Time Series Notes

Nathan Ueda

March 30, 2024

Contents

1 Time Series Basics

What is a time series?

- A *time series* is a sequence of observations of the same variable indexed in time order (i.e. monthly stock returns).
- Defining x_t as a r.v., a time series may be written

$$\{x_1, x_2, \ldots, x_T\} \text{ or } \{x_t\},\ t = 1, 2, \ldots, T$$

Modeling time series data

- The models impose structure, so when dealing with model selection, it is important to evaluate to see if the model captures the features you believe to be present in the data.

2 Stationarity and Ergodicity

2.1 Basics and Importance

- Time series are typically not i.i.d. (i.e. if GNP is unusually high today, GNP will likely be unusually high tomorrow). So instead, we need different desirable properties for time series data. Two of the most important of these properties are stationarity and ergodicity.
- Stationarity is a property for time series with some time-invariant behavior.
- Ergodicity is a property for times series that expresses the idea that the effect of the past on the future eventually dies out.
- Time series with these properties are easy to estimate and have desirable properties.
- Time series that are nonstationarity and nonergodic require a different set of techniques.
- Under the assumption of i.i.d r.v.s, we had LLN and CLT. For r.v.s that are not i.i.d. and are instead autocorrelated, stationarity and ergodicity are similar results.
- In practice, by saying the time series is ergodic, it is typically implied it is both stationary and ergodic.
- If a times series is stationary, it does not imply ergodicity; however, a time series that is ergodic is always stationary.
- Takeaway: Ensure your time series is stationary and ergodic.

2.2 Strongly Stationarity and Weakly Stationarity

- A process $\{x_t\}$ is *strongly stationary* or *strictly stationary* if all aspects of its behavior are unchanged by shifts in time. More formally, it is defined as the requirement that for every m and n, the distributions of $\{x_1, \ldots, x_n\}$ and $\{x_{1+m}, \ldots, x_{n+m}\}$ are the same, that is, the joint probability distribution of a sequence of n observations does not depend on their time origin.

- A process $\{x_t\}$ is *weakly stationary* or *covariance stationary* if its mean, variance, and covariance are unchanged by time shifts. More formally, it is defined as weakly stationary if
  - First moment is a finite constant: $E(x_t) = \mu$
  - Second moment is a finite constant: $Var(x_t) = \sigma^2$
  - $\text{Cov}(x_t, x_s) - \gamma(|t - s|)$

  In other words, the mean and variance do not change with time and the covariance between two observations depends on the time distance between them (aka having same number of observations), not the specific points.

- Strong stationarity does not imply weak stationarity and weak stationarity does not imply strong stationarity.

- Stationarity is important as a stationary process can be modeled with relatively few paramters.

2.3 Testing for Stationarity

- When a times series is observed, a natural question is whether it appears to be stationary.

- **Time series plot**: Looking at a *time series plot* (plot of the series in chronological order) may be useful. If the time series is a stationary series, it should show some random oscillation around some fixed level, a phenomenon called *mean reversion*. If the series wanders without returning repeatedly to some fixed level, then the series should not be modeled as a stationary process.

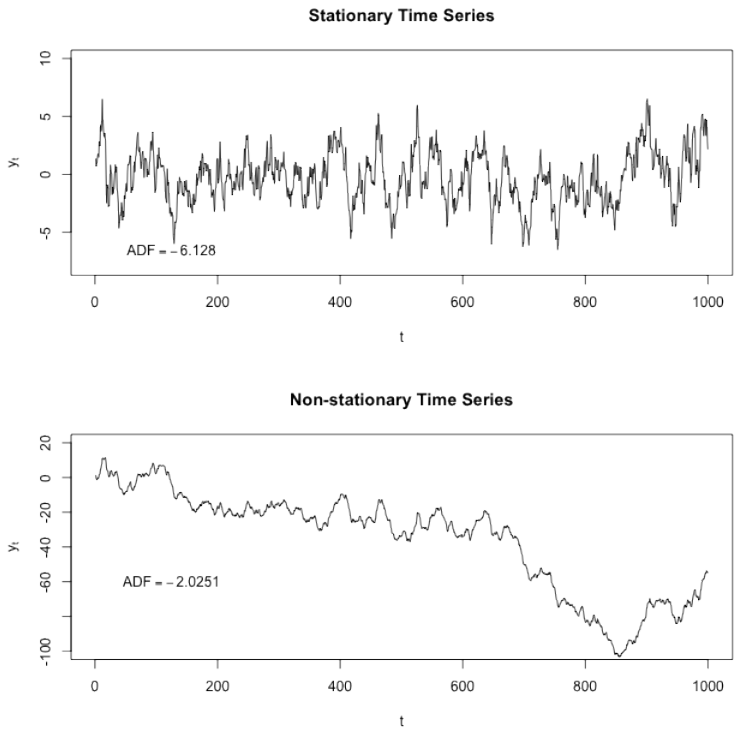

Figure 1: Stationary vs. non-stationary time series

- If all the roots of a characteristic polynomial for an AR process are greater than 1 in absolute value, then the AR process is stationary.
- Any finite-order MA(q) process is stationary and ergodic.

2.4 Ergodicity

- A stochastic process $\{x_t\}$ is ergodic if any two random variables sufficiently far apart in time are essentially independent. In other words, $\{x_t\}$ is ergodic if x_t and x_{t-j} are close to uncorrelated if j is large enough.

2.5 Law of Large Numbers and Central Limit Theorem for an Autocorrelated Process

- The *ergodic theorem* is a LLM for an autocorrelated process that states, if $\{x_t\}$ is stationary and ergodic, then
$$\bar{x}_T = \frac{1}{T}\sum_{t=1}^{T} x_t \to \mathrm{E}(x_t) \text{ as } T \to \infty$$
- If $\{x_t\}$ is stationary and ergodic, then the asymptotic distribution of the sample mean is normal

3 White Noise

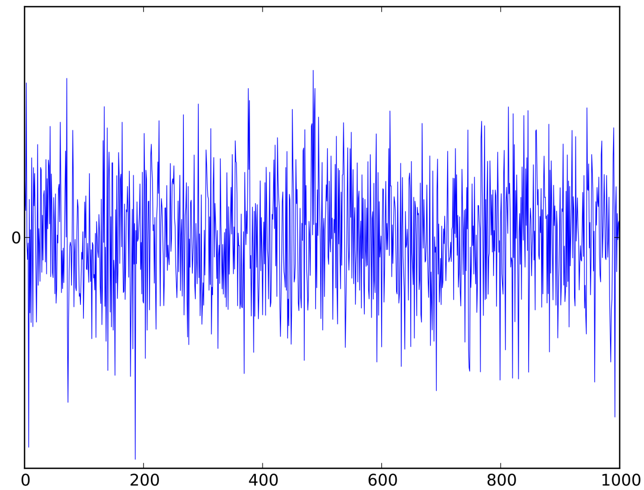

Figure 2: White noise

- The building block for time series models is the *white noise process*, denoted ϵ_t.
- In the least general case,

$$\epsilon_t \sim \text{i.i.d. } N(0, \sigma_\epsilon^2)$$

The assumption of i.i.d. have the following implications

1. No predictability: Past values of a white noise process contain no information to predict future values. Therefore, the best predictor is its mean, which is the same prediction without observing past values. More formally,

$$E(\epsilon) = E(\epsilon_t | \epsilon_{t-1}, \epsilon_{t-2} \ldots) = 0$$

2. No autocorrelation: Each observation is independent each other. More formally,

$$E(\epsilon_t \epsilon_{t-j}) = \text{Cov}(\epsilon_t \epsilon_{t-j}) = 0$$

3. Conditional homoskedasticity: The conditional variance is a constant, and is the same variance without observing past values. More formally,

$$\text{Var}(\epsilon_t) = \text{ Var}(\epsilon_t | \epsilon_{t-1}, \epsilon_{t-2} \ldots) = \sigma_\epsilon^2$$

4 Autocovariance and Autocorrelation

4.1 Autocovariance

- Autocovariance: Specifies the covariance between the value of a process at two times.

- Covariance: A nonstandardized measure to quantify the relationship between two variables. Can take on any value. A positive value means the variables tend to move in the same direction, a negative values means the variables tend to move in the opposite direction, and a zero value means they are independent and don't move in relation to each other.
- The *autocovariance* of a series x_t is defined as

$$\gamma_j = \text{Cov}(x_t, x_{t-j})$$

- $\gamma(h) = \gamma(-h)$ since what is important is the the space between the two observations, rather than the exact observations themselves.
- $\gamma_0 = \sigma^2$

4.2 Autocorrelation

- Autocorrelation (serial correlation): The degree of correlation of the same variable between different time intervals.
- Correlation: A standardized measure to quantify the relationship between two variables. Can take on a value inclusively between -1 and 1.
- A time series with autocorrleation implies that, predictive power (i.e. knowing the price of a stock today helps forecast its price tomorrow).
- The *autocorrelation* of a series x_t is defined as

$$\rho_j = \frac{\gamma_j}{\gamma_0}$$

4.3 Testing for autocorrelation

ACF Plots

- Show a correlation between a time series and lagged versions of itself.
- The plot also includes test bounds used to test the null hypothesis that an autocorrelation coefficient is 0. The null is rejected if the sample autocorrelatioon is outside the bounds, The usual level of the test is 0.05, so one can expect to see about 1 out of 20 samples outside the bounds simply by chance.

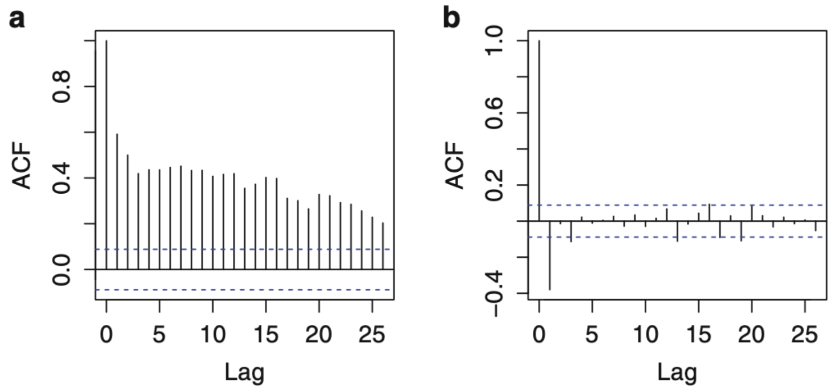

Figure 3: Sample ACF plots of (a) one-month inflation rate (decays slowly, indicating either nonstationarity or long-memory dependence) and (b) changes in the inflation rate (decays to 0 quickly, indicating that it is stationary).

Ljung-Box Test

- A *simultaneous test* is one that tests whether a group of null hypotheses area all true versus the alternative that at least one of them is false.
- The *Ljung-Box test* is a simultaneous test where the null is $\rho_1 = \rho_2 = \ldots = \rho_K$ for some K.
- If the Ljung-Box test rejects, then we conclude that one or more of $\rho_1, \rho_2, \ldots, \rho_K$ is nonzero.

5 Lag and Difference Operators

- These operators are useful in describing some time series models.

5.1 Lag Operators

- The lag operator moves the index back one unit in time.
- The lag operator L is defined

$$Lx_t = x_{t-1}$$

$$L^j x_t = x_{t-j}$$

- Also commonly referred to as a backwards operator.

5.2 Difference Operators

- Defined as $\Delta = 1 - L$ so that

$$\Delta x_t = x_t - Lx_t = x_t - x_{t-1}$$

- Δ^k is called the k-th order differencing operator.

$$\Delta^k x_t = (1 - L)^k x_t$$

6 Autoregressive Processes

- Time series models with correlation can be constructed from white noise.
- The simplest correlated stationary processess are *autoregressive processeses*, where $\{x_t\}$ is modeled as a weighted average of past observations plus a white noise "error."
- The term *autoregression* refers to the regression of the process on its own past values.
- The ACF of AR processes declines geometrically.

6.1 AR(1) Models

- Let $\epsilon_1, \epsilon_2, \ldots$ be $\mathrm{WN}(0, \sigma_\epsilon^2)$. Then $\{x_t\}$ is an AR(1) process if, for some constant parameter ϕ

  $$x_t = \phi x_{t-1} + \epsilon_t$$

  for all t.
- ϕx_{t-1} may be thought of as representing the memory of the past observation into the present value of the process. This is what we believe we can model and predict.
- ϕ determines the amount of feedback, where a larger absolute value results in more feedback.
- ϵ_t represents the effect of new information that cannot be modeled, hence, it is represented as white noise. This is what we cannot model nor predict.
- Properties of a Stationary AR(1) Process
  - $E(x_t) = \mu$
  - $\mathrm{Var}(x_t) = \gamma_0 = \sigma_x^2 = \frac{\sigma_{\epsilon_t}^2}{1- \ phi^2}$
- The ACF of an AR(1) process depends only on one parameter, ϕ. This parsimony comes at the cost that the ACF has only a very limited range of shapes. If the ACF does not behave in one of these shapes, the AR(1) model is not suitable.

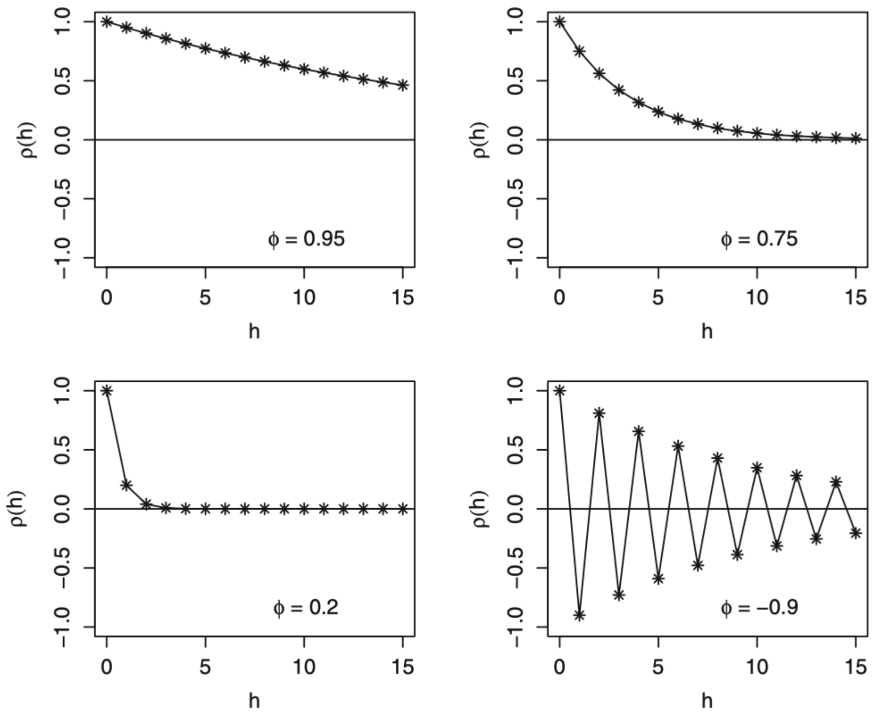

Figure 4: ACF of AR(1) processes with ϕ equal to 0.95, 0.75, 0.2, and -0.9.

- When $|\phi| < 1$ the process is stationary.
- When $\phi = 1$ the process is not stationary and takes the form of a random walk. Recall a random walk, in each period, takes a step random step that is i.i.d. from its previous steps.
- When $|\phi| > 1$ the process is non-stationary and has explosive behavior.
- A non-zero value of ϕ mean that there is some information in the previous observation, but a small value of ϕ means the prediction will not be very accurate.

6.2 AR(p) Models

- A more flexible adaptation of an AR model that is still parsimonious and regresses on the p past values.
- Let $\epsilon_1, \epsilon_2, \ldots$ be WN$(0, \sigma_\epsilon^2)$. Then $\{x_t\}$ is an AR(p) process if, for constant parameters $\phi_1, \phi_2, \ldots, \phi_{t-p}$
$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \ldots + \phi_p x_{t-p} + \epsilon_t$$
- If the AR(p) model fits the time series well, then the residuals should look like white noise. The residual autocorrelation can be detected examining the sample ACF of the residuals and using the Ljung-Box test. Any significant residual autocorrelation is a sign the AR(p) model does not fit well.
- A problem with AR models is that they often need a rather large value of p to fit a dataset.

7 Moving Average Processes

- With AR models, feeding past values into the current value has the effect of having at least some correlation at all lags. Sometimes data show correlation only at short lags, in these cases a MA process may be a suitable alternative.

- A process $\{x_t\}$ is a *moving average process* if $\{x_t\}$ can be expressed as a weighted average (moving average) of the past values of the white noise process $\{\epsilon_t\}$

7.1 MA(1)

- The MA(1) (moving average of order 1) process is

$$x_t = \epsilon_t + \theta\epsilon_{t-1}$$

  where, as before, the ϵ_t are weak WN$(0, \sigma_\epsilon^2)$.

7.2 MA(q)

- The MA(q) (moving average of order q) process is

$$x_t = \epsilon_t + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \ldots + \theta_q\epsilon_{t-q}$$

  where, as before, the ϵ_t are weak WN$(0, \sigma_\epsilon^2)$.

8 ARMA Processes

- Stationary time series with complex autocorrelation behavior are ofren more parsimoniously modeled by mixed AR and MA (ARMA) processes rather than by a pure AR or pure MA process.

8.1 ARMA(p, q)

- An ARMA(p, q) model combines both AR and MA terms, and is defined by the equation

$$x_t = \phi_1 x_{t-1} + \ldots + \phi_p x_{t-p} + \epsilon_t + \theta_1\epsilon_{t-1} + \ldots + \theta_q\epsilon_{t-q}$$

  which shows how $\{x_t\}$ depends on lagged values of itself and lagged values of the white noise process.

9 ARIMA Processes

- Often the first or perhaps second differences of nonstationarity time series are stationary. *Autoregressive integrated moving average* (ARIMA) processed include stationary as well as nonstationary processes.

- A time series $\{x_t\}$ is said to be an ARIMA(p, d, q) process if $\Delta^d x_t$ is ARMA(p, q).

- An ARIMA(p, d, q) is stationary if $d = 0$, otherwise its difference of order d or above are stationary.

- An ARIMA$(p, 0, q)$ is the same as an ARMA(p, q).
- A process is $I(d)$ if it is stationary after being differenced d times.
- An ARIMA(p, d, q) process has d unit roots, therefore we want to difference the process d times to get rid of the unit roots.

10 Impulse Response Functions

- The idea of an IRF is to enact a single shock to ϵ_t (in period t) and, via the IRF, see how the shock affects $x_{t+1}, x_{t+2}, \ldots$
- It allows us to start thinking about causes and effects.
- For an AR(1) process, notice that the effect of the shock never dies when $\phi = 1$, and it dies out quicker and quicker as we move from $\phi = 0.95$ to $\phi = 0.5$. This makes sense as the closer to 1 $|\phi|$ is, the more useful information there is in the previous observation in forecasting the current observation.

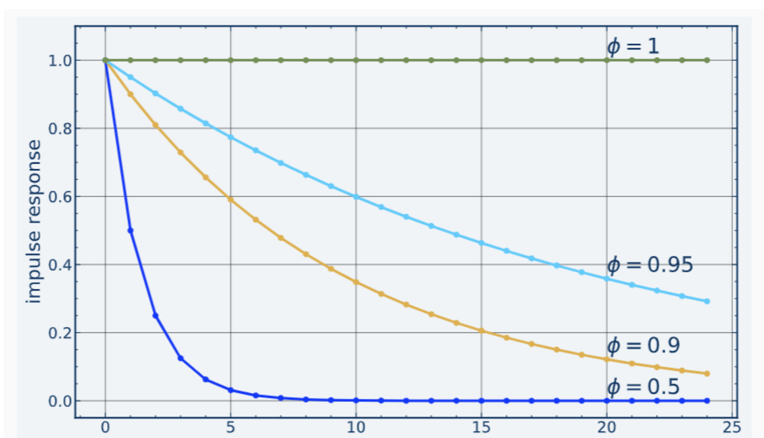

Figure 5: IRF on an AR(1) Process

11 Unit Root Tests

- Determining whether a time series is best modeled as stationary or nonstationary can be difficult, unit root tests aid in this process.
- Recall the definition of an ARMA(p, q) process

$$x_t = \phi_1 x_{t-1} + \ldots + \phi_p x_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q}$$

- The condition of $\{x_t\}$ to be stationary is that all roots of the polynomial

$$1 - \phi_1 x - \ldots - \phi_p x^p$$

have absolute values greater than one.

- If there is a unit root, that is, a root with an absolute value equal to 1, then the ARMA process is nonstationary and behaves much like a random walk. This is called the unit root case.
- Unit root tests are used to decide if an AR model have an absolute root equal to 1.
- A popular unit root test is the augmented Dickey-Fuller test (ADF test). In this test
  - H_0: there is a unit root (the process is nonstationary)
  - H_1: the process is stationary

12 Regression

13 Estimation for AR Models

- AR models can be estimated by OLS. Recal OLS is defined

$$y = X\beta + e$$

  where

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^{n\times 1},\ X = \begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{bmatrix} \in \mathbb{R}^{n\times k},\ e = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} \in \mathbb{R}^{n\times 1}$$

  where k is the number of regressors.

- Since true errors are not observed, MA models (which depend upon the true errors), MA models are not estimated by OLS.

13.1 OLS for AR(p) Process

- Recall the definition of an AR(p) process

$$z_t = \alpha + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \ldots + \phi_p z_{t-p} + \epsilon_t,\ t = 1, \ldots, T$$

- We let

$$y_i = z_t$$
$$x_i = [1, z_{t-1}, \ldots, z_{t-p}]'$$
$$\beta = [\alpha, \phi_1, \ldots, \phi_p]'$$

  where

$$y = \begin{bmatrix} z_{p+1} \\ z_{p+2} \\ \vdots \\ z_T \end{bmatrix},\ X = \begin{bmatrix} 1 & z_p & \ldots & z_1 \\ 1 & z_{p+1} & \ldots & z_2 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & z_{T-1} & \ldots & z_{T-p} \end{bmatrix},\ e = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

14 Misc.

- Returns are closer to i.i.d. than prices and overall exhibit more attractive statistical qualities than prices, therefore making it more sensible to study returns rather than prices.

15 Definitions

- Homoskedasticity: A condition in which the variance of the residual is constant, that is, the error term does not vary much. In other words, the variance of the data points is roughly the same for all data points.