Conclusions and Extensions of this Course

Jared Fisher

Lecture 12

Announcements/Schedule

- Today: review/conclusions/extensions of this semester
- Friday, April 30: Homework 6 is due by 11:59pm PDT. No formal lab, but OH for Q&A on HW6, project, etc.
- Monday, May 10: Final Project Report and Forecasts due. Please review specifications on assignment doc! For example, there's a six page limit. For those interested, I'll share the rubric in the coming days so you'll know exactly what we're looking for.

Grading policies

- New Grading Policy: Homework drop may be used on a project checkpoint instead.
- New Grading Policy: Contact me if you are concerned with failing/not passing and we'll work out a late homework/checkpoint submission option (maximum score of C- or P).

Midterm 2 Information

- Online was out of 65, written out of 35.
- Online version A: mean 45.93, stdev 10.1
- Online version B: mean 45.52, stdev 11.08
- Written portion: mean 32.51, stdev 5.05

Today's topics

- Big Picture
- Review/synthesize frequency domain
- Extensions and "what's next" in Time Series. (not part of HW6)

Big Picture

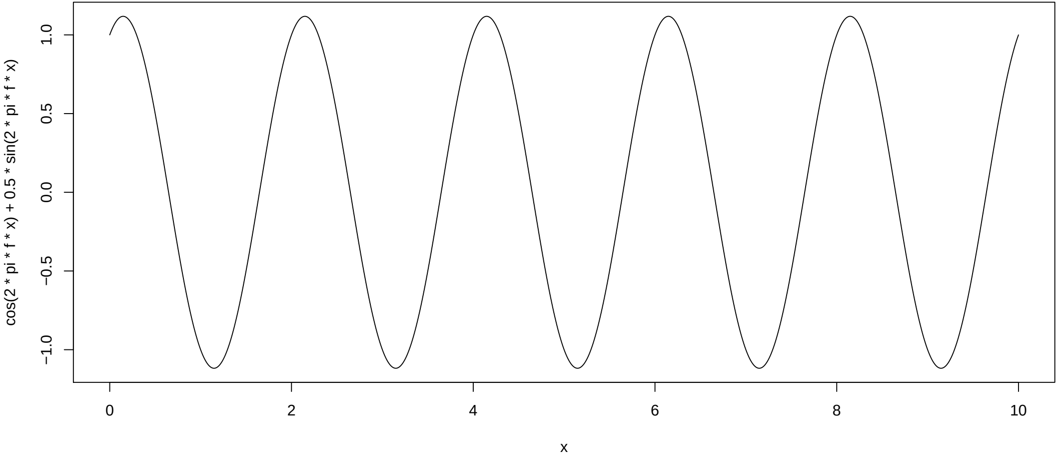
Big Picture

- Our general model:
$$Y_t = signal_t + noise_t$$
- $signal_t$ may be a deterministic function of time (but you can use the time series data to estimate parameters) or a filter, etc.
- $noise_t$ is some stationary process (i.e. some ARMA or SARIMA model for us!)

Sinusoid

$$g(t) = A\cos(2\pi ft) + B\sin(2\pi ft)$$

For $f = 0.5$:



Sinusoids in our Big Picture model

- It may be helpful to include sinusoids in the trend/seasonality!
- As the signal is a function of t, we'll need to find A, B, and f in

$$A\cos(2\pi ft) + B\sin(2\pi ft).$$

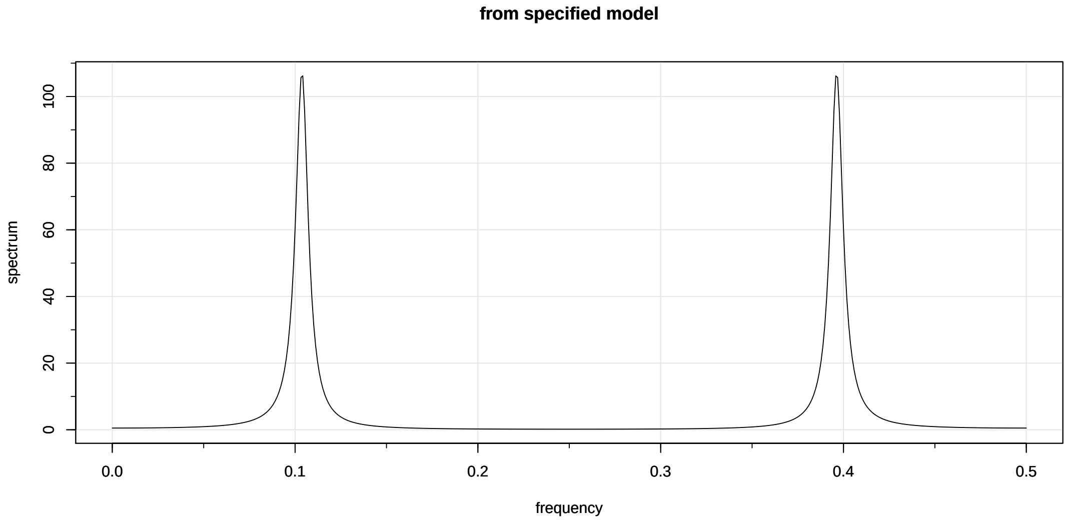
- lm() and other statistical tools are really good at estimating coefficients like A and B, so we still need to determine f.
- The periodogram and spectral density can help us identify the frequency f.

What does the spectral density help with in this class?

- Can help identify important frequencies without the worries of leakage that the periodogram has. Instead, we have to consider how it's being estimated.
- The power transfer function can help us with intuition about how ARMA fits to the noise: Remember the spectral density of ARMA process X_t is equal to the filter's power transfer function times a constant (the spectral density of white noise)

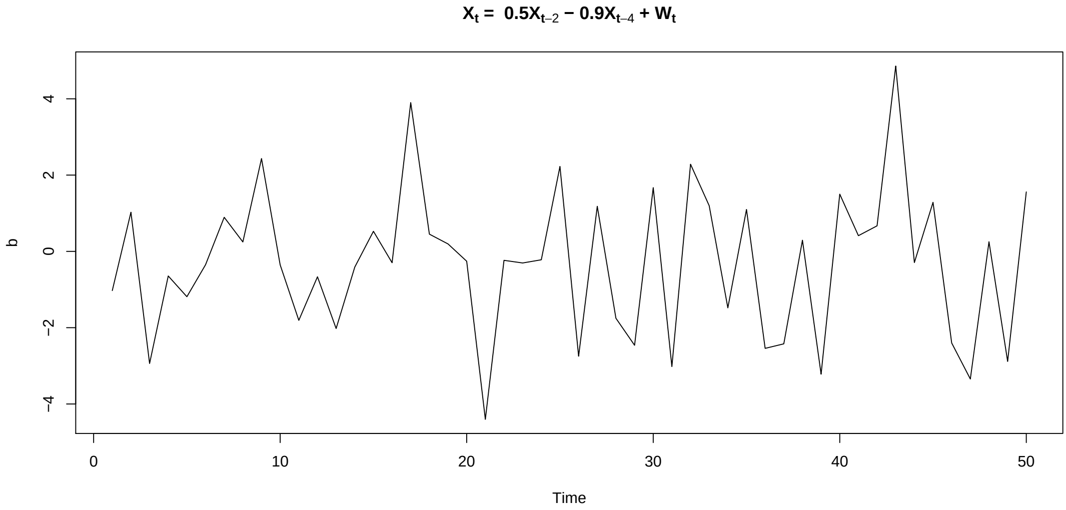
HW6 Guidance: "Oscillate"

- Note that ARMA can look seasonal, e.g. $X_t = 0.5X_{t-2} - 0.9X_{t-4} + W_t$



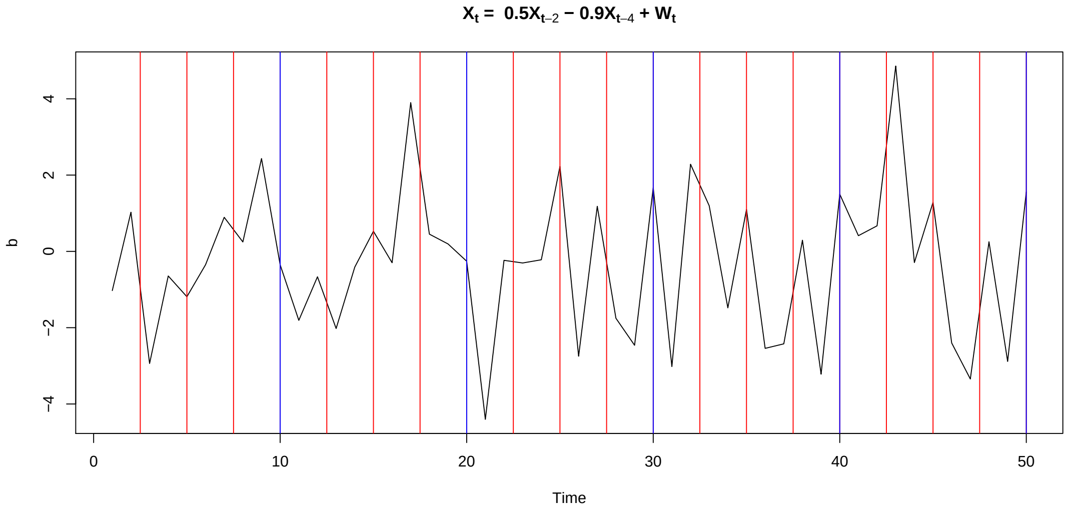
HW6 Guidance: “Oscillate”

- Note how this simulated process “oscillates” with frequencies 0.1 and 0.4/periods of 10 and 2.5.



HW6 Guidance: "Oscillate"

- Note how this simulated process "oscillates" with frequencies 0.1 and 0.4/periods of 10 and 2.5.



Is the frequency domain's information a gray area then?

- If a process is stationary, then the frequency domain is not telling us the deterministic seasonal function s_t, but giving us information about the stationary process (e.g. ARMA)
- If a process is not stationary due to deterministic seasonality, then yes, we use the frequency domain to choose the periods in s_t.
- So, yes, there can be a gray area about which information the frequency domain is giving you, not unlike some cases with the ACF plot. Thus, we should cross examine each of our diagnostic plots together.

To the code!

Let's look at this difference with the Hare dataset

Brief Break

Extensions

Extensions

- The purpose here is to show you what's on the horizon if this class were to continue.
- Some of your questions this semester pertain to models outside of our current deterministic + ARMA setup. We will address some of these briefly here!

Good Question Y'all Have asked:

- How can we use information from another variable ("covariate")?
- For example, if we have a time series $\{Y_t\}$ and $\{X_t\}$, can we have

$$Y_t = f(X_t) + noise_t?$$

One Option: Lagged Regression

- We can view this as a filter:

$$Y_t = \sum_{j=-\infty}^{\infty} \beta_j X_{t-j} + W_t$$

- However, the filters we have looked at previous essentially treat the weights a_j as known.
- Here the weights/coefficients β_j need to be estimated.
- See section 4.8 in TSA4e for more info!

Lagged Regression for Forecasting

- If forecasting, we probably don't want to use future X_t's to estimate Y_t:

$$Y_t = \sum_{j=1}^{p} \beta_j X_{t-j} + W_t$$

  and we only look at the first few (p) lags of X_t.

- Note that W_t could still have an ARMA process, and now we're looking at lagged values of both X_t's and Y_t's

Vector Autoregression

- If we're using lagged values of both X_t and Y_t, we may be interested in a joint model:
$$\begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \begin{pmatrix} \phi_1 \\ \beta_1 \end{pmatrix}^T \begin{pmatrix} Y_{t-1} \\ X_{t-1} \end{pmatrix} + \begin{pmatrix} W_t \\ V_t \end{pmatrix}$$
- Of course this model can include multiple lags
$$\begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \begin{pmatrix} \phi_1 \\ \beta_1 \end{pmatrix}^T \begin{pmatrix} Y_{t-1} \\ X_{t-1} \end{pmatrix} + \cdots + \begin{pmatrix} \phi_p \\ \beta_p \end{pmatrix}^T \begin{pmatrix} Y_{t-p} \\ X_{t-p} \end{pmatrix} + \begin{pmatrix} W_t \\ V_t \end{pmatrix}$$
- The new challenge here that we don't have just a noise variable, but a noise vector $\begin{pmatrix} W_t \\ V_t \end{pmatrix}$, which will need a multivariate distribution/process!

Vector Autoregression

- For notation, we usually collapse Y_t and X_t into a single multivariate variable $X_t = \begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix}$, and the same for the coefficients.
- This yields more compact notation for the model:

$$X_t = \beta_1 X_{t-1} + \cdots + \beta_p X_{t-p} + W_t$$

  where everything is now a vector!
- An extension: what if we have multiple covariates, i.e. lots of X's, for predicting Y? Just expand the vectors.

Time-varying

- What if the coefficients change over time?

$$X_t = \beta_{1t} X_{t-1} + \cdots + \beta_{pt} X_{t-p} + W_t$$

- Then we can model the β's too!
- One way is through a state space model

State Space Model

- Disclaimer: these are a much broader class of model than just for time-varying coefficients.
- Also called a dynamic linear model (DLM)
- In addition to the observation equation

$$X_t = \beta_{1t} X_{t-1} + \cdots + \beta_{pt} X_{t-p} + W_t$$

we model the coefficients as one large vector

$$B_t = \xi B_{t-1} + \Omega_t.$$

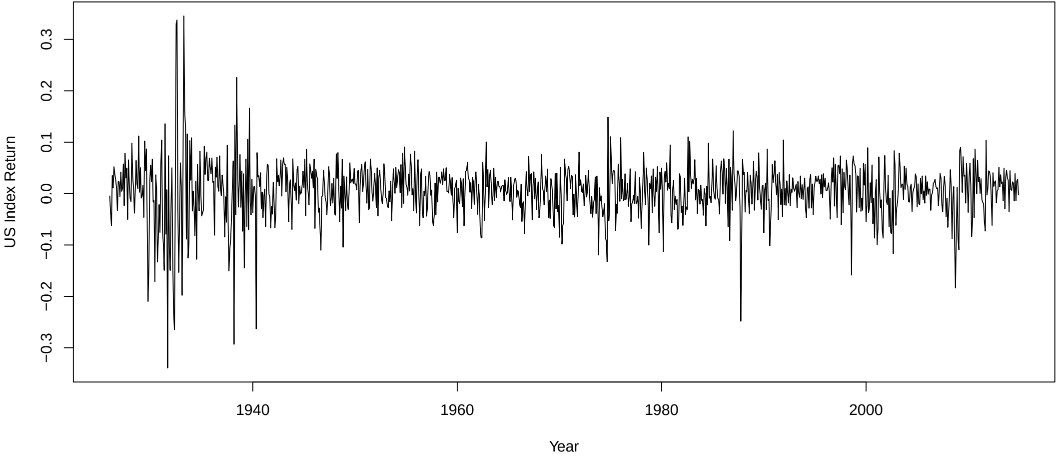
- Yes, this does look like an autoregression on the coefficients! It allows them to vary "gradually" over time.

State Space Model

- More detail is given in Chapter 6 of TSA4e
- The Kalman Filter (Property 6.1) is a particularly famous example of this class of model

Good Questions Y'all Have asked:

- What do we do if the heteroscedasticity has fluctuations in variance, not only increasing?



Stochastic Volatiltiy

- In other words: ARMA assumes constant variance. What if it's not?
- Object of interest is usually the return of a financial asset or growth rate of an object:

$$r_t = \frac{X_t - X_{t-1}}{X_{t-1}}$$

- An example of a stochastic volatility model: if we ignore the signal (i.e. assume constant signal/constant mean), and assume x_t is a hidden volatiltiy process,

$$r_t = \beta exp(x_t/2)\epsilon_t$$
$$x_t = \phi x_{t-1} + w_t$$

where w_t is guassian noise and ϵ_t is iid nois

GARCH

- Object of interest is usually the return of a financial asset or growth rate of an object:
$$r_t = \frac{X_t - X_{t-1}}{X_{t-1}}$$
- Let $\epsilon_t \sim N(0,1)$
- ARCH(1):
$$r_t = \sigma_t \epsilon_t, \qquad \epsilon_t \sim N(0,1)$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2$$
- Note that this assumes there is no signal (mean is zero everywhere!)

GARCH

- GARCH(1,1):

$$r_t = \sigma_t \epsilon_t, \qquad \epsilon_t \sim N(0,1)$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

- GARCH(p,q):

$$r_t = \sigma_t \epsilon_t$$
$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^{p} \alpha_j r_{t-j}^2 + \sum_{k=1}^{q} \beta_k \sigma_{t-k}^2$$

fin

Thanks for a great semester!