

系统工程第 1 次作业

张博睿 自 75 2017011537

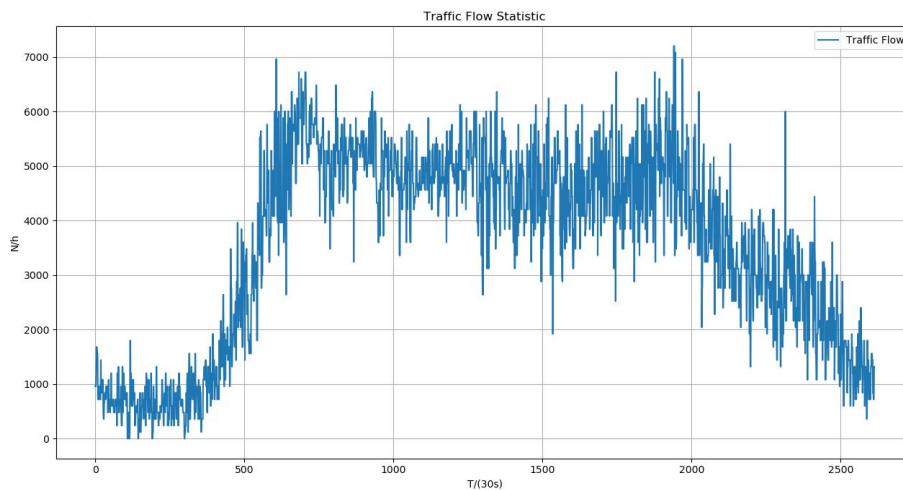
给定一个时间序列数据，data.mat，每个数据点表示 30 秒时间内国内某高速公路上车的流量，单位已经被转换成（辆/小时）

要求：

1. (2 point)使用 matlab 或者 python 对数据作可视化，包含横纵坐标、标题、网格等；

解：

使用 python 进行数据可视化结果如下：



2. (3 point)自学移动平均法。实现移动平均法，选择 $N = 5, 30$ ，分别画出平滑后的流量变化曲线；

解：

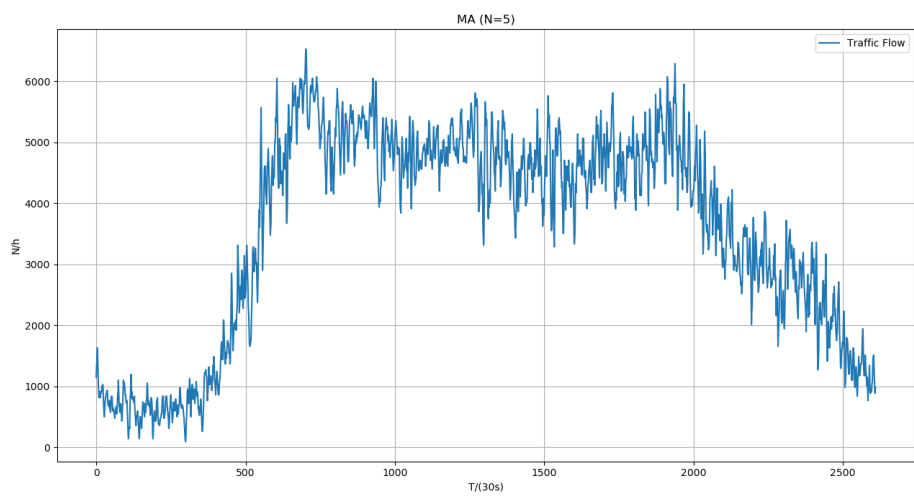
简单移动平均的计算公式为

$$F_t = \frac{A_{t-1} + A_{t-2} + \cdots + A_{t-n}}{n} \quad (1)$$

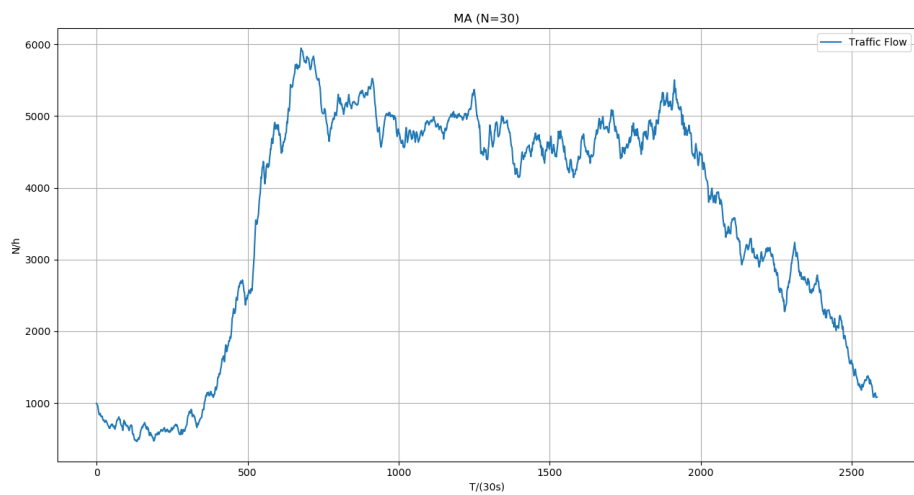
其中， F_t 是对下一期的预测值， n 是移动平均的时期个数， A_{t-a} 前 a 期的实际值。

流量变化曲线如下：

$N = 5$



$N = 30$



3. (3 point)自学指数平滑法。实现指数平滑法，选择指数 $\alpha = 0.2, 0.05$ ，分别画出平滑后的流量变化曲线；

解：

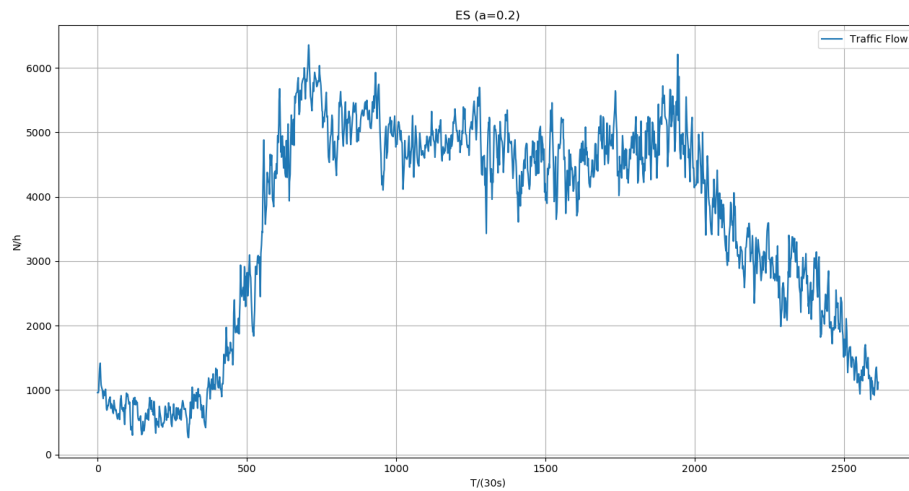
指数平滑法的基本公式为

$$S_t = a \cdot y_t + (1 - a)S_{t-1} \quad (2)$$

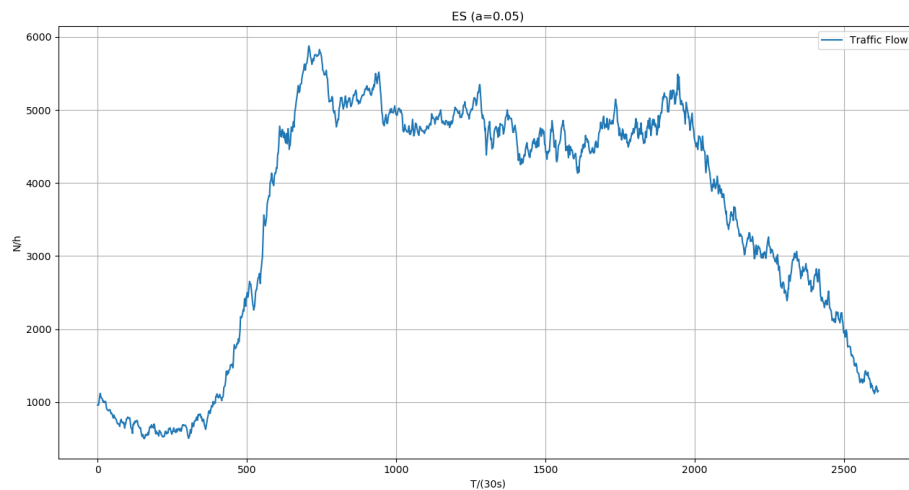
其中， S_t 表示 t 时刻的平滑值， y_t 表示 t 时刻的实际值， a 为平滑常数。

流量变化曲线如下：

$$a = 0.2$$



$$a = 0.05$$



4. (2 point)请大家推导上述两种方法的增量形式

解:

(1) 移动平均法

由移动平均法计算公式

$$F_t = \frac{A_{t-1} + A_{t-2} + \cdots + A_{t-n}}{n} \quad (1)$$

$$F_{t+1} = \frac{A_t + A_{t-1} + \cdots + A_{t-n+1}}{n} \quad (3)$$

由上述两个等式消元后得到

$$F_{t+1} = F_t + \frac{A_t - A_{t-n}}{n} \quad (4)$$

(2) 指数平均法

在一般使用过程, 指数平均法使用的是增量形式的公式

$$S_t = a \cdot y_t + (1 - a)S_{t-1} \quad (2)$$

通过递推展开的方式, 可以发现

$$\begin{aligned} S_t &= ay_t + (1 - a)S_{t-1} \\ &= ay_t + (1 - a)(ay_{t-1} + (1 - a)S_{t-2}) \\ &\dots \\ &= ay_t + (1 - a)ay_{t-1} + \cdots (1 - a)^i ay_{t-i} + \cdots \end{aligned}$$

可以看到, 通过这种平滑方式, 使得距离现在越远的实际值影响越小。

5. (BONUS 2 point): 请大家自学 ARIMA, 实现并使用 ARIMA 对数据作平滑处理, 画出平滑前后的流量变化曲线

解:

ARIMA(p, d, q)模型的公式为

$$(1 - L^d)X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \cdots + \alpha_p X_{t-p} + \epsilon_t + \beta_1 \epsilon_{t-1} + \cdots + \beta_q \epsilon_{t-q}$$

其中L表示滞后算子

由于本次作业中使用ARIMA模型进行数据的平滑, 而非数据的拟合。因此, 在编写代码的时候并非学习模型的参数, 而是类似其他方法, 人为设定权值(默认为均匀分布)。画出的平滑曲线如下

