working with time series

Autoregressive Moving Average (ARMA) Models

A combination of Autoregressive Models and Moving Average Models → ARMA(p,q)

$$X_t = \mathcal{E}_t + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} \qquad \text{AR(p)}$$

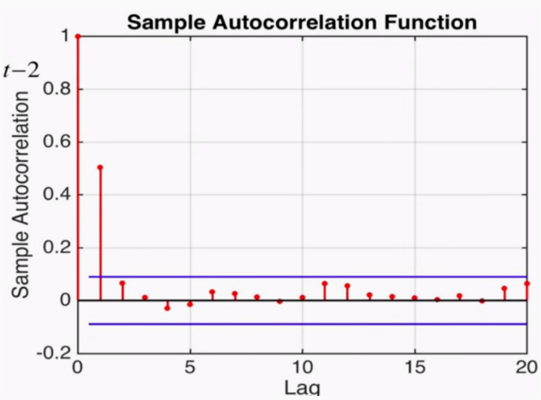
$$X_t = \mu + \mathcal{E}_t + \theta_1 \mathcal{E}_{t-1} + \cdots + \theta_q \mathcal{E}_{t-q} \qquad \text{MA(q)}$$

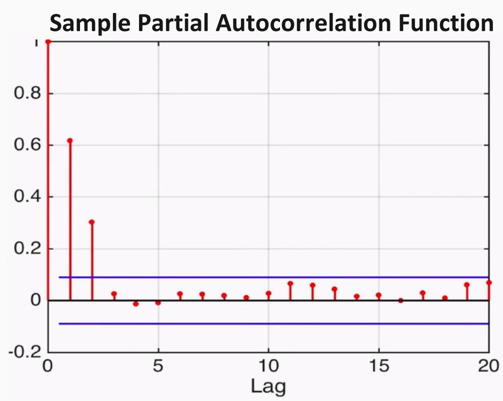
for example: the price of the product today depends on the price yesterday, the day before, etc. but also depends on the announcements of the company's welfare today, the day before, etc.

if p and q are known (order of both parts), regression can be performed to obtain all parameters

*Remember: pure MA models experience a sharp cutoff for the Autocorrelation Function
*Remember: pure AR models experience a sharp cutoff for the Partial Autocorrelation Function

Example: $X_t = \mathcal{E}_t + 0.6\mathcal{E}_{t-1} + 0.4\mathcal{E}_{t-2}$

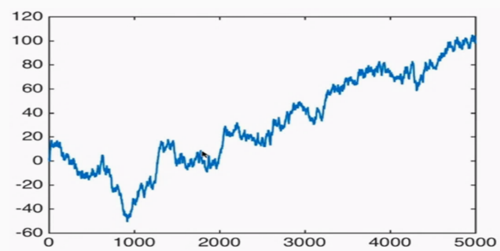




*However: if both the Autocorrelation and Partial Autocorrelation functions experience slow exponential decay, both the AR and MA terms are needed

Differencing

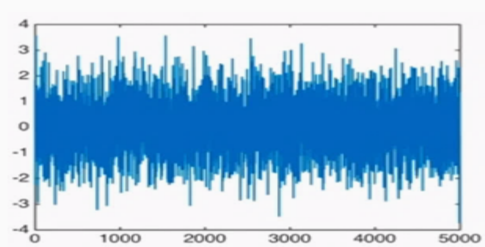
illustrated as 'a random walk' $X_{t-1} + \mathcal{E}_t$, with yesterday's position's represented as X_{t-1} and the random step for today's position (noise) denoted as $\mathcal{E}_t$.



a 'random' walk illustrated to the right is **not stationary**; indicating prior models are not applicable alone. computation of the 'difference series': $X_t - X_{t-1}$

(e.g. the difference between today and yesterday)

$$X_t - X_{t-1} = X_{t-1} + \mathcal{E}_t - X_{t-1} = \mathcal{E}_t$$



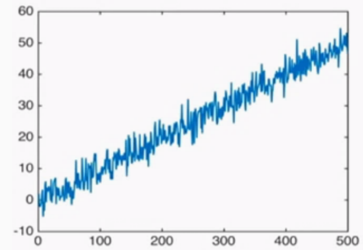
application of differencing creates a series that **is stationary.** Therefore, nonstationary data might benefit from computing the difference series prior to modeling (bottom graphic) an integrative moving average example:

$$\underbrace{X_{t-1} + \mathcal{E}_t}_{\text{'random walk' terms}} + \underbrace{\theta_1 \mathcal{E}_{t-1}}_{\text{MA(1) term}} \qquad \rightarrow \text{ also not stationary}$$

the above appears as both a Moving Average and a Random Walk; regardless it is still not stationary

$$X_t - X_{t-1} = X_{t-1} + \mathcal{E}_t + \theta_1 \mathcal{E}_{t-1} - X_{t-1} = \mathcal{E}_t + \theta_1 \mathcal{E}_{t-1} \rightarrow \text{stationary}$$

the above applies Differencing and results in simply a Moving Average Series to the order of 1 (MA(1)); this series is in fact stationary and can modeled effectively.



Differencing Continued

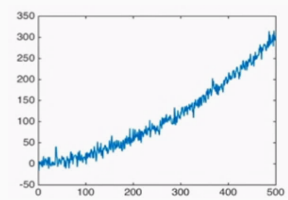
assuming a linear trend plus a stationary series:

$$X_t = \beta_0 + \beta_1 t + Z_t \longrightarrow$$

$$X_t - X_{t-1} = \cancel{\beta_0} + \beta_1 t + Z_t - (\cancel{\beta_0} + \beta_1 (t-1) + Z_{t-1})$$

$$= \cancel{\beta_1 t} - \beta_1 \cancel{(t-1)} + (Z_t - Z_{t-1})$$

$$= \underbrace{\beta_1}_{\text{constant}} + \underbrace{(Z_t - Z_{t-1})}_{\text{difference series of a stationary series}} \quad \rightarrow \quad \text{stationary}$$



the above is similar to taking a derivative (e.g. discrete derivative): $\nabla X_t = \beta_1 + \nabla Z_t$

assuming a quadratic trend plus a stationary series:

$$X_t = \beta_0 + \beta_1 t + \beta_1 t^2 + Z_t$$

$$X_t - X_{t-1} = (\cancel{\beta_0} + \beta_1 t + \cancel{\beta_2 t^2}) - (\cancel{\beta_0} + \beta_1 \cancel{(t-1)} + \beta_2 \cancel{(t-1)^2} + \nabla Z_t)$$

$$= \underbrace{(\beta_1 - \beta_2)}_{\text{constant}} + \underbrace{2\beta_2 t}_{\text{linear trend}} + \underbrace{\nabla Z_t}_{\text{difference series of a stationary series}}$$

$$\nabla X_t = \text{constant} + 2\beta_1 t + \nabla Z_t \rightarrow \quad \text{linear trend + stationary}$$

the above is addressed by taking yet another difference:

$$\nabla^2 X_t = 2\beta_1 + \nabla^2 Z_t \rightarrow \quad \text{stationary}$$

Differencing allows techniques for stationary timeseries' to model nonstationary series'.
Higher order trends turn into stationary models through repeated differencing.

Autoregressive Integrated Moving Average Models (ARIMA Models)

If a Difference Series results in an ARMA model, the original series is an ARIMA model.

$$Y_t = \nabla^d X_t$$

The original series is denoted as X_t; the difference ∇^d is taken to get to Y_t ;

In order for X_t to be an ARIMA model, Y_t has to be an ARMA Model

$$Y_t = \underbrace{\mu}_{\text{constant}} + \underbrace{\phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p}}_{\text{Autoregressive Function}} - \underbrace{\mathcal{E}_t - \theta_1 \mathcal{E}_{t-1} - \cdots - \theta_q \mathcal{E}_{t-q}}_{\text{Moving Average Function}}$$

In general, if the Difference Series is either a purely Moving Average Model or the Difference Series is a purely Autoregressive Model, then the original series is still an ARIMA Model:

$\mathrm{ARIMA}(d, p, q) \rightarrow d =$ order of differences; $p =$ order of AR Model; $q =$ order of MA Model

Properties of ARIMA Models

- As long as the difference series of any order is an ARMA model, the model is an AIMA model; differences can continuously be taken in testing if an ARMA model exists.
- Differences can be computed easily; take differences until the model becomes stationary
- The Autocorrelation function can be used to determine the order of a purely MA model, or the Partial Autocorrelation to determine the order of a purely AR model.