1. 小波软硬阈值折衷法——随机噪声

通常情况下，我们在从设备上采集到的信号都是具有一定的噪声的，大多数情况下，可认

为这种噪声为高斯白噪声。被噪声污染的信号=干净的信号+噪声。 为什么要使用阈值：由

于信号在空间上(或者时间域)是有一定连续性的，因此在小波域，有效信号所产生的小波系

数其模值往往较大；而高斯白噪声在空间上(或者时间域)是没有连续性的，因此噪声经过小

波变换，在小波阈仍然表现为很强的随机性，通常仍认为是高斯白噪的。 那么就得到这样

一个结论：在小波域，有效信号对应的系数很大，而噪声对应的系数很小。

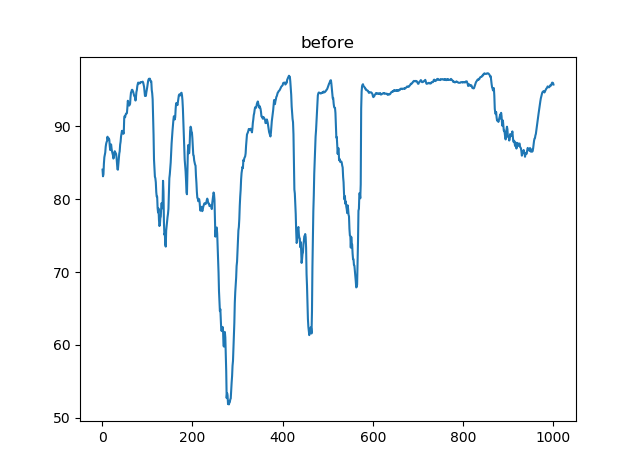


图1-小波去噪前信号

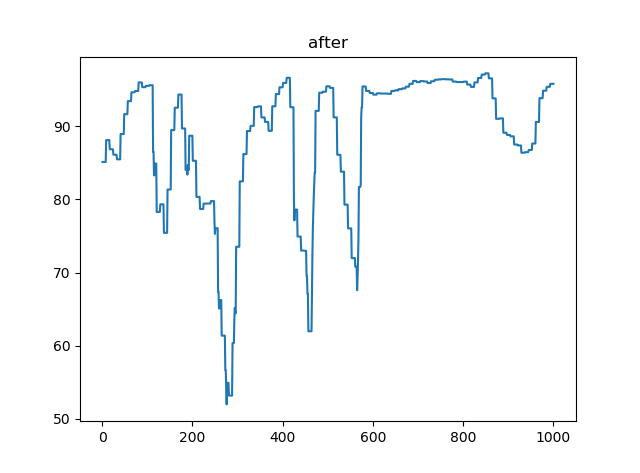


图2-小波去噪后信号

2.回归和聚类---异常值检测

 当遇到一组数据中需要剔除一下有明显异常的值，需要将其剔除。异常值：指样本中的个别值，其数值明显偏离它（或他们）所属样本的其余观测值，也称异常数据。异常是在数据集中与众不同的数据，使人怀疑这些数据并非随机偏差，而是产生于完全不同的机制。聚类算法对异常的定义：异常是聚类嵌于其中的背景噪声。异常检测算法对异常的定义：异常是既不属于聚类也不属于背景噪声的点。它的行为与正常的行为有显著的不同。在某个季节里，某一天的气温很高或很低，这个温度数据就是一个异常。它用来发现“小的模式”(相对于聚类)，即数据集中间显著不同于其它数据的对象。

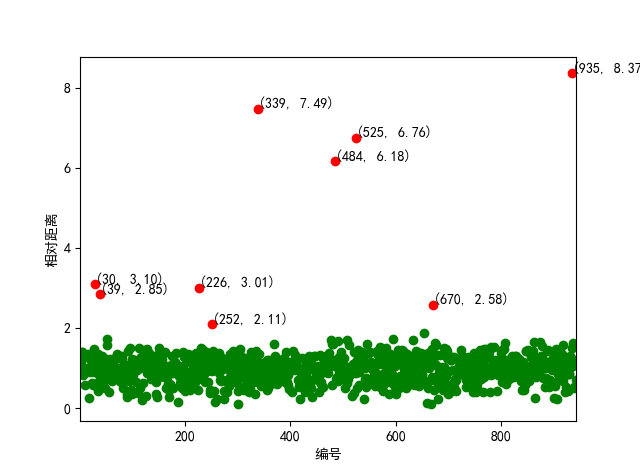


图3-聚类算法检测异常值

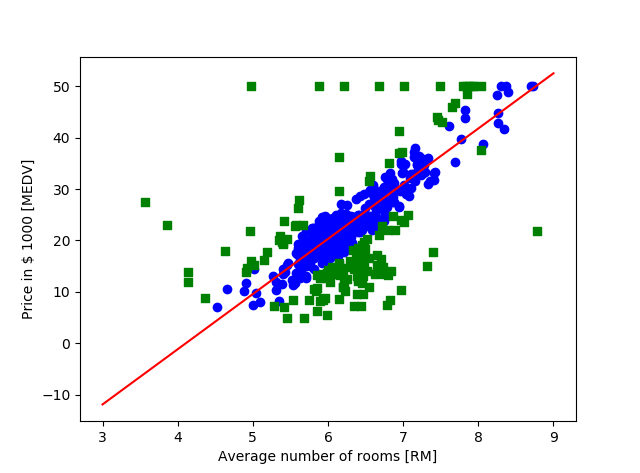


图4-回归检测异常值（设定为线性回归）