4.10 Lagged Regression Models

One of the intriguing possibilities offered by the coherence analysis of the relation between the SOI and Recruitment series discussed in Example 4.18 would be extending classical regression to the analysis of lagged regression models of the form

$$y_t = \sum_{r=-\infty}^{\infty} \beta_r x_{t-r} + v_t, \tag{4.122}$$

where v_t is a stationary noise process, x_t is the observed input series, and y_t is the observed output series. We are interested in estimating the filter coefficients β_r relating the adjacent lagged values of x_t to the output series y_t.

We assume that the inputs and outputs have zero means and are jointly stationary with the 2×1 vector process $(x_t, y_t)'$ having a spectral matrix of the form

$$f(\omega) = \begin{pmatrix} f_{xx}(\omega) & f_{xy}(\omega) \\ f_{yx}(\omega) & f_{yy}(\omega) \end{pmatrix}. \tag{4.123}$$

Here, $f_{xy}(\omega)$ is the cross-spectrum relating the input x_t to the output y_t, and $f_{xx}(\omega)$ and $f_{yy}(\omega)$ are the spectra of the input and output series, respectively. Generally, we observe two series, regarded as input and output and search for regression functions $\{\beta_t\}$ relating the inputs to the outputs. We assume all autocovariance functions satisfy the absolute summability conditions of the form (4.30).

Then, minimizing the mean squared error

$$MSE = E\left(y_t - \sum_{r=-\infty}^{\infty} \beta_r x_{t-r}\right)^2 \tag{4.124}$$

leads to the usual orthogonality conditions

$$E\left[\left(y_t - \sum_{r=-\infty}^{\infty} \beta_r x_{t-r}\right) x_{t-s}\right] = 0 \tag{4.125}$$

for all $s = 0, \pm1, \pm2, \ldots$. Taking the expectations inside leads to the normal equations

$$\sum_{r=-\infty}^{\infty} \beta_r \, \gamma_{xx}(s-r) = \gamma_{yx}(s) \tag{4.126}$$

for $s = 0, \pm1, \pm2, \ldots$. These equations might be solved, with some effort, if the covariance functions were known exactly.

$$\int_{-1/2}^{1/2} \sum_{r=-\infty}^{\infty} \beta_r \, e^{2\pi i\omega(s-r)} \, f_{xx}(\omega) \, d\omega = \int_{-1/2}^{1/2} e^{2\pi i\omega s} B(\omega) f_{xx}(\omega) \, d\omega,$$

where

$$B(\omega) = \sum_{r=-\infty}^{\infty} \beta_r \, e^{-2\pi i\omega r} \tag{4.127}$$

is the Fourier transform of the regression coefficients β_t. Now, because $\gamma_{yx}(s)$ is the inverse transform of the cross-spectrum $f_{yx}(\omega)$, we might write the system of equations in the frequency domain, using the uniqueness of the Fourier transform, as

$$B(\omega) f_{xx}(\omega) = f_{yx}(\omega), \tag{4.128}$$

which then become the analogs of the usual normal equations. Then, we may take

$$\widehat{B}(\omega_k) = \frac{\widehat{f}_{yx}(\omega_k)}{\widehat{f}_{xx}(\omega_k)} \tag{4.129}$$

as the estimator for the Fourier transform of the regression coefficients, evaluated at some subset of fundamental frequencies $\omega_k = k/M$ with $M << n$. Generally, we assume smoothness of $B(\cdot)$ over intervals of the form $\{\omega_k + \ell/n;\ \ell = -(L-1)/2, \dots, (L-1)/2\}$. The inverse transform of the function $\widehat{B}(\omega)$ would give $\widehat{\beta}_t$, and we note that the discrete time approximation can be taken as

$$\widehat{\beta}_t = M^{-1} \sum_{k=0}^{M-1} \widehat{B}(\omega_k) e^{2\pi i\omega_k t} \tag{4.130}$$

for $t = 0, \pm1, \pm2, \dots, \pm(M/2-1)$. If we were to use (4.130) to define $\widehat{\beta}_t$ for $|t| \geq M/2$, we would end up with a sequence of coefficients that is periodic with a period of M. In practice we define $\widehat{\beta}_t = 0$ for $|t| \geq M/2$ instead. Problem 4.32 explores the error resulting from this approximation.

Example 4.24 Lagged Regression for SOI and Recruitment

The high coherence between the SOI and Recruitment series noted in Example 4.18 suggests a lagged regression relation between the two series. A natural direction for the implication in this situation is implied because we feel that the sea surface temperature or SOI should be the input and the Recruitment series should be the output. With this in mind, let x_t be the SOI series and y_t the Recruitment series.

Although we think naturally of the SOI as the input and the Recruitment as the output, two input-output configurations are of interest. With SOI as the input, the model is

$$y_t = \sum_{r=-\infty}^{\infty} a_r x_{t-r} + w_t$$

whereas a model that reverses the two roles would be

$$x_t = \sum_{r=-\infty}^{\infty} b_r y_{t-r} + v_t,$$

where w_t and v_t are white noise processes. Even though there is no plausible environmental explanation for the second of these two models, displaying both possibilities helps to settle on a parsimonious transfer function model.

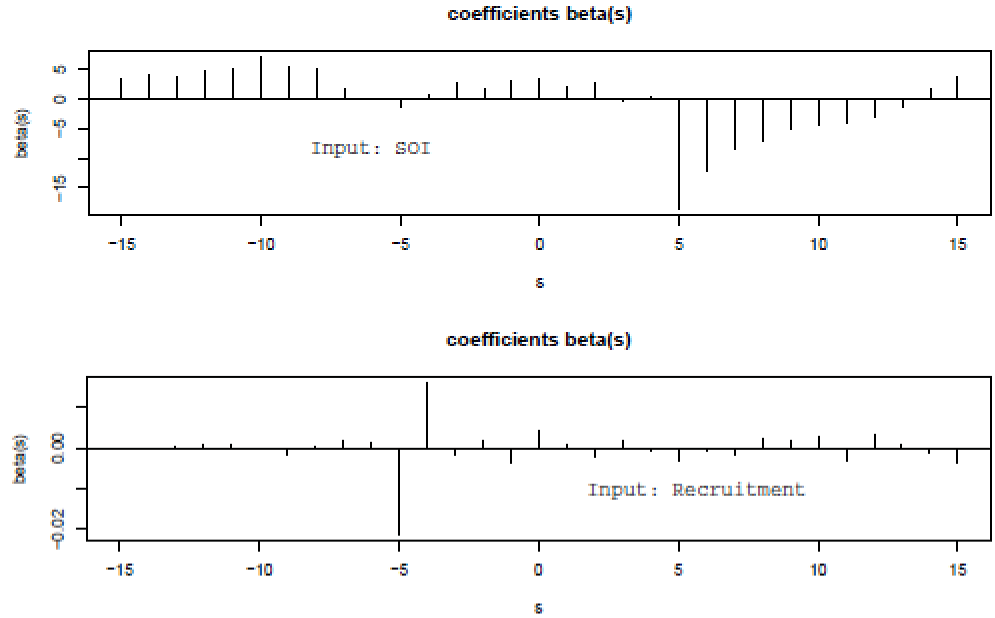

Fig. 4.27. Estimated impulse response functions relating SOI to Recruitment (top) and Recruitment to SOI (bottom) $L = 15, M = 32$.

Based on the script LagReg (see Appendix R, §R.1), the estimated regression or impulse response function for SOI, with $M = 32$ and $L = 15$ is

```
LagReg(soi, rec, L=15, M=32, threshold=6)
         lag s    beta(s)
    [1,]     5 -18.479306
    [2,]     6 -12.263296
    [3,]     7  -8.539368
    [4,]     8  -6.984553
The prediction equation is
rec(t) = alpha + sum_s[ beta(s)*soi(t-s) ], where alpha = 65.97
MSE = 414.08
```

approximately exponential and a possible model is

$$y_t = 66 - 18.5x_{t-5} - 12.3x_{t-6} - 8.5x_{t-7} - 7x_{t-8} + w_t.$$

If we examine the inverse relation, namely, a regression model with the Recruitment series y_t as the input, the bottom of Figure 4.27 implies a much simpler model,

```
LagReg(rec, soi, L=15, M=32, inverse=TRUE, threshold=.01)
```

```
          lag s      beta(s)
     [1,]     4  0.01593167
     [2,]     5 -0.02120013
The prediction equation is
soi(t) = alpha + sum_s[ beta(s)*rec(t+s) ], where alpha = 0.41
MSE = 0.07
```

depending on only two coefficients, namely,

$$x_t = .41 + .016y_{t+4} - .02y_{t+5} + v_t.$$

Multiplying both sides by $50B^5$ and rearranging, we have

$$(1 - .8B)y_t = 20.5 - 50B^5x_t + \epsilon_t,$$

where ϵ_t is white noise, as our final, parsimonious model.

The example shows we can get a clean estimator for the transfer functions relating the two series if the coherence $\widehat{\rho}^2_{xy}(\omega)$ is large. The reason is that we can write the minimized mean squared error (4.124) as

$$MSE = E\left[\left(y_t - \sum_{r=-\infty}^{\infty} \beta_r x_{t-r}\right)y_t\right] = \gamma_{yy}(0) - \sum_{r=-\infty}^{\infty} \beta_r \gamma_{xy}(-r),$$

using the result about the orthogonality of the data and error term in the Projection theorem. Then, substituting the spectral representations of the autocovariance and cross-covariance functions and identifying the Fourier transform (4.127) in the result leads to

$$\begin{aligned} MSE &= \int_{-1/2}^{1/2} [f_{yy}(\omega) - B(\omega) f_{xy}(\omega)]\, d\omega \\ &= \int_{-1/2}^{1/2} f_{yy}(\omega)[1 - \rho^2_{yx}(\omega)] d\omega, \end{aligned} \tag{4.131}$$

where $\rho^2_{yx}(\omega)$ is just the squared coherence given by (4.87). The similarity of (4.131) to the usual mean square error that results from predicting y from x is obvious. In that case, we would have

$$E(y - \beta x)^2 = \sigma^2_y(1 - \rho^2_{xy})$$

for jointly distributed random variables x and y with zero means, variances σ_x^2 and σ_y^2, and covariance $\sigma_{xy} = \rho_{xy}\sigma_x\sigma_y$. Because the mean squared error in (4.131) satisfies $MSE \geq 0$ with $f_{yy}(\omega)$ a non-negative function, it follows that the coherence satisfies

$$0 \leq \rho_{xy}^2(\omega) \leq 1$$

for all ω. Furthermore, Problem 4.33 shows the squared coherence is one when the output are linearly related by the filter relation (4.122), and there is no noise, i.e., $v_t = 0$. Hence, the multiple coherence gives a measure of the association or correlation between the input and output series as a function of frequency.

4.11 Signal Extraction and Optimum Filtering

A model closely related to regression can be developed by assuming again that

$$y_t = \sum_{r=-\infty}^{\infty} \beta_r x_{t-r} + v_t, \tag{4.132}$$

but where the βs are known and x_t is some unknown random signal that is uncorrelated with the noise process v_t. In this case, we observe only y_t and are interested in an estimator for the signal x_t of the form

$$\widehat{x}_t = \sum_{r=-\infty}^{\infty} a_r y_{t-r}. \tag{4.133}$$

In the frequency domain, it is convenient to make the additional assumptions that the series x_t and v_t are both mean-zero stationary series with spectra $f_{xx}(\omega)$ and $f_{vv}(\omega)$, often referred to as the signal spectrum and noise spectrum, respectively. Often, the special case $\beta_t = \delta_t$, in which δ_t is the Kronecker delta, is of interest because (4.132) reduces to the simple signal plus noise model

$$y_t = x_t + v_t \tag{4.134}$$

in that case. In general, we seek the set of filter coefficients a_t that minimize the mean squared error of estimation, say,

$$MSE = E\left[\left(x_t - \sum_{r=-\infty}^{\infty} a_r y_{t-r}\right)^2\right]. \tag{4.135}$$

This problem was originally solved by Kolmogorov (1941) and by Wiener (1949), who derived the result in 1941 and published it in classified reports during World War II.

We can apply the orthogonality principle to write

$$E\left[\left(x_t - \sum_{r=-\infty}^{\infty} a_r y_{t-r}\right) y_{t-s}\right] = 0$$

for $s = 0, \pm 1, \pm 2, \ldots$, which leads to

$$\sum_{r=-\infty}^{\infty} a_r \gamma_{yy}(s-r) = \gamma_{xy}(s),$$

to be solved for the filter coefficients. Substituting the spectral representations for the autocovariance functions into the above and identifying the spectral densities through the uniqueness of the Fourier transform produces

$$A(\omega) f_{yy}(\omega) = f_{xy}(\omega), \tag{4.136}$$

where $A(\omega)$ and the optimal filter a_t are Fourier transform pairs for $B(\omega)$ and β_t. Now, a special consequence of the model is that (see Problem 4.23)

$$f_{xy}(\omega) = \overline{B(\omega)} f_{xx}(\omega) \tag{4.137}$$

and

$$f_{yy}(\omega) = |B(\omega)|^2 f_{xx}(\omega) + f_{vv}(\omega), \tag{4.138}$$

implying the optimal filter would be Fourier transform of

$$A(\omega) = \frac{\overline{B(\omega)}}{\left(|B(\omega)|^2 + \frac{f_{vv}(\omega)}{f_{xx}(\omega)}\right)}, \tag{4.139}$$

where the second term in the denominator is just the inverse of the signal to noise ratio, say,

$$\text{SNR}(\omega) = \frac{f_{xx}(\omega)}{f_{vv}(\omega)}. \tag{4.140}$$

$$a_t^M = M^{-1} \sum_{k=0}^{M-1} A(\omega_k) e^{2\pi i \omega_k t} \tag{4.141}$$

as the estimated filter function. It will often be the case that the form of the specified frequency response will have some rather sharp transitions between regions where the signal-to-noise ratio is high and regions where there is little signal. In these cases, the shape of the frequency response function will have ripples that can introduce frequencies at different amplitudes. An aesthetic solution to this problem is to introduce tapering as was done with spectral estimation in (4.61)-(4.68). We use below the tapered filter $\tilde{a}_t = h_t a_t$ where h_t is the cosine taper given in (4.68). The squared frequency response of the resulting filter will be $|\tilde{A}(\omega)|^2$, where

$$\tilde{A}(\omega) = \sum_{t=-\infty}^{\infty} a_t h_t e^{-2\pi i \omega t}. \tag{4.142}$$

The results are illustrated in the following example that extracts the El Niño component of the sea surface temperature series.

Example 4.25 Estimating the El Niño Signal via Optimal Filters

Figure 4.5 shows the spectrum of the SOI series, and we note that essentially two components have power, the El Niño frequency of about .02 cycles per month (the four-year cycle) and a yearly frequency of about .08 cycles per month (the annual cycle). We assume, for this example, that we wish to preserve the lower frequency as signal and to eliminate the higher order frequencies, and in particular, the annual cycle. In this case, we assume the simple signal plus noise model

$$y_t = x_t + v_t,$$

so that there is no convolving function β_t. Furthermore, the signal-to-noise ratio is assumed to be high to about .06 cycles per month and zero thereafter. The optimal frequency response was assumed to be unity to .05 cycles per point and then to decay linearly to zero in several steps. Figure 4.28 shows the coefficients as specified by (4.141) with $M = 64$, as well as the frequency response function given by (4.142), of the cosine tapered coefficients; recall Figure 4.9, where we demonstrated the need for tapering to avoid severe

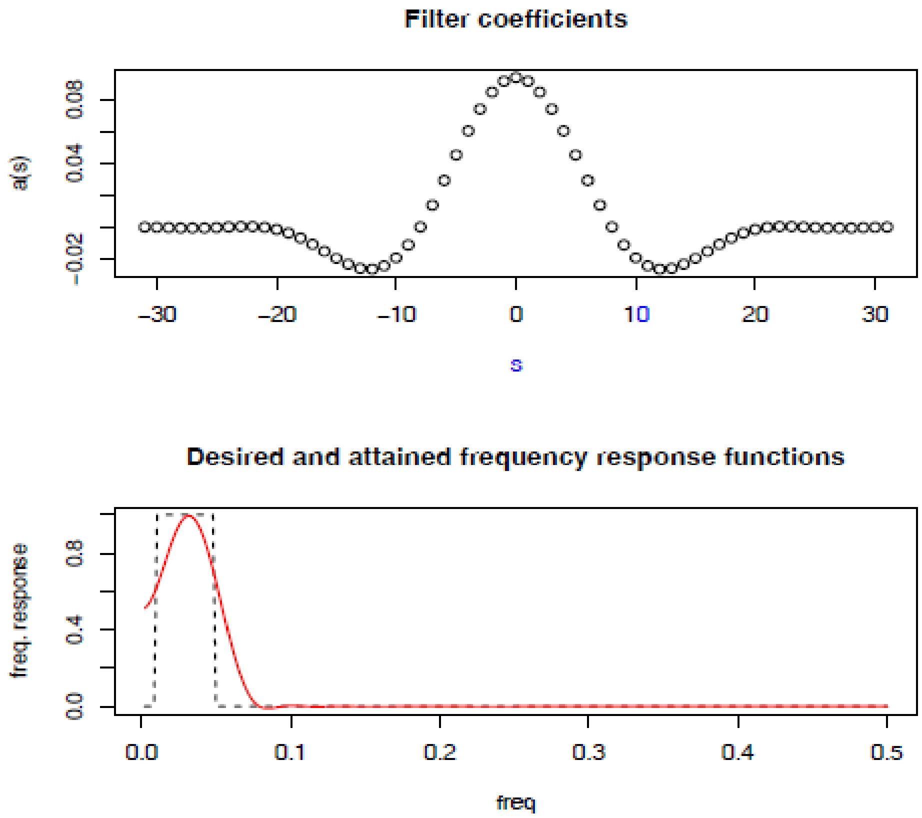

Fig. 4.28. Filter coefficients (top) and frequency response functions (bottom) for designed SOI filters.

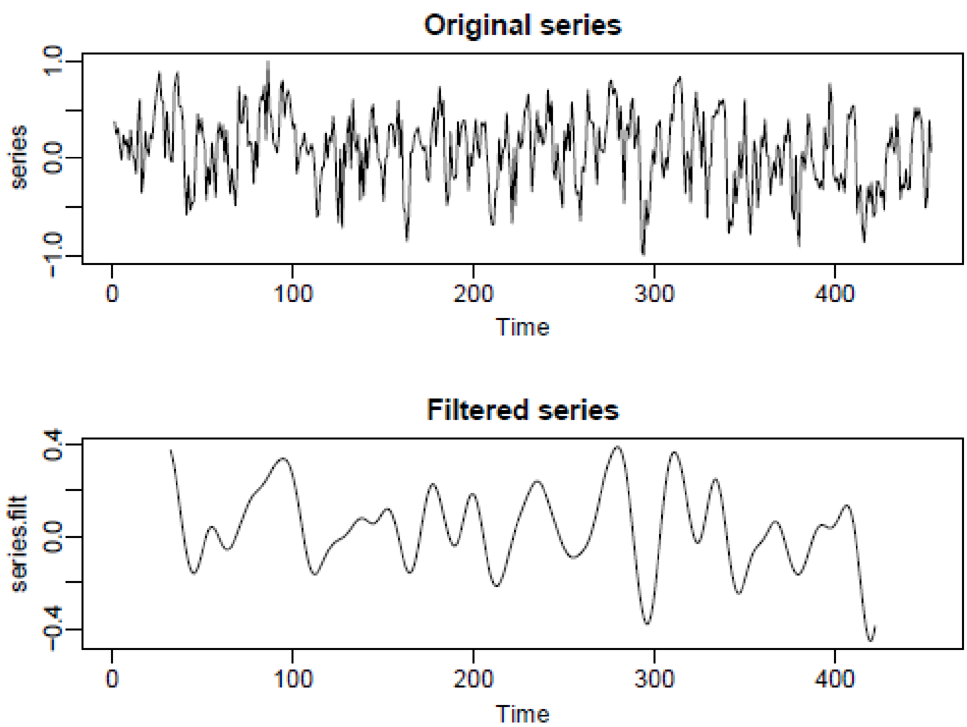

Fig. 4.29. Original SOI series (top) compared to filtered version showing the estimated El Niño temperature signal (bottom).