Statistics 5350/7110
Forecasting

Lecture 4
Estimating the ACF

Professor Stine

Admin Issues

- Questions?
  - TA office hours posted

- Assignment
  - A1 is due next Thursday
  - We'll discuss in class next Tuesday, so get started before then. Come with your questions ready.

- Quick review
  - Second-order stationarity $E(X_t) = \mu$, $\text{Cov}(X_s, X_t) = \gamma_x(s - t)$
  - What could have happened, not just what did happen...
  - Wold representation, white noise

$$X_t = \sum_{j=0}^{\infty} \psi_j w_{t-j}, \quad \psi_0 = 1$$

  - Examples (found in the revised Lecture_3.Rmd file)

Today's Topics

- Estimates of
  - Mean
  - Autocovariance and autocorrelation

- Multiple time series
  - Leading and lagging series
  - Cross-correlation

- Non-stationary data
  - Spurious correlations
  - Transformations to stationarity

Estimating the Mean

Textbook §2.3

- Estimating the mean of a stationary process
  - Sample mean is the usual choice, but is it the best estimator?

$$\overline{X} = \frac{1}{n}\sum_{t=1}^{n} X_t$$

- Properties of the average
  - Variance of the average depends on the autocovariance of the process

$$n^2 \text{Var}\left(\overline{X}\right) = \text{Var}\left(\sum_{t=1}^{n} X_t\right) = \sum_{s=1}^{n}\sum_{t=1}^{n}\gamma(s-t)$$

  - Summing up the covariances gives

$$\text{Var}\left(\overline{X}\right) = \frac{1}{n^2}\sum_{h=-n}^{n}\left(n - |h|\right)\gamma_x(h)$$

Equation 2.20

  - Generalized least squares provides an alternative estimator which has smaller MSE, but you have to know the autocovariances $\gamma_x(h)$ in order to compute it!

Estimating the Mean

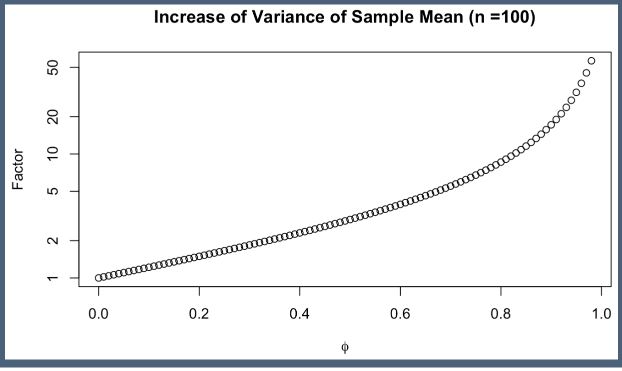
- Effects of covariances on variance of the mean
- White noise
  - Because uncorrelated, the ariance of the mean is familiar, namely

$$\text{Var}(\overline{X}) = \frac{\sigma_w^2}{n}$$

- Autoregression with coefficient ϕ
  - General expression is "messy"

$$\text{Var}(\overline{X}) = \frac{1}{n^2}\sum_{s=1}^{n}\sum_{t=1}^{n}\gamma(s-t) = \frac{\sigma_w^2}{n(1-\phi^2)}\underbrace{\frac{2}{n}\left(\frac{n}{2} + (n-1)\phi + (n-2)\phi^2 + \cdots + \phi^{n-1}\right)}$$

  - Numerical example (Lecture_4.Rmd)
    How much larger is the variance of the mean compared to the usual calculation that would use $\text{Var}(X_t)/n$?



$0 < \phi$

What happens if there's negative autocorrelation???

Estimating Autocorrelations

- Estimating the autocovariances
  - Use the sample mean

$$\hat{\gamma}_x(h) = \frac{1}{n}\sum_{t=1}^{n-h}(X_{t+h}-\overline{X})(X_t-\overline{X}) \qquad h = 0,1,2,\ldots,n-1$$

  - Biased estimator since fewer summands as h increases
- Estimating the autocorrelation
  - Use the estimated autocovariances

$$\text{Corr}(X,Y) = \frac{\text{Cov}(X,Y)}{SD(X)SD(Y)}$$

$$\hat{\rho}_x(h) = \frac{\hat{\gamma}_x(h)}{\hat{\gamma}_x(0)}$$

Equation 2.21

- Properties of the estimated autocorrelation
  - Like the sample mean, properties of $\hat{\rho}_x(h)$ depend on the properties of the true process
  - The estimated $\hat{\rho}_x(h)$ are generally more autocorrelated than the process itself!

$$\text{Cov}(\hat{\rho}_x(r+h),\hat{\rho}_x(r)) \approx \frac{1}{n}\sum_{j=-\infty}^{\infty}\rho_x(j+h)\,\rho_x(j) \quad \text{and} \quad \text{Var}(\hat{\rho}_x(r)) \approx \frac{1}{n}\sum_{j=-\infty}^{\infty}\rho_x(j)^2$$

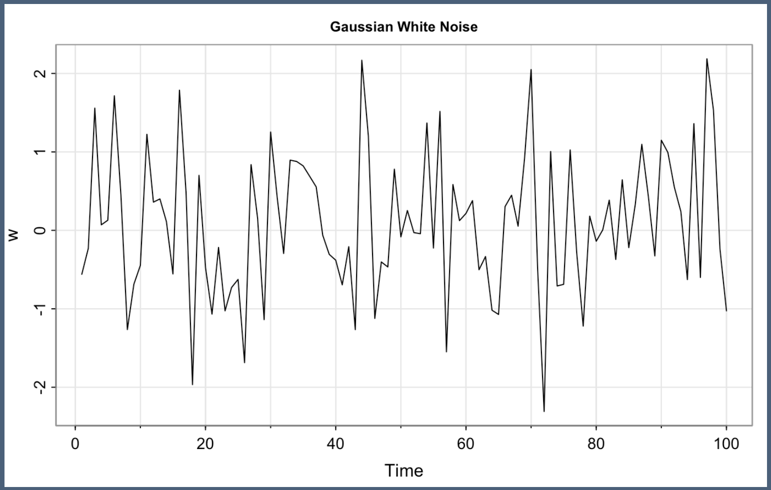
See Property 2.28

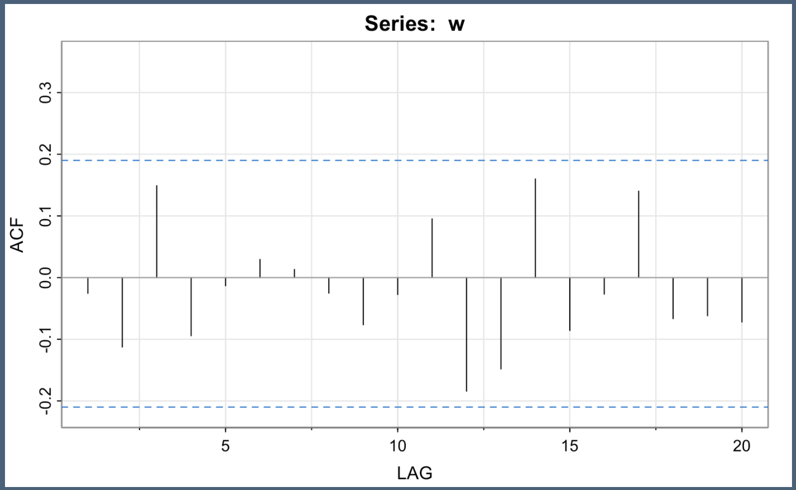
Estimated ACF

- White noise
  - Standard error of estimated autocorrelation if X_t is white noise is

$$\text{std error}\left(\hat{\rho}(h)\right) \approx \frac{1}{\sqrt{n}}$$

Property 2.28



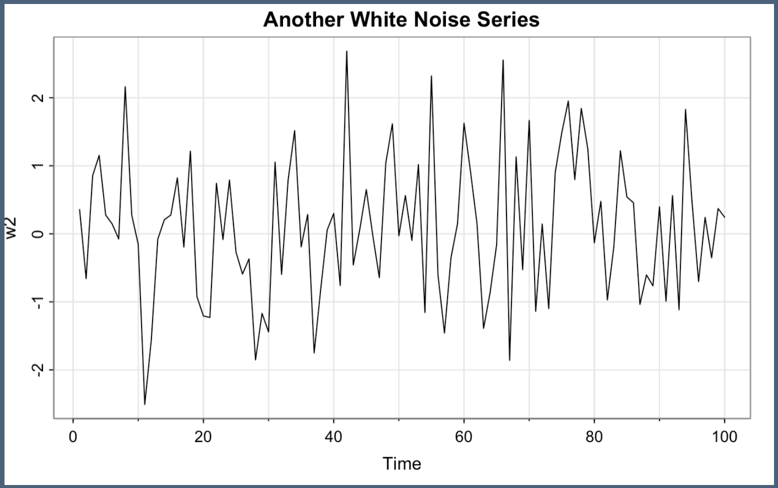


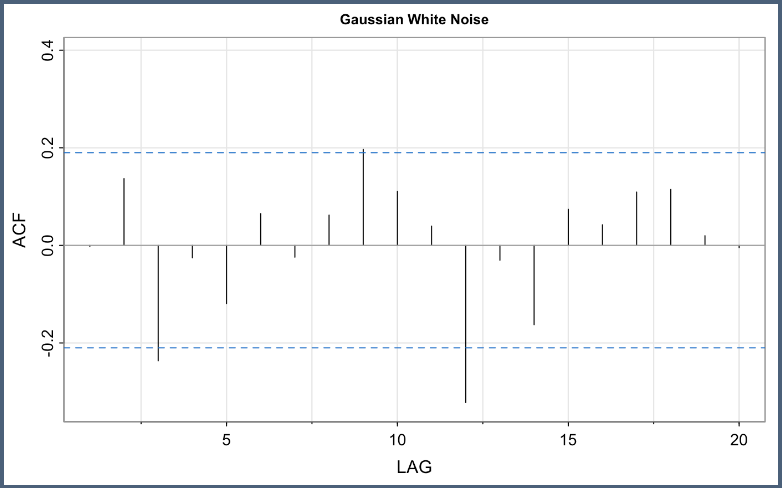
Estimated ACF

- Generate more white noise
  - Standard error of estimated autocorrelation is

$$\text{std error}\left(\hat{\rho}(h)\right) \approx \frac{1}{\sqrt{n}}$$

  - Three of the estimated autocorrelations are outside standard error lines





Why might this happen?

Multiplicity

- What does it mean for a statistical test to be "statistically significant"?
- Example: two-sided test of the mean, H_0: $\mu = 0$ vs H_a $\mu \neq 0$
  - Assume required assumptions hold, along with normality of the data

$$X_i \sim N(\mu, \sigma^2), \quad i = 1, 2, \ldots, n$$

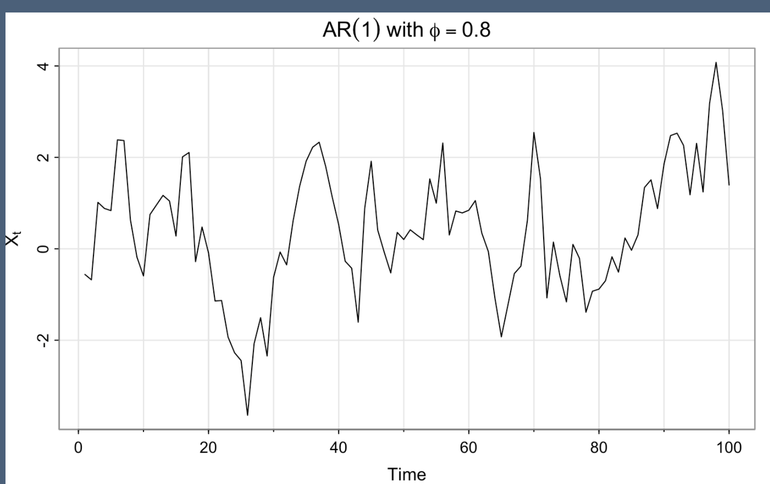
  - Assuming we know Var$(X_i) = \sigma^2$, the most powerful test statistic is $z = \sqrt{n}\,\overline{X}/\sigma \sim N(0,1)$
- Significance
  - Reject H_0 if $|z| > 1.96$
  - If the null hypothesis H_0 holds, then $P(1.96 < |z|) = 0.05$
- Multiple tests
  - Suppose we test H_0 with independent samples from the same population and $\mu = 0$.
  - 10 samples: P(At least one test rejects) = 1 - P(none reject) = $1 - 0.95^{10} \approx 0.40$
  - 25 samples: P(At least one test rejects) = 1 - P(none reject) = $1 - 0.95^{25} \approx 0.72$
  - The more you test, the more likely you will declare a statistically significant result.

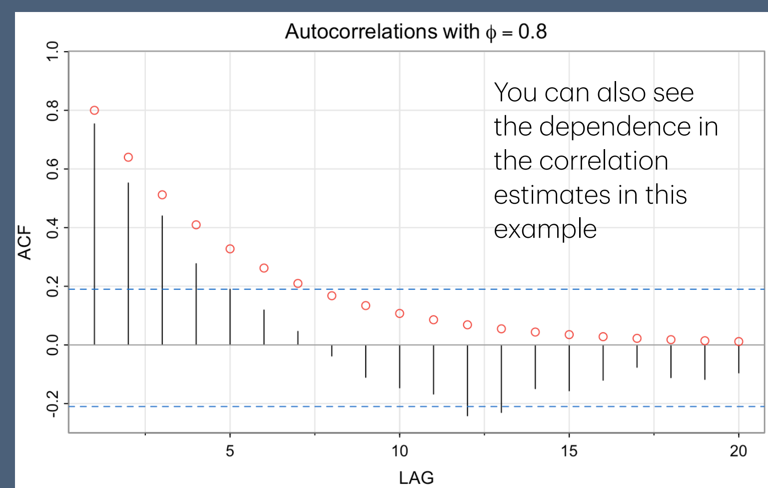
Estimated ACF

- Autoregression
  - True correlations for AR(1) are $\rho(h) = \phi^{|h|}$
  - Standard error of estimated autocorrelation depends on the process

$$\text{std error}(\hat{\rho}_x(r)) \approx \sqrt{\frac{1}{n}\sum_{j=-\infty}^{\infty}\rho_x(j)^2} = \sqrt{\frac{1}{n}\sum_{j=-\infty}^{\infty}\phi^{2|j|}} = \sqrt{\frac{1+\phi^2}{n(1-\phi^2)}}$$

  - Software places standard error bars as if white noise at $\pm 2/\sqrt{n}$





Multiple Time Series

- Joint stationarity
  - Means, variances and covariances are invariant of time origin
- Cross-correlation function
  - Correlation between two time series at different lags
  - Unlike ρ(h), the cross-covariance and cross-correlation need not be symmetric

$$\gamma_{x,y}(h) = \text{Cov}(X_{t+h}, Y_t) \qquad \left[= \gamma_{y,x}(-h) = \text{Cov}(Y_{t-h}, X_t) \right]$$

Definition 2.30

$$\rho_{x,y}(h) = \frac{\gamma_{x,y}(h)}{\sqrt{\gamma_x(0)\,\gamma_y(0)}}$$

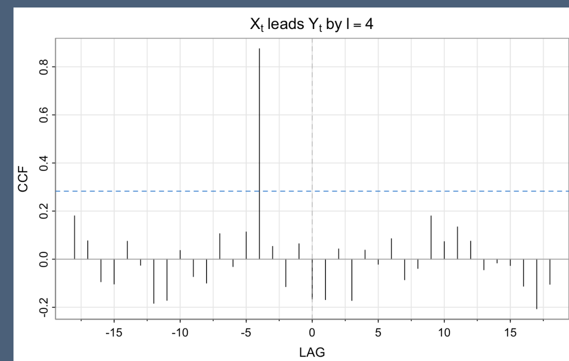
  - Example: If X_t and w_t are ind. white noise, what is γ_{xy} if the process is $Y_t = \alpha + \beta X_{t-4} + w_t$?
- Estimates
  - Formed as in the case of $\hat{\rho}(h)$ (Put hats on γ in the above definition)

$$\hat{\rho}_{x,y}(h) = \frac{\hat{\gamma}_{x,y}(h)}{\sqrt{\hat{\gamma}_x(0)\,\hat{\gamma}_y(0)}}$$

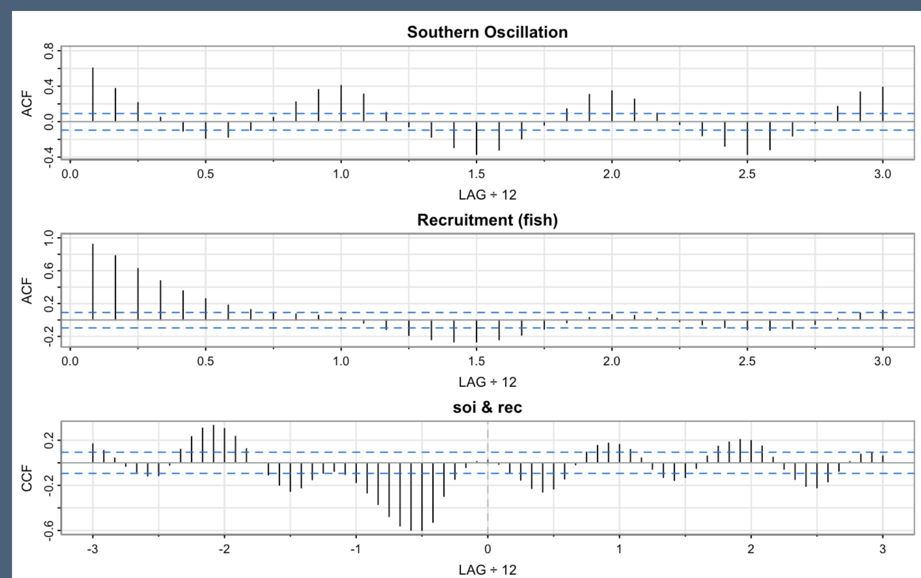
Text examples emphasize problem when have two time series.

Example of Cross-Correlation

- Simulated process on prior slide

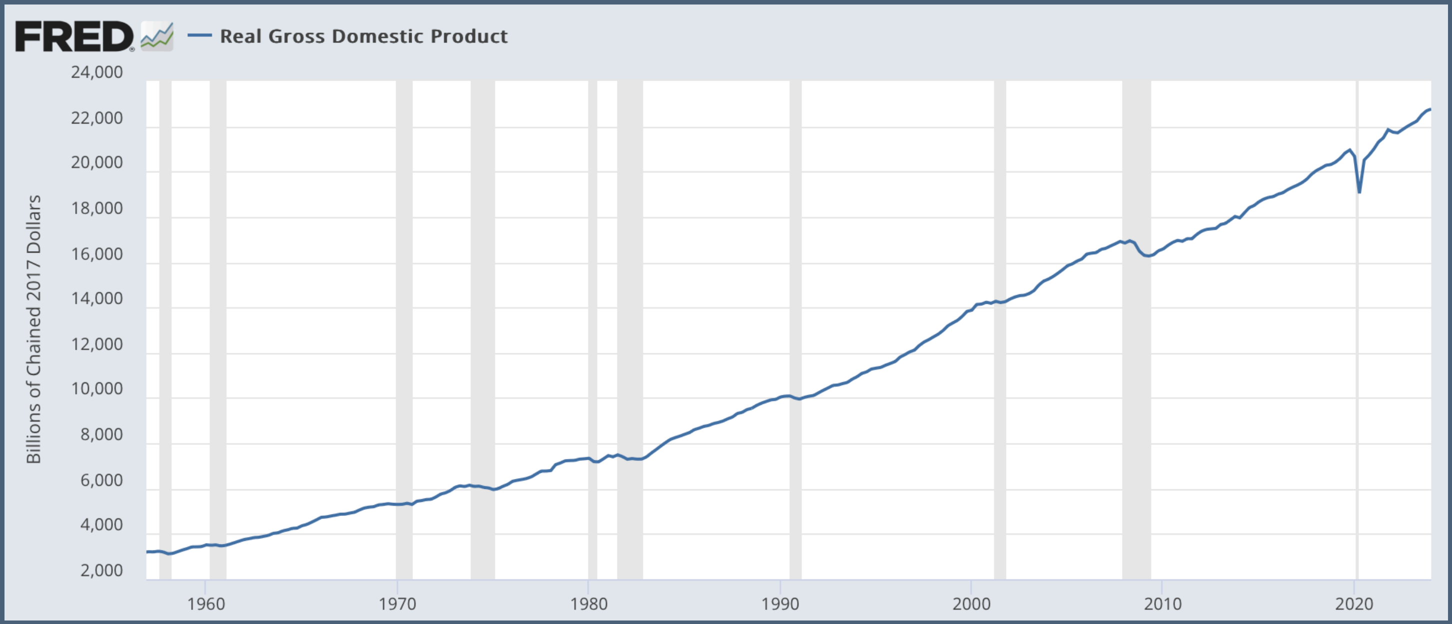


- Data: Southern oscillation and fish population
  - Example 2.32
  - SOI leads recruitment by about six months
  - Lagged correlation is hiding important aspect of association (see Lecture_4.Rmd)



Macroeconomic Correlations

- Common to see high correlations among macroeconomic variables
  - Many measure the size of the economy, producing collinearity
- Any time series with strong growth will appear correlated with GDP
  - Evidently not a stationary time series

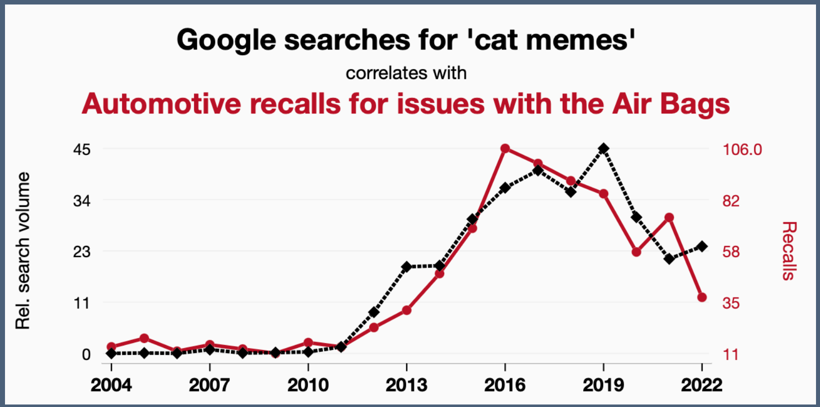


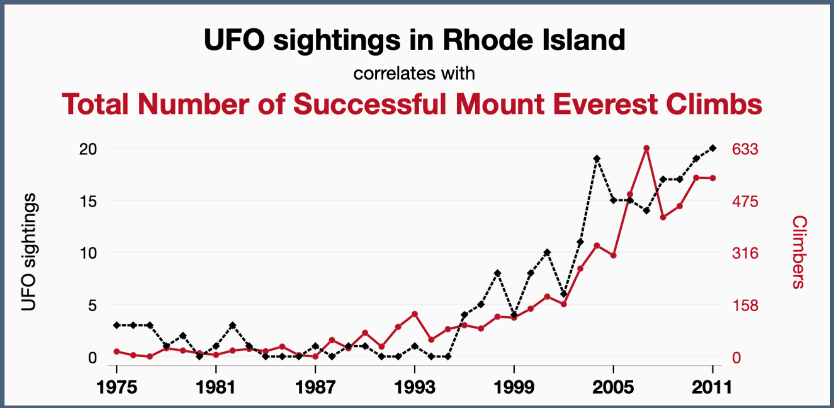
FRED is great source for economic times series

Spurious Correlation

- Website
  - Vast collection of examples, mostly pairs of time series
  - https://www.tylervigen.com/spurious-correlations

The site also offers AI explanations for the observed phenomena.





- How's this happen?
  - Multiplicity: Statistics rewards persistence
  - Non-stationarity: Any generally increasing time series will be correlated with US GDP

Converting to Stationarity

- Why
  - Stationary processes permit simpler analysis, both in theory and practice
  - Reduces likelihood of spurious correlation induced by trends
- Test case
  - Suppose that $X_t = \alpha + \rho X_{t-1} + \beta t + w_t$
  - Can we distinguish whether $\rho = 1$ with $\beta = 0$ (random walk) from $\rho = 0$ with $\beta \neq 0$?
- Two approaches
  - Estimate the trend using regression (and then work with residuals)
  - Difference the data, forming $X_t - X_{t-1} = \nabla X_t \qquad (= \Delta X_t$ or with $BX_t = X_{t-1}, (1-B)X_t)$
- Each approach has advantages and disadvantages
  - Subtract regression trend: familiar, but inference for $\hat{\rho}$ and $\hat{\beta}$ is nonstandard if $\rho = 1$
  - Difference: robust to specification errors (inefficient), but lose expression for the trend
  - General consensus: difference economic time series
  - "Testing for unit roots"

Issue will come up again later

What's next?

- Regression analysis
  - Review of least squares regression modeling
  - Regression diagnostics
  - Emphasis on plots and assumptions
  - Calculations in R

- In context of time series forecasting
  - Use in capturing trends such as see in macro data
  - Estimating effects of leading indicators
  - Analysis of unexplained (residual) variation