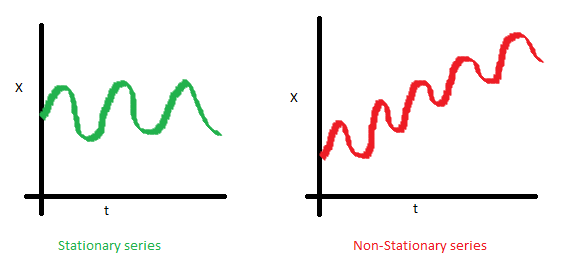
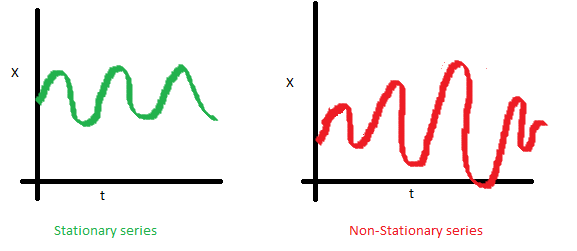
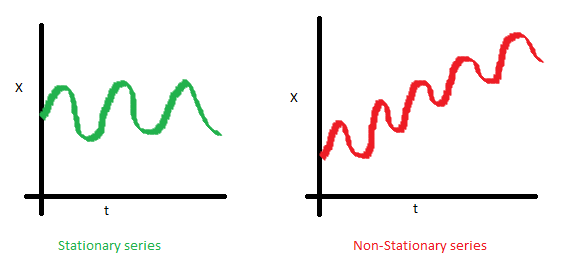
시계열 분석 : 시간의 흐름에 따라 변화하는 수치형 데이터를 분석하는 기법

트렌드(기울기)/분산/계절성이 불변하는 규칙적 시계열, 이 중 하나라도 변화하는 불규칙적 시계열 데이터로 나눌 수 있음

* 시계열 분석은 불규칙적 데이터를 특정 기법/모델을 적용해 규칙적 패턴을 적용하여 예측할 수 있도록 함







분석 기법 종류 : 전통적 분석 기법(ARIMA)와 기계학습 및 딥러닝 (LSTM, RNN) 기법이 있음

ARIMA 분석 기법 분석

AR 기법 (AutoRegression) : 자기회귀모델, 자기 자신의 과거값으로 현재값을 예측하는 모델

 - AR(1) 모델

직전 자기의 상태(X(t-1)에 w를 곱하고 b(상수항) 과 e(t) – white noise follows N(0,1)에 u를 ㄱㅂ한 것을 더한 것.

 - AR(n) 모델 : n 이전의 자기의 값부터 영향을 끼침

MA 기법 (Moving Average) 기법 : 이동평균모델, 이전 항의 오차(변동값, e(t))로 현재 상태 예측

 - MA(1) 모델

이전항의 오차 (e(t-1))이 트렌드이므로 이에 맞춘 값을 추론한다는 의미

ARMA 기법 : AR과 MA를 모두 결합한 기법

 - ARMA(1,1) 모델 == ARIMA(1,0,1)

ARIMA 기법 : 차분을 통해 비정상적 데이터를 정상적으로 변화 시킨 후에 AR or MA, 혹은 둘 다를 적용시킨 기법 (Autoregressive Integrated Moving Average)

 차분 x

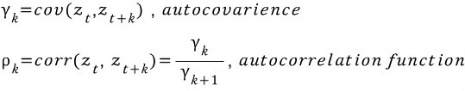
 1차 차분

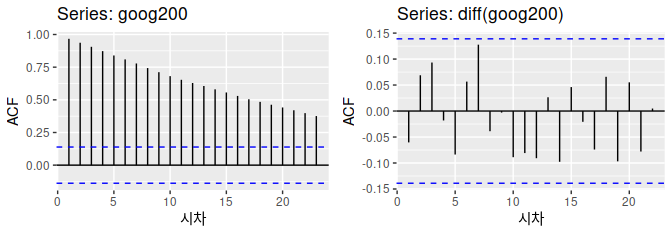
 2차 차분

ARIMA(p,d,q)의 차수를 결정하는 것이 중요해짐

이를 찾는 방법으로 ACF(AutoCorrelation Function) – q값 찾음, PACF(Partial ACF) – p값 찾음을 활용

ACF : K 시간 차이를 둔 Y값의 상관 관계 (y(t) & y(t-k))

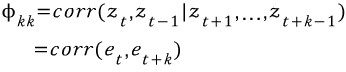




(좌) : 무차분한 원래값 (우) : 1차 차분한 값

파이썬에서 제공하는 파란선 안에 들어올 때까지 차분을 해야 한다. (보통 차분은 max 2)

PACF : K 시간 차이를 둔 et값의 상관 관계 (e(t) & e(t-k))의 cor



X와 y에서의 시간의 효과를 제거한 후의 상관계수 \*

시계열 분해란 :

추세/순환/계절성을 제거하여 불규칙 요소만 남기는 것.

덧셈 분해와 곱셈 분해로 나뉨. -> 곱셈분해는 로그 변환을 통해 덧셈 분해로 변환.

.\* 덧셈/곱셈 나누는 기준은?

ACF/PACF가 아래의 기준에 맞을 때까지 차분(N차까지) 로그 변환(최대 2회) 해보기

+ ADF(Augmented Dickey-Fuller) test 도 통과 해야함

H0 : 자료에 단위근이 존재한다. (정상성 만족X), H1 : 자료가 정상성을 만족한다.

(p,d,q) 정해지면 이를 모수로 한 arima 모델의 잔차 acf pacf 를 그려, 정상성을 따르는지 확인해야 한다.

모델 평가 : AIC/BIC 최소인 모델 -> p와 q가 작을수록 좋음

시계열 분해란 :

https://leedakyeong.tistory.com/entry/%EC%8B%9C%EA%B3%84%EC%97%B4-%EB%B6%84%ED%95%B4%EB%9E%80Time-Series-Decomposition-%EC%8B%9C%EA%B3%84%EC%97%B4-%EB%B6%84%EC%84%9D%EC%9D%B4%EB%9E%80-%EC%8B%9C%EA%B3%84%EC%97%B4-%EB%8D%B0%EC%9D%B4%ED%84%B0%EB%9E%80-%EC%B6%94%EC%84%B8Trend-%EC%88%9C%ED%99%98Cycle-%EA%B3%84%EC%A0%88%EC%84%B1Seasonal-%EB%B6%88%EA%B7%9C%EC%B9%99-%EC%9A%94%EC%86%8CRandom-Residual

정상성 판별 후 ACF/PACF 그래프 판별법 :

https://leedakyeong.tistory.com/entry/ARIMA%EB%9E%80-ARIMA-%EB%B6%84%EC%84%9D%EA%B8%B0%EB%B2%95-AR-MA-ACF-PACF-%EC%A0%95%EC%83%81%EC%84%B1%EC%9D%B4%EB%9E%80

날씨 시계열 데이터로 ARIMA 적용하기

https://leedakyeong.tistory.com/entry/Python-%EB%82%A0%EC%94%A8-%EC%8B%9C%EA%B3%84%EC%97%B4-%EB%8D%B0%EC%9D%B4%ED%84%B0Kaggle%EB%A1%9C-ARIMA-%EC%A0%81%EC%9A%A9%ED%95%98%EA%B8%B0

## python 패키지 설치 오류시 (ssl/https 에러)

pip install --trusted-host pypi.python.org --trusted-host pypi.org --trusted-host files.pythonhosted.org ortools

# 계수 설정법

https://towardsdatascience.com/identifying-ar-and-ma-terms-using-acf-and-pacf-plots-in-time-series-forecasting-ccb9fd073db8

출처 : https://yamalab.tistory.com/112

<https://support.minitab.com/ko-kr/minitab/18/help-and-how-to/modeling-statistics/time-series/how-to/autocorrelation/interpret-the-results/autocorrelation-function-acf/>

<https://support.minitab.com/ko-kr/minitab/18/help-and-how-to/modeling-statistics/time-series/how-to/partial-autocorrelation/interpret-the-results/partial-autocorrelation-function-pacf/>

<https://velog.io/@cha-suyeon/%EC%8B%9C%EA%B3%84%EC%97%B4-%EB%8D%B0%EC%9D%B4%ED%84%B0-%EC%A0%95%EC%83%81%EC%84%B1Stationarity-%EC%B0%A8%EB%B6%84differencing-%ED%99%95%EB%A5%A0%EB%B3%B4%ED%96%89-2%EC%B0%A8-%EC%B0%A8%EB%B6%84>

AIC/BIC 설명 : <https://otexts.com/fppkr/arima-estimation.html>

Q1. (P)ACF 그래프에서 파란색 선의 의미 : 이 선을 넘어가는 (Partial) AutoCorrelation가 y(0)과 유의한 관계를 갖는다. (0이 아님) 즉, t=6에서 acf가 파란색 선을 넘어가면 y는 6의 시차를 갖고 상관관계를 갖는 것이고, 이를 ma 식을 세워서 반영해야 하는 것이다.

