Statistics 5350/7110
Forecasting

Lecture 22
Periodicity in Time Series

Professor Robert Stine

Preliminaries

- Questions?

- Assignments
  - Assignment 5 due on Thursday

- Quick review
  - SARIMA models:
    Multiplicative ARIMA models for seasonal data
  - Combining regression with SARIMA
    Calendar examples

Today's Topics

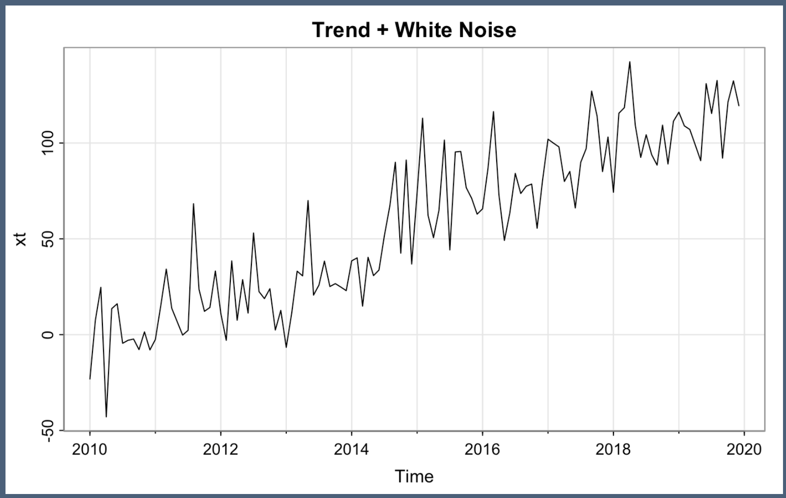
Text, §6.1-6.2

- Seasonal models with deterministic periodicity
  - We saw this in the housing example
- Fitting periodic functions
  - Forming predictors from sines and cosines
  - Regression estimates
- Periodogram
  - Efficient computation of a regression with n predictors
  - Fourier frequencies: Uncorrelated explanatory variables
  - Fast Fourier transform

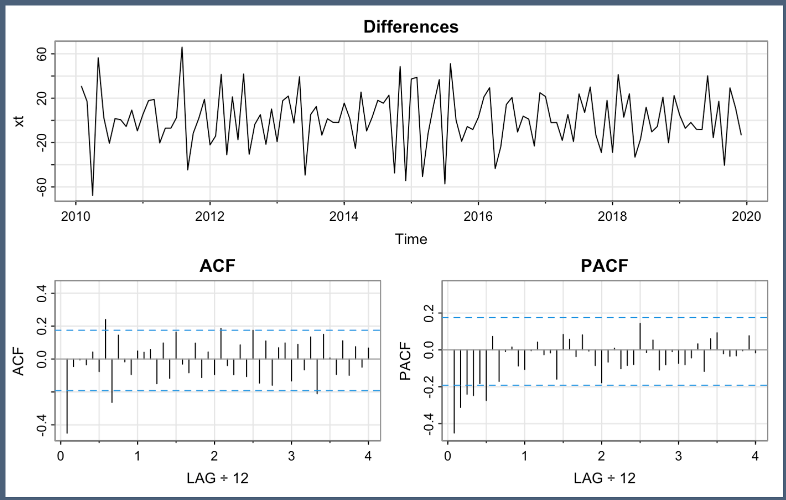
Differencing and Deterministic Trends

Differencing a Linear Trend

- Model
  - Suppose process is a linear trend plus white noise
$$X_t = \alpha + \beta t + w_t$$
  - You difference it, as if a random walk
$$(1 - B) X_t = \beta + (1 - B) w_t$$



- ACF and PACF of differences
  - ACF appears to "cut off"
  - PACF drops off gradually
  - Suggests an IMA(1,1) model



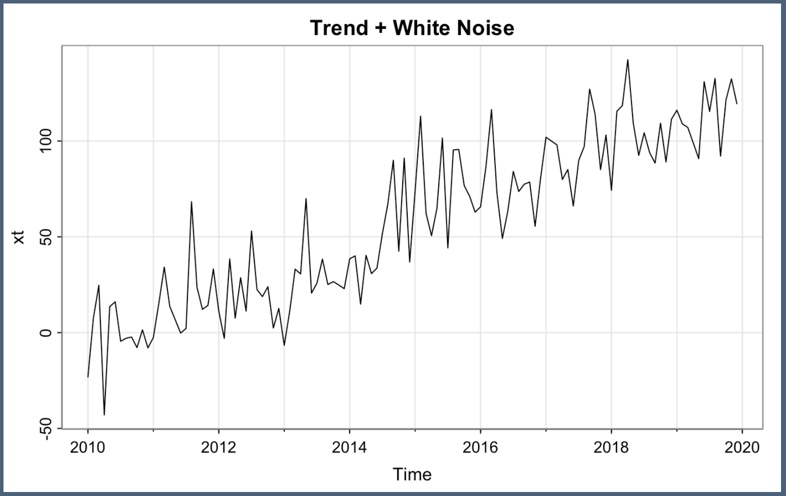
Differencing a Linear Trend

- Model
  - Linear trend plus white noise

$$X_t = \alpha + \beta t + w_t$$

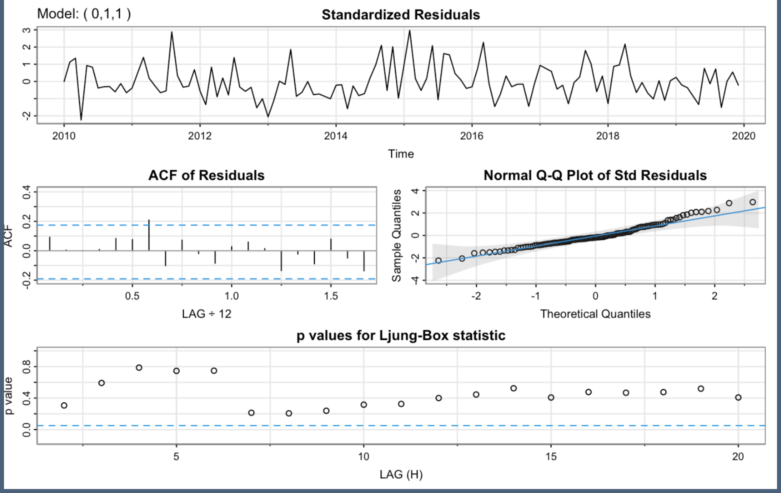
  - Differencing yields

$$(1 - B)\,X_t = \beta + (1 - B)\,w_t$$



- Estimation results
  - IMA(1,1) model has $\hat{\theta}_1 = -1$
  - Fine diagnostics, though not invertible
  - Indicates X_t is non-stochastic trend plus white noise —> Fit a regression!

	Estimate	SE	t.value	p.value
ma1	-1.0000	0.0333	-30.0417	0
constant	1.1079	0.0491	22.5815	0



Differencing a Seasonal Trend

- Model
  - Process is a seasonal trend plus AR(1) noise

$$X_t = \alpha + S_t + Z_t, \quad (1 - \phi B) Z_t = w_t$$

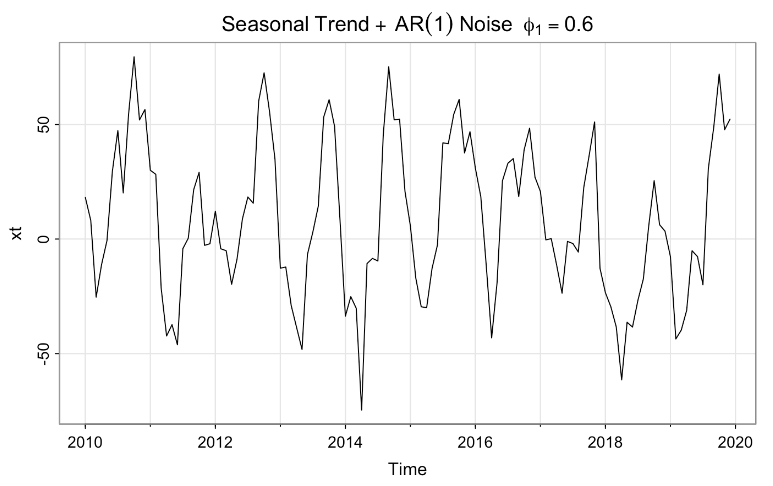
where S_t is a deterministic seasonal pattern

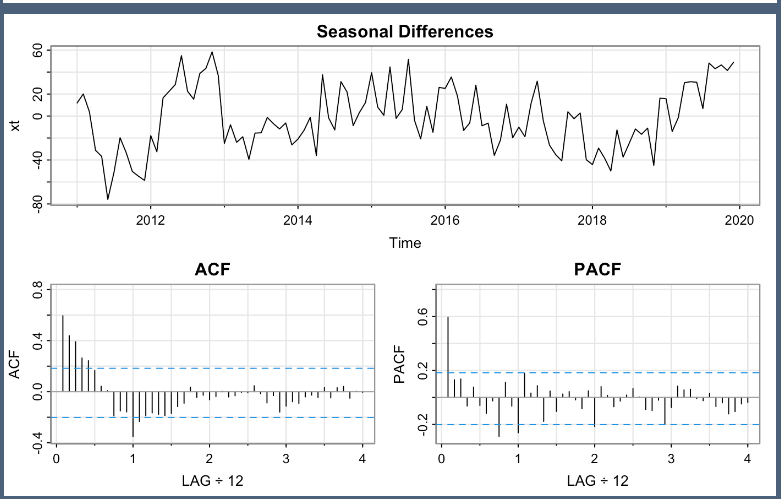
$$S_t - S_{t-12} = 0$$

  - You seasonally difference it, as if stochastic trend

$$(1 - B^{12})(1 - \phi B) X_t = (1 - B^{12}) w_t$$

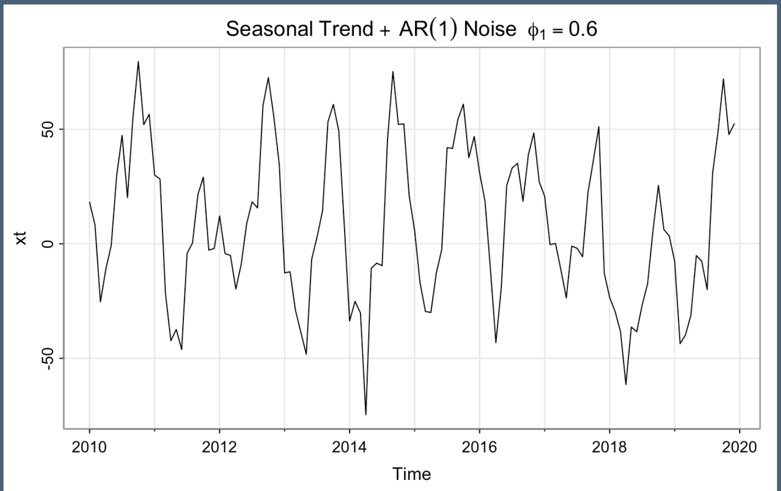
- ACF and PACF of process
  - ACF drops off as in AR(1) or AR(2), seasonal?
  - PACF has spike then drops off: cut off or decay?
  - Try to fit an ARMA(1,1) seasonally





Differencing a Seasonal Trend

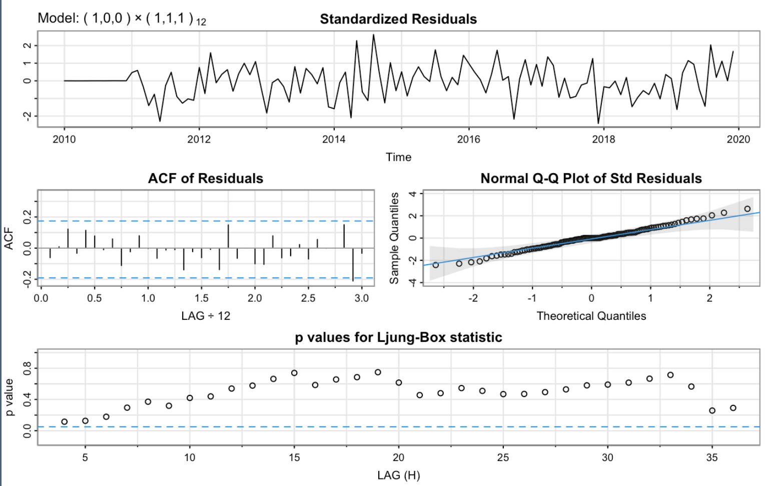
- Model
  - Seasonal trend plus AR(1) noise
    $$X_t = \alpha + S_t + z_t, \quad (1-\phi B) z_t = w_t$$
    with $S_t - S_{t-12} = 0$
  - Seasonally differenced
    $$(1-B^{12})(1-\phi B)X_t = (1-B^{12})w_t$$



- Estimation results
  - Fit initial model as SARIMA(1,0,0)(1,1,1)$_{12}$
  - Model has $\hat{\Theta}_1 = -1$
  - Fine diagnostics, though not invertible

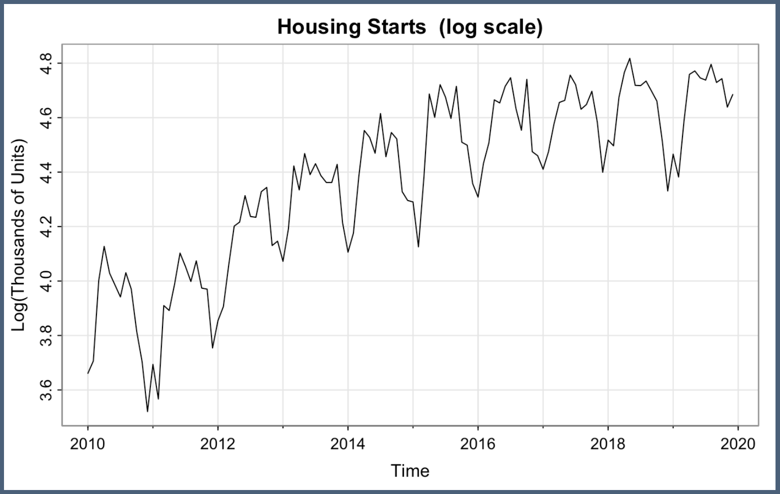
```
ar1         0.5856 0.0791  7.4066  0.0000
sar1        0.0349 0.1113  0.3138  0.7543
sma1       -1.0000 0.2046 -4.8886  0.0000
constant   -0.1304 0.1063 -1.2272  0.2225

sigma^2 estimated as 281.5911 on 104 degrees of freedom
```

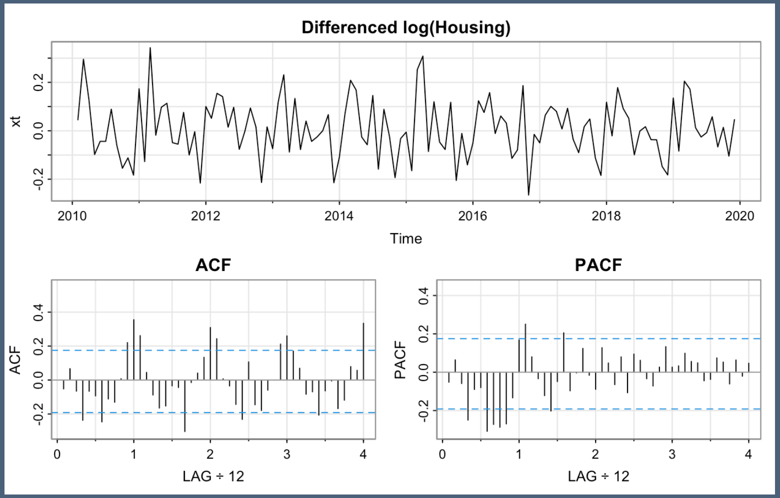


Back to the Housing Example

- Monthly housing starts
  - 10 years, 2010-2019
  - log scale
  - Trend appears stochastic, so difference
  - Stochastic seasonality too?



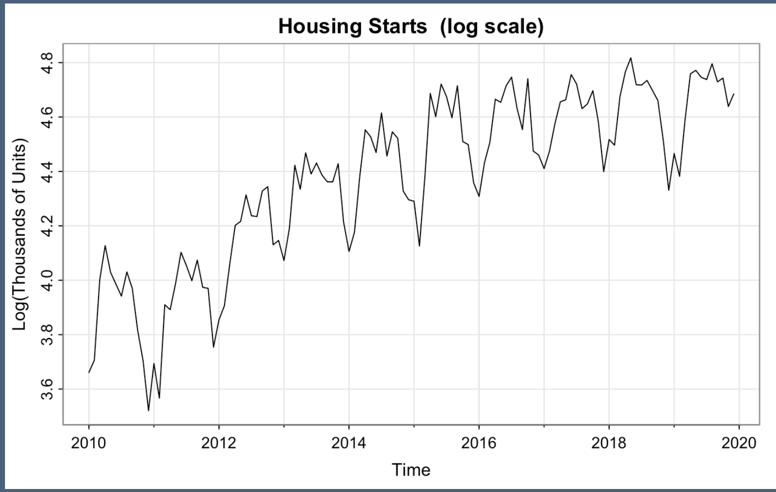
- Identification from ACF/PACF
  - ACF has persistent seasonal correlation
    -> seasonal difference
  - PACF drops off
    -> MA sort of dependence?
  - Start with “familiar” model SARIMA(0,1,1)(0,1,1)$_{12}$

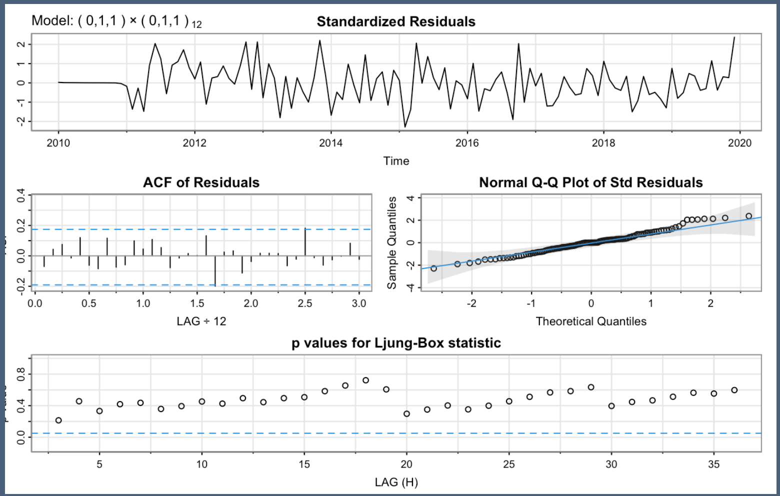


Back to the Housing Example

- Estimation results
  - Consistent with deterministic seasonal trend
  - Residual diagnostics look okay

```
      Estimate     SE t.value p.value
ma1    -0.6387 0.0669 -9.5504       0
sma1   -1.0000 0.1801 -5.5523       0
sigma^2 estimated as 0.005661648 on 105
```





- What to do?
  - Software is fine with the fitted model (though model is not invertible)

    -> Leave it alone (forecasts work fine)
  - Revise model to have a deterministic seasonal trend (periodic) fit with regression.

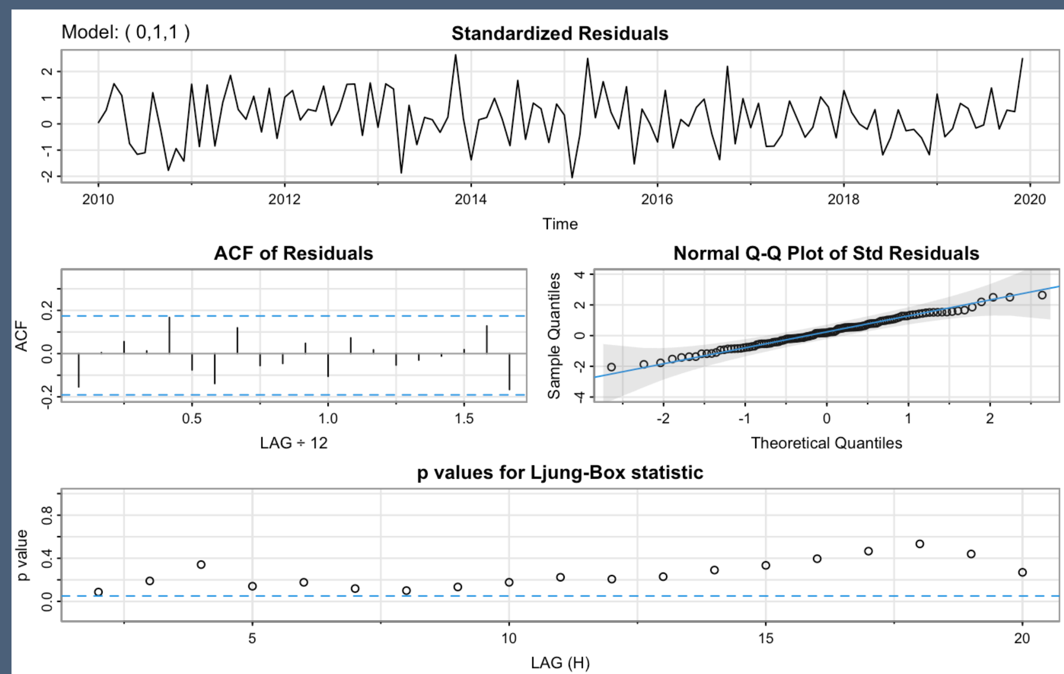
Deterministic Seasonal Trend

- Regression features
  - Use monthly dummy variables to create seasonal trend for which $(1 - B^{12})\, S_t = 0$
  - Retain differencing for overall trend.
  - Overall fit and forecasts comparable to SARIMA model

Close to MA1 estimate in SARIMA mode

	Estimate	SE	t.value	p.value
ma1	-0.5710	0.0628	-9.0918	0.0000
m_Feb	0.0005	0.0267	0.0173	0.9862
m_Mar	0.1977	0.0281	7.0374	0.0000
m_Apr	0.3043	0.0291	10.4401	0.0000
m_May	0.3066	0.0299	10.2601	0.0000
m_Jun	0.3174	0.0303	10.4603	0.0000
m_Jul	0.3059	0.0305	10.0177	0.0000
m_Aug	0.2608	0.0305	8.5603	0.0000
m_Sep	0.2662	0.0301	8.8358	0.0000
m_Oct	0.2333	0.0295	7.9064	0.0000
m_Nov	0.1163	0.0286	4.0671	0.0001
m_Dec	-0.0017	0.0274	-0.0629	0.9499

sigma^2 estimated as 0.005440313 on 107



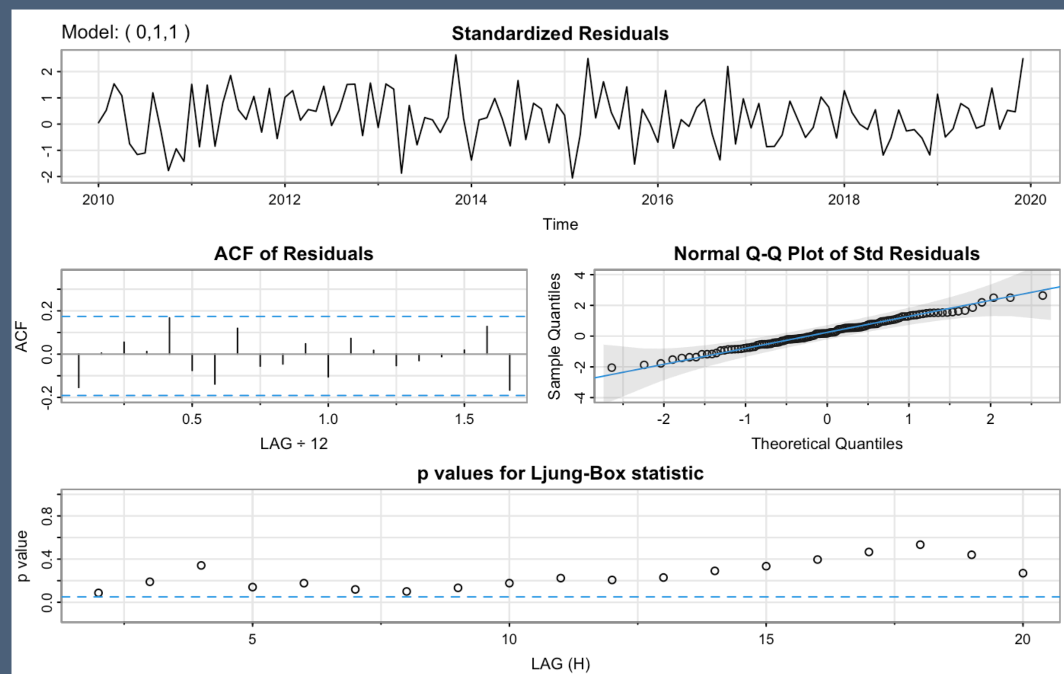
Deterministic Seasonal Trend

- Different approach to regression features
  - Create features from sines and cosines, so that $(1 - B^{12})\, S_t = 0$

    Cosine/Sine with period 12, period 6, period 4, period 3, period 2.4, and period 2
  - Use differencing for overall trend. Fit and forecasts match the prior regression

		Estimate	SE	t.value	p.value
	ma1	-0.5710	0.0628	-9.0924	0.0000
period 12	C1	-0.1524	0.0107	-14.2551	0.0000
	S1	-0.0484	0.0109	-4.4481	0.0000
period 6	C2	-0.0305	0.0083	-3.6802	0.0004
	S2	-0.0568	0.0083	-6.8162	0.0000
period 4	C3	0.0020	0.0078	0.2631	0.7930
	S3	-0.0078	0.0078	-1.0094	0.3151
period 3	C4	0.0026	0.0076	0.3451	0.7307
	S4	0.0230	0.0076	3.0344	0.0030
period 2.4	C5	-0.0092	0.0075	-1.2172	0.2262
	S5	0.0063	0.0075	0.8401	0.4027
period 2	C6	0.0065	0.0053	1.2292	0.2217

sigma^2 estimated as 0.00544031 on 107



Periodic Predictors using Sines and Cosines

Periodic Behavior

- Widespread phenomena are periodic
  - Astronomical data
    variable star magnitude, quasars
  - Climate
    annual temperature, ozone levels
  - Medical data
    pulse, EKG
  - Business cycles are somewhat periodic

- Questions
  - Does a time series show periodic variation?
  - At what frequencies?
  - Should white noise have periodic features?

Sinusoidal Functions

- Sinusoid model
  - Define

$$X_t = R\,\cos(2\pi\,\nu\,t + \varphi) + w_t$$

  - ν is the frequency (reciprocal of the period)
  - R is the amplitude of the cosine wave.
  - φ is the phase of the cosine (does not always start at 1 when t = 0)
- Least squares estimates
  - Estimate v, R, and φ with regression
  - Key from trigonometry

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

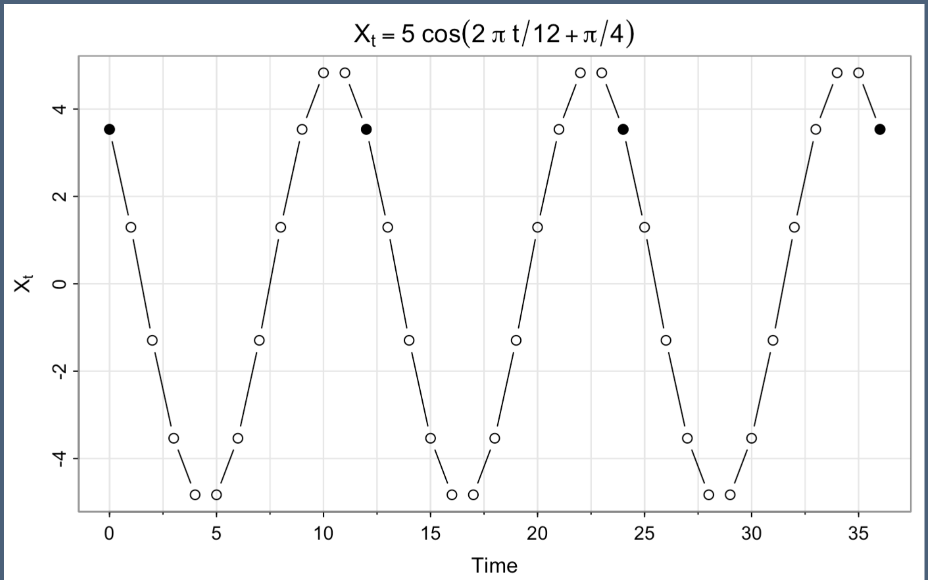
  - Hence

$$X_t = A\,\cos(2\pi\nu t) + B\,\sin(2\pi\nu t) + w_t$$

where

$$A = R\,\cos(\varphi),\ B = -R\,\sin(\varphi)$$

Text writes ω rather than v for the frequency. I just cannot do that!



Random Phase Model

Not covered in lecture

- Sinusoid as stationary process
  - Treat phase as a random v variable

$$X_t = \underbrace{(R\ \cos\varphi)}_{A}\cos(2\pi\nu\, t) + \underbrace{(R\ \sin\varphi)}_{B}\sin(2\pi\nu\, t)$$

  - Suppose phase is uniformly distributed on -π to π. Then the means are zero:

$$E\,A = R\,E(\cos\varphi) = 0, \quad E\,B = R\,E(\sin\varphi) = 0$$

  - The variance doesn't depend on *ν*,

$$\text{Var}(A) = E\,A^2 = R^2\,E(\cos^2\varphi) = R^2\,\pi = \sigma^2, \quad \text{Var}(B) = \sigma^2$$

  - And the cosine and sine terms are uncorrelated

See Appendix C.5

$$\text{Cov}(A, B) = R^2\,E\left(\cos(\varphi)\ \sin(\varphi)\right) = R^2/2\,\sin(2\varphi) = 0$$

- Covariances
  - Covariance depends only on the time difference (2 π ν = ω)

$$\text{Cov}(X_t, X_s) = E\left[\left(A\cos(\omega\, t) + B\sin(\omega\, t)\right)\,\left(A\cos(\omega\, s) + B\sin(\omega\, s)\right)\right]$$

$$= \sigma^2\left(\cos(\omega\, t)\cos(\omega\, s) + \sin(\omega\, t)\sin(\omega\, s)\right) = \sigma^2\cos(\omega(t-s))$$

Equation 6.3

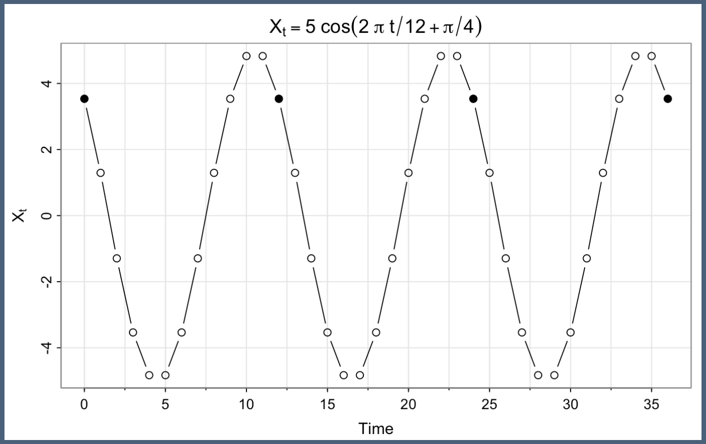
What's the frequency?

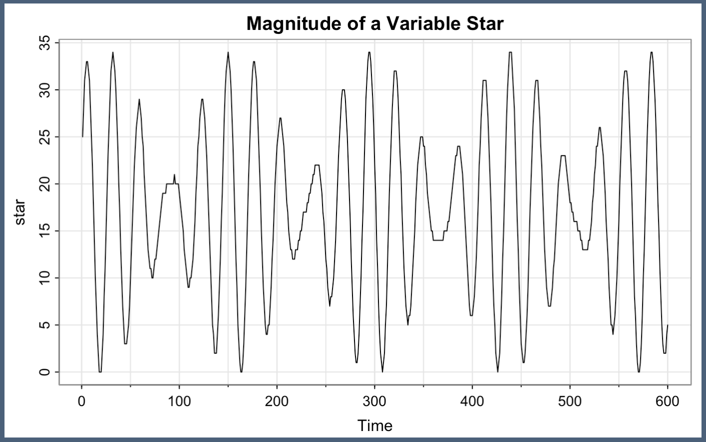
- Frequency doesn't come from regression
  - Simple estimator
  - Divide the number of peaks by the length of the time series

$$\hat{\nu} = 3/36 = 1/12$$

- Prominent frequencies are easy to estimate
  - Standard error on the order of n (rather than √n)

- Example: Magnitude of variable star
  - Daily data from Whittaker and Robinson (1923)
  - Main frequency is **21**/600 ≈ 0.035, about 28.6 day period
  - Clearly not a single sinusoid:
    What other frequencies are present?





time series included in astsa

Estimating a Periodic Function

- Least squares estimates
  - Two coefficients for each frequency

$$X_t = R\cos(2\pi\nu t + \varphi) = A\ \cos(2\pi\nu t) + B\ \sin(2\pi\nu t)$$

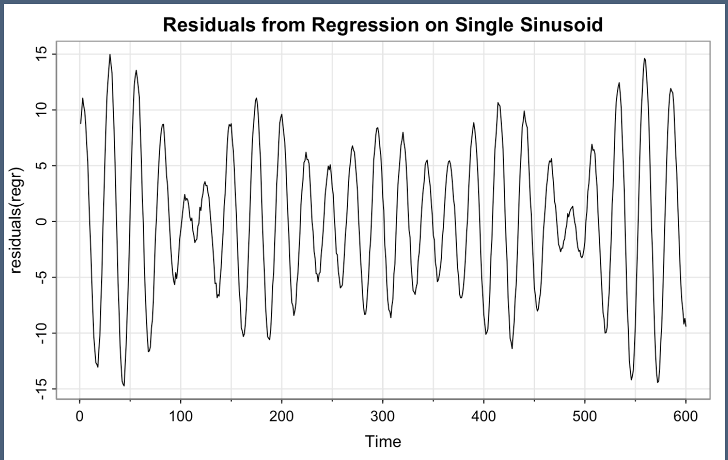
  where

$$A = R\ \cos(\varphi),\ B = -\ R\ \sin(\varphi)$$

  - Converting from A, B back to R and φ is a little harder
    - Since $\cos^2 + \sin^2 = 1$, $R^2 = A^2 + B^2$
    - $\tan(2\pi\varphi) = -B/A$, solved in R as φ = atan(-B/A)

Trig appendix C.5, p 234

- Variable star example
  - Data isn't a single sinusoid
  - Residuals from regression of star magnitudes on cos/sine with frequency 21/600 is another sine wave.
  - Counting peaks suggests second frequency 25/600



Estimating a Periodic Function

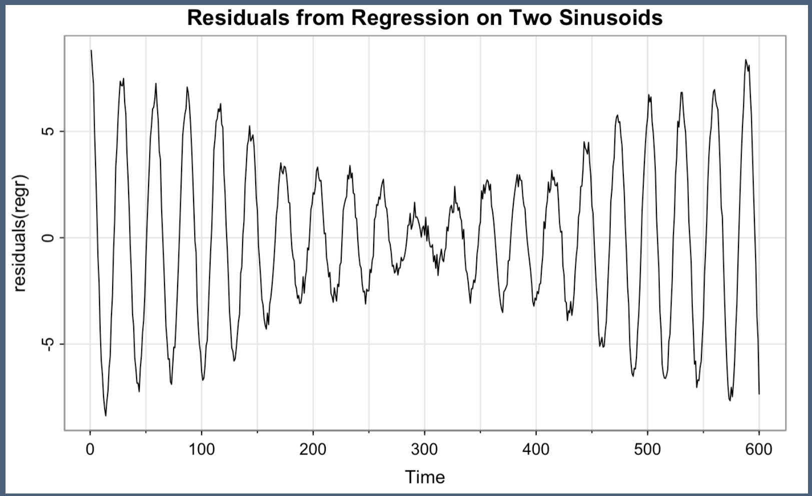
- Least squares estimates
  - Sum of sinusoids

$$X_t = \sum_j A_j \cos(2\pi \nu_j t) + B_j \sin(2\pi \nu_j t)$$

  - Magnitude of each given by $R_j^2 = A_j^2 + B_j^2$.

- Variable star example
  - Try with frequencies set to ν_1=21/600, ν_2=25/600
  - Still not quite right: sinusoid remains

- Integer periods?
  - Periods of these are 600/21≈28.6 and 600/25=24.
  - Round these to integer periods



See Bloomfield, Chapter 3 for further discussion of finding these frequencies

Estimating a Periodic Function

- Least squares estimates
  - Sum of sinusoids

$$X_t = \sum_j A_j \cos(2\pi\nu_j t) + B_j \sin(2\pi\nu_j t)$$

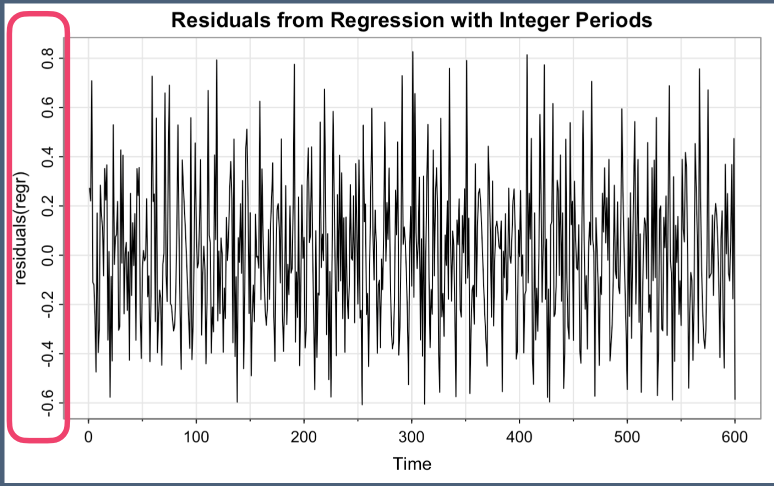
  - Magnitude of each given by $R_j^2 = A_j^2 + B_j^2$.

$$\begin{aligned}\hat{R}_1^2 &= \hat{A}_1^2 + \hat{B}_1^2 \approx 6.07^2 + 7.99^2 \approx 100.6 \\ \hat{R}_2^2 &= \hat{A}_2^2 + \hat{B}_2^2 \approx 1.83^2 + 6.84^2 \approx 50.2\end{aligned}$$

- Variable star example
  - Try with frequencies set to $\nu_1=1/29$ and $\nu_2=1/24$. Wow!

```
Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)               17.08580    0.01238  1380.4   <2e-16 ***
cos(2 * pi * tt * freq1)   6.06876    0.01753   346.2   <2e-16 ***
sin(2 * pi * tt * freq1)   7.98648    0.01755   455.2   <2e-16 ***
cos(2 * pi * tt * freq2)  -1.83495    0.01753  -104.7   <2e-16 ***
sin(2 * pi * tt * freq2)   6.84309    0.01754   390.1   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3031 on 595 degrees of freedom
Multiple R-squared:  0.9989,    Adjusted R-squared:  0.9989
```



Finding Periodic Components

The periodogram

Hidden Periodicity

- Does a time series contain periodic components?
  - Easy to identify periodic functions in cases like the variable star
    - natural in context
    - obvious in data
  - What about periodicities in other time series?
    - stationary process
    - seasonal data
    - residuals from models that supposedly capture seasonality
- Diagnostic
  - Last application is an important diagnostic
  - Do “deseasonalized” data have periodic components?
  - Do residuals from an SARIMA model have periodic components?
- Broad approach
  - Compute the amount of energy (amplitude) of all the possible sinusoids … the periodogram

Periodogram

- Why estimate just a couple of frequencies
  - If n is an even number, we can write with no error

$$X_t = \sum_{j=0}^{n/2} a_j \cos(2\pi\, t\, j/n) + b_j \sin(2\pi\, t\, j/n)$$

$$a_k = (2/n) \sum x_t \cos(2\pi t k/n)$$
$$b_k = (2/n) \sum x_t \sin(2\pi t k/n)$$

Equation 6.8

  - End cases are special since sin(0) = sin(k π) = 0 so $b_0 = b_{n/2} = 0$ (cos(0) = 1, cos(k π)=±1)
  - Scaled periodogram

$$P(k/n) = a_k^2 + b_k^2 = R_k^2$$

Definition 6.3

- Fourier frequencies
  - Sine/cosines with frequencies $v_j = j/n$, for j = 0, 1, …, n/2
  - Highest frequency is $v_{n/2} = 1/2$, a.k.a. the folding frequency or Nyquist frequency

Supplemental slides discuss aliasing

- Asides
  - Linear transformation (a rotation) of time series (X_1, …, X_n) into coefficients (a_0, a_1, b_1, …, $a_{n/2}$)
  - If n is highly composite (e.g. $n = 2^m$), then can compute in O(n log n) using FFT

Properties of Periodogram

- Suppose time series is white noise
  - Coefficients

$$a_k = (2/n) \sum w_t \cos(2\pi tk/n) \quad \text{and} \quad b_k = (2/n) \sum w_t \sin(2\pi tk/n)$$

  - Expected values are easy

$$E\, a_k = E\, b_k = 0$$

  - Variances and covariances are also easy – if you remember basic trigonometry

Appendix
Property C.3
page 234

$$\text{Var}(a_k) = \text{Var}(b_k) = \frac{4\,\sigma_w^2}{n^2} \underbrace{\sum_{t=1}^{n} \sin^2(2\pi kt/n)}_{n/2} = \frac{2\,\sigma_w^2}{n} \qquad k = 1, 2, \ldots, n/2-1$$

$$\text{Cov}(a_k, a_j) = 0\ (j \neq k) \quad \text{and} \quad \text{Cov}(a_k, b_j) = 0$$

- Periodogram
  - Define as

$$I(k/n) = \frac{n}{4}\left(a_k^2 + b_k^2\right) \approx \sigma_w^2 \frac{N(0,1)^2 + N(0,1)^2}{2} = \sigma_w^2 \frac{\chi_2^2}{2}$$

Proportional to an exponential random variable

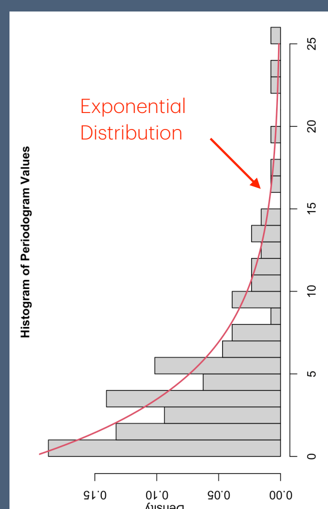
Properties of Periodogram

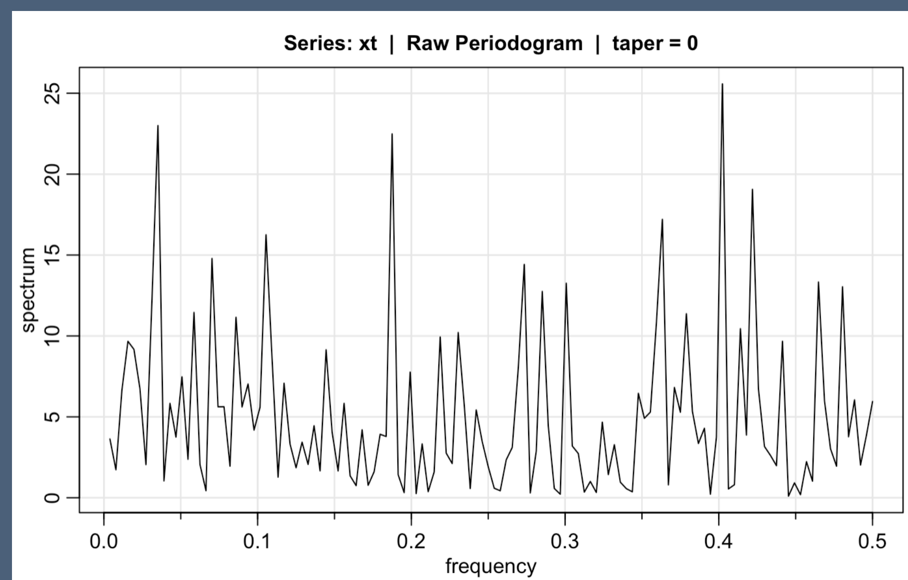
- Periodogram of white noise
  - Proportional to exponential r.v.

$$I(k/n) = \frac{n}{4}\left(a_k^2 + b_k^2\right) \sim \sigma_w^2 \frac{\chi_2^2}{2}$$

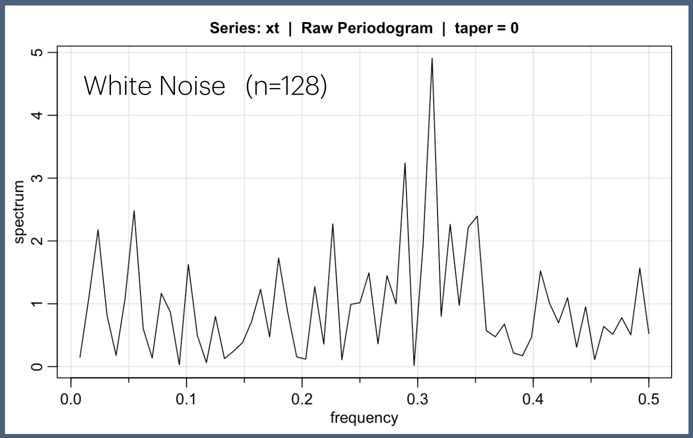
  - Approximately independent values sampled from skewed distribution: Very noisy appearance

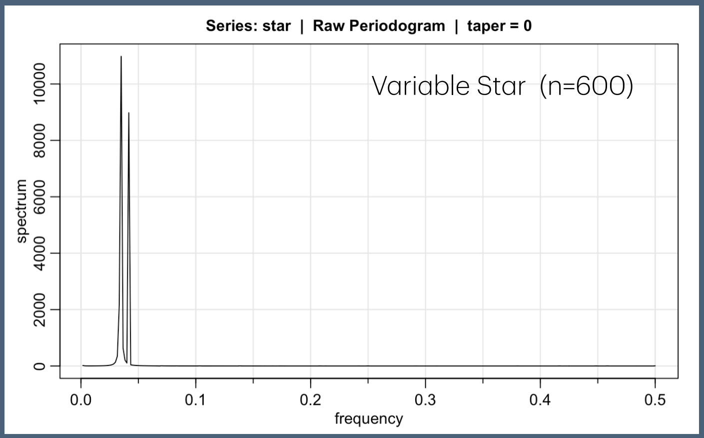
- Example
  - n = 256
  - White noise, $\sigma_w^2 = 5$

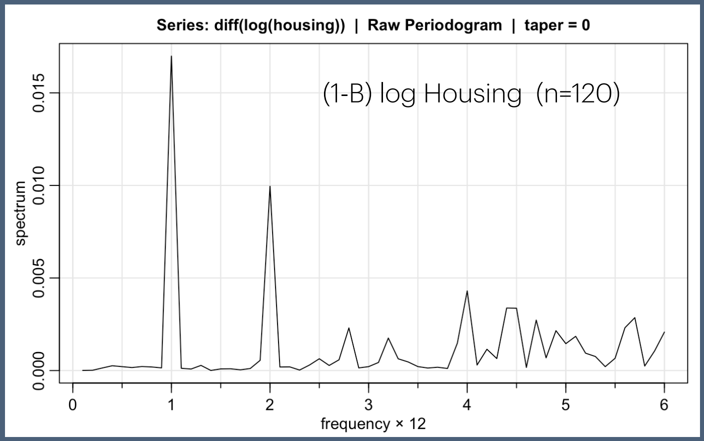


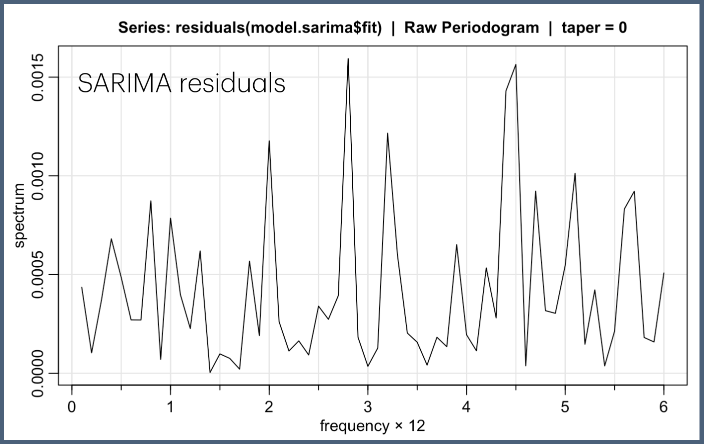


Examples of Periodogram









What's next?

- Sampling properties of the periodogram
  - Are peaks in white noise important or just sampling variation?
  - Key diagnostic:
    - For residuals from a model, what should the periodogram show?
    - For deseasonalized data from FRED, what should the periodogram show?

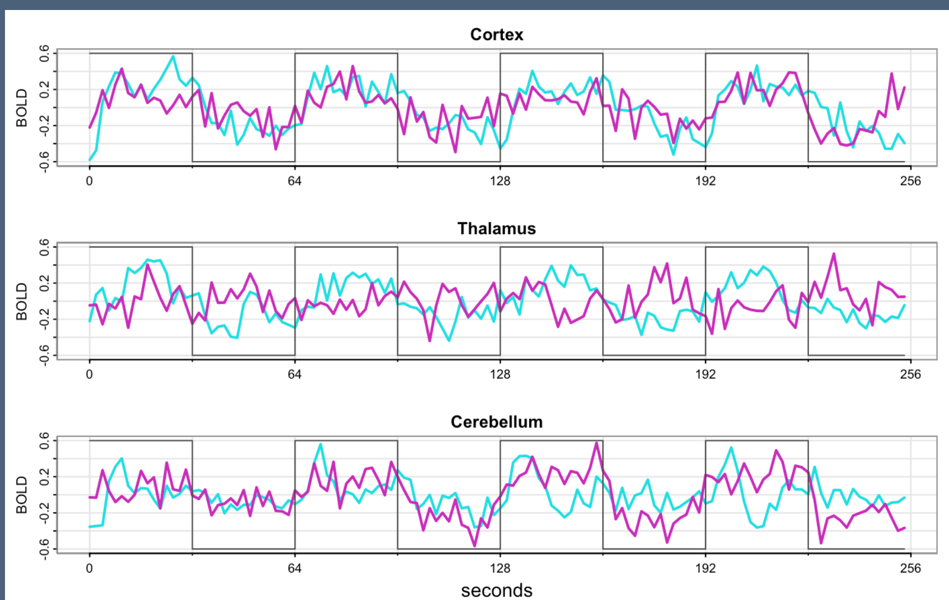
- Spectral density and covariances
  - Periodic representation of any stationary processes
  - Smoothing the periodogram
  - Connecting time and frequency-domain analyses

Supplements

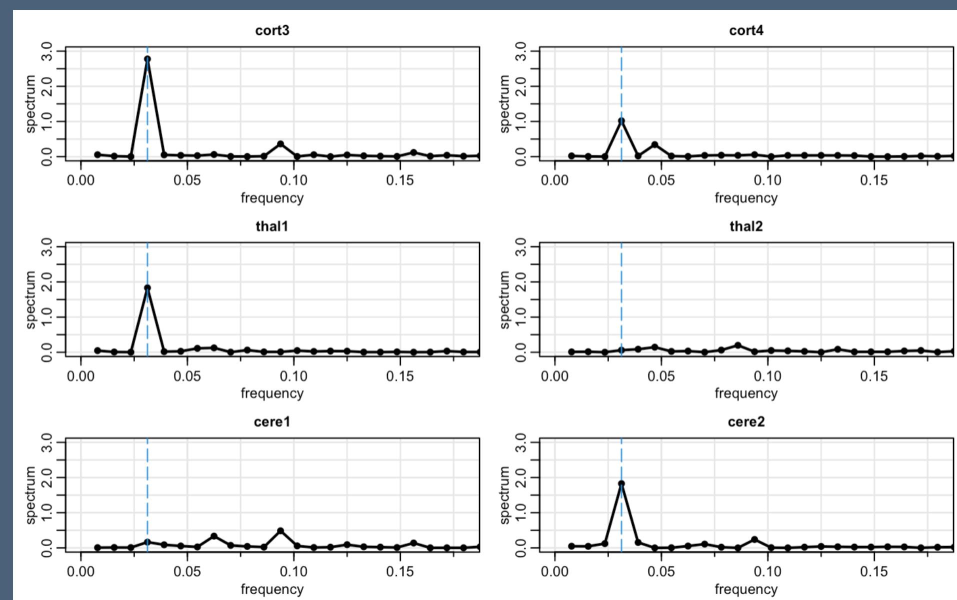
Not covered, but might be useful in your field

Not covered: More Examples

- fMRI
  - Signal indicated by square waves (Example 1.6) with period 64 seconds
  - Caution: Data sampled 1 point every two seconds. Max observable frequency is 1/2δ = 1/4 seconds
  - Expect signal response at period 32 measurements
  - Signal is evident in 4 output series, but are these significant?



BOLD = blood oxygen level dependent

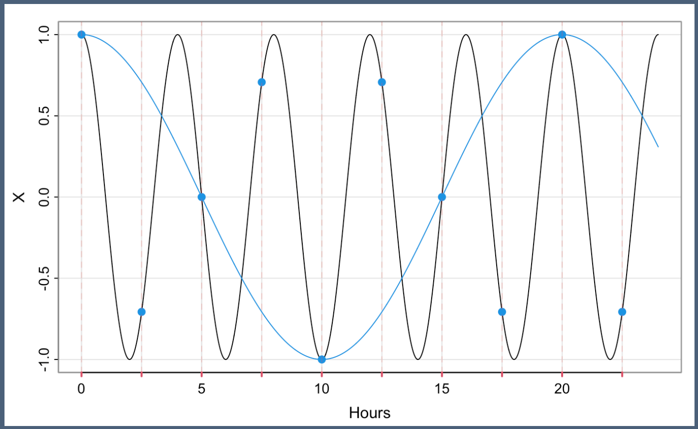


Periodograms shown on zoomed scale

Aliasing and Nyquist Frequency

- Behavior under sampling
  - We observe a discrete time series, equally spaced observations of a continuous process
  - Highest frequency we can estimate is 1/2 since we need at least 2 observations within a cycle
  - 1/2 is known as the Nyquist or folding frequency
  - Sinusoids with frequencies higher than 1/2 appear at a lower frequency
- Example
  - Continuous process has period 4 hours, or frequency 1/4
  - Process is sampled every 2.5 hours
  - Discrete time series has cycle at lower frequency 1/20
- Math
  - Sample continuous process at times t = 0, δ, 2δ, 3δ...
  - Continuous process has frequency 1/2δ < f < 1/δ
  - Define frequency f = 1/δ – f′ where 0 < f′ < 1/2δ
  - Oscillation at f aliased to oscillation at f′

≈ Figure 6.1

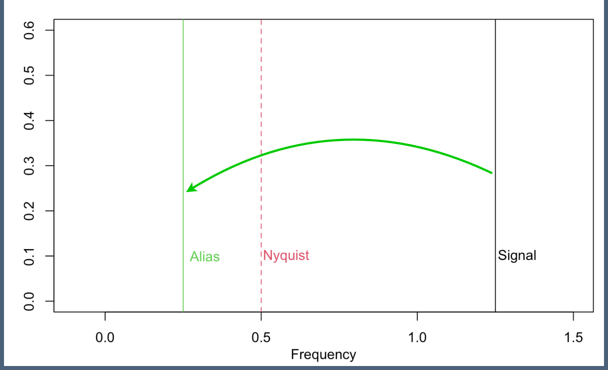


$$
\begin{aligned}
\cos(2\pi\, t\, f) &= \cos\left(2\pi t \left(1/\delta - f'\right)\right) \\
&= \cos\left(2\pi (k\delta) \left(1/\delta - f'\right)\right) \\
&= \cos\left(2\pi (k - t\, f')\right) \\
&= \cos(2\pi k)\cos(2\pi t\, f') + \sin(2\pi k)\sin(2\pi t\, f') \\
&= \cos(2\pi\, t\, f')
\end{aligned}
$$

Aliasing Example

- Variable star
  - Brightness of a variable star is a continuous process
  - Magnitude has period of 0.8 days, frequency f = 1.25 cycles per day
- Discrete time series
  - You observe the magnitude each night at 12 am
  - Sampling interval $\delta = 1$
  - Nyquist (folding) frequency is 1/2
- Alias
  - Aliasing will happen since 1/2 < f
  - Resulting discrete time series has sinusoid at frequency 0.25

$$\begin{aligned}\cos(2\pi\, t\, 1.25) &= \cos\left(2\pi t(1+1/4)\right)\\ &= \cos(2\pi t)\cos(2\pi t/4) - \sin(2\pi t)\sin(2\pi t/4)\\ &= \cos(2\pi\, t/4)\end{aligned}$$



- Remedy
  - Sample at a higher rate, or apply a low-pass filter prior to sampling