STAT 626: Outline of Lecture 19

Spectral Density Func., DFT, Periodogram (§4.3, §4.4)

1. There are two major approaches to time series analysis, both have the goal of reducing the data to WN or making them **uncorrelated**:

   Time-Domain Approach: Relies on the correlations, ACF and PACF. Finite parameter ARMA models reduce the data to WN. It is more popular in economics and social sciences, works better for smaller sample sizes (about 80 or more), and the results are more interpretable in these areas of application and useful for forecasting.

   Spectral-Domain Approach: Relies on the **spectral density function** which is the Fourier series (transform) of the ACF. The finite discrete Fourier transform (DFT) makes the data (nearly) uncorrelated. The approach is more popular in engineering, physical and biological sciences, requires larger sample sizes (it is non-parametric in nature) about 1000, and the results are more interpretable in these areas of application, because of their inherent interest in detection of period, cycle, oscillation, etc.

   NOTE: State-space modeling or Kalman filtering is closely related to the first method. It uses the idea of Gram-Schmidt orthogonaliztaion in the framework of vactor autoregression (VAR), see Chap. 6 for more details.

The Theory or Population Concepts

2. **Spectral Density Function (§4.3)**: For a stationary time series with the autocovariance function $\{\gamma(h)\}$, its *spectral density function (sdf)* is defined as the Fourier series (transform) of the ACF

$$f(\omega) = \sum_{h=-\infty}^{\infty} \gamma(h) e^{-2\pi i \omega h} \geq 0, \quad \omega \in [0, 0.5]$$

provided that $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$.

The sdf is a population parameter or quantity, we need to internalize its shape for various models, just like what we did for the shapes of ACF and PACF of various MA, AR, ARMA models.

3. **What is the statistical meaning of the sdf?**
At each frequency ω, as a nonnegative quantity, $f(\omega)$ can be viewed as the **variance (power)** of that frequency.
This can be seen using the inverse of the above Fourier series, where,

$$\gamma(h) = \int_{-.5}^{.5} e^{2\pi i \omega h} f(\omega) d\omega, \quad h = 0, 1, 2, \ldots,$$

and for $h = 0$:

$$\gamma(0) = \mathrm{Var}(x_t) = \int_{-.5}^{.5} f(\omega) d\omega = \text{Sum of Variances (Powers) at various frequencies.}$$

Surprisingly, this looks like a sort of ANOVA for a time series.

4. What is the spectral density function (sdf) of
A white noise?

$$f(\omega) = \sigma_w^2, \quad \omega \in [0, .5].$$

What is the sdf of other time series models?

Filter Theorem: Property 4.7, p.222

Analogue of

$$\text{Var}(aX) = a^2\text{Var}(X),$$

for X a random variable and a constant.

5. An MA(q): $x_t = w_t + \theta_1 w_{t-1} + \ldots + \theta_q w_{t-q}, \quad \theta_q \neq 0$?

$$f(\omega) = \sigma_w^2 |A(\omega)|^2 = \sigma_w^2 |\sum_{k=0}^{q} \theta_k e^{-2\pi i \omega k}|^2, \quad \omega \in [0, .5],$$

where

$$A(\omega) = \sum_{k=0}^{q} \theta_k e^{-2\pi i \omega k},$$

is called the *frequency response function* of the filter.

6. A linear process: $x_t = \sum_{j=-\infty}^{+\infty} \psi_j w_{t-j}$?

Recall that it is stationary with the *autocovariance function*

$$\gamma(h) = \sigma_w^2 \sum_{j=-\infty}^{\infty} \psi_{j+h}\psi_j.$$

7. An AR(p) Model:

$$x_t = \phi_1 x_{t-1} + \ldots + \phi_p x_{t-p} + w_t, \quad \phi_p \neq 0?$$

8. ARMA(p,q)?

The Sample Concepts: Hidden Periodocity

9. **Examples 2.8, 2.9: Trig. Regression and Periodogram**

$$x_t = A\cos(2\pi\omega_j t + \phi) + w_t = \beta_1 \cos 2\pi\omega_j t + \beta_2 \sin 2\pi\omega_j t + w_t, \quad t = 1, \ldots, n,$$

   where $\omega_j = 2\pi/n$, is the fundamental Fourier frequency.

10. Connection between SSE, DFT and Periogogram

   Writing the above linear model in matrix form, the normal equations for the LS estimation of the parameters are,

$$Z'Z\beta = Z'x,$$

   where the design matrix Z is $n \times 2$ and $Z'Z = \text{diag}(n/2, n/2)$, due to orthogonality of its columns.

   What is the first column of Z? Second column?

   Furthermore,

$$\text{SSE}(\omega_j) = \sum x_t^2 - x'Z(Z'Z)^{-1}Z'x,$$

   where basic calculations reveal that the entries of $Z'x$ are the sine and cosine transforms of the time series data $x_1, \ldots, x_n$.

11. **DFT and Periodogram (§4.4)**: Finally,

$$\text{SSE}(\omega_j) = \sum x_t^2 - 2I(\omega_j), \quad j = 0, 1, \ldots, m,$$

   where the periodogram is the square of the absolute value of the DFT:

$$I(\omega_j) = |d(\omega_j)|^2.$$

4.4 Periodogram and Discrete Fourier Transform

We are now ready to tie together the periodogram, which is the sample-based concept presented in §4.2, with the spectral density, which is the population-based concept of §4.3.

Definition 4.1 *Given data $x_1, \ldots, x_n$, we define the* **discrete Fourier transform (DFT)** *to be*

$$d(\omega_j) = n^{-1/2} \sum_{t=1}^{n} x_t e^{-2\pi i \omega_j t} \tag{4.18}$$

for $j = 0, 1, \ldots, n-1$, where the frequencies $\omega_j = j/n$ are called the **Fourier** *or* **fundamental frequencies**.

If n is a highly composite integer (i.e., it has many factors), the DFT can be computed by the fast Fourier transform (FFT) introduced in Cooley and Tukey (1965). Also, different packages scale the FFT differently, so it is a good idea to consult the documentation. R computes the DFT defined in (4.18) without the factor $n^{-1/2}$, but with an additional factor of $e^{2\pi i \omega_j}$ that can be ignored because we will be interested in the squared modulus of the DFT. Sometimes it is helpful to exploit the inversion result for DFTs which shows the linear transformation is one-to-one. For the inverse DFT we have,

$$x_t = n^{-1/2} \sum_{j=0}^{n-1} d(\omega_j) \mathrm{e}^{2\pi i \omega_j t} \tag{4.19}$$

for $t = 1, \ldots, n$. The following example shows how to calculate the DFT and its inverse in R for the data set $\{1, 2, 3, 4\}$; note that R writes a complex number $z = a + ib$ as `a+bi`.

```
(dft = fft(1:4)/sqrt(4))
  [1]  5+0i  -1+1i  -1+0i  -1-1i
(idft = fft(dft, inverse=TRUE)/sqrt(4))
  [1] 1+0i  2+0i  3+0i  4+0i
(Re(idft))  # keep it real
  [1] 1  2  3  4
```

We now define the periodogram as the squared modulus[5] of the DFT.

Definition 4.2 *Given data $x_1, \ldots, x_n$, we define the* **periodogram** *to be*

$$I(\omega_j) = |d(\omega_j)|^2 \tag{4.20}$$

for $j = 0, 1, 2, \ldots, n-1$.

[5] Recall that if $z = a + ib$, then $\bar{z} = a - ib$, and $|z|^2 = z\bar{z} = a^2 + b^2$.

Note that $I(0) = n\bar{x}^2$, where $\bar{x}$ is the sample mean. In addition, because $\sum_{t=1}^{n} \exp(-2\pi it\frac{j}{n}) = 0$ for $j \neq 0$,[6] we can write the DFT as

$$d(\omega_j) = n^{-1/2} \sum_{t=1}^{n} (x_t - \bar{x}) e^{-2\pi i\omega_j t} \tag{4.21}$$

for $j \neq 0$. Thus, for $j \neq 0$,

$$\begin{aligned} I(\omega_j) = |d(\omega_j)|^2 &= n^{-1} \sum_{t=1}^{n} \sum_{s=1}^{n} (x_t - \bar{x})(x_s - \bar{x}) e^{-2\pi i\omega_j (t-s)} \\ &= n^{-1} \sum_{h=-(n-1)}^{n-1} \sum_{t=1}^{n-|h|} (x_{t+|h|} - \bar{x})(x_t - \bar{x}) e^{-2\pi i\omega_j h} \\ &= \sum_{h=-(n-1)}^{n-1} \widehat{\gamma}(h) e^{-2\pi i\omega_j h} \end{aligned} \tag{4.22}$$

where we have put $h = t - s$, with $\widehat{\gamma}(h)$ as given in (1.34).[7]

Recall, $P(\omega_j) = (4/n)I(\omega_j)$ where $P(\omega_j)$ is the scaled periodogram defined in (4.6). Henceforth we will work with $I(\omega_j)$ instead of $P(\omega_j)$. In view of (4.22), the periodogram, $I(\omega_j)$, is the sample version of $f(\omega_j)$ given in (4.12). That is, we may think of the periodogram as the "sample spectral density" of x_t.

It is sometimes useful to work with the real and imaginary parts of the DFT individually. To this end, we define the following transforms.

Definition 4.3 *Given data $x_1, \dots, x_n$, we define the* **cosine transform**

$$d_c(\omega_j) = n^{-1/2} \sum_{t=1}^{n} x_t \cos(2\pi\omega_j t) \tag{4.23}$$

and the **sine transform**

$$d_s(\omega_j) = n^{-1/2} \sum_{t=1}^{n} x_t \sin(2\pi\omega_j t) \tag{4.24}$$

where $\omega_j = j/n$ for $j = 0, 1, \dots, n-1$.

We note that $d(\omega_j) = d_c(\omega_j) - i\, d_s(\omega_j)$ and hence

$$I(\omega_j) = d_c^2(\omega_j) + d_s^2(\omega_j). \tag{4.25}$$

We have also discussed the fact that spectral analysis can be thought of as an analysis of variance. The next example examines this notion.

[6] $\sum_{t=1}^{n} z^t = z\frac{1-z^n}{1-z}$ for $z \neq 1$.

[7] Note that (4.22) can be used to obtain $\widehat{\gamma}(h)$ by taking the inverse DFT of $I(\omega_j)$. This approach was used in Example 1.27 to obtain a two-dimensional ACF.

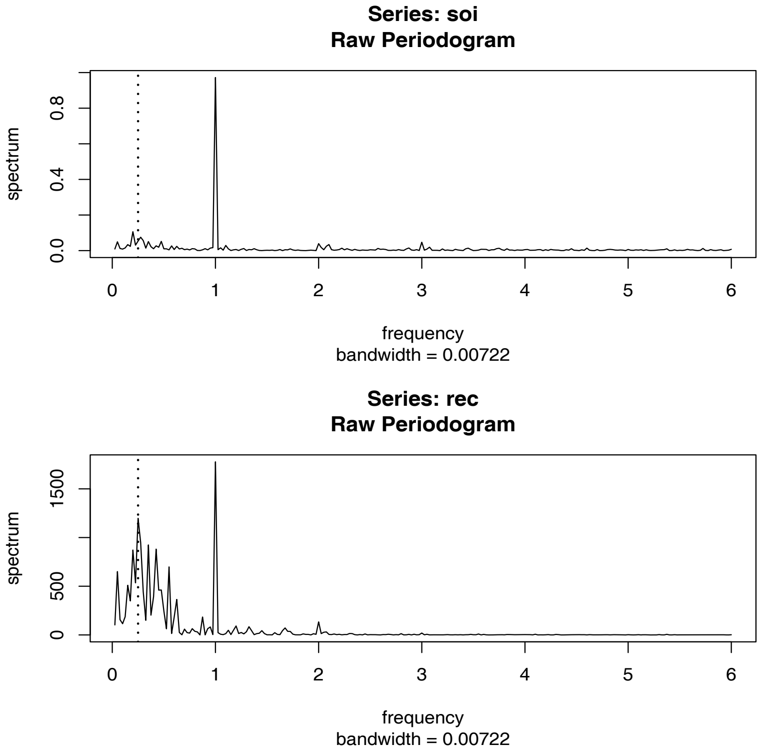

Fig. 4.4. Periodogram of SOI and Recruitment, $n = 453$ ($n' = 480$), where the frequency axis is labeled in multiples of $\Delta = 1/12$. Note the common peaks at $\omega = 1\Delta = 1/12$, or one cycle per year (12 months), and $\omega = \frac{1}{4}\Delta = 1/48$, or one cycle every four years (48 months).

Example 4.10 Periodogram of SOI and Recruitment Series

Figure 4.4 shows the periodograms of each series, where the frequency axis is labeled in multiples of $\Delta = 1/12$. As previously indicated, the centered data have been padded to a series of length 480. We notice a narrow-band peak at the obvious yearly (12 month) cycle, $\omega = 1\Delta = 1/12$. In addition, there is considerable power in a wide band at the lower frequencies that is centered around the four-year (48 month) cycle $\omega = \frac{1}{4}\Delta = 1/48$ representing a possible El Niño effect. This wide band activity suggests that the possible El Niño cycle is irregular, but tends to be around four years on average. We will continue to address this problem as we move to more sophisticated analyses.

Noting $\chi_2^2(.025) = .05$ and $\chi_2^2(.975) = 7.38$, we can obtain approximate 95% confidence intervals for the frequencies of interest. For example, the periodogram of the SOI series is $I_S(1/12) = .97$ at the yearly cycle. An approximate 95% confidence interval for the spectrum $f_S(1/12)$ is then

$$[2(.97)/7.38,\ 2(.97)/.05] = [.26,\ 38.4],$$

which is too wide to be of much use. We do notice, however, that the lower value of .26 is higher than any other periodogram ordinate, so it is safe to say that this value is significant. On the other hand, an approximate 95% confidence interval for the spectrum at the four-year cycle, $f_S(1/48)$, is

$$[2(.05)/7.38,\ 2(.05)/.05] = [.01,\ 2.12],$$

which again is extremely wide, and with which we are unable to establish significance of the peak.

We now give the R commands that can be used to reproduce Figure 4.4. To calculate and graph the periodogram, we used the `spec.pgram` command in R. We note that the value of Δ is the reciprocal of the value of `frequency` used in `ts()` when making the data a time series object. If the data are not time series objects, `frequency` is set to 1. Also, we set `log="no"` because R will plot the periodogram on a $\log_{10}$ scale by default. Figure 4.4 displays a `bandwidth` and by default, R tapers the data (which we override in the commands below). We will discuss bandwidth and tapering in the next section, so ignore these concepts for the time being.

```
par(mfrow=c(2,1))
soi.per = spec.pgram(soi, taper=0, log="no")
abline(v=1/4, lty="dotted")
rec.per = spec.pgram(rec, taper=0, log="no")
abline(v=1/4, lty="dotted")
```

The confidence intervals for the SOI series at the yearly cycle, $\omega = 1/12 = 40/480$, and the possible El Niño cycle of four years $\omega = 1/48 = 10/480$ can be computed in R as follows:

```
soi.per$spec[40]   # 0.97223; soi pgram at freq 1/12 = 40/480
soi.per$spec[10]   # 0.05372; soi pgram at freq 1/48 = 10/480
#  conf intervals  -  returned value:
U = qchisq(.025,2)     #  0.05063
L = qchisq(.975,2)     #  7.37775
2*soi.per$spec[10]/L   #  0.01456
2*soi.per$spec[10]/U   #  2.12220
2*soi.per$spec[40]/L   #  0.26355
2*soi.per$spec[40]/U   # 38.40108
```

The example above makes it clear that the periodogram as an estimator is susceptible to large uncertainties, and we need to find a way to reduce the variance. Not surprisingly, this result follows if we think about the periodogram, $I(\omega_j)$ as an estimator of the spectral density $f(\omega)$ and realize that it is the sum of squares of only two random variables for any sample size. The solution to this dilemma is suggested by the analogy with classical statistics where we look for independent random variables with the same variance and average the squares of these common variance observations. Independence and equality of variance do not hold in the time series case, but the covariance structure of the two adjacent estimators given in Example 4.9 suggests that for neighboring frequencies, these assumptions are approximately true.

4.5 Nonparametric Spectral Estimation

To continue the discussion that ended the previous section, we introduce a frequency band, $\mathcal{B}$, of $L << n$ contiguous fundamental frequencies, centered around frequency $\omega_j = j/n$, which is chosen close to a frequency of interest, ω. For frequencies of the form $\omega^* = \omega_j + k/n$, let

$$\mathcal{B} = \left\{\omega^* \colon \ \omega_j - \frac{m}{n} \le \omega^* \le \omega_j + \frac{m}{n}\right\}, \tag{4.44}$$

where

$$L = 2m + 1 \tag{4.45}$$

is an odd number, chosen such that the spectral values in the interval $\mathcal{B}$,

$$f(\omega_j + k/n), \quad k = -m, \ldots, 0, \ldots, m$$

are approximately equal to $f(\omega)$. This structure can be realized for large sample sizes, as shown formally in §C.2.

We now define an averaged (or smoothed) periodogram as the average of the periodogram values, say,

$$\bar{f}(\omega) = \frac{1}{L}\sum_{k=-m}^{m} I(\omega_j + k/n), \tag{4.46}$$

over the band $\mathcal{B}$. Under the assumption that the spectral density is fairly constant in the band $\mathcal{B}$, and in view of (4.41) we can show that under appropriate conditions,[11] for large n, the periodograms in (4.46) are approximately distributed as independent $f(\omega)\chi_2^2/2$ random variables, for $0 < \omega < 1/2$, as long as we keep L fairly small relative to n. This result is discussed formally in §C.2. Thus, under these conditions, $L\bar{f}(\omega)$ is the sum of L approximately independent $f(\omega)\chi_2^2/2$ random variables. It follows that, for large n,

$$\frac{2L\bar{f}(\omega)}{f(\omega)} \ \dot{\sim}\ \chi^2_{2L} \tag{4.47}$$

where $\dot{\sim}$ means *is approximately distributed as*.

In this scenario, where we smooth the periodogram by simple averaging, it seems reasonable to call the width of the frequency interval defined by (4.44),

[11] The conditions, which are sufficient, are that x_t is a linear process, as described in Property 4.4, with $\sum_j \sqrt{|j|}\,|\psi_j| < \infty$, and w_t has a finite fourth moment.

$$B_w = \frac{L}{n}, \tag{4.48}$$

the bandwidth.[12] The concept of the bandwidth, however, becomes more complicated with the introduction of spectral estimators that smooth with unequal weights. Note (4.48) implies the degrees of freedom can be expressed as

$$2L = 2B_w n, \tag{4.49}$$

or twice the time-bandwidth product. The result (4.47) can be rearranged to obtain an approximate $100(1-\alpha)\%$ confidence interval of the form

$$\frac{2L\bar{f}(\omega)}{\chi^2_{2L}(1-\alpha/2)} \le f(\omega) \le \frac{2L\bar{f}(\omega)}{\chi^2_{2L}(\alpha/2)} \tag{4.50}$$

for the true spectrum, $f(\omega)$.

Many times, the visual impact of a spectral density plot will be improved by plotting the logarithm of the spectrum instead of the spectrum (the log transformation is the variance stabilizing transformation in this situation). This phenomenon can occur when regions of the spectrum exist with peaks of interest much smaller than some of the main power components. For the log spectrum, we obtain an interval of the form

$$\begin{aligned}\big[\log\bar{f}(\omega) + \log 2L - \log\chi^2_{2L}(1-\alpha/2),&\\ \log\bar{f}(\omega) + \log 2L - \log\chi^2_{2L}(\alpha/2)\big].&\end{aligned} \tag{4.51}$$

We can also test hypotheses relating to the equality of spectra using the fact that the distributional result (4.47) implies that the ratio of spectra based on roughly independent samples will have an approximate $F_{2L,2L}$ distribution. The independent estimators can either be from different frequency bands or from different series.

If zeros are appended before computing the spectral estimators, we need to adjust the degrees of freedom and an approximation is to replace $2L$ by $2Ln/n'$. Hence, we define the adjusted degrees of freedom as

$$df = \frac{2Ln}{n'} \tag{4.52}$$

[12] The bandwidth value used in R is based on Grenander (1951). The basic idea is that bandwidth can be related to the standard deviation of the weighting distribution. For the uniform distribution on the frequency range $-m/n$ to m/n, the standard deviation is $L/n\sqrt{12}$ (using a continuity correction). Consequently, in the case of (4.46), R will report a bandwidth of $L/n\sqrt{12}$, which amounts to dividing our definition by $\sqrt{12}$. Note that in the extreme case $L = n$, we would have $B_w = 1$ indicating that everything was used in the estimation; in this case, R would report a bandwidth of $1/\sqrt{12}$. There are many definitions of bandwidth and an excellent discussion may be found in Percival and Walden (1993, §6.7).

and use it instead of $2L$ in the confidence intervals (4.50) and (4.51). For example, (4.50) becomes

$$\frac{df\,\bar{f}(\omega)}{\chi^2_{df}(1-\alpha/2)} \le f(\omega) \le \frac{df\,\bar{f}(\omega)}{\chi^2_{df}(\alpha/2)}. \tag{4.53}$$

A number of assumptions are made in computing the approximate confidence intervals given above, which may not hold in practice. In such cases, it may be reasonable to employ resampling techniques such as one of the parametric bootstraps proposed by Hurvich and Zeger (1987) or a nonparametric local bootstrap proposed by Paparoditis and Politis (1999). To develop the bootstrap distributions, we assume that the contiguous DFTs in a frequency band of the form (4.44) all came from a time series with identical spectrum $f(\omega)$. This, in fact, is exactly the same assumption made in deriving the large-sample theory. We may then simply resample the L DFTs in the band, with replacement, calculating a spectral estimate from each bootstrap sample. The sampling distribution of the bootstrap estimators approximates the distribution of the nonparametric spectral estimator. For further details, including the theoretical properties of such estimators, see Paparoditis and Politis (1999).

Before proceeding further, we pause to consider computing the average periodograms for the SOI and Recruitment series, as shown in Figure 4.5.

Example 4.11 Averaged Periodogram for SOI and Recruitment

Generally, it is a good idea to try several bandwidths that seem to be compatible with the general overall shape of the spectrum, as suggested by the periodogram. The SOI and Recruitment series periodograms, previously computed in Figure 4.4, suggest the power in the lower El Niño frequency needs smoothing to identify the predominant overall period. Trying values of L leads to the choice $L = 9$ as a reasonable value, and the result is displayed in Figure 4.5. In our notation, the bandwidth in this case is $B_w = 9/480 = .01875$ cycles per month for the spectral estimator. This bandwidth means we are assuming a relatively constant spectrum over about $.01875/.5 = 3.75\%$ of the entire frequency interval (0, 1/2). To obtain the bandwidth, $B_w = .01875$, from the one reported by R in Figure 4.5, we can multiply $.065\Delta$ (the frequency scale is in increments of Δ) by $\sqrt{12}$ as discussed in footnote 12 on page 197.

The smoothed spectra shown in Figure 4.5 provide a sensible compromise between the noisy version, shown in Figure 4.4, and a more heavily smoothed spectrum, which might lose some of the peaks. An undesirable effect of averaging can be noticed at the yearly cycle, $\omega = 1\Delta$, where the narrow band peaks that appeared in the periodograms in Figure 4.4 have been flattened and spread out to nearby frequencies. We also notice, and have marked, the appearance of harmonics of the yearly cycle, that is, frequencies of the form $\omega = k\Delta$ for $k = 1, 2, \ldots$. Harmonics typically occur when a periodic component is present, but not in a sinusoidal fashion; see Example 4.12.

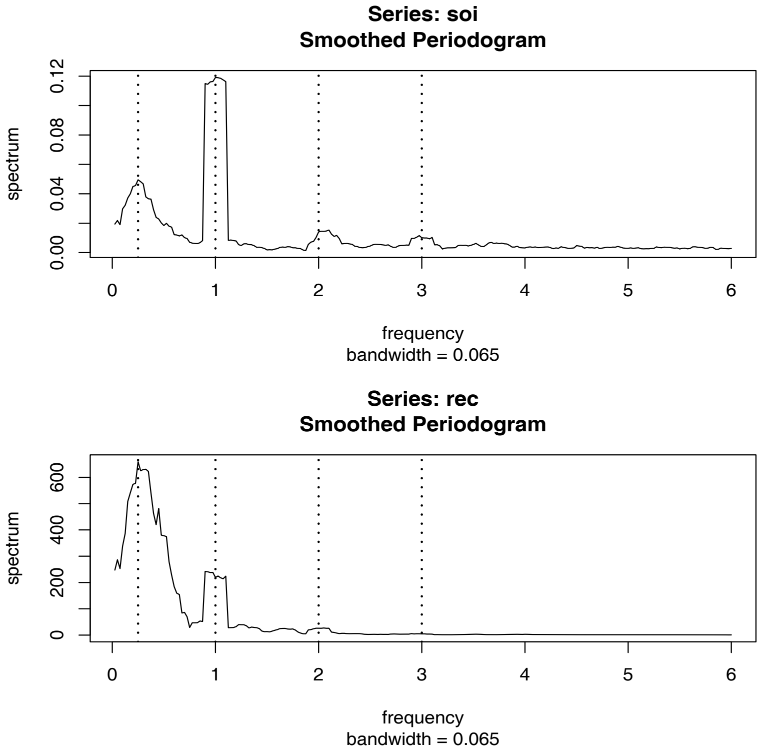

Fig. 4.5. The averaged periodogram of the SOI and Recruitment series $n = 453$, $n' = 480$, $L = 9$, $df = 17$, showing common peaks at the four year period, $\omega = \frac{1}{4}\Delta = 1/48$ cycles/month, the yearly period, $\omega = 1\Delta = 1/12$ cycles/month and some of its harmonics $\omega = k\Delta$ for $k = 2, 3$.

Figure 4.5 can be reproduced in R using the following commands. The basic call is to the function `spec.pgram`. To compute averaged periodograms, use the Daniell kernel, and specify m, where $L = 2m + 1$ ($L = 9$ and $m = 4$ in this example). We will explain the kernel concept later in this section, specifically just prior to Example 4.13.

```
1 par(mfrow=c(2,1))
2 k = kernel("daniell", 4)
3 soi.ave = spec.pgram(soi, k, taper=0, log="no")
4 abline(v=c(.25,1,2,3), lty=2)
5 # Repeat above lines using rec in place of soi on line 3
6 soi.ave$bandwidth       # 0.0649519 = reported bandwidth
7 soi.ave$bandwidth*(1/12)*sqrt(12)        # 0.01875 = Bw
```

The adjusted degrees of freedom are $df = 2(9)(453)/480 \approx 17$. We can use this value for the 95% confidence intervals, with $\chi^2_{df}(.025) = 7.56$ and $\chi^2_{df}(.975) = 30.17$. Substituting into (4.53) gives the intervals in Table 4.1 for the two frequency bands identified as having the maximum power. To

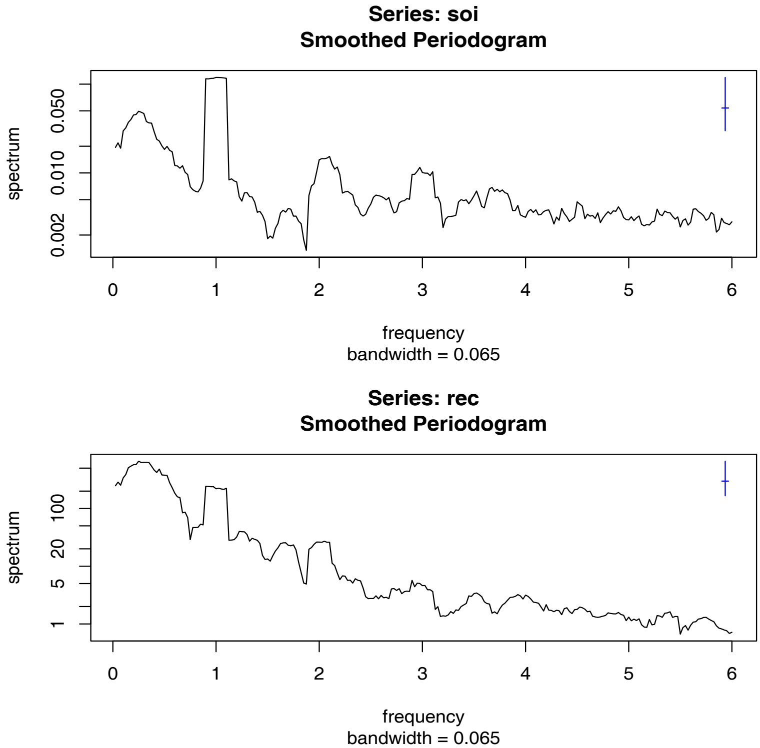

Fig. 4.6. Figure 4.5 with the average periodogram ordinates plotted on a $\log_{10}$ scale. The display in the upper right-hand corner represents a generic 95% confidence interval.

examine the two peak power possibilities, we may look at the 95% confidence intervals and see whether the lower limits are substantially larger than adjacent baseline spectral levels. For example, the El Niño frequency of 48 months has lower limits that exceed the values the spectrum would have if there were simply a smooth underlying spectral function without the peaks. The relative distribution of power over frequencies is different, with the SOI having less power at the lower frequency, relative to the seasonal periods, and the recruit series having relatively more power at the lower or El Niño frequency.

The entries in Table 4.1 for SOI can be obtained in R as follows:

```
df = soi.ave$df          #  df = 16.9875  (returned values)
U = qchisq(.025, df)     #  U = 7.555916
L = qchisq(.975, df)     #  L = 30.17425
soi.ave$spec[10]         #  0.0495202
soi.ave$spec[40]         #  0.1190800
# intervals
df*soi.ave$spec[10]/L     # 0.0278789
df*soi.ave$spec[10]/U     # 0.1113333
```

Table 4.1. Confidence Intervals for the Spectra of the SOI and Recruitment Series

Series	ω	Period	Power	Lower	Upper
SOI	1/48	4 years	.05	.03	.11
	1/12	1 year	.12	.07	.27
Recruits	1/48	4 years	6.59	3.71	14.82
$\times 10^2$	1/12	1 year	2.19	1.24	4.93

```
df*soi.ave$spec[40]/L    # 0.0670396
df*soi.ave$spec[40]/U    # 0.2677201
# repeat above commands with soi replaced by rec
```

Finally, Figure 4.6 shows the averaged periodograms in Figure 4.5 plotted on a $\log_{10}$ scale. This is the default plot in R, and these graphs can be obtained by removing the statement `log="no"` in the `spec.pgram` call. Notice that the default plot also shows a generic confidence interval of the form (4.51) (with log replaced by $\log_{10}$) in the upper right-hand corner. To use it, imagine placing the tick mark on the averaged periodogram ordinate of interest; the resulting bar then constitutes an approximate 95% confidence interval for the spectrum at that frequency. We note that displaying the estimates on a log scale tends to emphasize the harmonic components.

Example 4.12 Harmonics

In the previous example, we saw that the spectra of the annual signals displayed minor peaks at the harmonics; that is, the signal spectra had a large peak at $\omega = 1\Delta = 1/12$ cycles/month (the one-year cycle) and minor peaks at its harmonics $\omega = k\Delta$ for $k = 2, 3, \dots$ (two-, three-, and so on, cycles per year). This will often be the case because most signals are not perfect sinusoids (or perfectly cyclic). In this case, the harmonics are needed to capture the non-sinusoidal behavior of the signal. As an example, consider the signal formed in Figure 4.7 from a (fundamental) sinusoid oscillating at two cycles per unit time along with the second through sixth harmonics at decreasing amplitudes. In particular, the signal was formed as

$$
\begin{aligned}
x_t = {} & \sin(2\pi 2t) + .5\sin(2\pi 4t) + .4\sin(2\pi 6t) \\
& + .3\sin(2\pi 8t) + .2\sin(2\pi 10t) + .1\sin(2\pi 12t)
\end{aligned} \tag{4.54}
$$

for $0 \le t \le 1$. Notice that the signal is non-sinusoidal in appearance and rises quickly then falls slowly.

A figure similar to Figure 4.7 can be generated in R as follows.

```
t = seq(0, 1, by=1/200)
amps = c(1, .5, .4, .3, .2, .1)
x = matrix(0, 201, 6)
for (j in 1:6) x[,j] = amps[j]*sin(2*pi*t*2*j)
x = ts(cbind(x, rowSums(x)), start=0, deltat=1/200)
```

small. Finally, we note that the coherence is persistent at the seasonal harmonic frequencies.

This example may be reproduced using the following R commands.

```
sr=spec.pgram(cbind(soi,rec),kernel("daniell",9),taper=0,plot=FALSE)
sr$df                    # df = 35.8625
f = qf(.999, 2, sr$df-2)  # = 8.529792
C = f/(18+f)              # = 0.318878
plot(sr, plot.type = "coh", ci.lty = 2)
abline(h = C)
```

4.8 Linear Filters

Some of the examples of the previous sections have hinted at the possibility the distribution of power or variance in a time series can be modified by making a linear transformation. In this section, we explore that notion further by defining a linear filter and showing how it can be used to extract signals from a time series. The linear filter modifies the spectral characteristics of a time series in a predictable way, and the systematic development of methods for taking advantage of the special properties of linear filters is an important topic in time series analysis.

A linear filter uses a set of specified coefficients a_j, for $j = 0, \pm 1, \pm 2, \ldots$, to transform an input series, x_t, producing an output series, y_t, of the form

$$y_t = \sum_{j=-\infty}^{\infty} a_j x_{t-j}, \qquad \sum_{j=-\infty}^{\infty} |a_j| < \infty. \tag{4.99}$$

The form (4.99) is also called a convolution in some statistical contexts. The coefficients, collectively called the *impulse response function*, are required to satisfy absolute summability so y_t in (4.99) exists as a limit in mean square and the infinite Fourier transform

$$A_{yx}(\omega) = \sum_{j=-\infty}^{\infty} a_j \, e^{-2\pi i \omega j}\,, \tag{4.100}$$

called the *frequency response function*, is well defined. We have already encountered several linear filters, for example, the simple three-point moving average in Example 4.16, which can be put into the form of (4.99) by letting $a_{-1} = a_0 = a_1 = 1/3$ and taking $a_t = 0$ for $|j| \geq 2$.

The importance of the linear filter stems from its ability to enhance certain parts of the spectrum of the input series. To see this, assuming that x_t is stationary with spectral density $f_{xx}(\omega)$, the autocovariance function of the filtered output y_t in (4.99) can be derived as

$$
\begin{aligned}
\gamma_{yy}(h) &= \operatorname{cov}(y_{t+h}, y_t) \\
&= \operatorname{cov}\left(\sum_r a_r x_{t+h-r}, \sum_s a_s x_{t-s}\right) \\
&= \sum_r \sum_s a_r \gamma_{xx}(h-r+s) a_s \\
&= \sum_r \sum_s a_r \left[\int_{-1/2}^{1/2} \mathrm{e}^{2\pi i\omega(h-r+s)} f_{xx}(\omega) d\omega\right] a_s \\
&= \int_{-1/2}^{1/2} \left(\sum_r a_r \mathrm{e}^{-2\pi i\omega r}\right)\left(\sum_s a_s \mathrm{e}^{2\pi i\omega s}\right) \mathrm{e}^{2\pi i\omega h} f_{xx}(\omega)\ d\omega \\
&= \int_{-1/2}^{1/2} \mathrm{e}^{2\pi i\omega h} |A_{yx}(\omega)|^2 f_{xx}(\omega)\ d\omega,
\end{aligned}
$$

where we have first replaced $\gamma_{xx}(\cdot)$ by its representation (4.11) and then substituted $A_{yx}(\omega)$ from (4.100). The computation is one we do repeatedly, exploiting the uniqueness of the Fourier transform. Now, because the left-hand side is the Fourier transform of the spectral density of the output, say, $f_{yy}(\omega)$, we get the important filtering property as follows.

Property 4.7 Output Spectrum of a Filtered Stationary Series

The spectrum of the filtered output y_t in (4.99) is related to the spectrum of the input x_t by

$$f_{yy}(\omega) = |A_{yx}(\omega)|^2\ f_{xx}(\omega), \tag{4.101}$$

where the frequency response function $A_{yx}(\omega)$ is defined in (4.100).

The result (4.101) enables us to calculate the exact effect on the spectrum of any given filtering operation. This important property shows the spectrum of the input series is changed by filtering and the effect of the change can be characterized as a frequency-by-frequency multiplication by the squared magnitude of the frequency response function. Again, an obvious analogy to a property of the variance in classical statistics holds, namely, if x is a random variable with variance σ_x^2, then $y = ax$ will have variance $\sigma_y^2 = a^2\sigma_x^2$, so the variance of the linearly transformed random variable is changed by multiplication by a^2 in much the same way as the linearly filtered spectrum is changed in (4.101).

Finally, we mention that Property 4.3, which was used to get the spectrum of an ARMA process, is just a special case of Property 4.7 where in (4.99), $x_t = w_t$ is white noise, in which case $f_{xx}(\omega) = \sigma_w^2$, and $a_j = \psi_j$, in which case

$$A_{yx}(\omega) = \psi(\mathrm{e}^{-2\pi i\omega}) = \theta(\mathrm{e}^{-2\pi i\omega})\ /\ \phi(\mathrm{e}^{-2\pi i\omega}).$$

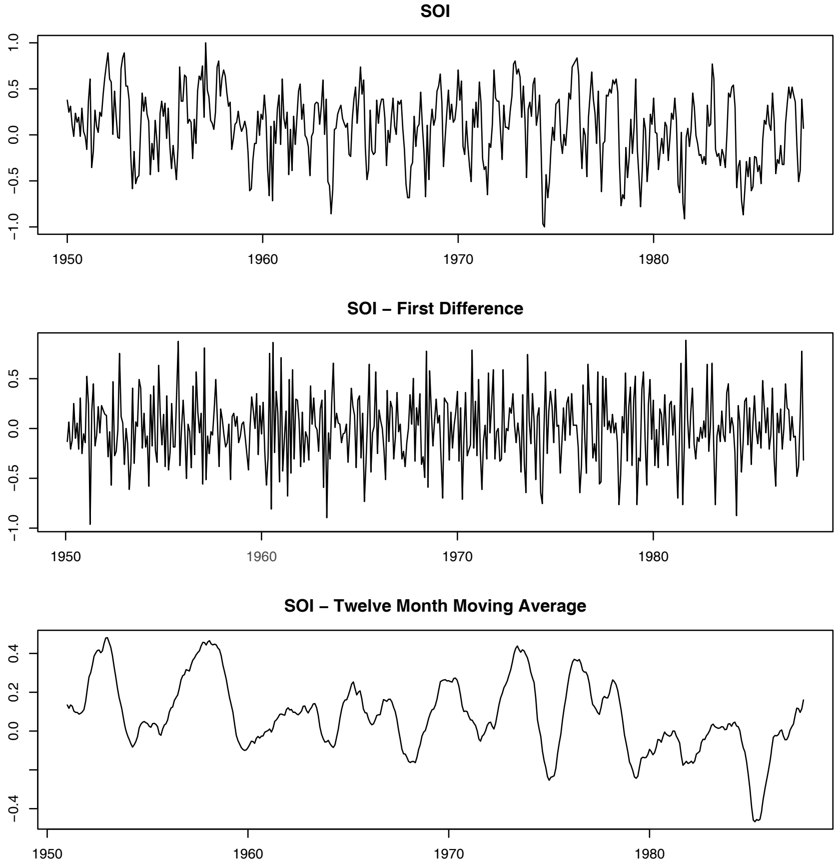

Fig. 4.14. SOI series (top) compared with the differenced SOI (middle) and a centered 12-month moving average (bottom).

Example 4.19 First Difference and Moving Average Filters

We illustrate the effect of filtering with two common examples, the first difference filter

$$y_t = \nabla x_t = x_t - x_{t-1}$$

and the symmetric moving average filter

$$y_t = \tfrac{1}{24}\big(x_{t-6} + x_{t+6}\big) + \tfrac{1}{12}\sum_{r=-5}^{5} x_{t-r},$$

which is a modified Daniell kernel with $m = 6$. The results of filtering the SOI series using the two filters are shown in the middle and bottom panels of Figure 4.14. Notice that the effect of differencing is to roughen the series because it tends to retain the higher or faster frequencies. The centered

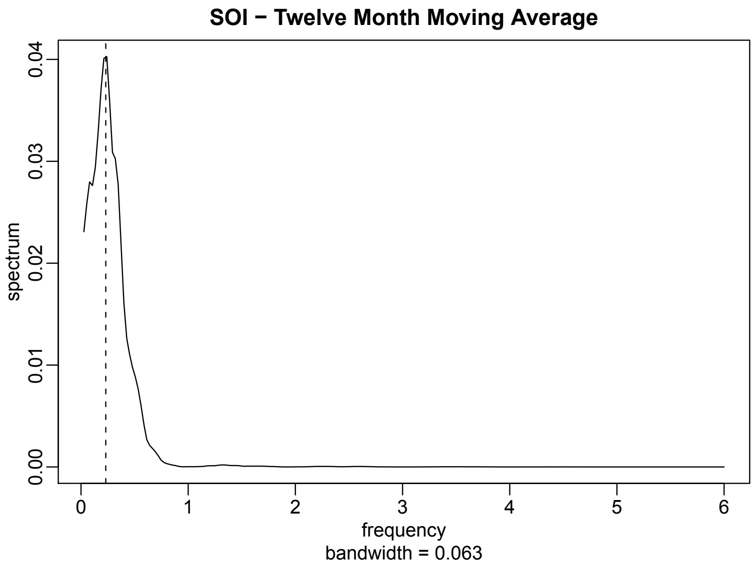

Fig. 4.15. Spectral analysis of SOI after applying a 12-month moving average filter. The vertical line corresponds to the 52-month cycle.

moving average smoothes the series because it retains the lower frequencies and tends to attenuate the higher frequencies. In general, differencing is an example of a *high-pass filter* because it retains or passes the higher frequencies, whereas the moving average is a *low-pass filter* because it passes the lower or slower frequencies.

Notice that the slower periods are enhanced in the symmetric moving average and the seasonal or yearly frequencies are attenuated. The filtered series makes about 9 cycles in the length of the data (about one cycle every 52 months) and the moving average filter tends to enhance or extract the signal that is associated with El Niño. Moreover, by the low-pass filtering of the data, we get a better sense of the El Niño effect and its irregularity. Figure 4.15 shows the results of a spectral analysis on the low-pass filtered SOI series. It is clear that all high frequency behavior has been removed and the El Niño cycle is accentuated; the dotted vertical line in the figure corresponds to the 52 months cycle.

Now, having done the filtering, it is essential to determine the exact way in which the filters change the input spectrum. We shall use (4.100) and (4.101) for this purpose. The first difference filter can be written in the form (4.99) by letting $a_0 = 1, a_1 = -1$, and $a_r = 0$ otherwise. This implies that

$$A_{yx}(\omega) = 1 - \mathrm{e}^{-2\pi i\omega},$$

and the squared frequency response becomes

$$|A_{yx}(\omega)|^2 = (1 - \mathrm{e}^{-2\pi i\omega})(1 - \mathrm{e}^{2\pi i\omega}) = 2[1 - \cos(2\pi\omega)]. \tag{4.102}$$

The top panel of Figure 4.16 shows that the first difference filter will attenuate the lower frequencies and enhance the higher frequencies because the multiplier of the spectrum, $|A_{yx}(\omega)|^2$, is large for the higher frequencies and small for the lower frequencies. Generally, the slow rise of this kind of filter does not particularly recommend it as a procedure for retaining only the high frequencies.

For the centered 12-month moving average, we can take $a_{-6} = a_6 = 1/24$, $a_k = 1/12$ for $-5 \le k \le 5$ and $a_k = 0$ elsewhere. Substituting and recognizing the cosine terms gives

$$A_{yx}(\omega) = \tfrac{1}{12}\Big[1 + \cos(12\pi\omega) + 2\sum_{k=1}^{5}\cos(2\pi\omega k)\Big]. \tag{4.103}$$

Plotting the squared frequency response of this function as in Figure 4.16 shows that we can expect this filter to cut most of the frequency content above .05 cycles per point. This corresponds to eliminating periods shorter than $T = 1/.05 = 20$ points. In particular, this drives down the yearly components with periods of $T = 12$ months and enhances the El Niño frequency, which is somewhat lower. The filter is not completely efficient at attenuating high frequencies; some power contributions are left at higher frequencies, as shown in the function $|A_{yx}(\omega)|^2$ and in the spectrum of the moving average shown in Figure 4.3.

The following R session shows how to filter the data, perform the spectral analysis of this example, and plot the squared frequency response curve of the difference filter.

```
par(mfrow=c(3,1))
plot(soi)  # plot data
plot(diff(soi))  # plot first difference
k = kernel("modified.daniell", 6)  # filter weights
plot(soif <- kernapply(soi, k))    # plot 12 month filter
dev.new()
spectrum(soif, spans=9, log="no")  # spectral analysis
abline(v=12/52, lty="dashed")
dev.new()
w = seq(0, .5, length=500)  # frequency response
FR = abs(1-exp(2i*pi*w))^2
plot(w, FR, type="l")
```

The two filters discussed in the previous example were different in that the frequency response function of the first difference was complex-valued, whereas the frequency response of the moving average was purely real. A short derivation similar to that used to verify (4.101) shows, when x_t and y_t are related by the linear filter relation (4.99), the cross-spectrum satisfies

$$f_{yx}(\omega) = A_{yx}(\omega) f_{xx}(\omega),$$

so the frequency response is of the form

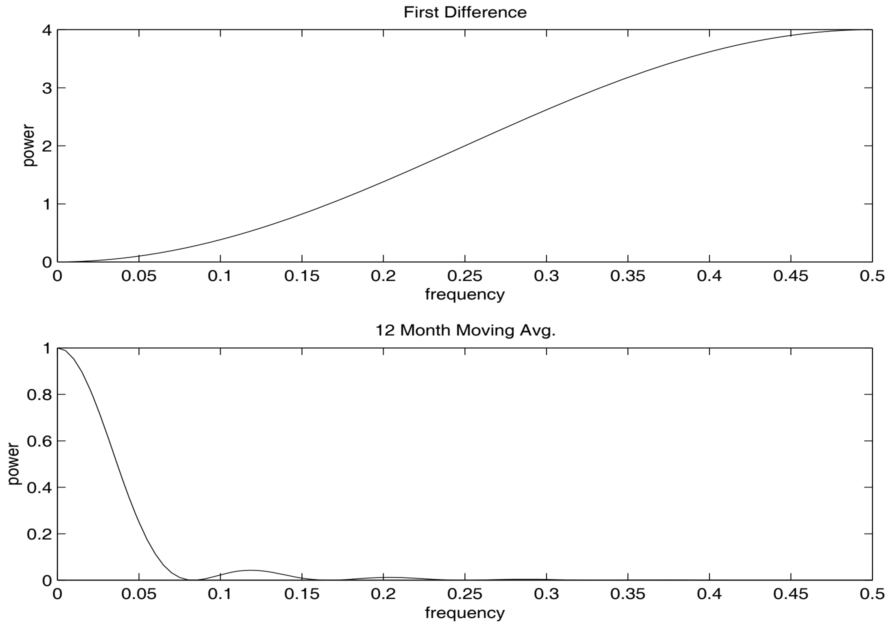

Fig. 4.16. Squared frequency response functions of the first difference and 12-month moving average filters.

$$A_{yx}(\omega) = \frac{f_{yx}(\omega)}{f_{xx}(\omega)} \tag{4.104}$$

$$= \frac{c_{yx}(\omega)}{f_{xx}(\omega)} - i\frac{q_{yx}(\omega)}{f_{xx}(\omega)}, \tag{4.105}$$

where we have used (4.81) to get the last form. Then, we may write (4.105) in polar coordinates as

$$A_{yx}(\omega) = |A_{yx}(\omega)| \ \exp\{-i \ \phi_{yx}(\omega)\}, \tag{4.106}$$

where the amplitude and phase of the filter are defined by

$$|A_{yx}(\omega)| = \frac{\sqrt{c_{yx}^2(\omega) + q_{yx}^2(\omega)}}{f_{xx}(\omega)} \tag{4.107}$$

and

$$\phi_{yx}(\omega) = \tan^{-1}\left(-\frac{q_{yx}(\omega)}{c_{yx}(\omega)}\right). \tag{4.108}$$

A simple interpretation of the phase of a linear filter is that it exhibits time delays as a function of frequency in the same way as the spectrum represents the variance as a function of frequency. Additional insight can be gained by considering the simple delaying filter

$$y_t = Ax_{t-D},$$

where the series gets replaced by a version, amplified by multiplying by A and delayed by D points. For this case,

$$f_{yx}(\omega) = A\mathrm{e}^{-2\pi i\omega D} f_{xx}(\omega),$$

and the amplitude is $|A|$, and the phase is

$$\phi_{yx}(\omega) = -2\pi\omega D,$$

or just a linear function of frequency ω. For this case, applying a simple time delay causes phase delays that depend on the frequency of the periodic component being delayed. Interpretation is further enhanced by setting

$$x_t = \cos(2\pi\omega t),$$

in which case

$$y_t = A\cos(2\pi\omega t - 2\pi\omega D).$$

Thus, the output series, y_t, has the same period as the input series, x_t, but the amplitude of the output has increased by a factor of $|A|$ and the phase has been changed by a factor of $-2\pi\omega D$.

Example 4.20 Difference and Moving Average Filters

We consider calculating the amplitude and phase of the two filters discussed in Example 4.19. The case for the moving average is easy because $A_{yx}(\omega)$ given in (4.103) is purely real. So, the amplitude is just $|A_{yx}(\omega)|$ and the phase is $\phi_{yx}(\omega) = 0$. In general, symmetric $(a_j = a_{-j})$ filters have zero phase. The first difference, however, changes this, as we might expect from the example above involving the time delay filter. In this case, the squared amplitude is given in (4.102). To compute the phase, we write

$$\begin{aligned}
A_{yx}(\omega) &= 1 - e^{-2\pi i\omega} = e^{-i\pi\omega}(e^{i\pi\omega} - e^{-i\pi\omega}) \\
&= 2ie^{-i\pi\omega}\sin(\pi\omega) = 2\sin^2(\pi\omega) + 2i\cos(\pi\omega)\sin(\pi\omega) \\
&= \frac{c_{yx}(\omega)}{f_{xx}(\omega)} - i\frac{q_{yx}(\omega)}{f_{xx}(\omega)},
\end{aligned}$$

so

$$\phi_{yx}(\omega) = \tan^{-1}\left(-\frac{q_{yx}(\omega)}{c_{yx}(\omega)}\right) = \tan^{-1}\left(\frac{\cos(\pi\omega)}{\sin(\pi\omega)}\right).$$

Noting that

$$\cos(\pi\omega) = \sin(-\pi\omega + \pi/2)$$

and that

$$\sin(\pi\omega) = \cos(-\pi\omega + \pi/2),$$

we get

$$\phi_{yx}(\omega) = -\pi\omega + \pi/2,$$

and the phase is again a linear function of frequency.