

09

SARIMA 실습

[SARIMA 개념]

이전 시간에 ARIMA모형에서 데이터 과적합과 test데이터에 대한 설명력이 낮았던 문제를 SARIMA로 해결할 수 있습니다.

00:00

9.5 계절성 시계열 데이터에 대한 분석 : SARIMA

SARIMA: 데이터가 지닌 계절성(주기에 따라 유사한 양상으로 변화하는 모양)까지 고려한 ARIMA 모델

9.5.1 데이터 로드

```
In [16]: import pandas as pd
import matplotlib.pyplot as plt

data = pd.read_csv('../data/sarima_data.csv', names = ['day', 'price'])
data['day'] = pd.to_datetime(data['day'], format='%Y-%m-%d')
data.set_index('day', inplace=True)
s_data = data
```

```
In [17]: s_data
```

```
Out[17]:
```

day	price
2013-01-01	3794
2013-02-01	3863
2013-03-01	5190

[SARIMA 모형 해석]

03:11

```
model = SARIMAX(training.price.values,
                 order=(2,1,2), seasonal_order=(1,1,1,12),
                 trends='ct',
                 enforce_stationarity=False,
                 enforce_invertibility=False)

results = model.fit()
results.summary()
```

C:\Users\W\jang\anaconda3\envs\WADP_Class\lib\site-packages\statsmodels\base\model.py:568: ConvergenceWarning: Maximum Likelihood optimization failed to converge. Check mle_retvals
"Check mle_retvals", ConvergenceWarning)

```
Out[16]:
```

SARIMAX Results

Dep. Variable:	y	No. Observations:	42			
Model:	SARIMAX(2, 1, 2)x(1, 1, 1, 12)	Log Likelihood:	-90.141			
Date:	Thu, 09 Dec 2021	AIC:	198.281			
Time:	00:44:30	BIC:	204.033			
Sample:	0	HQIC:	197.749			
			- 42			
Covariance Type:	opg					
	coef	std err	z	P> z	[0.025	0.975]
Intercept	-266.3161	2243.254	-0.119	0.905	-4863.012	4130.380
drift	8.4685	52.534	0.161	0.872	-94.497	111.434
ar.L1	-0.8374	1.085	-0.772	0.440	-2.963	1.289

[SARIMA 시각화 및 해석]

04:16

