Análise Harmônica de Séries Temporais

Airlane Pereira Alencar

15 de Março de 2019

Índice

Contexto

Em geral, na área de sinais, a análise Harmônica também é chamada de análise de Fourier e é utilizada na solução de equações diferenciais.
Ex: Equações de Calor e de Ondas

Em análise de séries temporais, corresponde ao ajuste de modelos lineares que são combinações de senos e cossenos (MT2018).
Os coeficientes estimados são as transformadas discretas de Fourier da série.

- Conhecidas as frequências, estimar amplitudes e fases;
- Estimar também frequências.

Modelos com uma periodicidade conhecida

$$Z_t = \mu + R\cos(\omega t + \phi) + e_t \qquad (1)$$
$$Z_t = \mu + A\cos(\omega t) + B\sin(\omega t) + e_t \qquad (2)$$

$e_t \sim RB$
Para ω conhecida, temos que (2) é modelo linear e (1) não é.

Trigonometria! $\cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b)$

$$R\cos(\omega t + \phi) = R\cos(\phi)\cos(\omega t) - R\sin(\phi)\sin(\omega t)$$

$A = R\cos(\phi)$; $B = -R\sin(\phi)$
$R = A^2 + B^2$ e $-\frac{B}{A} = \frac{\sin(\phi)}{\cos(\phi)} = tg(\phi)$

Estimação por Mínimos Quadrados

Minimizo

$$SQR(\mu, A, B) \;=\; \sum_{t=1}^{N}[Z_t - \mu - A\cos(\omega t) - B\sin(\omega t)]^2$$

Matricialmente $\mathbf{Z} = \mathbf{W}\theta + \mathbf{e} \Rightarrow \widehat{\theta} = (\mathbf{W}^T\mathbf{W})^{-1}\mathbf{W}^T\mathbf{Z}$

$$\mathbf{W} = \begin{pmatrix} 1 & \cos(\omega 1) & \sin(\omega 1) \\ 1 & \cos(\omega 2) & \sin(\omega 2) \\ \vdots & \vdots & \vdots \\ 1 & \cos(\omega N) & \sin(\omega N) \end{pmatrix} \theta = \begin{pmatrix} \mu \\ A \\ B \end{pmatrix}$$

$$\mathbf{W}^{-1}\mathbf{W} = \begin{pmatrix} N & \sum_{t=1}^{N}\cos(\omega t) & \sum_{t=1}^{N}\sin(\omega t) \\ \sum_{t=1}^{N}\cos(\omega t) & \sum_{t=1}^{N}\cos^2(\omega t) & \sum_{t=1}^{N}\cos(\omega t)\sin(\omega t) \\ \sum_{t=1}^{N}\sin(\omega t) & \sum_{t=1}^{N}\cos(\omega t)\sin(\omega t) & \sum_{t=1}^{N}\sin^2(\omega t) \end{pmatrix}$$

Soma de senos Li2001, Hall Knight1952

$$cos(a - b/2) - cos(a + b/2) = 2sen(a)sen(b/2), \quad a = x, 2x, \ldots, Nx, \quad b/2 = x/2$$

$$\left[cos\left(\frac{x}{2}\right) - cos\left(\frac{3x}{2}\right)\right] + \left[cos\left(\frac{3x}{2}\right) - cos\left(\frac{5x}{2}\right)\right] + \ldots + \left[cos\left(\left(N - \frac{1}{2}\right)x\right) - cos\left(\left(N + \frac{1}{2}\right)x\right)\right] = 2sen\left(\frac{x}{2}\right)[senx + sen(2x) + \ldots + sen(Nx)]$$

$$\sum_{k=1}^{N} sen(kx) = \frac{cos(x/2) - cos((N + 1/2)x)}{2sen(x/2)} \overset{x=\frac{2\pi k}{N}}{=} \frac{cos(\pi k/N) - cos(2\pi k + \frac{\pi k}{N})}{2sen(2\pi k/N)} = 0$$

Soma de senos Li2001, Hall Knight1952

Analogamente

$$\sum_{k=1}^{N} cos(kx) = \frac{sen((N+1/2)x) - sen(x/2)}{2sen(x/2)} \overset{x=\frac{2\pi k}{N}}{=} 0$$

$$cos(2x) = cos^2(x) - sen^2(x) = 1 - 2sen^2(x) = -1 + 2cos^2(x)$$

$$\sum_{k=1}^{N} cos^2(kx) = \sum_{k=1}^{N} \frac{cos(2kx)+1}{2} = \sum_{k=1}^{N} \frac{cos(k2x)}{2} + \frac{N}{2} \overset{x=\frac{2\pi j}{N}}{=}$$

$$= \frac{sen((\frac{2N}{2}+\frac{1}{2})2\frac{2\pi j}{N}) - sen(\frac{2\pi j}{N})}{2sen(x/2)} + \frac{N}{2} = \frac{N}{2}$$

Estimação por Mínimos Quadrados

Para frequências fundamentais $\omega = \frac{2\pi k}{N}, k = 1, \ldots, \lfloor \frac{N}{2} \rfloor$

$$\mathbf{W}^{-1}\mathbf{W} = \begin{pmatrix} N & 0 & 0 \\ 0 & N/2 & 0 \\ 0 & 0 & N/2 \end{pmatrix}$$

$$\begin{aligned}
\widehat{\mu} &= \overline{Z} \\
\widehat{A} &= \frac{2}{N}\sum_{t=1}^{N} Z_t cos(\omega t), \quad \widehat{A} = \frac{2}{N}\sum_{t=1}^{N} Z_t(-1)^t, \omega = \pi \\
\widehat{B} &= \frac{2}{N}\sum_{t=1}^{N} Z_t sen(\omega t), \quad \widehat{B} = 0, \omega = \pi \\
\widehat{R}^2 &= \widehat{A}^2 + \widehat{B}^2 \qquad \widehat{\phi} = arctg - \frac{\widehat{B}}{\widehat{A}}
\end{aligned}$$

Estimação por Mínimos Quadrados

Para frequências não fundamentais, há resultado aproximado.

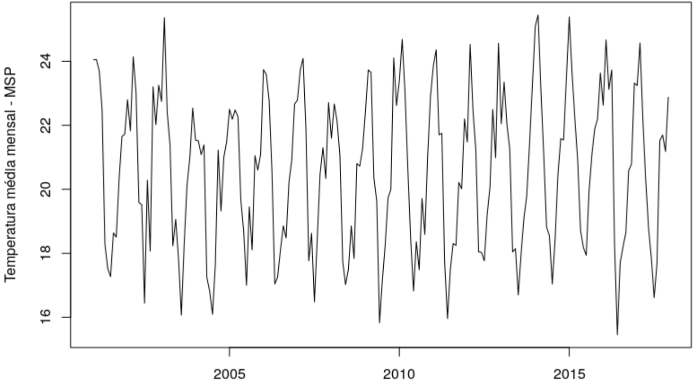
$$
\begin{aligned}
\widehat{\mu} &= \overline{Z} \\
\tilde{A} &= \frac{2}{N}\sum_{t=1}^{N}(Z_t - \bar{Z})cos(\omega t), \\
\tilde{B} &= \frac{2}{N}\sum_{t=1}^{N}(Z_t - \bar{Z})sen(\omega t), \\
\tilde{R}^2 &= \tilde{A}^2 + \tilde{B}^2 \qquad \widehat{\phi} = arctg - \frac{\tilde{B}}{\tilde{A}}
\end{aligned}
$$

Soma de Quadrados de Resíduos

$$
\begin{aligned}
SQR &= \sum_{t=1}^{N}(Z_t - \overline{Z} - \widehat{A}cos(\omega t) - \widehat{B}sen(\omega t))^2 \\
&= \sum_{t=1}^{N}(Z_t - \overline{Z})^2 - 2\sum_{t=1}^{N}(Z_t - \overline{Z})[\widehat{A}cos(\omega t) + \widehat{B}sen(\omega t)] + \\
&\quad + \sum_{t=1}^{N}[\widehat{A}cos(\omega t) + \widehat{B}sen(\omega t)]^2 = \\
&= \sum_{t=1}^{N}(Z_t - \overline{Z})^2 - N(\widehat{A}^2 + \widehat{B}^2) + \frac{N}{2}(\widehat{A}^2 + \widehat{B}^2) \\
&= \sum_{t=1}^{N}(Z_t - \overline{Z})^2 - \frac{N}{2}(\widehat{A}^2 + \widehat{B}^2) = SQT - \frac{N}{2}\widehat{R}^2
\end{aligned}
$$

$\frac{N}{2}\widehat{R}^2$ corresponde à parcela da var. total explicada pela freq ω.

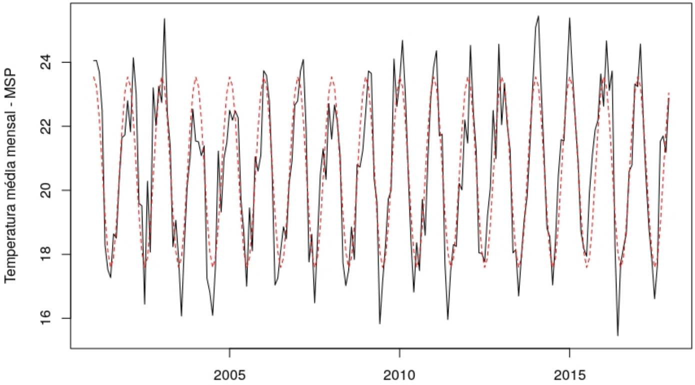
Temperatura no Mun. São Paulo desde 2001



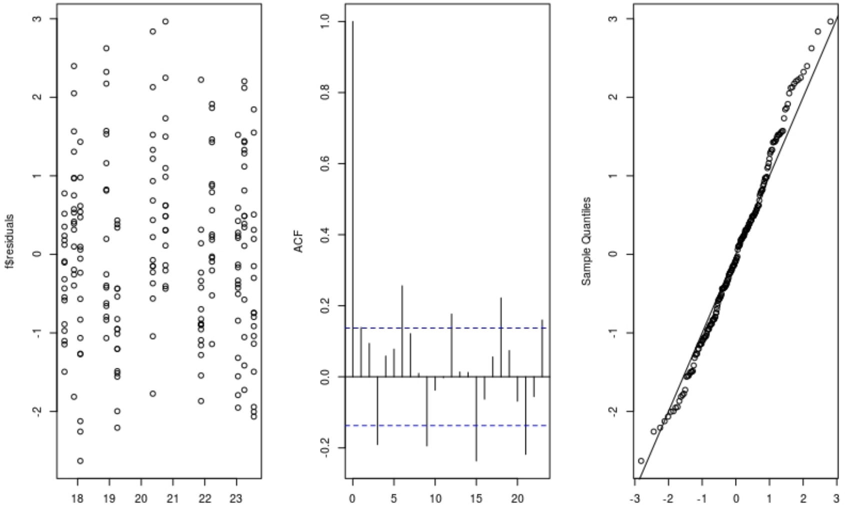
Fonte: INMET - Mirante de Santana   Ciclo: 12 meses: $\omega = 2\pi/12$

```
lm(formula = temper ~ cost + sint)
               Estimate Std. Error t value       p
(Intercepto) 20.56472    0.07924  259.52   <2e-16 ***
cos(2pit/12)  2.47914    0.11207   22.12   <2e-16 ***
sin(2pit/12)  1.66090    0.11207   14.82   <2e-16 ***

Residual standard error: 1.132 on 201 degrees of freedom
Multiple R-squared:  0.7791,  Adjusted R-squared:  0.7769
F-statistic: 354.5 on 2 and 201 DF,  p-value: < 2.2e-16

>(f$coef[2]^2 + f$coef[3]^2) * 204/2   #SQExpl
908.2798
> sum((temper-mean(temper))^2)        # SQT
1165.761

> (f$coef[2]^2 + f$coef[3]^2) * 204/2/
sum((temper-mean(temper))^2)
0.7791307
```

Temperatura no Mun. São Paulo desde 2001



Temperatura no Mun. São Paulo desde 2001



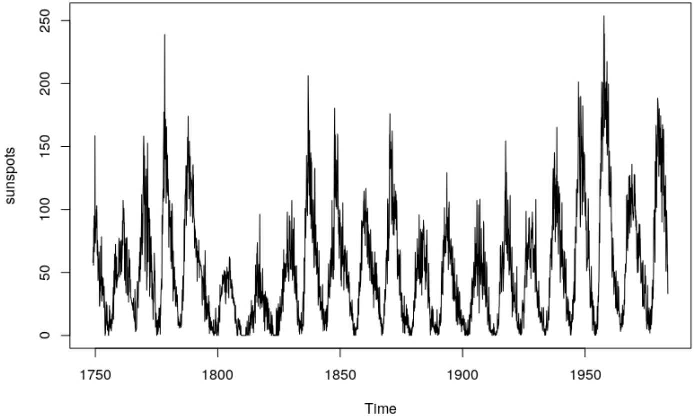
Frequência desconhecida

- Encontrar frequência ω que maximiza $\widehat{R}^2(\omega) = \widehat{A}^2(\omega) + \widehat{B}^2(\omega)$.
- Consideramos $\tilde{R}^2(\omega) = \tilde{A}^2(\omega) + \tilde{B}^2(\omega)$ mesmo para freq. não fundamentais.
- Max. $\tilde{R}^2$ equivale a maximizar o periodograma:

$$
\begin{aligned}
I(\omega) &= \frac{N}{8\pi}\tilde{R}^2(\omega) \\
&= \frac{1}{2\pi N}\left[\sum_{t=1}^{N}[(Z_t - \overline{Z})cos(\omega t)]^2 + [(Z_t - \overline{Z})sen(\omega t)]^2\right]
\end{aligned}
$$

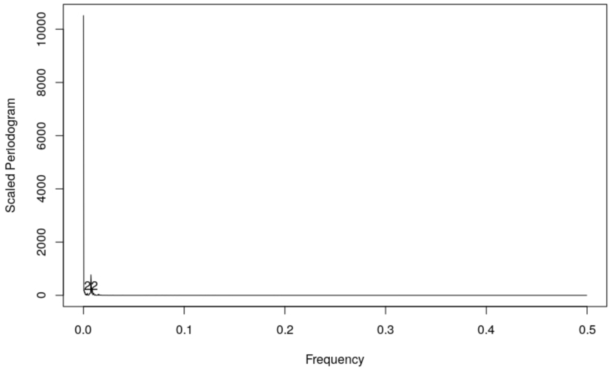
Manchas Solares

Número médio mensal de manchas solares de 1749 a 1983 (N=2820)



Fonte: Swiss Federal Observatory, Zurich até 1960 e depois Tokyo

Manchas Solares



I = abs(fft(sunspots)) ^{2} /2820 I(1:1410)

2820/22 = 128.1818 meses (2820/22)/12 = 10.68 anos

Observações

- Para várias frequências ω, basta analisar $\widehat{R}(\omega)$ devido à ortogonalidade.
- Pode usar regressão não linear e estimar ω (MT)
- Posso ter modelos com tendências, mudanças de tendências, heterocedasticidade e sazonalidade com erros ARMA.

Testes baseados no Periodograma

$$Z_t = \sum_{i=1}^{K} R_i cos(\omega_i t + \phi_i) + e_i, t = 1, \ldots, N \qquad (3)$$

com K, R_i e ω_i constantes e $\phi \sim U(-\pi, \pi)$ indep. e $e_t \sim RB$.

Teste de Fisher (1929). Z_t Gaussiano (MT, 2018, p.382)
$H_0 : R_i = 0, \forall i$ (sem periodicidade)
Estatística $g = \frac{max I_j^{(N)}}{\sum_{j=1}^{[N/2]} I_j^{(N)}}$
Para N ímpar, sob H_0, temos:
$P(g > a) = n(1-a)^{n-1} - \binom{n}{2}(1-2a)^{n-1} + \cdots + (-1)^x \binom{n}{x}(1-xa)^{n-1}$,
com $n = [N/2]$ e x o maior inteiro menor que $1/a$.
Tenho que encontrar a, para ter $P(g > a(\alpha)) = \alpha$.

Testes baseados no Periodograma

Uma boa aproximação é obtida só com: $P(g > a) = n(1-a)^{n-1}$.

Whittle(1952) sugeriu extensão para testar a segunda maior ordenada do periodograma, omitindo-se o maior valor do periodograma:

$$g = \frac{I^{(2)}}{\sum_{j=1}^{[N/2]} I_j^{(N)} - I^{(1)}}$$

pode considerar a 3a maior ordenada,...

Priestley(1981) apresenta teste mais geral em que e_t não precisa ser ruído branco mas sim processo estacionário.

```
(g = I[22]/sum(I[1:1410]))
[1] 0.05397431
N*(1-g)^(N-1)
[1] 3.316e-65
```

Teste de Fisher em fisher.g.test na library(GeneCycle) do R.

Estimação do espectro

O periodograma é estimador assint. não viesado da **densidade espectral** ou **espectro**:

$$f(\omega) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} \gamma(\tau) e^{-i\omega\tau}$$

que caracteriza o processo estacionário como faz a FACv $\gamma()$.

- Essa densidade só está definida quando $\sum_{\tau=-\infty}^{\infty} |\gamma(\tau)| < \infty$.
- **Propriedades**: O espectro $f(\omega)$ é função não negativa, uniformemente contínua, par, e de período 2π.
- É possível obter a FAC usando a transformada inversa:

$$\gamma(\tau) = \int_{-\pi}^{\pi} f(\omega) e^{i\omega\tau} d\omega, \tau = 0, \mp 1, \ldots$$

- Decomposição da variância $\gamma(0) = \int_{-\pi}^{\pi} f(\omega) d\omega$

ANOVA Espectral p.189 SS2016

SS define o periodograma para as frequências fundamentais $\omega_j = j/N$:

$$
\begin{aligned}
I^*(\omega) &= |d(\omega_j)|^2 = |n^{-1/2}\sum_{t=1}^{N} x_t e^{-2\pi i\omega_j t}|^2 \\
&= \left| n^{-1/2}\sum_{t=1}^{N} x_t cos(2\pi\omega_j t) - i\ n^{-1/2}\sum_{t=1}^{N} x_t sen(2\pi\omega_j t)\right|^2 \\
&= |d_c(\omega_j) - id_s(\omega_j)|^2 = d_c(\omega_j)^2 + d_s(\omega_j)^2
\end{aligned}
$$

N ímpar para simplificar e m=(N-1)/2.

$$
X_t = a_0 + \sum_{t=1}^{m}[a_j cos(2\pi\omega_j t) + b_j sen(2\pi\omega_j t)] \tag{4}
$$

$$
\begin{aligned}
a_0 &= \bar{X} && (5)\\
a_j &= \frac{2}{N}\sum_{t=1}^{N} x_t cos(2\pi\omega_j t) = \frac{2}{\sqrt{N}} d_c(\omega_j) && (6)\\
b_j &= \frac{2}{N}\sum_{t=1}^{N} x_t sen(2\pi\omega_j t) = \frac{2}{\sqrt{N}} d_s(\omega_j) && (7)
\end{aligned}
$$

Decomposição

$$
\begin{aligned}
x_t - \bar{x} &= \frac{2}{\sqrt{N}}\sum_{j=1}^{m}[d_c(\omega_j)cos(2\pi\omega_j t) + d_s(\omega_j)sen(2\pi\omega_j t)]\\
\sum_{t=1}^{N}(x_t - \bar{x})^2 &= 2\sum_{j=1}^{m}[d_c^2(\omega_j) + d_s^2(\omega_j)] = 2I(\omega_j)
\end{aligned}
$$

Source	df	SS	MS
ω_1	2	$2I(\omega_1)$	$I(\omega_1)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
ω_m	2	$2I(\omega_m)$	$I(\omega_m)$
Total	n-1	$\sum_{t=1}^{N}(x_t - \bar{x})^2$	

Para $x_t iid(0, \sigma^2)$, assintoticamente $d_c(\omega_{j:n}) \sim AN(0, \sigma/2)$ e $\frac{2I(\omega_j)}{\sigma^2} \sim \chi^2_2$. (SS2016 Teo C.7)

Referências

Time series analysis

- Morettin e Toloi
- Shumway and Stoffer
- Wei