

刹车性能分布研究及异常检测

该部分分析基于北京朝阳艾瑞特2号的3-5月的日常行驶数据进行，主要涉及：

- 从三个月的所有行驶数据中按照一定逻辑定位出刹车数据，并对于这些刹车数据做基本的统计分析；
- 利用减速过程中加速度达到最大程度的加速度均值作为衡量刹车性能的参数，利用多个刹车数据作为样本，估计出该参数应有的分布，再用置信区间做异常值检测；
- 用kernel ridge对 $P(\text{acc}|\text{v}, \text{cmd} = -57)$ 和 $P(s|\text{v}, \text{cmd} = -57)$ 进行拟合；
- 使用可视化的方式呈现出关于时间刹车性能各项参数的分布的变化。

1. 刹车数据提取

对于刹车数据的提取主要按照一下逻辑：

- 自某一帧发出为-57的command且车辆行驶速度高于某一给定阈值起始，直到车辆持续发出-57的刹车命令到车辆刹车到速度小于某一阈值截止（某一接近0的正数）。
- 滤除数据中存在的一些异常数据，如“加速度观测一直为0”（传感器非正常工作状态）和“system_status & 1 != 1”车辆处于人工接管状态等。
- 同时由于在4月中旬前后，刹车命令的底层逻辑所有更改，因此样本中，既包含起始即为-57的刹车，也包含指令为[0, -20, -40, -57]的刹车。

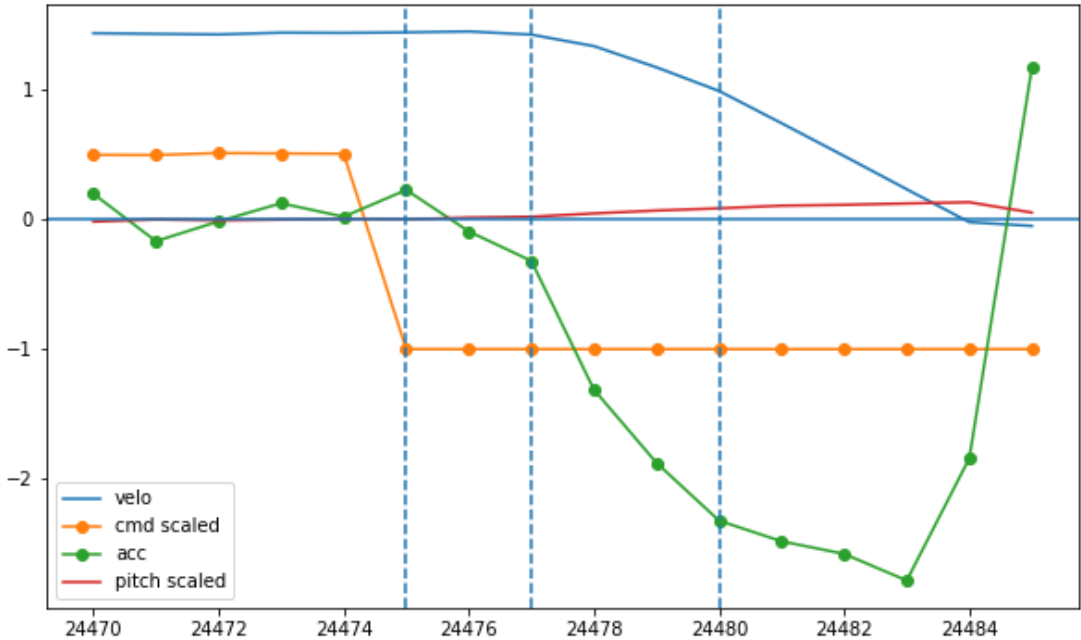
2. 刹车各阶段分析

整个刹车的过程可分为三个阶段：

1. 延迟阶段：从车辆控制系统发出车刹车指令，到机械结构实际开始执行刹车间存在一定的延迟。
2. 推程阶段：从车辆实际开始执行刹车，推刹车板并逐步增大减速度到减速度最大的过程。
3. 稳定刹车阶段：指车辆减速度稳定在最大值周围，持续进行有效刹车直到速度为0的过程。

先验经验和一些实验结果告诉我们，前两个阶段很可能都是一时间恒定的过程，因此我采取了一些方式来估计两个阶段的时间。由于对于延迟的事后分析方法，很大程度是heuristic的，因此我主要采用了“补偿延迟并极小化cmd与acceleration的RMSE”和数据可视化的方法。

结果得到第一阶段的持续时间大都在200ms-300ms，第二阶段的持续时间基本在300ms。（如下，为某一刹车阶段的示意图，其中横轴的一个刻度为100ms）



3. 刹车性能参数

分析某一刹车阶段的刹车性能，我们需要从数据中提取出一些关键的参数来分析。

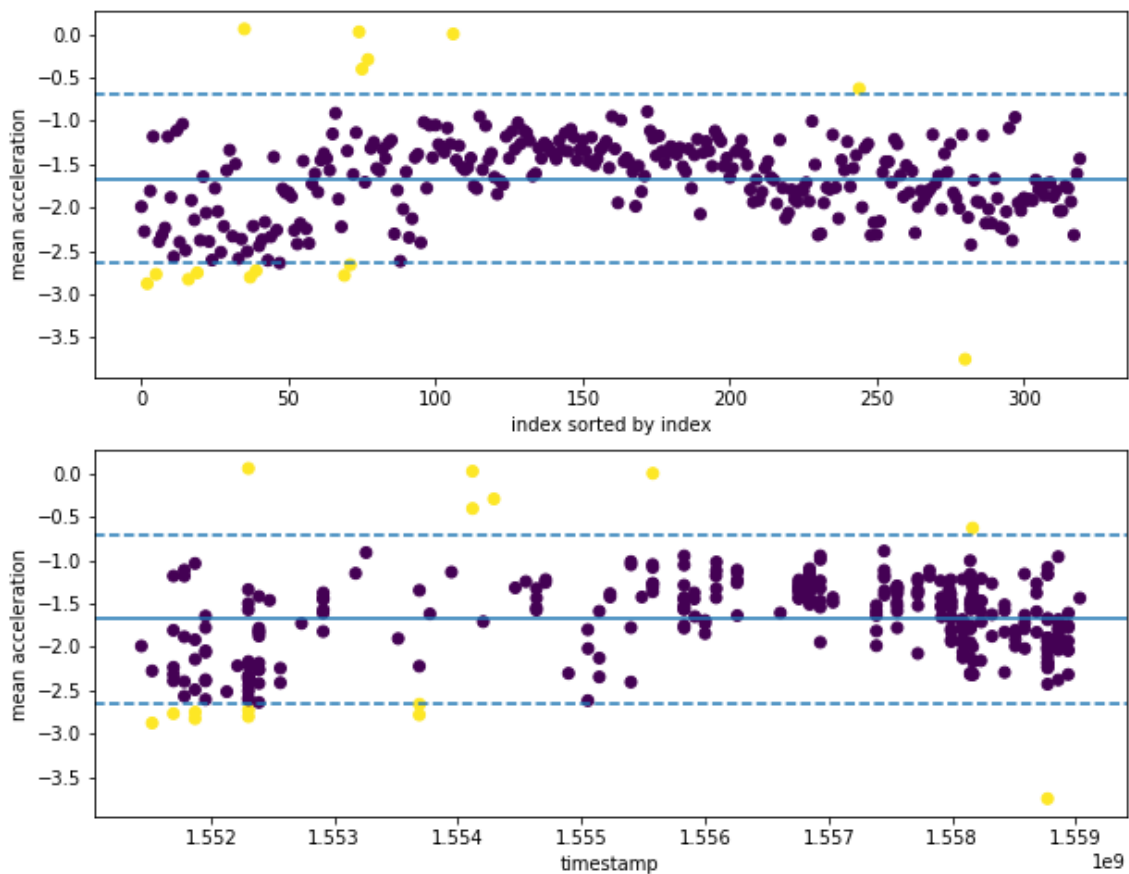
- 刹车减速度：刹车提供的制动能力的主要体现，为了更加准确地得到整个刹车过程中的稳定加速度，我采用了刹车三个阶段中的“稳定刹车阶段”的平均减速度作为估计；如果整个刹车过程太短，没达到第三阶段，则通过第二阶段中的最后一帧的减速度作为估计。
- 刹车距离：通过起始刹车速度和刹车截止速度过程中的速度值做积分，得到刹车距离的估计值。
- 刹车时间：刹车截止时间减去刹车开始时间，由于我们截取数据的频率并不高，因此该参数才在较大误差，因此不建议做定量分析。

4. 异常检测

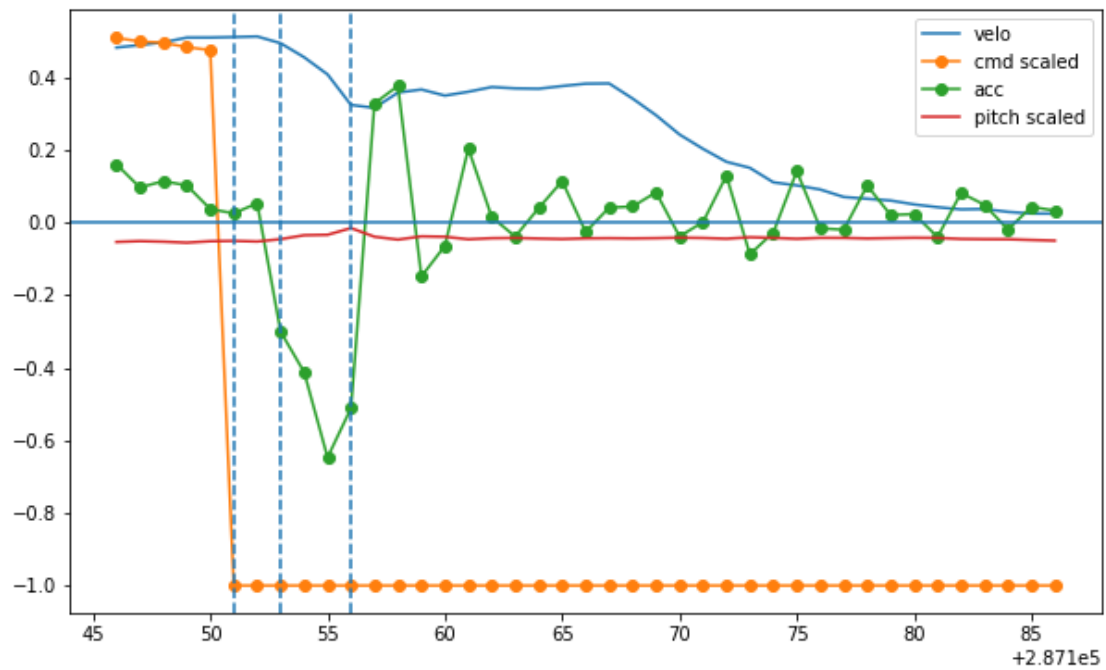
对于刹车样本的异常值的检测，我们假设给定坡度 (pitch) 和初始速度 (v) 及刹车的指令 (cmd==57)，正常性能的车辆的减速度应当服从某一正态分布。

然后我们可以通过假设检验/置信区间的方式，对各个样本的“刹车减速度”做t检验进行异常值检测。即假如我们的正态分布假设成立，我们将以95%的置信度保证刹车减速度会落在我们得到的95%置信区间内。

如果某一观测值偏离了分布的中心，我们则有理由相信，假设不再成立，即该观测并不来自于我们假设的正态分布，车辆的刹车可能出现了系统性上的改变。



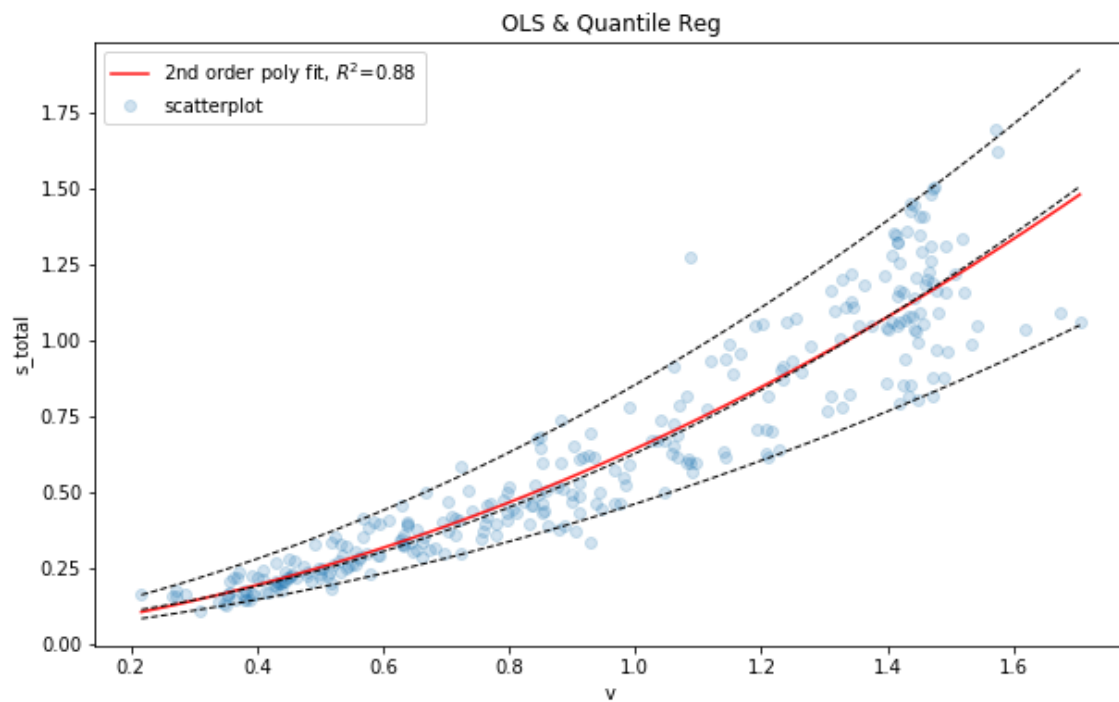
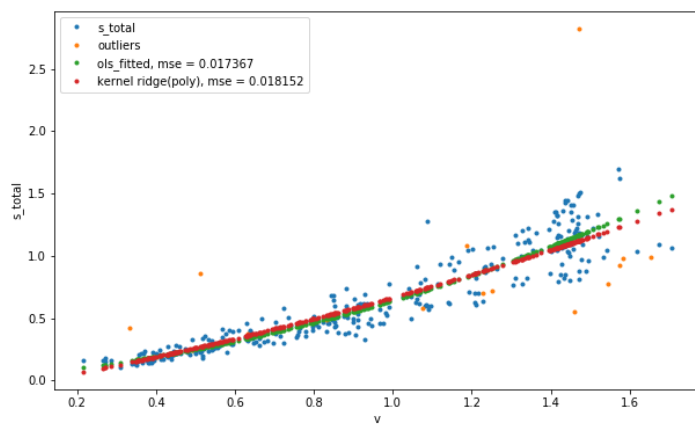
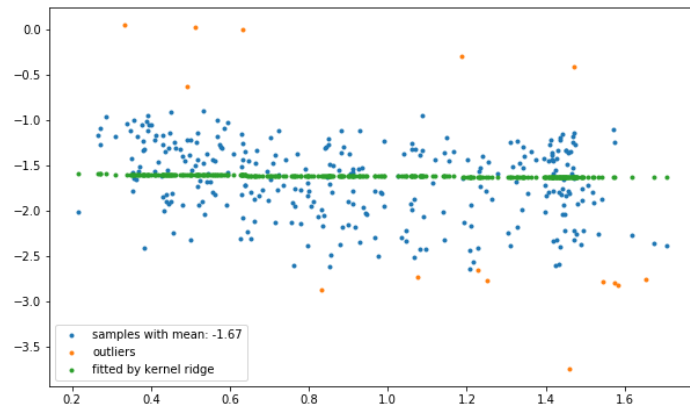
如图，是通过置信区间得到的3-5月样本中的异常值(黄色散点)的示意图（基于95%置信区间，2个样本标准差）。其中一个异常点的刹车具体过程如下，



我们可以看到，虽然车辆控制指令一直在发出-57的刹车指令，车辆却没有正常的稳定减速度，导致车辆的刹车距离与刹车时间很长。

5. 分布拟合

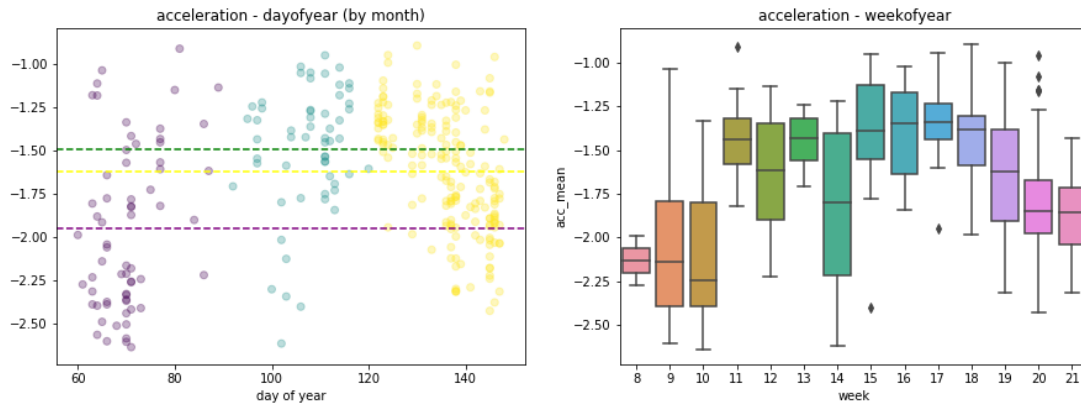
我还对刹车过程中的一些分布进行了拟合，以更加准确的方式来具体化各参数间的联合/条件分布。我对于 $P(\text{acc}|\text{v}, \text{cmd}=-57)$ 和 $P(\text{s}|\text{v}, \text{cmd}=-57)$ 分别利用了ridge kernel进行了拟合，同时对 $P(\text{s}|\text{v}, \text{cmd}=-57)$ 也使用了非参中的分位数回归(可作为另一种异常值检验的参考)。得到的结果如下



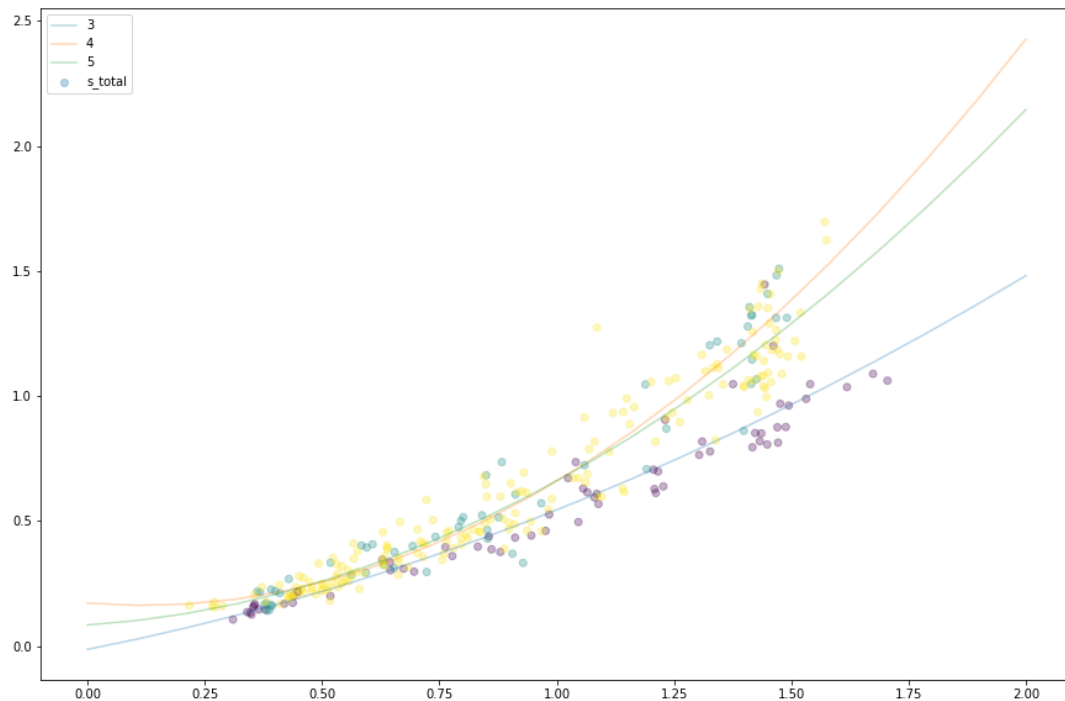
由于该部分并非我们分析的重点，因此对于模型超参数的选取采用了简单的grid选取。从结果中我们可以看到减速度的分布和初始速度并无太大关系，且刹车距离的分布呈现方差逐渐增大的初始速度的平方关系。

6. Time decay及其他分析

该分析的另一个方面旨在统计刹车性能的参数随着时间的迁移是否存在明显的time decay。我采用了统计中分析MCMC常用的stationary plot来观测是否存在明显的变化趋势。



如图右的stationary plot并不能看出随着时间变化刹车加速度存在着什么明显的变化趋势。但是早期的数据和后期的数据好像也并不来自同一分布。



然后，我对三个月的数据分别进行了多项式回归(2次)，发现三月数据的“刹车距离-初始速度”的函数关系似乎和四月与五月有所不同。

此外，其他的一些分析内容还包括了刹车时间的变化趋势图(但由于初始速度不同，参考意义不大)和pitch与刹车加速度的关联分析。更多的具体内容可以参考gitlab中的notebook file。