



گزارش کار تمرین اول
درس یادگیری ماشین

استاد درس: دکتر کمندی

دانشجویان گروه:

محیا معتمدی ۸۱۰۸۹۷۰۵۳

مینو احمدی ۸۱۰۸۹۷۰۳۲

دانشکده علوم مهندسی، دانشگاه تهران

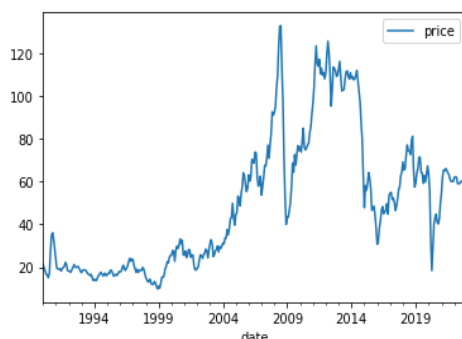
بهار ۱۴۰۱

سوال ۱)

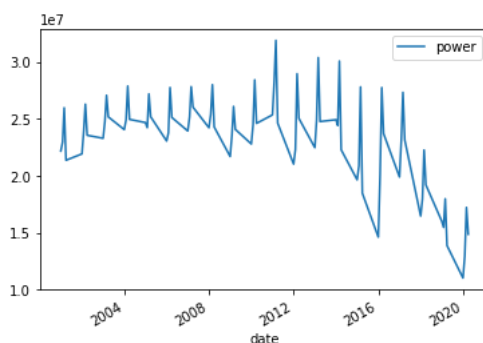
در این سوال ابتدا دیتا را از فایل csv خواندیم و سپس لیبل مورد نظر را تغییر دادیم و ستون نشان دهنده زمان را با استفاده از `pd.to_datetime()` از حالت تشخیص به صورت رشته به فرمت تاریخ تبدیل کردیم.

سپس در مرحله بعدی خود دیتاهای داده شده برای `Brent Spot price` و `Coal Power` را رسم کردیم نمودار `BSP` برای دیتاست اول و `CP` برای دیتاست دوم است.

BSP:



CP:



در مرحله ی بعدی به کمک تابع `adfuller()` که تست Augmented Dickey-Fuller (ADF) را انجام میدهد ایستایی سری زمانی ها را بررسی میکنیم. خروجی این تابع برای هر دو دیتاست در زیر آمده است.

BSP

```
(-2.3662084811231265, 0.15150178029715095, 1, 394, {'1%': -3.4470566389664703, '5%': -2.8689037160476016, '10%': -2.570692663557422}, 2212.551893789197)
```

CP

```
(1.976302909694134, 0.9986413567364828, 10, 69, {'1%': -3.528889992207215, '5%': -2.9044395987933362, '10%': -2.589655654274312}, 2117.1250184879855)
```

برای آنکه سری زمانی ما ایستا باشد باید عدد ابی شده (مقدار اول) از تمام سه مقدار بنفش کمتر باشد که اینطور نیست پس نتیجه میشود که هر دو سری زمانی ما نا ایستا هستند و باید هر دو آن ها را ایستا کنیم.

روش differencing :

در گام بعدی حال که سری ها نا ایستا هستند با استفاده از دستور `np.diff()` بر روی دیتاست ها و با تعیین کردن $n=1$ (نشان دهنده آن است که با یک با تفاضل گیری این روش اجرا شده است) سری های زمانی را ایستا میکنیم سپس برای اطمینان از ایستایی آن ها دوباره از دستور `adfuller()` استفاده میکنیم تا ببینیم که آیا سری ایستا شده است یا خیر. نتیجه تست در زیر آمده است.

BSP with differencing

```
(-8.282210848987841, 4.488030141441116e-13, 17, 377, {'1%': -3.4478152097574224, '5%': -2.8692371153451903, '10%': -2.5708704003405356}, 2247.8084279227455)
```

CP with differencing

```
(-4.319019758926173, 0.00041170373591297105, 12, 66, {'1%': -3.5335601309235605, '5%': -2.9064436883991434, '10%': -2.590723948576676}, 2095.199309475299)
```

همان طور که میبینیم اعداد مشخص شده با رنگ ابی (مقادیر اول) از سه عددی که با رنگ بنفش مشخص شده اند کوچکتر شده است در حالی که قبل از این روش این طور نبود پس ما توانستیم تا با اعمال روش differencing سری زمانی های خود را ایستا کنیم. با توجه به اینکه $n=1$ در دستور `diff()` قرار داده شده بود و با همین عدد سری های ما ایستا شدند پس نتیجه میگیریم که حداقل یک مرحله تفاضل گیری برای ایستا کردن این دو سری زمانی کافی بود (این عدد برای سری زمانی های دیگر امکان دارد بیشتر باشد).

روش moving average :

در گام بعدی حال که سری ها نا ایستا هستند با استفاده از دستور `data.rolling(window=num).mean` بر روی دیتاست ها و با تعیین کردن عدد های مختلف برای `window` (نشان دهنده آن است که با میانگین گیری چندتا چندتا از دیتا این روش اجرا شده است) سری های زمانی را ایستا میکنیم سپس

برای اطمینان از ایستایی آن ها دوباره از دستور `adfuller()` استفاده میکنیم تا ببینیم که آیا سری ایستا شده است یا خیر. نتیجه تست برای `window` های مختلف در زیر آمده است.

BSP with moving average:

```
ADF_result_BSP_moving_window50 (-1.8024304170167433, 0.37925572838986443, 14, 332, {'1%': -3.4502011472639724, '5%': -2.8702852297358983, '10%': -2.5714292194077513}, -454.09911307262894)
ADF_result_BSP_moving_window30 (-1.5572125624271391, 0.5050434809680446, 16, 350, {'1%': -3.4491725955218655, '5%': -2.8698334971428574, '10%': -2.5711883591836733}, -90.52833370050371)
ADF_result_BSP_moving_window100 (-1.5572125624271391, 0.5050434809680446, 16, 350, {'1%': -3.4491725955218655, '5%': -2.8698334971428574, '10%': -2.5711883591836733}, -90.52833370050371)
```

همان طور که میبینیم اعداد مشخص شده با رنگ ابی (مقادیر اول) از سه عددی که با رنگ بنفش مشخص شده اند برای دیتا BSP کوچکتر نشده است و نشان میدهد که دیتای ما با اعمال روش moving avrage سری زمانی BSP ایستا نشده است. با توجه به اینکه `window` در دستور `rolling(window=num).mean()` عدد های مختلفی را قرار دادیم این کار را به کمک `for` و تغییر عدد `window` انجام دادیم:

```
for i in range(1,100):
    BSPbest = BSP['price'].rolling(window=i).mean()
    BSPbest.dropna(inplace = True)
    BSPbest_dataframe = pd.DataFrame(BSPbest)
    BSPbest_values = BSPbest_dataframe.iloc[:,0].values
    ADF_result_BSP_moving_windowbest = adfuller(BSPbest_values)
    print('ADF_result_BSP_moving_window',i,":",ADF_result_BSP_moving_windowbest)

BSP50 = BSP['price'].rolling(window=50).mean()
```

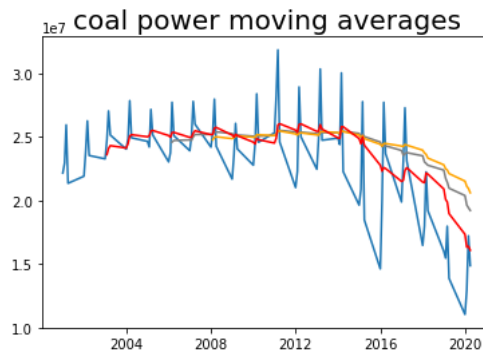
با مشاهده خروجی ها متوجه شدیم که دیتای ما تا `window=50` با این روش ایستا نشد چنانچه سری زمانی ما با اعدادی بالاتر از این مقدار ایستا بشود باز هم چون دیتای زیادی را به دلیل افزایش عدد `window` از دست داده ایم پس ایستا سازی مناسب نخواهد بود چون با این کار عملا مقدار بسیار زیادی از داده ها را حذف کرده ایم.

price of Brent Spot Price moving averages



CP with moving average:

ADF_result_CP_moving_window23 (0.33671910600274074, 0.978967820998611, 11, 46, {'1%': -3.5812576580093696, '5%': -2.9267849124681518, '10%': -2.6015409829867675}, 1193.8493518455205)
 ADF_result_CP_moving_window30 (2.9487608389165634, 1.0, 5, 45, {'1%': -3.584828853223594, '5%': -2.9282991495198907, '10%': -2.6023438271604937}, 985.0381194085551)
 ADF_result_CP_moving_window10 (1.1903279956671196, 0.995918372810919, 11, 59, {'1%': -3.5463945337644063, '5%': -2.911939409384601, '10%': -2.5936515282964665}, 1597.3820933109353)



همان طور که میبینیم اعداد مشخص شده با رنگ ابی (مقادیر اول) از سه عددی که با رنگ بنفش مشخص شده اند برای دیتا CP کوچکتر شده است در حالی که قبل از این روش این طور نبود پس ما توانستیم تا با اعمال روش moving average سری زمانی CP را ایستا کنیم. با توجه به اینکه window در دستور `rolling(window=num).mean()` عدد های مختلفی را قرار دادیم این کار را به کمک `for` و تغییر عدد `window` انجام دادیم:

با مقایسه خروجی ها عدد ۲۳ برای `window` کمترین مقدار داده ابی رنگ را به همراه دارد به طوری که داده ایستا نیز شده است:

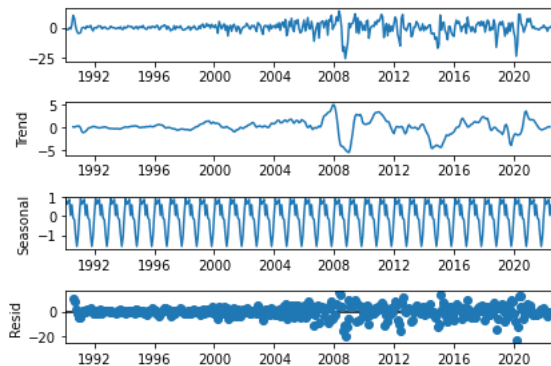
ADF_result_CP_moving_window 23 : (0.33671910600274074, 0.978967820998611, 11, 46, {'1%': -3.5812576580093696, '5%': -2.9267849124681518, '10%': -2.6015409829867675}, 1193.8493518455205)

سوال ۲)

در این بخش ما نمودار های decomposition ، خود همبستگی و خود همبستگی جزئی را رسم کردیم.

در بخش decomposition به کمک دستور `seasonal_decompose(bsp/cp_diff, model='additive')` و رسم نتایج آن نمودار ها به شکل زیر حاصل شدند.

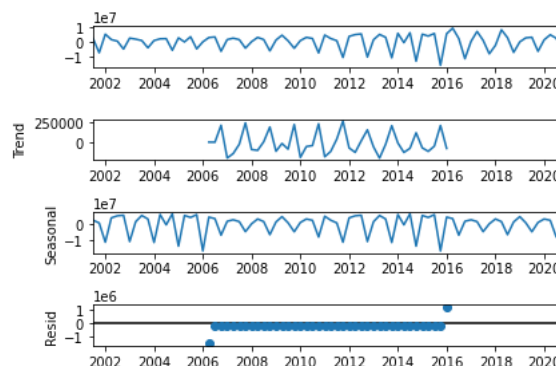
BSP:



تحلیل:

ترند آن ثابت نیست ولی حالت فصلی دارد و کاملاً افت و خیز دارد و مانده آن در نمودار resid آورده شده است.

CP:

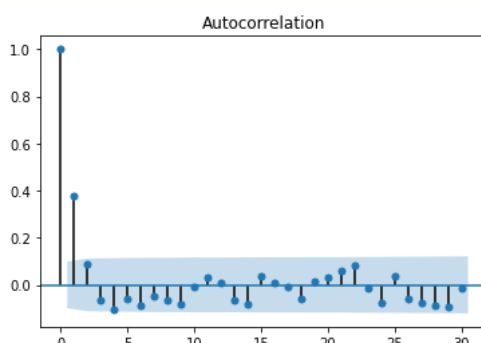


تحلیل: در این بخش میبینیم که ترند آن تقریباً ثابت است و حالت فصلی دارد ولی بازه های فصلی آن نسبت به دیتا قبلی بزرگ تر است و مانده آن در نمودار resid آورده شده است.

در بخش خود همبستگی از تابع

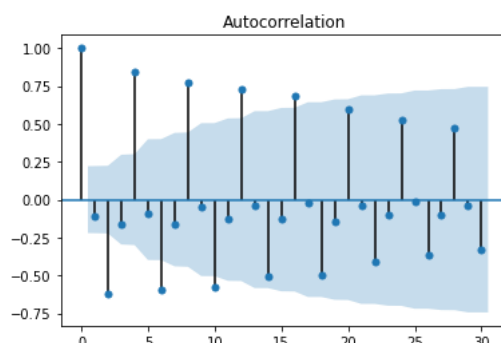
`sm.graphics.tsa.plot_acf(BSP/CP_diff, lags=30)` استفاده کردیم
و مقدار lag آن را با عدد ۳۰ تعیین کردیم و نتیجه را رسم کردیم نمودار های زیر
حاصل شد.

BSP:



تحلیل: برای بررسی نمودار ما فقط خود همبستگی هایی که از سطح معنا دار بالاتر هستند را بررسی میکنیم و آن هایی که بالای سطح معنا دار هستند، برای ما مهم است. در اینجا نشان داده شده که لگ اول تاثیر مثبت دارد و باقی داده ها تاثیر چندانی ندارند چون زیر سطح معنا دارند. به طور کلی ما در این نمودار ارتباط لگ ها را با هم مشاهده میکنیم این ارتباط میتواند مستقیم یا معکوس باشد و داده ها تاثیر مثبت یا منفی داشته باشند.

CP:



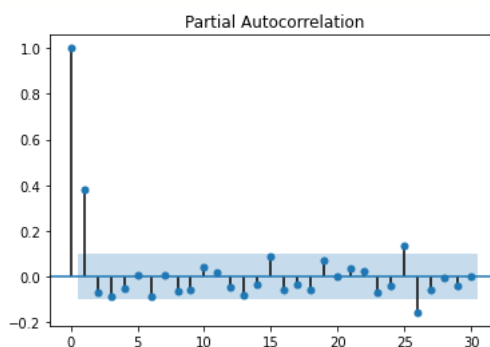
تحلیل: در این شکل مبینیم که لگ های پنجم ، هشتم ، دوازدهم ، شانزدهم باهم ارتباط مستقیم دارند ولی ارتباطشان با لگ های دوم ، ششم و دهم عکس میباشد لگ چهارم

با لگ شانزدهم آخرین ارتباط قوی مستقیم و با لگ دهم آخرین ارتباط قوی معکوس را دارد چراکه ما بقی در زیر سطح معنا دارند و تاثیر چندانی ندارند.

در بخش بعدی برای بررسی خودهمبستگی جزی از دستور
`sm.graphics.tsa.plot_pacf(BSP_diff, lags=30, method="ywm")`

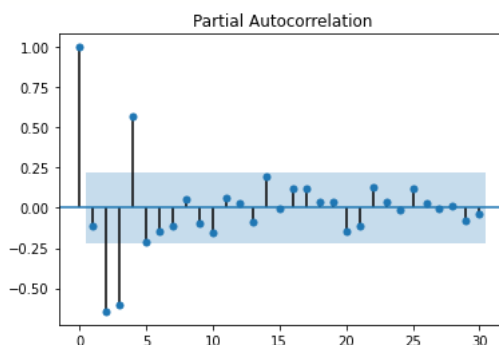
استفاده کردیم و سپس نتایج حاصل شده را رسم کردیم. در خود همبستگی جزی وابستگی غیر مستقیم حذف میشود. رابطه بین دو داده طوری بیان میشود که با هیچ لگ کوچتری قابل بیان نیست.

BSP:



تحلیل: در این نمودار دیتا اول تاثیر مثبت دارد و مابقی زیر سطح معنا دارند و تاثیری ندارند در ۲۵ و ۲۶ دیتا های بالای سطح معنا دار داریم و دیتا ۲۶ با ۲۵ در ارتباط است که دیتای در ۲۵ تاثیر مثبت و دیتا ۲۶ تاثیر منفی دارد.

CP:



تحلیل: در این نمودار نیز دیتا سه با دو در ارتباط هست و هر دو تاثیر منفی دارند ولی دیتا چهارم تاثیر مثبت دارد و با سومی در ارتباط است (در جزی با داده قبلی در ارتباط هست نه قبل تر هاش) مابقی در زیر سطح معنا دارند و تاثیر زیادی ندارند.

سوال ۳)

در این سوال ابتدا کل داده ها را با نسبت ۸۰ به ۲۰ به دو دسته داده های آموزش و آزمایش تقسیم کردیم و سپس به کمک داده های آموزش مدل های و MA و AR و ARIMA و ARMA را آموزش دادیم سپس با تست کردن داده های آزمایش نتایج و میزان مطلوبیت و صحت مدل ها را بررسی کرده و به کمک تابع پردیکت دیتا های مجهول را پیش بینی میکنیم.

AR:

تعداد لگ ها را از نمودار خود هبستگی جزئی استخراج میکنیم برای BSP میتوانیم هم ۶ و هم ۳ را بگذاریم که ۶ نتیجه بهتری میداد که ما ۶ را گذاشتیم. همین طور برای CP با توجه به نمودار ان مقدار لگ را ۳ میگذاریم.

BSP AutoReg model:

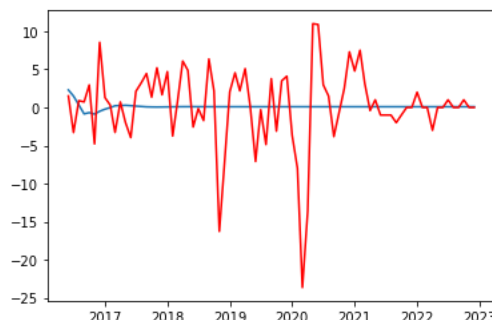
AutoReg Model Results						
=====						
Dep. Variable:	price	No. Observations:		316		
Model:	AutoReg(6)	Log Likelihood		-893.841		
Method:	Conditional MLE	S.D. of innovations		4.325		
Date:	Fri, 22 Apr 2022	AIC		1803.683		
Time:	14:53:10	BIC		1833.575		
Sample:	08-01-1990	HQIC		1815.633		
	- 05-01-2016					
=====						
	coef	std err	z	P> z	[0.025	0.975]

const	0.0728	0.246	0.296	0.767	-0.409	0.554
price.L1	0.3855	0.056	6.840	0.000	0.275	0.496
price.L2	0.0181	0.060	0.299	0.765	-0.100	0.136
price.L3	-0.0350	0.060	-0.580	0.562	-0.153	0.083
price.L4	-0.0718	0.060	-1.188	0.235	-0.190	0.047
price.L5	0.0419	0.061	0.690	0.490	-0.077	0.161
price.L6	-0.1271	0.057	-2.237	0.025	-0.239	-0.016
Roots						
=====						
	Real	Imaginary		Modulus	Frequency	

AR.1	1.1379	-0.6167j		1.2943	-0.0790	
AR.2	1.1379	+0.6167j		1.2943	0.0790	
AR.3	-1.1881	-0.7974j		1.4309	-0.4059	
AR.4	-1.1881	+0.7974j		1.4309	0.4059	
AR.5	0.2149	-1.4989j		1.5143	-0.2273	
AR.6	0.2149	+1.4989j		1.5143	0.2273	

مقدار عدد های زیر ستون مربوط به P باید کمتر از 0.05 باشد که برای price.l6 این عدد برابر 0.025 است و کمتر از 0.05 میباشد

BSP prediction with AR:



در شکل بالا خط قرمز نشان دهنده ی داده های ایستا شده و خط آبی نشانه داده های پردیکت شده هستند.

CP AutoReg model:

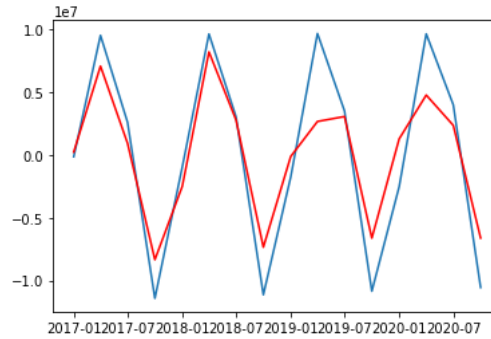
AutoReg Model Results						
=====						
Dep. Variable:	power	No. Observations:	62			
Model:	AutoReg(3)	Log Likelihood	-931.566			
Method:	Conditional MLE	S.D. of innovations	1741613.819			
Date:	Fri, 22 Apr 2022	AIC	1873.133			
Time:	14:53:11	BIC	1883.521			
Sample:	04-01-2002	HQIC	1877.188			
	- 10-01-2016					
=====						
	coef	std err	z	P> z	[0.025	0.975]

const	7.888e+04	2.27e+05	0.348	0.728	-3.66e+05	5.24e+05
power.L1	-0.9191	0.052	-17.796	0.000	-1.020	-0.818
power.L2	-0.9727	0.052	-18.852	0.000	-1.074	-0.872
power.L3	-0.9496	0.053	-18.065	0.000	-1.053	-0.847
Roots						
=====						
	Real	Imaginary	Modulus	Frequency		

AR.1	0.0148	-0.9995j	0.9996	-0.2476		
AR.2	0.0148	+0.9995j	0.9996	0.2476		
AR.3	-1.0539	-0.0000j	1.0539	-0.5000		

مقدار عدد های زیر ستون مربوط به P باید کمتر از 0.05 باشد که برای power.L3 این عدد برابر 0.000 است و کمتر از 0.05 میباشد

CP prediction with AR:



در شکل بالا خط قرمز نشان دهنده ی داده های ایستا شده و خط آبی نشانه داده های پردیکت شده هستند.

MA:

برای moving average ما از کتاب خانه Arima استفاده کردیم که با استفاده از آن میتوان هم AR,MA,AMRA,ARIMA را پیاده سازی کنیم در ARIMA در قسمت order سه ورودی میگیرد p, d, q که P لگ ها اتو رگرشن است و p درجه differ و q لگ های moving average است. در moving average ، q را عدد داده و دوتای دیگر را صفر میگذاریم. تعداد لگ های moving average از نمودار auto correlation متوجه میشویم که برای BSP ما هم میتوانستیم ۴ و هم ۶ را بگذاریم که ۶ نتیجه بهتری داد.

```
RIMA(bsp_train, order=(0,0,6)).fit()
```

برای CP هم ۳ و ۴ و ۶ را میتوان گذاشت چون likelihood سه از همه کمتر بود ما سه را قرار دادیم.

```
ARIMA(cp_train, order=(0,0,3)).fit()
```

BSP MA model:

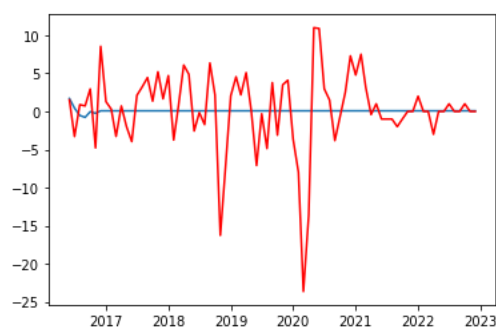
```

=====
SARIMAX Results
=====
Dep. Variable:          price      No. Observations:          316
Model:                 ARIMA(0, 0, 6)  Log Likelihood          -909.526
Date:                 Fri, 22 Apr 2022  AIC              1835.052
Time:                 14:53:11         BIC              1865.098
Sample:              02-01-1990       HQIC             1847.055
                             - 05-01-2016
Covariance Type:      opg
=====
              coef      std err          z      P>|z|      [0.025      0.975]
-----
const          0.0812      0.340         0.239      0.811      -0.586      0.748
ma.L1          0.3858      0.042         9.177      0.000       0.303      0.468
ma.L2          0.1682      0.035         4.819      0.000       0.100      0.237
ma.L3          0.0051      0.041         0.124      0.901      -0.075      0.085
ma.L4         -0.0957      0.050        -1.933      0.053      -0.193      0.001
ma.L5         -0.0420      0.043        -0.977      0.329      -0.126      0.042
ma.L6         -0.1105      0.048        -2.324      0.020      -0.204     -0.017
sigma2        18.4999      1.057        17.510      0.000      16.429     20.571
=====
Ljung-Box (L1) (Q):          0.00      Jarque-Bera (JB):          130.89
Prob(Q):                    0.95      Prob(JB):              0.00
Heteroskedasticity (H):      16.92      Skew:                 -0.57
Prob(H) (two-sided):         0.00      Kurtosis:              5.94
=====

```

مقدار عدد های زیر ستون مربوط به P باید کمتر از 0.05 باشد که برای ma.L6 این عدد برابر 0.020 است و کمتر از 0.05 میباشد

BSP Prediction with MA:



در شکل بالا خط قرمز نشان دهنده ی داده های ایستا شده و خط آبی نشانه داده های پردیکت شده هستند.

CP MA model:

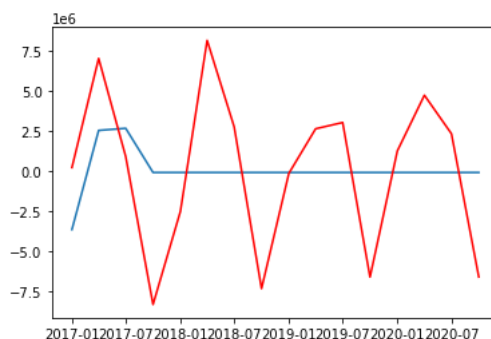
```

=====
SARIMAX Results
=====
Dep. Variable:          power    No. Observations:          62
Model:                  ARIMA(0, 0, 3)    Log Likelihood          -1038.025
Date:                   Fri, 22 Apr 2022    AIC                    2086.050
Time:                   14:53:11    BIC                    2096.686
Sample:                 07-01-2001    HQIC                   2090.226
                        - 10-01-2016
Covariance Type:        opg
=====
              coef      std err          z      P>|z|      [0.025      0.975]
-----
const      -7.787e+04    2.44e+05     -0.320     0.749    -5.55e+05     4e+05
ma.L1       -0.0400      0.259      -0.154     0.877     -0.548     0.468
ma.L2       -0.2991      0.284     -1.052     0.293     -0.856     0.258
ma.L3       -0.5922      0.278     -2.131     0.033     -1.137     -0.048
sigma2      2.922e+13      0.003     9.8e+15     0.000     2.92e+13     2.92e+13
=====
Ljung-Box (L1) (Q):          8.45    Jarque-Bera (JB):          3.97
Prob(Q):                    0.00    Prob(JB):          0.14
Heteroskedasticity (H):      4.43    Skew:          -0.61
Prob(H) (two-sided):         0.00    Kurtosis:         3.23
=====

```

مقدار عدد های زیر ستون مربوط به P باید کمتر از 0.05 باشد که برای ma.L6 این عدد برابر 0.033 است و کمتر از 0.05 میباشد.

CP Prediction with MA:



در شکل بالا خط قرمز نشان دهنده ی داده های ایستا شده و خط آبی نشانه داده های پردیکت شده هستند.

ARMA:

در این مدل هم از AR و هم از MA استفاده میکنیم پس برای ان p,q را مقدار دادیم ولی d را صفر میگذاریم و از داده های ایستا شده استفاده کردیم .

میتوانیم از عدد های کمتر از AR و MA هم جواب میگیریم که برای BSP ما ۲ و ۳ قرار دادیم

```
ARIMA(bsp_train, order=(2,0,3)).fit()
```

ولی برای CP عدد ها ۲ و ۱ را گذاشتیم.

```
ARIMA(cp_train, order=(2,0,1)).fit()
```

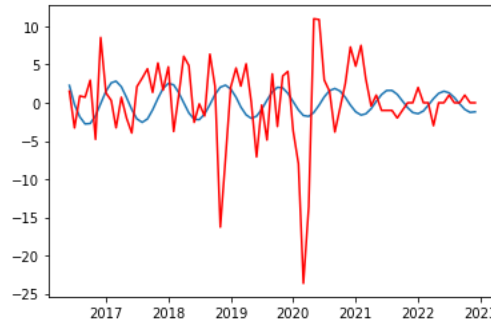
BSP ARMA model:

SARIMAX Results						
=====						
Dep. Variable:	power		No. Observations:		62	
Model:	ARIMA(2, 0, 1)		Log Likelihood		-1045.033	
Date:	Fri, 22 Apr 2022		AIC		2100.066	
Time:	14:53:12		BIC		2110.702	
Sample:	07-01-2001		HQIC		2104.242	
	- 10-01-2016					
Covariance Type:	opg					
=====						
	coef	std err	z	P> z	[0.025	0.975]

const	-7.787e+04	2.1e+04	-3.716	0.000	-1.19e+05	-3.68e+04
ar.L1	-0.2131	0.041	-5.206	0.000	-0.293	-0.133
ar.L2	-0.6015	0.056	-10.710	0.000	-0.712	-0.491
ma.L1	-0.8541	0.042	-20.155	0.000	-0.937	-0.771
sigma2	3.205e+12	0.002	1.65e+15	0.000	3.21e+12	3.21e+12
=====						
Ljung-Box (L1) (Q):	12.70		Jarque-Bera (JB):		0.79	
Prob(Q):	0.00		Prob(JB):		0.67	
Heteroskedasticity (H):	4.40		Skew:		0.17	
Prob(H) (two-sided):	0.00		Kurtosis:		3.44	
=====						

مقدار عدد های زیر ستون مربوط به P باید کمتر از 0.05 باشد در این بخش هم باید برای AR و هم برای MA هر دو کمتر از 0.05 باشند که برای ma.L1 این عدد برابر ۰.۰۰ است و هم چنین برای ar.L2 برابر ۰.۰۰ است که هر دو کمتر از 0.05 هستند.

BSP Prediction with ARMA:



در شکل بالا خط قرمز نشان دهنده ی داده های ایستا شده و خط آبی نشانه داده های پردیکت شده هستند.

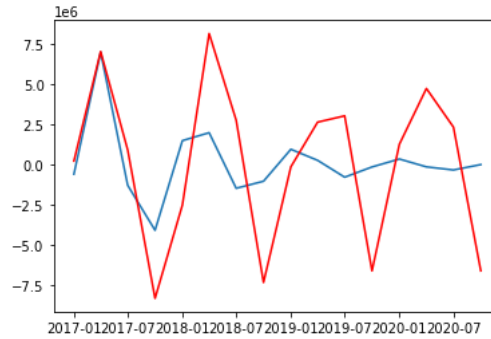
CP ARMA model:

SARIMAX Results						
=====						
Dep. Variable:	power	No. Observations:	62			
Model:	ARIMA(2, 0, 1)	Log Likelihood	-1045.033			
Date:	Fri, 22 Apr 2022	AIC	2100.066			
Time:	14:53:12	BIC	2110.702			
Sample:	07-01-2001	HQIC	2104.242			
	- 10-01-2016					
Covariance Type:	opg					
=====						
	coef	std err	z	P> z	[0.025	0.975]

const	-7.787e+04	2.1e+04	-3.716	0.000	-1.19e+05	-3.68e+04
ar.L1	-0.2131	0.041	-5.206	0.000	-0.293	-0.133
ar.L2	-0.6015	0.056	-10.710	0.000	-0.712	-0.491
ma.L1	-0.8541	0.042	-20.155	0.000	-0.937	-0.771
sigma2	3.205e+12	0.002	1.65e+15	0.000	3.21e+12	3.21e+12
=====						
Ljung-Box (L1) (Q):	12.70	Jarque-Bera (JB):	0.79			
Prob(Q):	0.00	Prob(JB):	0.67			
Heteroskedasticity (H):	4.40	Skew:	0.17			
Prob(H) (two-sided):	0.00	Kurtosis:	3.44			
=====						

مقدار عدد های زیر ستون مربوط به P باید کمتر از 0.05 باشد در این بخش هم باید برای AR و هم برای MA هر دو کمتر از 0.05 باشند که برای ma.L1 این عدد برابر ۰.۰۰ است و هم چنین برای ar.l2 برابر ۰.۰۰ است که هر دو کمتر از 0.05 هستند.

CP Prediction with ARMA:



در شکل بالا خط قرمز نشان دهنده ی داده های ایستا شده و خط آبی نشانه داده های پردیکت شده هستند.

ARIMA:

چون در این مدل ما به d مقدار می‌دهیم داده ها می‌توانند ایستا شده نباشند و ما برای اینکه ببینیم چگونه با داده غیر ایستا برخورد میکند که از داده های اصلی استفاده میکنیم و به دو بخش `train` و `test` تقسیم کرده و برای BSP همان عدد های ۲ و ۳ قرار می‌دهیم فقط این بار مرتبه `diff` را ۱ می‌گذاریم.

```
ARIMA(bsp_train_arima, order=(2,1,3)).fit()
```

در CP نیز باقی عدد ها مثل حالت `arma` است فقط مرتبه `diff` را این بار ۱ قرار می‌دهیم.

```
ARIMA(cp_train_arima, order=(2,1,1)).fit()
```

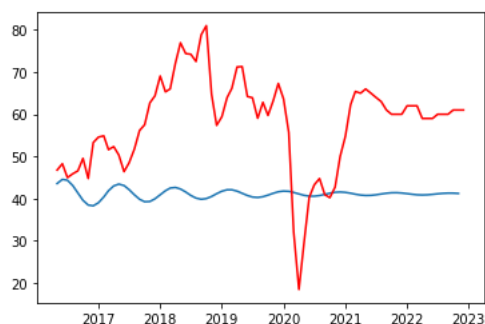
BSP ARIMA model:

SARIMAX Results						
=====						
Dep. Variable:	price	No. Observations:	316			
Model:	ARIMA(2, 1, 3)	Log Likelihood	-906.746			
Date:	Fri, 22 Apr 2022	AIC	1825.493			
Time:	15:04:36	BIC	1848.008			
Sample:	01-01-1990	HQIC	1834.488			
	- 04-01-2016					
Covariance Type:	opg					
=====						
	coef	std err	z	P> z	[0.025	0.975]

ar.L1	1.6038	0.049	32.828	0.000	1.508	1.700
ar.L2	-0.9264	0.047	-19.871	0.000	-1.018	-0.835
ma.L1	-1.2499	0.072	-17.374	0.000	-1.391	-1.109
ma.L2	0.4582	0.079	5.818	0.000	0.304	0.613
ma.L3	0.2317	0.052	4.494	0.000	0.131	0.333
sigma2	18.4999	0.963	19.201	0.000	16.612	20.388
=====						
Ljung-Box (L1) (Q):	0.07	Jarque-Bera (JB):	144.32			
Prob(Q):	0.79	Prob(JB):	0.00			
Heteroskedasticity (H):	17.25	Skew:	-0.48			
Prob(H) (two-sided):	0.00	Kurtosis:	6.17			

مقدار عدد های زیر ستون مربوط به P باید کمتر از 0.05 باشد در این بخش هم باید برای AR و هم برای MA هر دو کمتر از 0.05 باشند که برای ma.L3 این عدد برابر ۰.۰۰ است و هم چنین برای ar.L2 برابر ۰.۰۰ است که هر دو کمتر از 0.05 هستند.

BSP Prediction with ARIMA:



در شکل بالا خط قرمز نشان دهنده ی داده های ایستا شده و خط آبی نشانه داده های پردیکت شده هستند.

CP ARIMA model:

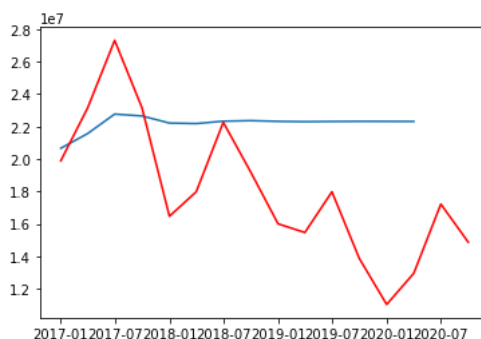
```

=====
SARIMAX Results
=====
Dep. Variable:          power      No. Observations:          64
Model:                 ARIMA(2, 1, 1)  Log Likelihood          -1051.093
Date:                 Fri, 22 Apr 2022  AIC                2110.185
Time:                 16:06:34      BIC                2118.758
Sample:              01-01-2001      HQIC              2113.557
                   - 10-01-2016
Covariance Type:      opg
=====
              coef      std err          z      P>|z|      [0.025      0.975]
-----
ar.L1          0.1647        0.067        2.441        0.015        0.032        0.297
ar.L2         -0.3475        0.021       -16.509        0.000       -0.389       -0.306
ma.L1         -0.4647        0.089        -5.192        0.000       -0.640       -0.289
sigma2        3.729e+12    1.8e-14    2.07e+26        0.000    3.73e+12    3.73e+12
=====
Ljung-Box (L1) (Q):          0.19      Jarque-Bera (JB):          15.02
Prob(Q):                    0.66      Prob(JB):              0.00
Heteroskedasticity (H):      1.57      Skew:                  0.52
Prob(H) (two-sided):        0.31      Kurtosis:              5.16
=====

```

مقدار عدد های زیر ستون مربوط به P باید کمتر از 0.05 باشد در این بخش هم باید برای AR و هم برای MA هر دو کمتر از 0.05 باشند که برای ma.L1 این عدد برابر ۰.۰۰ است و هم چنین برای ar.l2 برابر ۰.۰۰ است که هر دو کمتر از 0.05 هستند.

CP Prediction with ARIMA:



در شکل بالا خط قرمز نشان دهنده ی داده های ایستا شده و خط آبی نشانه داده های پردیکت شده هستند.

