

# MTSA: 作业三

魏沐昊 502023370047

January 15, 2024

## 1 X11 Decomposition & STL Decomposition

1.X11分解是一种经典的时间序列分解方法，用于分析具有明显季节性的时间序列数据。它是美国统计局开发的一种算法，常用于季节调整和季节性分析。X11分解的过程包括以下几个步骤：

- (1).移动平均：对原始数据进行移动平均，平滑数据以减少随机波动；
- (2).趋势估计：使用一种回归模型（如线性回归）估计数据的长期趋势；
- (3).季节性调整：从移动平均数据中减去趋势估计，得到季节性成分；
- (4).季节调整：对季节性成分进行进一步调整，以消除季节性的影响；
- (5).周期调整：对具有周期性(非季节性)的数据进行调整；
- (6).不规则成分：剩余的部分被认为是不规则成分，即除趋势、季节性和周期性之外的部分。

代码实现：利用python库statsmodels.tsa.seasonal.STL来求解趋势。

2.STL分解是一种非常灵活和广泛应用的时间序列分解方法。它采用局部加权回归(Loess)方法，能够处理各种类型的时间序列数据，包括具有不稳定季节性和非线性趋势的数据。STL分解的过程包括以下几个步骤：

- (1).季节性分解：使用局部加权回归方法估计季节性成分，通过适应性窗口大小来处理不同季节的变化幅度；
- (2).趋势分解：使用局部加权回归方法估计趋势成分，通过适应性窗口大小来捕捉趋势的变化；
- (3).季节性调整：对原始数据进行季节性调整，将季节性成分从原始数据中减去；
- (4).残差：剩余部分被视为残差，表示原始数据中未被季节性和趋势解释的部分。

代码实现：利用python库statsmodels.tsa.x13.x13\_arima\_analysis来求解趋势，注意需要手动下载x12a.exe可执行文件。

## 2 ARIMA & ThetaModel & ResidualModel

1.ARIMA模型是一种广泛应用于时间序列预测和分析的经典模型。它结合了自回归(AR)和移动平均(MA)的概念，并考虑了时间序列数据的差分(Integrated)属性。ARIMA模型的核心思想是通过时间序列数据的历史值和误差进行建模，来预测未来的观测值。ARIMA模型由三个参数组成： $p$ 、 $d$ 和 $q$ ，分别表示自回归阶数、差分阶数和移动平均阶数。

(1).自回归(AR)部分：使用过去观测值的线性组合来预测当前观测值。AR部分的阶数( $p$ )表示使用的过去观测值的数量；

(2).差分(I)部分：对原始时间序列进行差分，以消除非平稳性。差分阶数( $d$ )表示进行差分的次数；

(3).移动平均(MA)部分：使用过去误差的线性组合来预测当前观测值。MA部分的阶数( $q$ )表示使用的过去误差的数量。

代码实现：调用statsmodels.tsa.ARIMA模型来。由于这个模型对实验的数据集拟合效果太差，很难达到收敛，需要很久的时间，这里只做了一部分实验。

2.Theta模型是一种用于时间序列预测的简单模型。它基于时间序列数据的简单线性回归模型，通过对误差项的建模来预测未来的观测值。Theta模型的核心思想是将时间序列数据建模为误差项的线性组合。

(1).线性回归：使用过去观测值的线性组合来预测当前观测值，类似于自回归模型；

(2).误差建模：Theta模型假设误差项服从均值为0的对称分布，如正态分布或对称三角分布。通过对误差项的建模，对未来观测值进行预测。

代码实现：调用statsmodels.tsa.forecasting.theta.ThetaModel。

3.残差模型(Residual Model)是一种用于建模和分析时间序列数据残差的方法。残差模型的目的是对残差序列建立一个合适的模型，以便更好地理解 and 解释数据中的未解释部分。

代码实现：通过集成不同模型时间序列不同部分的forecast内容可得到最终结果。(个人觉得这部分的代码还是挺简洁的，可以直接复用。)

## 3 Reproduction

上述算法均参照课上所述PPT构建，此处不再赘述也不再附上相关的图，具体在Chapter 3查看ThetaMethod相关内容，在Chapter 7查看X11、STL、ARIMA相关内容。可直接执行python run.py来复现实验结果,注意所有数据均经过Standalization变换。具体实验结果如下：

Data	Model	Decomposition	MSE	MAE
ETTh1	<u>TsfKNN</u>	moving_average	0.1868	0.2782
		differential_decomposition	1.6623	0.7373
		STL_decomposition	1.6148	0.7304
		X11_decomposition	1.6152	0.7304
ETTh2	<u>TsfKNN</u>	moving_average	0.1870	0.2785
		differential_decomposition	0.1876	0.2797
		STL_decomposition	0.1868	0.2782
		X11_decomposition	0.1869	0.2785

Figure 1: TsfKNN

Model	Data	MSE	MAE
<u>ThetaModel</u>	ETTh1	2.3237	0.8335
	ETTh2	6.0261	0.3266
<u>ResidualModel</u>	ETTh1	1.5927	0.6956
	ETTh2	1.5513	0.2683
ARIMA	ETTh1	1.7020	0.6324
	ETTh2	0.1775	0.2615

Figure 2: Other Methods