Sec+1+Homework+%231

January 12, 2024

[1]:
```
import pandas as pd
import statsmodels.api as sm
```

1 1.) Import Data from FRED

[3]:
```
data = pd.read_csv("TaylorRuleData.csv", index_col = 0)
```

[4]:
```
data.index = pd.to_datetime(data.index)
```

[7]:
```
data.dropna(inplace = True)
```

[8]:
```
data.head()
```

[8]:

	FedFunds	Unemployment	HousingStarts	Inflation
1959-01-01	2.48	6.0	1657.0	29.01
1959-02-01	2.43	5.9	1667.0	29.00
1959-03-01	2.80	5.6	1620.0	28.97
1959-04-01	2.96	5.2	1590.0	28.98
1959-05-01	2.90	5.1	1498.0	29.04

2 2.) Do Not Randomize, split your data into Train, Test Holdout

[10]:
```
split1 = int(len(data) * .6)
split2 = int(len(data) * .9)
data_in = data[:split1]
data_out = data[split1:split2]
data_hold = data[split2:]
```

[12]:
```
data_in.iloc[0:10,1:3]
```

[12]:

	Unemployment	HousingStarts
1959-01-01	6.0	1657.0
1959-02-01	5.9	1667.0
1959-03-01	5.6	1620.0
1959-04-01	5.2	1590.0
1959-05-01	5.1	1498.0

```
1959-06-01           5.0         1503.0
1959-07-01           5.1         1547.0
1959-08-01           5.2         1430.0
1959-09-01           5.5         1540.0
1959-10-01           5.7         1355.0
```

[14]:
```
X_in = data_in.iloc[:,1:]
y_in = data_in.iloc[:,0]
X_out = data_out.iloc[:,1:]
y_out = data_out.iloc[:,0]
X_hold = data_hold.iloc[:,1:]
y_hold = data_hold.iloc[:,0]
```

[15]:
```
# Add Constants
X_in = sm.add_constant(X_in)
X_out = sm.add_constant(X_out)
X_hold = sm.add_constant(X_hold)
```

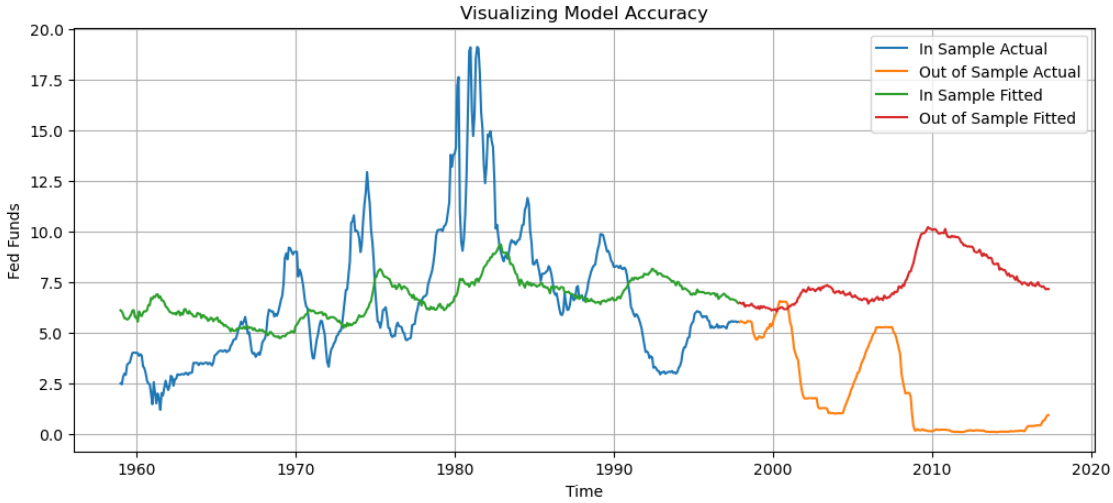
3 3.) Build a model that regresses FF~Unemp, HousingStarts, Inflation

[16]:
```
model1 = sm.OLS(y_in, X_in).fit()
```

4 4.) Recreate the graph fro your model

[17]:
```
import matplotlib.pyplot as plt
```

[32]:
```
plt.figure(figsize = (12,5))

###
plt.plot(y_in)
plt.plot(y_out)
plt.plot(model1.predict(X_in))
plt.plot(model1.predict(X_out))
###

plt.ylabel("Fed Funds")
plt.xlabel("Time")
plt.title("Visualizing Model Accuracy")
plt.legend(["In Sample Actual", "Out of Sample Actual", "In Sample Fitted",␣
 ↪"Out of Sample Fitted"])
plt.grid()
plt.show()
```



4.1 "All Models are wrong but some are useful" - 1976 George Box

5 5.) What are the in/out of sample MSEs

[22]:
```
from sklearn.metrics import mean_squared_error
```

[23]:
```
in_mse_1 = mean_squared_error(y_in, model1.predict(X_in))
out_mse_1 = mean_squared_error(y_out, model1.predict(X_out))
```

[25]:
```
print("Insample MSE : ", in_mse_1)
print("Outsample MSE : ", out_mse_1)
```

```
Insample MSE :  10.071422013168641
Outsample MSE :  40.36082783566856
```

6 6.) Using a for loop. Repeat 3,4,5 for polynomial degrees 1,2,3

[26]:
```
from sklearn.preprocessing import PolynomialFeatures
```

[27]:
```
#PolynomialFeatures?
```

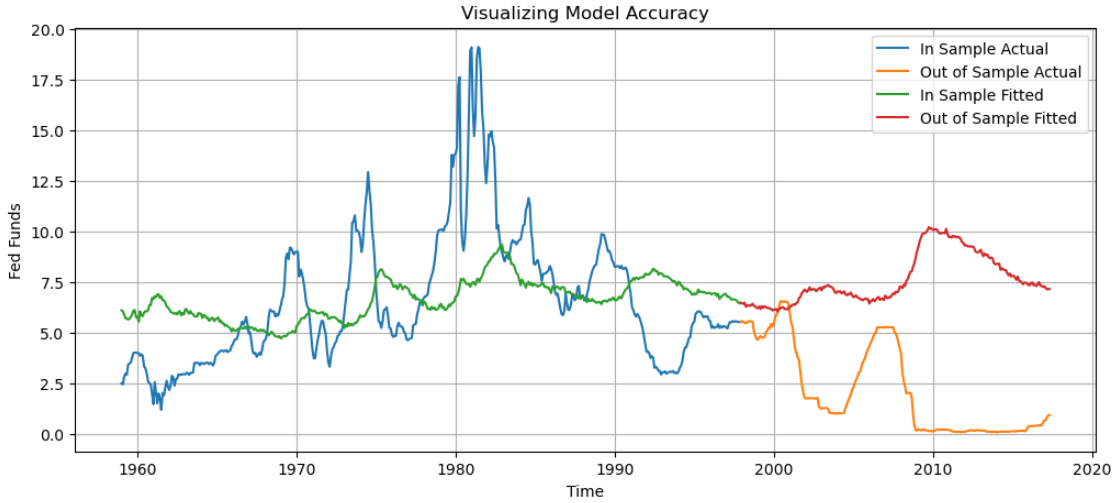
[28]:
```
degrees = 2
```

[30]:
```
max_degrees = 3
```

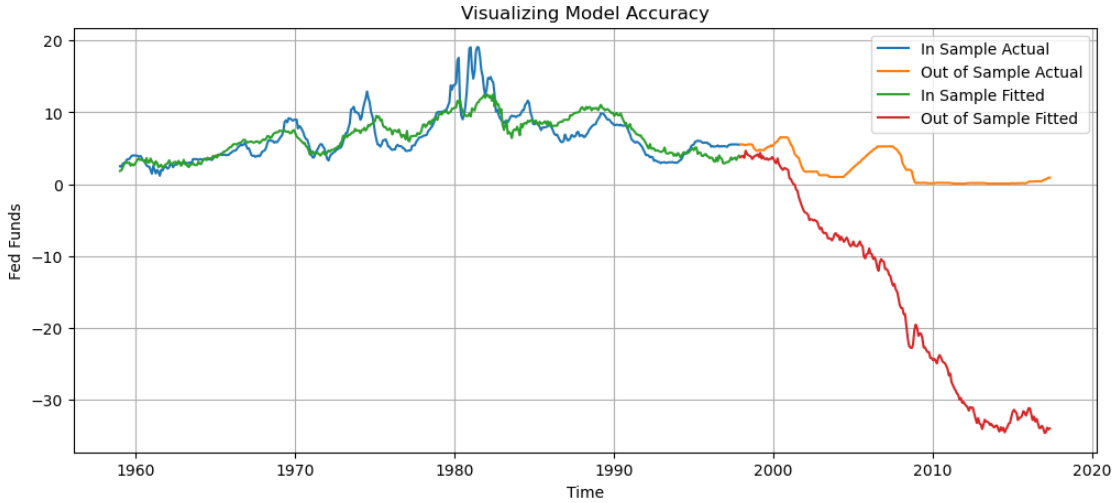
[39]:
```
for degrees in range(1,1+max_degrees):
    print("DEGREES:", degrees)
    poly = PolynomialFeatures(degree = degrees)
    X_in_poly = poly.fit_transform(X_in)
```

```
    X_out_poly = poly.transform(X_out)

    #Q3
    model1 = sm.OLS(y_in, X_in_poly).fit()

    #Q4
    plt.figure(figsize = (12,5))

    in_preds = model1.predict(X_in_poly)
    in_preds = pd.DataFrame(in_preds, index = y_in.index)
    out_preds = model1.predict(X_out_poly)
    out_preds = pd.DataFrame(out_preds, index = y_out.index)

    plt.plot(y_in)
    plt.plot(y_out)
    plt.plot(in_preds)
    plt.plot(out_preds)

    plt.ylabel("Fed Funds")
    plt.xlabel("Time")
    plt.title("Visualizing Model Accuracy")
    plt.legend(["In Sample Actual", "Out of Sample Actual", "In Sample Fitted",␣
↪"Out of Sample Fitted"])
    plt.grid()
    plt.show()

    #Q5
    in_mse_1 = mean_squared_error(y_in, model1.predict(X_in_poly))
    out_mse_1 = mean_squared_error(y_out, model1.predict(X_out_poly))
    print("Insample MSE : ", in_mse_1)
    print("Outsample MSE : ", out_mse_1)
    print("_______________________")
```

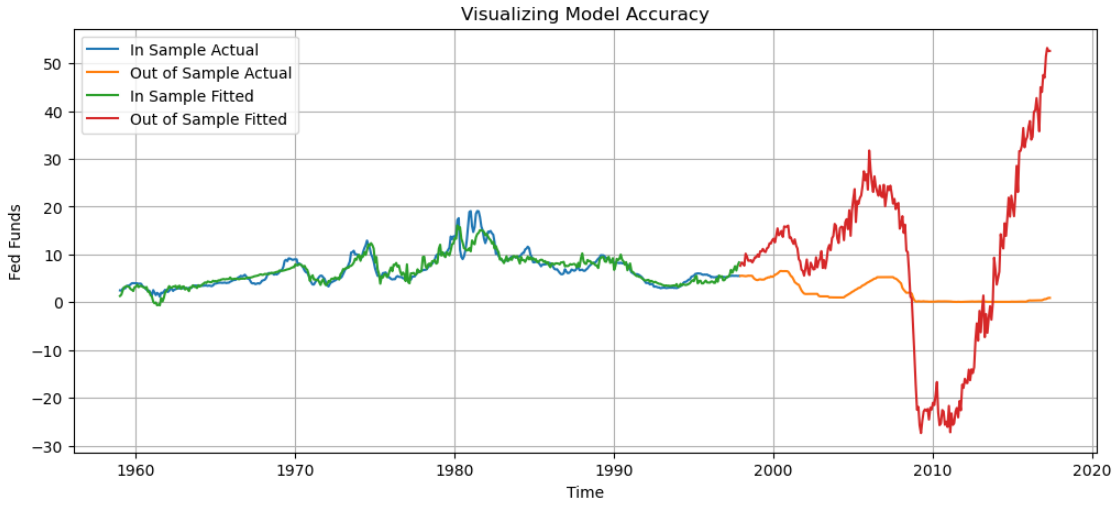
```
DEGREES: 1
```




```
Insample MSE :  10.071422013168641
Outsample MSE :  40.36082783566674

----------------------
DEGREES: 2
```




```
Insample MSE :  3.863477139276067
Outsample MSE :  481.44650990363203

----------------------
DEGREES: 3
```




```
Insample MSE :  1.872363627194615
Outsample MSE :  371.76618900618945
----------------------
```

7 7.) State your observations :

First model does not perform well. The second model performs significantly better. Model complexity is increasing, and our model is becoming more and more overfit. Variance is increasing in the models. Insample MSE is decreasing, and Outsample MSE increases and then decreases.

[ ]: