteristics. First, it is periodic: since $\cos(a + 2\pi) = \cos(a)$ and $\sin(a + 2\pi) = \sin(a)$, then $Y_{t+j2\pi/\omega} = Y_t$ for $|j| = 1, 2, \ldots$. So Y_t repeats itself with a period of $2\pi/\omega$. Second, the random components α and δ give Y a random amplitude (value of at its peak) and a random phase (value at $t = 0$). Thus, two realizations of Y will have different amplitudes and different timing of their peaks and troughs. Finally, the process is covariance stationary with $E(Y_t) = 0$ and $\lambda_k = \sigma^2 \cos(\omega k)$.

Adding together several components like Eqn. 1 produces a more interesting stochastic process:

$$Y_t = \sum_{j=1}^{n} [\alpha_j \cos(\omega_j t) + \delta_j \sin(\omega_j t)] \tag{2}$$

with $E(\alpha_j) = E(\delta_j) = 0$, for all j; $E(\alpha_j \delta_k) = 0$, for all j, k; $E(\alpha_j \alpha_k) = E(\delta_j \delta_k) = 0$, for $j \neq k$; and $E(\alpha_j^2) = E(\delta_j^2) = \sigma_j^2$. For this process, Y_t is the sum of n different uncorrelated components, each corresponding to a different frequency, and each with its own variance. (The variance of $\alpha_j \cos(\omega_j t) + \delta_j \sin(\omega_j t)$ is σ_j^2.) A calculation shows $E(Y_t) = 0$, $\lambda_k = \sum_{j=1}^{n} \sigma_j^2 \cos(\omega_j k)$, and the variance of Y_t is given by $var\,(Y_t) = \sum_{j=1}^{n} \sigma_j^2$.

So far, this has all been special. That is, two very special processes have been constructed. However, a fundamental result (Cramér 1942, Kolmorgorov 1940) shows that a generalization of this decomposition can be used to represent any covariance stationary process. The result, known as the 'Spectral Representation Theorem,' says that if Y_t is covariance stationary, then it can be represented as

$$Y_t = \int_0^{\pi} \cos(\omega t)\,\mathrm{d}\alpha(\omega) + \int_0^{\pi} \sin(\omega t)\,\mathrm{d}\delta(\omega)$$

where $d\alpha(\omega)$ and $d\delta(\omega)$ are zero mean random variables that are mutually uncorrelated, uncorrelated across frequency, and have variances that depend on frequency. Like the special example shown in Eqn. 2, this general representation decomposes Y_t into a set of strictly periodic components—each uncorrelated with the others, and each with its own variance. For example, processes with important seasonal components will have large values of the components corresponding to the seasonal frequency, processes for series with strong 'business cycle' components will have large variances at business cycle frequencies, etc.

Sometimes the spectral representation is written using complex numbers as:

$$Y_t = \int_{-\pi}^{\pi} e^{i\omega t} d\gamma(\omega) \tag{3}$$

where $d\gamma(\omega) = (1/\sqrt{2})[d\alpha(\omega) - i d\delta(\omega)]$ for $\omega \geqslant 0$ and $d\gamma(\omega) = d\gamma(-\omega)$ for $\omega < 0$. This representation simplifies notation for some calculations.

The spectral representation Eqn. 3 is a generalization of Eqn. 2 where the 'increments' $d\gamma(\omega)$ have orthogonality properties like those of α_j and δ_j. The variance of each component is frequency specific (like σ_j^2), and is summarized by a density function $S(\omega)$. Specifically, the variance of $d\gamma(\omega)$ is $E(d\gamma(\omega)\overline{d\gamma(\omega)}) = S(\omega)d\omega$. (The use of a density function to summarize the variance of $d\gamma(\omega)$ is not completely general and rules out processes with deterministic (or perfectly predictable) components. The references listed in the last section provide this more general result.)

You may wonder why the spectral representation doesn't use frequencies larger than π. The answer is that these are not needed to describe a process measured at discrete intervals. This discreteness means that periodic components associated with frequencies larger than π will look just like components associated with frequencies less than π. For example, a component with frequency 2π will have period of 1 and, because the series is measured only once every period, this component will appear to be constant—it will look just like (be 'aliased' with) a component that has a frequency of $\omega = 0$.

2.2 Relationship Between the Spectral Density and the Autocovariances

There is a one-to-one relationship between the spectral density function and the autocovariances of the process. The autocovariances follow directly from the spectral representation:

$$\begin{aligned}\lambda_k = E(Y_t Y_{t-k}) &= E(Y_t \bar{Y}_{t-k}) \\ &= E\left[\int e^{i\omega t}\, d\gamma(\omega) \int e^{-i\omega(t-k)}\, \overline{d\gamma(\omega)}\right] \\ &= \int e^{i\omega k} S(\omega)\, d\omega. \end{aligned} \tag{4}$$

The spectrum can be determining by inverting this relation:

$$S(\omega) = (2\pi)^{-1} \sum_{j=-\infty}^{\infty} \lambda_j e^{-i\omega j} \tag{5}$$

(To verify that Eqn. 5 is the inverse of Eqn. 4, use

$$\int_{-\pi}^{\pi} e^{i\omega k} d\omega = \left\{ \begin{array}{lll} 2\pi & \text{for} & k = 0 \\ 0 & \text{for} & k \neq 0 \end{array} \right\} \tag{6}$$

so that $(2\pi)^{-1}\{\int_{-\pi}^{\pi} \sum_{j=-\infty}^{\infty} \lambda_j e^{-i\omega(j-k)} d\omega\} = \lambda_k$.)

A simpler formula for the spectrum is

$$S(\omega) = (2\pi)^{-1}[\lambda_0 + 2\sum_{j=1}^{\infty} \lambda_j \cos(j\omega)]. \tag{7}$$

which follows from Eqn. 5, since $\lambda_j = Cov(Y_t Y_{t-j}) = Cov(Y_{t+j} Y_t) = \lambda_{-j}$.

2.3 A Summary of the Properties of the Spectrum

Summarizing the results presented in Sect. 2.2., the spectrum (or spectral density function) has 4 important properties.

(a) $S(\omega)d\omega$ can be interpreted as the variance of the cyclical component of Y corresponding to the frequency ω. The period of this component is $2\pi/\omega$.

(b) $S(\omega) \geqslant 0$. This follows, because $S(\omega)$ is a variance function.

(c) $S(\omega) = S(-\omega)$. This follows from the definition of $\gamma(\omega)$ in the spectral representation Eqn. 3, or from Eqn. 7, since $\cos(a) = \cos(-a)$. Because of this symmetry, plots of the spectrum are presented $0 \leqslant \omega \leqslant \pi$.

(d) $Var(Y) = \int_{-\pi}^{\pi} S(\omega)\, d\omega$.

2.4 The Spectrum of Building Permits

The time series plot of building permits (Fig. 1) shows that most of the variability in the series comes from two sources: seasonal variation over the year and business-cycle variability. This is evident in the estimated spectrum for the series, shown in Fig. 2. Most of the mass in the spectrum is concentrated around the seven peaks evident in the plot. (These peaks are sufficiently large that spectrum is plotted on a log scale.) The first peak occurs at frequency $\omega = 0.07$ corresponding to a period of 90 months. This represents the business cycle variability in the series. The other peaks occur at frequencies $2\pi/12$, $4\pi/12$, $6\pi/12$, $8\pi/12$, and π. These are peaks for the seasonal frequencies: the first corresponds to a period of 12 months, and the others are the seasonal 'harmonics' 6, 4, 3, and 2 months.

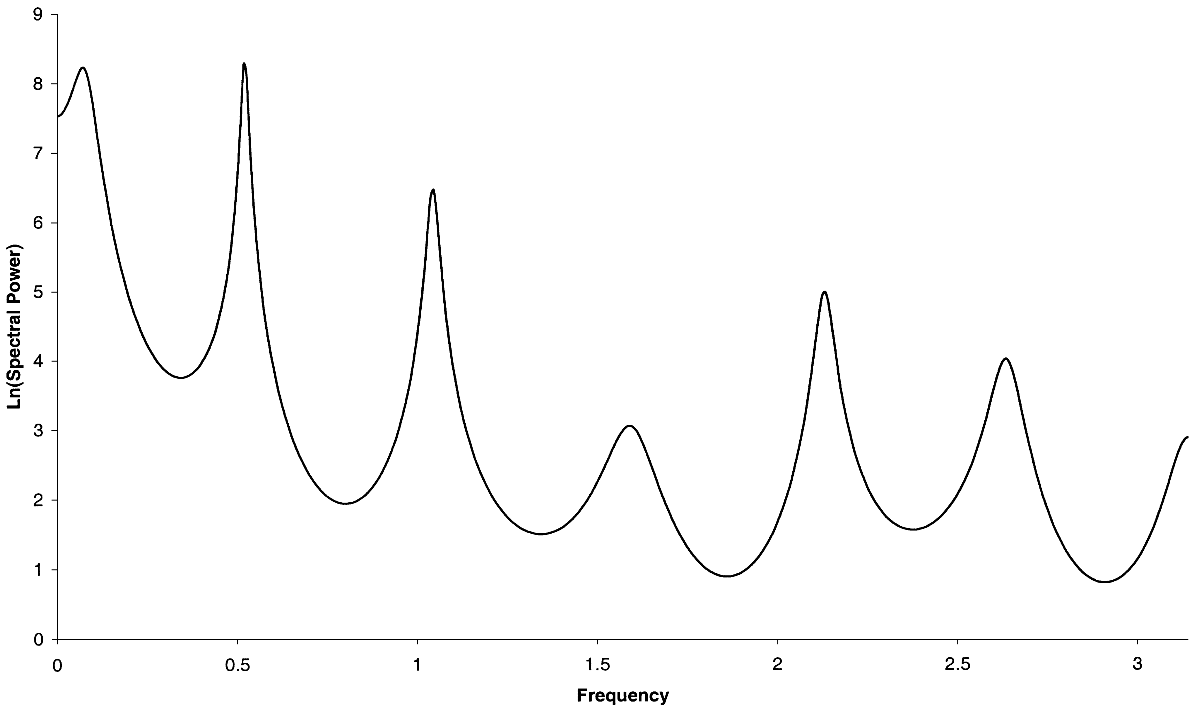

Figure 2
Spectrum of building permits

3. Spectral Properties of Moving Average Filters

3.1 Some General Results

Often, one time series (Y) is converted into another time series (X) through a moving average function such as:

$$X_t = \sum_{j=-r}^{s} c_j Y_{t-j}$$

For example, X_t might be the first difference of Y_t, i.e., $X_t = Y_t - Y_{t-1}$, or an annual moving average $X_t = \frac{1}{12}\sum_{j=0}^{11} Y_{t-j}$, or one of the more complicated moving averages that approximate official seasonal adjustment procedures (see *Time Series: Seasonal Adjustment*).

Let the 'lag' operator L shift time series back by one period (so that $LY_t = Y_{t-1}$, $L^2 Y_t = L(LY_t) = Y_{t-2}$, $L^{-1}Y_t = Y_{t+1}$, etc.), then this moving average can be represented as:

$$X_t = c(L)Y_t$$

where,

$$c(L) = c_{-r}L^{-r} + \cdots + c_{-1}L^{-1} + c_0 L^0 + c_1 L + \cdots + c_s L^s$$

The operator $c(L)$ is sometimes called a 'linear filter.'

How does $c(L)$ change the spectral properties of Y? If the spectral representation of Y_t is written as:

$$Y_t = \int_{-\pi}^{\pi} \mathrm{e}^{i\omega t}\, \mathrm{d}\gamma(\omega)$$

then X_t can be written as,

$$\begin{aligned} X_t &= \sum_{j=-r}^{s} c_j Y_{t-j} \\ &= \sum_{j=-r}^{s} c_j \int_{-\pi}^{\pi} \mathrm{e}^{i\omega(t-j)}\, \mathrm{d}\gamma(\omega) \\ &= \int_{-\pi}^{\pi} \mathrm{e}^{i\omega t} \left(\sum_{j=-r}^{s} c_j e^{-i\omega j} \right) \mathrm{d}\gamma(\omega) \\ &= \int_{-\pi}^{\pi} \mathrm{e}^{i\omega t} c(\mathrm{e}^{-i\omega})\, \mathrm{d}\gamma(\omega) \\ &= \int_{-\pi}^{\pi} \mathrm{e}^{i\omega \left(t - \frac{p(\omega)}{\omega}\right)} g(\omega)\, \mathrm{d}\gamma(\omega) \end{aligned} \tag{8}$$

where the last line uses the polar form of $c(\mathrm{e}^{-i\omega})$: $c(\mathrm{e}^{-i\omega}) = g(\omega)\mathrm{e}^{-ip(\omega)}$ where

$$g(\omega) = |c(\mathrm{e}^{i\omega})| \tag{9}$$

and

$$\rho(\omega) = \tan^{-1}\left[\frac{-\operatorname{Im}[c(\mathrm{e}^{-i\omega})]}{\operatorname{Re}[c(\mathrm{e}^{-i\omega})]}\right] \tag{10}$$

This representation for X_t shows that the filter $c(L)$ 'amplifies' each component of Y_t by the factor $g(\omega)$ and shifts it back in time by $\rho(\omega)/\omega$ time units.

The function $g(\omega)$ is called the filter *gain* (or sometimes the 'amplitude gain.' The function $\rho(\omega)$ is called the filter 'phase' and $g(\omega)^2 = [c(\mathrm{e}^{-i\omega})c(\mathrm{e}^{i\omega})]$ is called the 'power transfer function' of the filter.

3.2 Three Examples

3.2.1 Example one: lag filter. Suppose that $c(L) = L^2$. Then $c(\mathrm{e}^{-i\omega}) = \mathrm{e}^{-2i\omega}$,

$$g(\omega) = |c(\mathrm{e}^{i\omega})| = 1$$

and

$$\rho(\omega) = \tan^{-1}\left[\frac{\sin 2\omega}{\cos 2\omega}\right] = 2\omega$$

Thus, this filter does not alter the amplitude of any of the components, $g(\omega) = 1$, but shifts each component back in time by $\rho(\omega)/\omega = 2$ time periods.

3.2.2 Example two: first difference filter. Suppose that $c(L) = (1-L)$ so that X_t is the first difference of Y_t. Then $c(\mathrm{e}^{-i\omega}) = 1-\mathrm{e}^{-i\omega}$,

$$g(\omega) = (-\mathrm{e}^{-i\omega}+2-\mathrm{e}^{i\omega})^{\frac{1}{2}} = \sqrt{2(1-\cos(\omega))}$$

and

$$\rho(\omega) = \tan^{-1}\left[\frac{\sin(\omega)}{1-\cos(\omega)}\right]$$

Thus, both the gain and phase are frequency specific. For example, $g(0) = 0$ so that the first difference filter eliminates the lowest frequency component of the series (the level of the series is 'differenced out'), and $g(\pi) = \sqrt{2}$ so that the high frequency components are amplified (X_t is 'choppier' than Y_t).

3.2.3 Example three: census X-12 seasonal adjustment Filter. The official monthly seasonal adjustment procedure in the USA and several other countries (Census X-12) can be well approximated by a linear filter $c(L) = \sum_{j=-84}^{84} c_{|j|}L^j$ where the coefficients c_i are given in Wallis (1974). Since the filter is symmetric, $c_j = c_{-j}$, $c(\mathrm{e}^{i\omega})$ is real and so $\rho(\omega) = 0$. The function $g(\omega)$ is plotted in Fig. 3. The gain is nearly unity except near the seasonal frequencies, and so the filter can be interpreted as producing a new series that leaves the non-seasonal components unaltered, but eliminates the seasonal components. For a detailed discussion of seasonal adjustment procedures see *Time Series: Seasonal Adjustment*.

3.3 Band-pass Filters

The seasonal adjustment filter has the desirable properties of (essentially) zeroing out certain frequency components (the 'seasonal') and leaving the other components (the non-seasonals) unaltered. Filters with this characteristic are called 'band-pass filters,' since they 'pass' components associated with certain frequencies.

Band-pass filters for prespecified frequencies are easy to construct. For example, suppose that you want a filter with no phase shift that passes components between $0 \leqslant \omega \leqslant \underline{\omega}$. That is, you want to construct a filter, say $c(L)$ with gain $|g(\omega)| = 1$ for $0 \leqslant \omega \leqslant \underline{\omega}$ and $|g(\omega)| = 0$ for $\underline{\omega} < \omega \leqslant \pi$. (Assume that the corresponding negative frequencies are also passed.) Choosing a symmetric filter, $c_j = c_{-j}$, will make the phase 0 for all ω. In this case, $c(\mathrm{e}^{i\omega})$ is a real number, and so the gain is given by:

$$g(\omega) = \sum_{j=-\infty}^{\infty} c_{|j|}\mathrm{e}^{-i\omega j}$$

Now, using Eqn. 6, it is straightforward to verify the identity

$$\begin{aligned} c_j &= (2\pi)^{-1}\int_{-\pi}^{\pi} \mathrm{e}^{-i\omega j} \sum_{j=-\infty}^{\infty} c_j \mathrm{e}^{-i\omega j}\,\mathrm{d}\omega \\ &= (2\pi)^{-1}\int_{-\pi}^{\pi} \mathrm{e}^{i\omega j} c(\mathrm{e}^{i\omega})\,\mathrm{d}\omega \end{aligned}$$

Replacing $c(\mathrm{e}^{i\omega}) = g(\omega)$ with the target value of the gain for the band-pass filter and carrying out the integration yields

$$c_j = (2\pi)^{-1}\frac{1}{ij}\,\mathrm{e}^{i\omega j}\Big|_{-\underline{\omega}}^{\underline{\omega}} = \left\{\begin{array}{ll} \dfrac{1}{j\pi}\sin(\underline{\omega} j) & \text{for} \quad j \neq 0 \\ \underline{\omega}/\pi \quad \text{for} \quad j = 0 & \end{array}\right\}$$

The resulting filter $c(L)$ passes all components with frequencies between $0 \leqslant \omega \leqslant \underline{\omega}$. This means that the filter $1-c(L)$ passes everything except these frequencies. The filter $c(L)$ is called a low-pass filter (it passes low frequencies), and $1-c(L)$ is called a high-pass filter (it passes the high frequencies). Combinations of high- and low-pass filters can be used to construct band-pass filters for any set of frequencies. For example, a filter that passes components with frequencies between ω_1 to ω_2 (with $\omega_2 \geqslant \omega_1$) can be

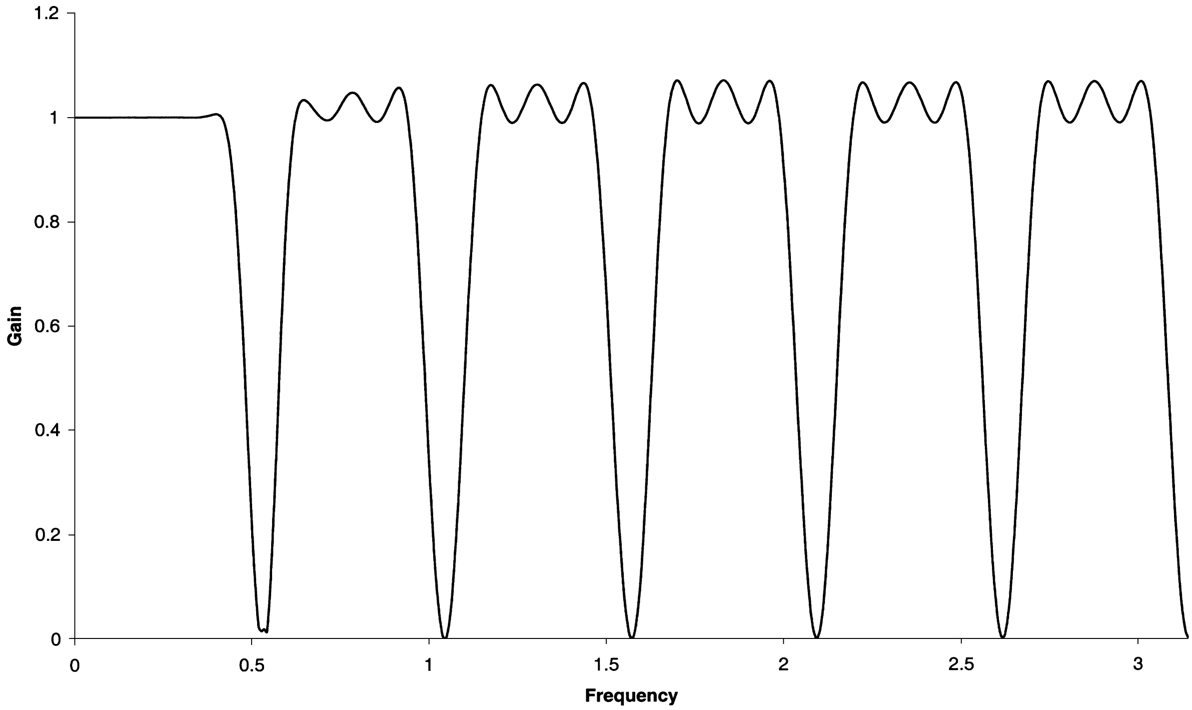

Figure 3
Gain of linear approximation to monthly × 12 seasonal adjustment filter

constructed as the difference of the low-pass filter for ω_2 and the low-pass filter for ω_1.

One practical problem with exact band-pass filters is that the coefficients c_j die out very slowly (at the rate $1/j$). This introduces important 'endpoint' problems when applying these filters to finite realizations of data, say $\{Y_t\}_{t=1}^{T}$. One approach is simply to truncate the filter at some point, for example to use $c_k(L) = \sum_{j=-k}^{k} c_j L^j$ and apply this filter to $\{Y_t\}_{t=k+1}^{T-k}$. An alternative procedure is to construct a minimum mean square error estimate of the infeasible band-pass values:

$$X_t = c(L)Y_t = \sum_{i=-\infty}^{\infty} c_i Y_{t-i},$$

as

$$E(X_t \mid \{Y_j\}_{j=1}^{T}) = \sum_{i=-\infty}^{\infty} c_i E(Y_{t-i} \mid \{Y_j\}_{j=1}^{T})$$

This can be accomplished by using backcasts and forecasts for the missing pre-sample and post-samples values of Y_t. The relative merits of these two approaches for band-pass filters is discussed in Baxter and King (1999), Geweke (1978) and Dagum (1980) discuss this problem in the context of seasonal adjustment.

4. *Spectra of Commonly Used Stochastic Processes*

Suppose that Y has spectrum $S_Y(\omega)$ and $X_t = c(L)Y_t$. What is the spectrum of X? As was shown in Eqn. 8, the frequency components of X are the frequency components of Y scaled by the factor $g(\omega)e^{-i\rho(\omega)}$, where $g(\omega)$ is the gain and $\rho(\omega)$ is the phase of $c(L)$. This means that spectra of X and Y are related by:

$$S_X(\omega) = g(\omega)^2 S_Y(\omega) = |c(e^{-i\omega})|^2 S_Y(\omega)$$

which follows from $|e^{-i\rho(\omega)}| = 1$ and the definition of $g(\omega)$.

Now, suppose that ε_t is a 'white noise' process, defined by the properties $E(\varepsilon_t) = 0$, $\lambda_0 = \sigma^2$, and $\lambda_k = 0$ for $k \neq 0$. The spectrum of ε is then easily calculated from Eqn. 7:

$$S_\varepsilon(\omega) = (2\pi)^{-1}\sigma^2$$

So, the spectrum of white noise is constant. (Which is why the process is called 'white' noise.)

Now suppose that $Y_t = c(L)\varepsilon_t$. Then:

$$S_Y(\omega) = |c(e^{-i\omega})|^2 S_\varepsilon(\omega) = |c(e^{-i\omega})|^2 (2\pi)^{-1}\sigma^2$$

This result can be used to determine the spectrum of any stationary ARMA process (see *Time Series:*

ARIMA Methods) for a detailed discussion of these models). If Y_t follows an ARMA process, then it can be represented as:

$$\phi(L)Y_t = \theta(L)\varepsilon_t$$

The autoregressive operator, $\phi(L)$ can be inverted to yield $Y_t = c(L)\varepsilon_t$ with $c(L) = \phi(L)^{-1}\theta(L)$. This means that:

$$S_y(\omega) = |c(\mathrm{e}^{-i\omega})|^2(2\pi)^{-1}\sigma^2 = \frac{2(\pi)^{-1}\sigma^2|\theta(\mathrm{e}^{-i\omega})|^2}{|\phi(\mathrm{e}^{-i\omega})|^2} \tag{11}$$

As an example, consider the AR(1) model:

$$Y_t = \phi Y_{t-1} + \varepsilon_t$$

equivalently written as,

$$(1-\phi L)Y_t = \varepsilon_t.$$

Applying Eqn. 11 yields

$$S_Y(\omega) = \frac{\sigma^2}{2\pi}\frac{1}{|1-\phi\mathrm{e}^{-i\omega}|^2} = \frac{\sigma^2}{2\pi}\frac{1}{(1+\phi^2-2\phi\cos(\omega))}$$

This spectrum is equal to $\sigma^2[2\pi(1+\phi^2-2\phi\cos(\omega))]^{-1}$ at $\omega = 0$. When $0 < \phi < 1$, it falls steadily as ω increases from 0 to π. This means that, relative to white noise, the low frequency components of the AR(1) are more important than the high frequency components. Thus, realizations of the series appear smoother than white noise.

5. *Spectral Estimation*

There are two general approaches to estimating spectra. Perhaps the simplest is to estimate an ARMA model for the series (as explained in *Time Series: ARIMA Methods*) and then compute the implied spectrum for this estimated model. Asymptotic approximations can then be used for statistical inference. Since the estimators of the ARMA parameters are asymptotically normally distributed and the spectrum is a smooth function of these parameters (see Eqn. 11), the estimated spectrum at any point ω will be asymptotically normal with a variance that can be computed using the δ-method (a mean-value expansion).

An interesting extension of this procedure is discussed in Berk (1974). He shows that, quite generally, a suitably long autoregression can be specified and the spectrum estimated using the estimated AR coefficients. The precise definition of 'suitably long' depends on the size of the available sample, with more terms included in larger samples. These 'autoregressive spectral estimators' are shown by Berk to be consistent and approximately normally distributed in large samples.

An alternative set of estimators are based on nonparametric methods. These estimators can be motivated by Eqn. 7 which shows the spectrum as a weighted sum of the autocovariances. This suggests an estimator of the form:

$$S(\omega) = (2\pi)^{-1}\sum_{k=-l_T}^{l_T} w_T(k)\hat{\lambda}_k$$

where $\hat{\lambda}_k$ is an estimator of λ_k, l_T are truncation values that depend on the sample size T, and $w_T(k)$ is a weight function that also depends on T. These estimators can be studied using nonparametric methods (see *Nonparametric Statistics: Asymptotics*).

A closely related set of nonparametric estimators are based on the 'periodogram'. The periodgram provides a frequency decomposition of the sample variance of a partial realization of the process, say $\{Y_t\}_{t=1}^T$. The periodgram is computed at frequencies $\omega_j = 2\pi j/T$, $j = 1, 2,\ldots, T/2$ (assuming T is even), and the value of the periodgram at frequency ω_j is:

$$p_j = \frac{2}{T}\left|\sum_{t=1}^{T}(y_t-\bar{y})\mathrm{e}^{-i\omega_j t}\right|^2$$

The periodgram has three interesting asymptotic sampling properties: (a) $E(p_j)/4\pi \to S_Y(\omega_j)$; (b) $var(p_j/2\pi) \to S_Y(\omega_j)^2$; and (c) $cov(p_j, p_k) = 0$ for $j \neq k$. Property (a) shows that a scaled version of the periodgram provides an asymptotically unbiased estimator of the spectrum, but since the variance doesn't approach zero (property (b)), the estimator is not consistent. The third property shows that averaging a set of periodgram ordinates around a given frequency reduces the variance of the estimator, although (from property (a)) it may introduce bias. By carefully selecting the averaging weights, consistent estimators with good finite sample mean square error properties can be constructed.

Some recent advances for nonparametric spectral estimators are developed in Andrews (1991).

See also: Nonparametric Statistics: Asymptotics; Time Series: ARIMA Methods; Time Series: Seasonal Adjustment

Bibliography

Andrews D W K 1991 Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* **59**: 817–58

Baxter M B, King R G 1999 Measuring business cycles: approximate band-pass filters for economic time series. *Review of Economics and Statistics* **81**(4): 575–93

Berk K N 1974 Consistent autoregressive spectral estimates. *Annals of Statistics* **2**(3): 489–502

Brockwell P J, Davis R A 1991 *Time Series: Theory and Methods*, 2nd edn. Springer, New York

Cramér H 1942 On harmonic analysis of certain function spaces. *Arkiv für Matematik, Astronomi och Fysik*. **28B**(12): 1–7

Dagum E B 1980 *The X-11 ARIMA Seasonal Adjustment Method*. Research Paper, Statistics, Canada

Geweke J F 1978 The revision of seasonally adjusted time series. *Proceedings of the Business and Economic Statistics Section—American Statistical Association*, pp. 320–25

Granger C W J, Newbold P 1977 *Forecasting Economic Time Series*. Academic Press, New York

Hamilton J 1994 *Time Series Analysis*. Princeton University Press, Princeton, NJ

Harvey A 1993 *Time Series Models*, 2nd edn. MIT Press, Cambridge, UK

Kolmorgorov A N 1940 Kurven in Hilbertschen Raum die gegenüber eine einparametrigen Gruppe von Bewegungen invariant sind. *C. R. (Doklady) de L'Academie des Sciences de l'IRSS, New Series* **26**: 6–9

Nerlove M, Grether D M, Carvalho J L 1979 *Analysis of Economic Time Series*. Academic Press, New York

Priestly M B 1981 *Spectral Analysis and Time Series*. Academic Press, New York

Wallis K F 1974 Seasonal adjustment and relations between variables. *Journal of the American Statistical Association* **69**: 18–31

M. W. Watson

Time Series: Economic Forecasting

Time-series forecasts are used in a wide range of economic activities, including setting monetary and fiscal policies, state and local budgeting, financial management, and financial engineering. Key elements of economic forecasting include selecting the forecasting model(s) appropriate for the problem at hand, assessing and communicating the uncertainty associated with a forecast, and guarding against model instability.

1. Time Series Models for Economic Forecasting

Broadly speaking, statistical approaches to economic forecasting fall into two categories: time-series methods and structural economic models. Time-series methods use economic theory mainly as a guide to variable selection, and rely on past patterns in the data to predict the future. In contrast, structural economic models take as a starting point formal economic theory and attempt to translate this theory into empirical relations, with parameter values either suggested by theory or estimated using historical data. In practice, time-series models tend to be small with at most a handful of variables, while structural models tend to be large, simultaneous equation systems which sometimes incorporate hundreds of variables (see *Economic Panel Data*; *Simultaneous Equation Estimation: Overview*). Time-series models typically forecast the variable(s) of interest by implicitly extrapolating past policies into the future, while structural models, because they rely on economic theory, can evaluate hypothetical policy changes. In this light, perhaps it is not surprising that time-series models typically produce forecasts as good as, or better than, far more complicated structural models. Still, it was an intellectual watershed when several studies in the 1970s (reviewed in Granger and Newbold 1986) showed that simple univariate time-series models could outforecast the large structural models of the day, a result which continues to be true (see McNees 1990). This good forecasting performance, plus the relatively low cost of developing and maintaining time-series forecasting models, makes time-series modeling an attractive way to produce baseline economic forecasts.

At a general level, time-series forecasting models can be written,

$$y_{t+h} = g(X_t, \theta) + \varepsilon_{t+h} \tag{1}$$

where y_t denotes the variable or variables to be forecast, t denotes the date at which the forecast is made, h is the forecast horizon, X_t denotes the variables used at date t to make the forecast, θ is a vector of parameters of the function g, and ε_{t+h} denotes the forecast error. The variables in X_t usually include current and lagged values of y_t. It is useful to define the forecast error in (1) such that it has conditional mean zero, that is, $E(\varepsilon_{t+h}|X_t) = 0$. Thus, given the predictor variables X_t, under mean-squared error loss the optimal forecast of y_{t+h} is its conditional mean, $g(X_t, \theta)$. Of course, this forecast is infeasible because in practice neither g nor θ are known. The task of the time-series forecaster therefore is to select the predictors X_t, to approximate g, and to estimate θ in such a way that the resulting forecasts are reliable and have mean-squared forecast errors as close as possible to that of the optimal infeasible forecast.

Time-series models are usefully separated into univariate and multivariate models. In univariate models, X_t consists solely of current and past values of y_t. In multivariate models, this is augmented by data on other time series observed at date t. The next subsections provide a brief survey of some leading time-series models used in economic forecasting. For simplicity attention is restricted to one-step ahead forecasts ($h = 1$). Here, the focus is on forecasting in a stationary environment; the issue of nonstationarity in the form of structural breaks or time varying parameters is returned to below.

1.1 Univariate Models

Univariate models can be either linear, so that g is linear in X_t, or nonlinear. All linear time-series models

International Encyclopedia of the Social & Behavioral Sciences ISBN: 0-08-043076-7