

Análise de Séries Temporais

Parte IV. Análise Espectral e Filtragem

Fernando Lucambio

Departamento de Estatística
Universidade Federal do Paraná

Dezembro, 2020

Vamos nos concentrar na abordagem do domínio da frequência para análise de séries temporais. Argumentamos que o conceito de regularidade de uma série pode ser melhor expresso em termos de variações periódicas do fenômeno subjacente que produziu a série.

Muitos dos exemplos vistos são séries temporais que são orientadas por componentes periódicos. Por exemplo, a gravação da fala no Exemplo I.3 contém uma mistura complicada de frequências relacionadas à abertura e ao fechamento da glote. O SOI mensal exibido no Exemplo I.5 contém duas periodicidades, um componente periódico sazonal de 12 meses e um componente El Niño de cerca de três a sete anos. De interesse fundamental é o período de retorno do fenômeno El Niño, que pode ter efeitos profundos no clima local.

Uma parte importante da análise de dados no domínio da frequência, bem como no domínio do tempo, é a investigação e exploração das propriedades do filtro linear invariante no tempo.

Essa transformação linear especial é usada de forma semelhante à regressão linear na estatística convencional e usamos muitos dos mesmos termos no contexto de séries temporais.

Muitas escalas de frequência coexistem, dependendo da natureza do problema. Por exemplo, no conjunto de dados da Johnson Johnson no Exemplo I.1, a frequência predominante de oscilação é de um ciclo por ano, 4 trimestres, ou $\omega = 0.25$ ciclos por observação. A frequência predominante nas séries de SOI e populações de peixes no Exemplo I.5 é também um ciclo por ano, mas isso corresponde a 1 ciclo a cada 12 meses ou $\omega = 1/12 = 0.083$ ciclos por observação.

Já encontramos a noção de periodicidade em numerosos exemplos nas Parte I, II e III. A noção geral de periodicidade pode ser tornada mais precisa introduzindo alguma terminologia. Para definir a taxa na qual uma série oscila, primeiro definimos um ciclo como um período completo de uma função seno ou cosseno definida ao longo de um intervalo de tempo unitário. Como no Exemplo I.12, consideramos o processo periódico

$$X_t = A \cos(2\pi\omega t + \phi),$$

para $t = 0, \pm 1, \pm 2, \dots$ onde ω é um índice de frequência, definido em ciclos por unidade de tempo com A determinando a altura ou amplitude da função e ϕ , chamada fase, determinando o ponto inicial da função cosseno. Podemos introduzir variação aleatória nesta série temporal permitindo a amplitude e fase para variar aleatoriamente.

Como discutido em exemplos, para fins de análise de dados, é mais fácil usar a identidade trigonométrica

$$\cos(\alpha \pm \beta) = \cos(\alpha) \cos(\beta) \mp \sin(\alpha) \sin(\beta),$$

e escrever a expressão acima como

$$X_t = U_1 \cos(2\pi\omega t) + U_2 \sin(2\pi\omega t),$$

onde $U_1 = A \cos(\phi)$ e $U_2 = -A \sin(\phi)$ são frequentemente considerados como variáveis aleatórias normalmente distribuídas.

Neste caso, a amplitude é $A = \sqrt{U_1^2 + U_2^2}$ e a fase é

$$\phi = \tan^{-1}(-U_2/U_1).$$

A partir desses fatos podemos mostrar que se, e somente se, A e ϕ são variáveis aleatórias independentes, onde $A^2 \sim \chi^2(2)$ e $\phi \sim U(-\pi, \pi)$, então U_1 e U_2 são variáveis aleatórias normais padronizadas.

Se assumirmos que U_1 e U_2 são variáveis aleatórias não correlacionadas com média 0 e variância σ^2 , então X_t é estacionária com média $E(X_t) = 0$ e, escrevendo

$$c_t = \cos(2\pi\omega t) \quad \text{e} \quad s_t = \sin(2\pi\omega t),$$

a função de autocovariância assume a forma

$$\gamma_X(h) = \text{Cov}(X_{t+h}, X_t) = \sigma^2 \cos(2\pi\omega h),$$

utilizando a relação $\cos(\alpha \pm \beta)$ mencionada acima e notando que $\text{Cov}(U_1, U_2) = 0$. Então, vemos que

$$\text{Var}(X_t) = \gamma_X(0) = \sigma^2.$$

O processo aleatório

$$X_t = U_1 \cos(2\pi\omega t) + U_2 \sin(2\pi\omega t),$$

é função da sua frequência ω . Para $\omega = 1$, a série faz um ciclo por unidade de tempo; para $\omega = 0.50$, a série faz um ciclo a cada duas unidades de tempo; para $\omega = 0.25$, a cada quatro unidades e assim por diante.

Em geral, para dados que ocorrem em pontos de tempo discretos, precisaremos de pelo menos dois pontos para determinar um ciclo, então a maior frequência de interesse é 0.5 ciclos por ponto. Essa frequência é chamada de frequência de dobra e define a frequência mais alta que pode ser vista na amostragem discreta. Frequências mais altas amostradas dessa maneira aparecerão em frequências mais baixas, chamadas aliases.

Um exemplo é o modo como uma câmera faz a amostragem de uma roda giratória em um automóvel em movimento em um filme, no qual a roda parece estar girando a uma taxa diferente, e às vezes para trás, o efeito da roda do vagão. Por exemplo, a maioria dos filmes são gravados a 24 quadros por segundo ou 24 Hertz. Se a câmera estiver filmando uma roda que esteja girando a 24 Hertz, a roda parecerá ficar parada.

Considere uma generalização que permite misturas de séries periódicas com múltiplas frequências e amplitudes,

$$X_t = \sum_{k=1}^q \left(U_{k1} \cos(2\pi\omega_k t) + U_{k2} \sin(2\pi\omega_k t) \right),$$

onde U_{k1} e U_{k2} , para $k = 1, 2, \dots, q$ são variáveis aleatórias de média zero não correlacionadas com variâncias σ_k^2 e as ω_k são frequências distintas.

Pode-se mostrar que a função de autocovariância do processo é

$$\gamma_X(h) = \sum_{k=1}^q \sigma_k^2 \cos(2\pi\omega h),$$

e notamos que a função de autocovariância é a soma de componentes periódicos com pesos proporcionais às variâncias σ_k^2 .

Portanto, X_t é um processo estacionário com média zero com variância

$$\gamma_X(0) = \text{Var}(X_t) = \sum_{k=1}^q \sigma_k^2,$$

exibindo a variância global como uma soma de variâncias de cada uma das partes componentes.

Exemplo IV.2. Estimação e periodograma.

Para qualquer amostra de série temporal x_1, x_2, \dots, x_n , onde n é ímpar, podemos escrever, exatamente

$$x_t = a_0 + \sum_{j=1}^{(n-1)/2} \left(a_j \cos(2\pi t j/n) + b_j \sin(2\pi t j/n) \right),$$

para $t = 1, \dots, n$ e coeficientes adequadamente escolhidos.

O ponto crucial aqui é que esta representação é exata para qualquer amostra. Portanto, a série periódica com múltiplas frequências e amplitudes, pode ser pensada como uma aproximação para a expressão acima, a ideia é que muitos dos coeficientes a serem adequadamente escolhidos podem ser próximos de zero.

Exemplo IV.2. Estimação e periodograma (continuação).

Usando os resultados de regressão, os coeficientes a_j e b_j são da forma $\sum_{t=1}^n x_t z_{tj} / \sum_{t=1}^n z_{tj}^2$, onde z_{tj} é um outro $\cos(2\pi t j/n)$ ou $\sin(2\pi t j/n)$. Pode-se demonstrar que $\sum_{t=1}^n z_{tj}^2 = n/2$, quando $j/n \neq 0, 1/2$ então os coeficientes de regressão podem ser escritos como $a_0 = \bar{x}$,

$$a_j = \frac{2}{n} \sum_{t=1}^n x_t \cos(2\pi t j/n) \quad \text{e} \quad b_j = \frac{2}{n} \sum_{t=1}^n x_t \sin(2\pi t j/n).$$

Em seguida, definimos o periodograma escalonado como sendo

$$P(j/n) = a_j^2 + b_j^2,$$

e é de interesse porque indica quais componentes de frequência são grandes em magnitude e quais componentes são pequenos.

Exemplo IV.2. Estimação e periodograma (continuação).

O periodograma escalado é simplesmente a variância amostral em cada componente de frequência e, conseqüentemente, é uma estimativa de σ_j^2 correspondente à senoide que oscila a uma frequência de $\omega_j = j/n$.

Essas frequências específicas são chamadas de frequências de Fourier ou frequências fundamentais.

Grandes valores de $P(j/n)$ indicam quais frequências $\omega_j = j/n$ são predominantes na série, enquanto valores pequenos de $P(j/n)$ podem estar associados a ruído. O periodograma foi introduzido em Schuster (1898).

Exemplo IV.2. Estimação e periodograma (continuação).

A transformada discreta de Fourier (DFT) é uma média ponderada de valor complexo dos dados fornecidos, dada por

$$\begin{aligned}d(j/n) &= \frac{1}{n} \sum_{t=1}^n x_t \exp(-2\pi i t j/n) \\ &= \frac{1}{n} \left(\sum_{t=1}^n x_t \cos(2\pi t j/n) - i \sum_{t=1}^n x_t \sin(2\pi t j/n) \right)\end{aligned}$$

para $j = 0, 1, \dots, n-1$, onde as frequências j/n são as frequências de Fourier ou frequências fundamentais.

Exemplo IV.2. Estimação e periodograma (continuação).

Devido a um grande número de redundâncias no cálculo, $d(j/n)$ pode ser calculado rapidamente usando a transformada rápida de Fourier (FFT). Observe que

$$|d(j/n)|^2 = \frac{1}{n} \left(\sum_{t=1}^n x_t \cos(2\pi t j/n) \right)^2 + \frac{1}{n} \left(\sum_{t=1}^n x_t \sin(2\pi t j/n) \right)^2$$

e é essa quantidade chamada de periodograma.

Podemos calcular o periodograma escalonado, utilizando o periodograma como

$$P(j/n) = \frac{4}{n} |d(j/n)|^2.$$

Exemplo IV.2. Estimação e periodograma (continuação).

O periodograma escalado de dados x_t simulados é mostrado na figura abaixo e identifica claramente os três componentes x_{t1} , x_{t2} e x_{t3} de x_t . Observe que

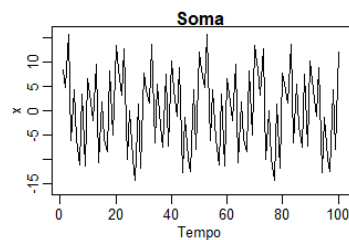
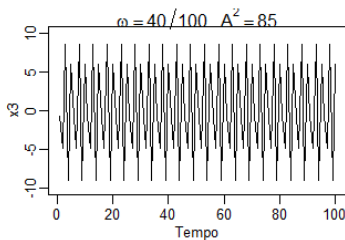
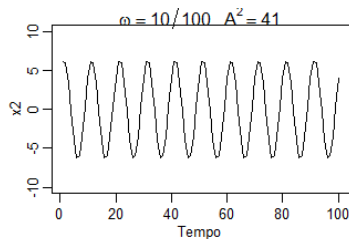
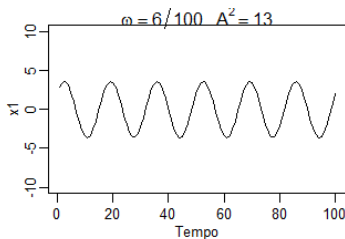
$$P(j/n) = P(1 - j/n), \quad j = 0, 1, \dots, n - 1,$$

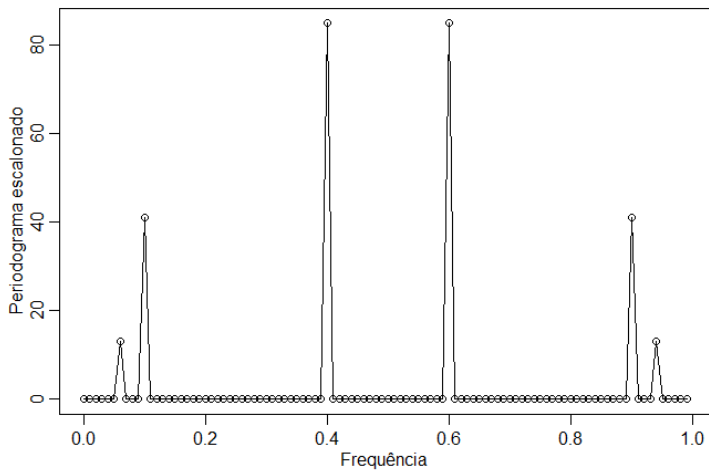
então há um efeito de espelhamento na frequência de dobramento de 1/2; conseqüentemente, o periodograma normalmente não é mostrado para frequências mais altas que a frequência de dobramento.

Além disso, note que as alturas do periodograma escalado mostrado na figura são $P(6/100) = P(94/100) = 13$, $P(10/100) = P(90/100) = 41$, $P(40/100) = P(60/100) = 85$, e $P(j/n) = 0$, caso contrário.

Assumindo que os dados simulados, x , foram retidos do exemplo anterior, o código R para reproduzir a figura acima é:

```
> x1 = 2*cos(2*pi*1:100*6/100) + 3*sin(2*pi*1:100*6/100)
> x2 = 4*cos(2*pi*1:100*10/100) + 5*sin(2*pi*1:100*10/100)
> x3 = 6*cos(2*pi*1:100*40/100) + 7*sin(2*pi*1:100*40/100)
> x = x1 + x2 + x3
> P = Mod(2*fft(x)/100)^2; Fr = 0:99/100
> par(mfrow=c(1,1),mar=c(4,3,1,1),mgp=c(1.6,.6,0))
> plot(Fr, P, type="o", xlab="Frequência",
      ylab="Periodograma escalonado", pch=19)
> grid()
```



Exemplo IV.3. Magnitude da Estrela.

Os dados na segunda figura embaixo são a magnitude de uma estrela tirada à meia-noite por 600 dias consecutivos. Os dados são retirados do texto clássico: *The Calculus of Observations, a Treatise on Numerical Mathematics*, por E.T. Whittaker e G. Robinson (1923, Blackie & Son, Ltd.).

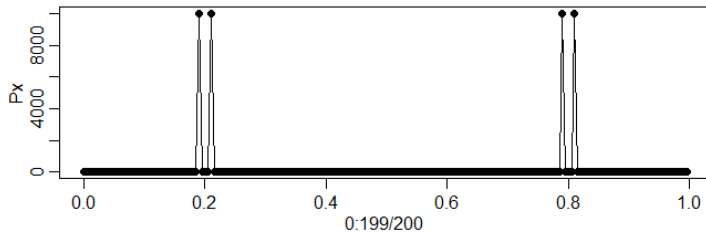
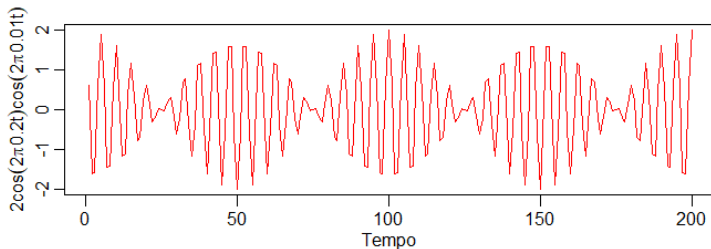
O periodograma para frequências inferiores a 0.08 também é exibido na figura; as ordenadas do periodograma para frequências superiores a 0.08 são essencialmente zero. Observe que o ciclo de $29 \approx 1/0.035$ dias e o ciclo de $24 \approx 1/0.041$ dias são os componentes periódicos mais proeminentes dos dados.

Exemplo IV.3. Magnitude da Estrela (continuação).

Podemos interpretar esse resultado conforme estamos observando um sinal modulado em amplitude. Por exemplo, suponha que estamos observando sinal-mais-ruído, $X_t = s_t + V_t$, onde $s_t = \cos(2\pi\omega t) \cos(2\pi\delta t)$ e δ é muito pequeno. Neste caso, o processo irá oscilar na frequência ω , mas a amplitude será modulada por $\cos(2\pi\delta t)$. Como

$$2 \cos(\alpha) \cos(\delta) = \cos(\alpha + \delta) + \cos(\alpha - \delta),$$

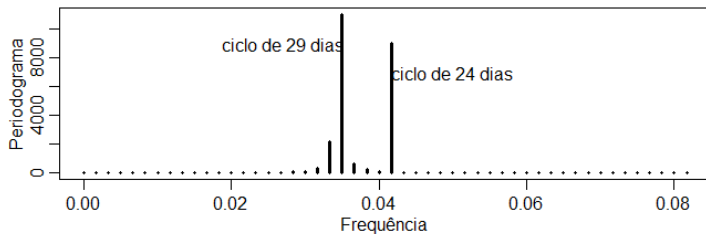
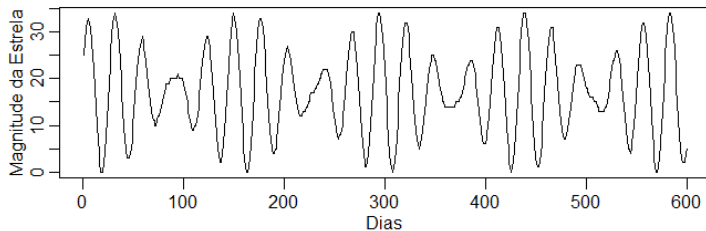
o periodograma dos dados gerados como X_t terá dois picos próximos um do outro em. Observe isso na figura a seguir:



Exemplo IV.3. Magnitude da Estrela (continuação).

Os comandos a seguir foram utilizados para gerar a figura acima, permitem entendermos melhor a informação fornecida pelo periodograma.

```
> t = 1:200
> x <- 2*cos(2*pi*.2*t)*cos(2*pi*.01*t)
> plot.ts(x, type="n", xlab="Tempo",
          ylab=expression(2*cos(2*pi*0.2*t)*cos(2*pi*.01*t)))
> lines(cos(2*pi*.19*t)+cos(2*pi*.21*t), col=2) # the same
> Px = Mod(fft(x))^2
> plot(0:199/200, Px, type='o', pch=19) # periodograma
```



Exemplo IV.3. Magnitude da Estrela (continuação).

```

> n = length(star)
> par(mfrow=c(2,1), mar=c(3,3,1,1), mgp=c(1.6,.6,0))
> plot(star, ylab="Magnitude da Estrela", xlab="Dias")
> Per = Mod(fft(star-mean(star)))^2/n
> Freq = (1:n -1)/n
> plot(Freq[1:50], Per[1:50], type='h', lwd=3,
       ylab="Periodograma", xlab="Frequência")
> u = which.max(Per[1:50]) # 22 freq=21/600=.035 ciclos/dia
> uu = which.max(Per[1:50][-u]) # 25 freq=25/600=.041 ciclos/dia
> 1/Freq[22]; 1/Freq[26] # período = dias/ciclo
[1] 28.57143
[1] 24
> text(.05, 7000, "ciclo de 24 dias")
> text(.027, 9000, "ciclo de 29 dias")

```


Nesta seção, definimos a ferramenta do domínio de frequência fundamental, a densidade espectral. Além disso, discutimos as representações espectrais para processos estacionários.

Teorema IV.2. A Densidade Espectral.

Se a função de autocovariância $\gamma(h)$, de um processo estacionário é tal que $\sum_h |\gamma(h)| < \infty$, então tem a representação

$$\gamma(h) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \exp(2\pi i \omega t) f(\omega) d\omega, \quad h = 0, \pm 1, \pm 2, \dots,$$

como a transformação inversa da densidade espectral,

$$f(\omega) = \sum_{h=-\infty}^{\infty} \gamma(h) \exp(-2\pi i \omega t), \quad -1/2 \leq \omega \leq 1/2.$$

Essa densidade espectral é o análogo da função de densidade; o fato de $\gamma(h)$ ser definido não negativo garante que $f(\omega) \geq 0$, para todo ω . Segue imediatamente do Teorema IV.2 que $f(\omega) = f(-\omega)$, verificando que a densidade espectral é uma função par.

Por causa da uniformidade, tipicamente traçamos apenas $f(\omega)$ para $0 \leq \omega \leq 1/2$. Mais ainda, escolhendo $h = 0$ na expressão acima, produz

$$\gamma(0) = \text{Var}(X_t) = \int_{-\frac{1}{2}}^{\frac{1}{2}} f(\omega) d\omega,$$

que expressa a variância total como densidade espectral integrada em todas as frequências. Mostramos mais adiante que um filtro linear pode isolar a variância em certos intervalos de frequência ou bandas.

Agora deve ficar claro que as funções de autocovariância e distribuição espectral contêm as mesmas informações. Essa informação, no entanto, é expressa de maneiras diferentes. A função de autocovariância expressa informações em termos de atrasos, enquanto a distribuição espectral expressa as mesmas informações em termos de ciclos.

Alguns problemas são mais fáceis de trabalhar quando consideramos as informações defasadas e nós tenderíamos a lidar com esses problemas no domínio do tempo. No entanto, outros problemas são mais fáceis de trabalhar quando consideramos informações periódicas e nós tenderíamos a lidar com esses problemas no domínio espectral.

Uma das possibilidades intrigantes oferecidas seria estender a regressão clássica à análise de modelos de regressão defasados da forma

$$Y_t = \sum_{r=-\infty}^{\infty} \beta_r X_{t-r} + V_t,$$

onde V_t é um processo de ruído estacionário, X_t é a série de entrada observada e Y_t é a série de saída observada. Estamos interessados em estimar os coeficientes de filtro β_r relacionando os valores defasados adjacentes de X_t à série de saída Y_t .

No caso das séries SOI e Recrutamento, podemos identificar a série de condução do El Niño, SOI, como a entrada, X_t e Y_t , a série de Recrutamento, como a saída. Em geral, haverá mais de uma série de entradas possíveis e podemos imaginar um vetor $q \times 1$ de séries de condução.

O modelo acima é útil sob vários cenários diferentes, correspondendo a diferentes suposições que podem ser feitas sobre os componentes.

Assumimos que as entradas e saídas têm médias zero e são conjuntamente estacionárias, de dimensão 2×1 , com o processo vetorial $(X_t, Y_t)^\top$ tendo uma matriz espectral da forma

$$f(\omega) = \begin{pmatrix} f_{X,X}(\omega) & f_{X,Y}(\omega) \\ f_{Y,X}(\omega) & f_{Y,Y}(\omega) \end{pmatrix}.$$

Aqui, $f_{X,Y}(\omega)$ é o espectro cruzado relacionando a entrada X_t com a saída Y_t , e $f_{X,X}(\omega)$ e $f_{Y,Y}(\omega)$ são os espectros das séries de entrada e saída, respectivamente. Geralmente, observamos duas séries, consideradas como entrada e saída e procuramos por funções de regressão para relacionar as entradas às saídas. Presumimos que todas as funções de autocovariância satisfazem as condições de soma absoluta.

Então, minimizando o erro quadrático médio

$$MSE = E \left(Y_t - \sum_{r=-\infty}^{\infty} \beta_r X_{t-r} \right)^2,$$

leva às condições usuais de ortogonalidade

$$E \left(\left(Y_t - \sum_{r=-\infty}^{\infty} \beta_r X_{t-r} \right) X_{t-s} \right) = 0$$

para todos os $s = 0, \pm 1, \pm 2, \dots$. Tomando as esperanças dentro
leva às equações normais

$$\sum_{r=-\infty}^{\infty} \beta_r \gamma_{X,X}(s-r) = \gamma_{Y,X}(s),$$

para $s = 0, \pm 1, \pm 2, \dots$.

Então, podemos escrever

$$\widehat{B}(\omega_k) = \frac{\widehat{f}_{Y,X}(\omega_k)}{\widehat{f}_{X,X}(\omega_k)},$$

como estimador da transformada de Fourier dos coeficientes de regressão, avaliados em algum subconjunto de frequências fundamentais $\omega_k = k/M$ com $M \ll n$.

A transformada inversa da função $\widehat{B}(\omega)$ nos daria $\widehat{\beta}_t$ e observamos que a aproximação discreta pode ser tomada como

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

para $t = 0, \pm 1, \pm 2, \dots, \pm(M/2 - 1)$. Se tivéssemos que usar a expressão acima para definir $\widehat{\beta}_t$ para $|t| \geq M/2$, acabaríamos com uma sequência de coeficientes que é periódica com um período de M . Na prática, definimos $\widehat{\beta}_t = 0$ para $|t| \geq M/2$.

Exemplo IV.24. Regressão retardada para SOI e Recrutamento.

A alta coerência entre as séries SOI e Recrutamento sugere uma relação de regressão defasada ou retardada entre as duas séries. Uma direção natural para a implicação nesta situação está implícita porque sentimos que a temperatura da superfície do mar ou SOI deve ser a entrada e a série de Recrutamento deve ser a saída. Com isto em mente, seja X_t a série SOI e Y_t a série Recrutamento.

Embora pensemos naturalmente no SOI como a entrada e o Recrutamento como a saída, duas configurações de entrada-saída são de interesse. Com o SOI como entrada, o modelo é

$$Y_t = \sum_{r=-\infty}^{\infty} a_r X_{t-r} + W_t.$$

Exemplo IV.24. Regressão retardada para SOI e Recrutamento.

Considerando que um modelo que inverte os dois papéis seria

$$X_t = \sum_{r=-\infty}^{\infty} b_r Y_{t-r} + V_t,$$

onde W_t e V_t são processos de ruído branco. Embora não exista uma explicação ambiental plausível para o segundo desses dois modelos, a exibição de ambas as possibilidades ajuda a estabelecer um modelo de função de transferência parcimoniosa, mas não será mostrado.

Baseado no script LagReg no pacote astsa, a função de regressão estimada ou resposta ao impulso para SOI, com $M = 32$ e $L = 15$ é:

Exemplo IV.24. Regressão retardada para SOI e Recrutamento.

```
> library(astsa)
> resultado = LagReg(soi, rec, L=15, M=32, threshold=6)
INPUT: soi OUTPUT: rec    L = 15    M = 32
```

The coefficients $\beta(0)$, $\beta(1)$, $\beta(2)$... $\beta(M/2-1)$ are

```
3.463141 2.088613 2.688139 -0.3515829 0.3717705 -18.47931
-12.2633 -8.539368 -6.984553 -4.978238 -4.526358 -4.223713
-3.239262 -1.372815 1.489903 3.744727
```

The coefficients $\beta(0)$, $\beta(-1)$, $\beta(-2)$... $\beta(-M/2+1)$ are

```
3.463141 2.835444 1.628129 2.726815 0.6330645 -1.256092
-0.05458373 1.722774 4.925481 5.440352 6.877381 5.141606
4.479202 3.796848 4.004762 3.184184
```

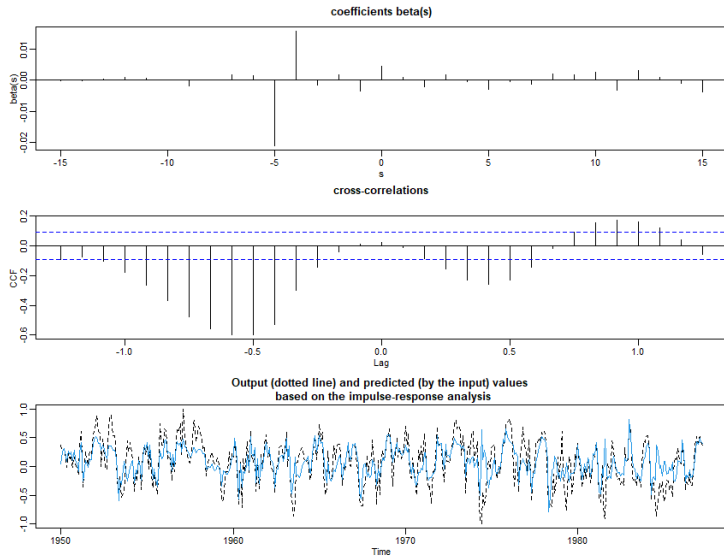
Exemplo IV.24. Regressão retardada para SOI e Recrutamento.

The positive lags, at which the coefficients are large in absolute value, and the coefficients themselves, are:

	lag	s	beta(s)
[1,]	5		-18.479306
[2,]	6		-12.263296
[3,]	7		-8.539368
[4,]	8		-6.984553

The prediction equation is

$\text{rec}(t) = \alpha + \sum_s [\text{beta}(s) \cdot \text{soi}(t-s)]$, where $\alpha = 65.96584$
MSE = 414.0847



Exemplo IV.24. Regressão retardada para SOI e Recrutamento.

Observe o pico negativo em um atraso de cinco pontos na figura acima. A queda após o desfasamento cinco parece ser aproximadamente exponencial e um possível modelo é

$$Y_t = 65.96584 - 18.479306X_{t-5} - 12.263296X_{t-6} \\ - 8.539368X_{t-7} - 6.984553X_{t-8} + W_t$$