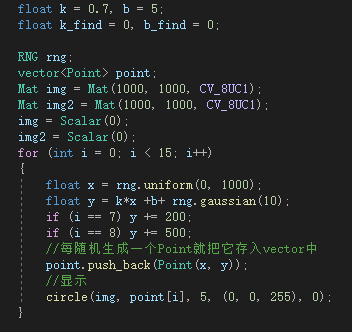
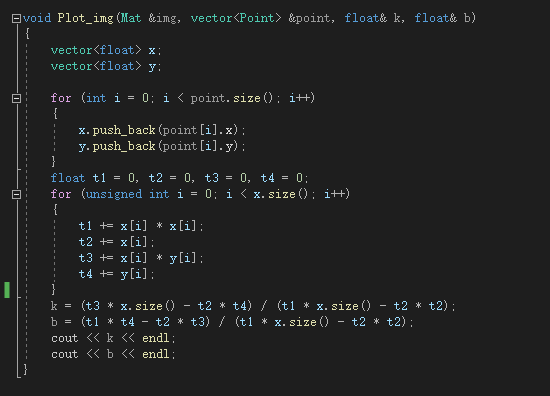
本程序采用最小二乘法进行直线拟合，通过计算点与拟合直线的距离进行离群点剔除。

具体流程结合代码进行展示：



这部分代码用于生成一组带噪音的点，并把第7和第8个点的y值增加200和500人为设置成离群点。



这部分代码用于生成拟合直线，采用最小二乘法，k和b的计算采用的是最小二乘法的公式。

直线拟合后未剔除离群点的拟合直线如图1：

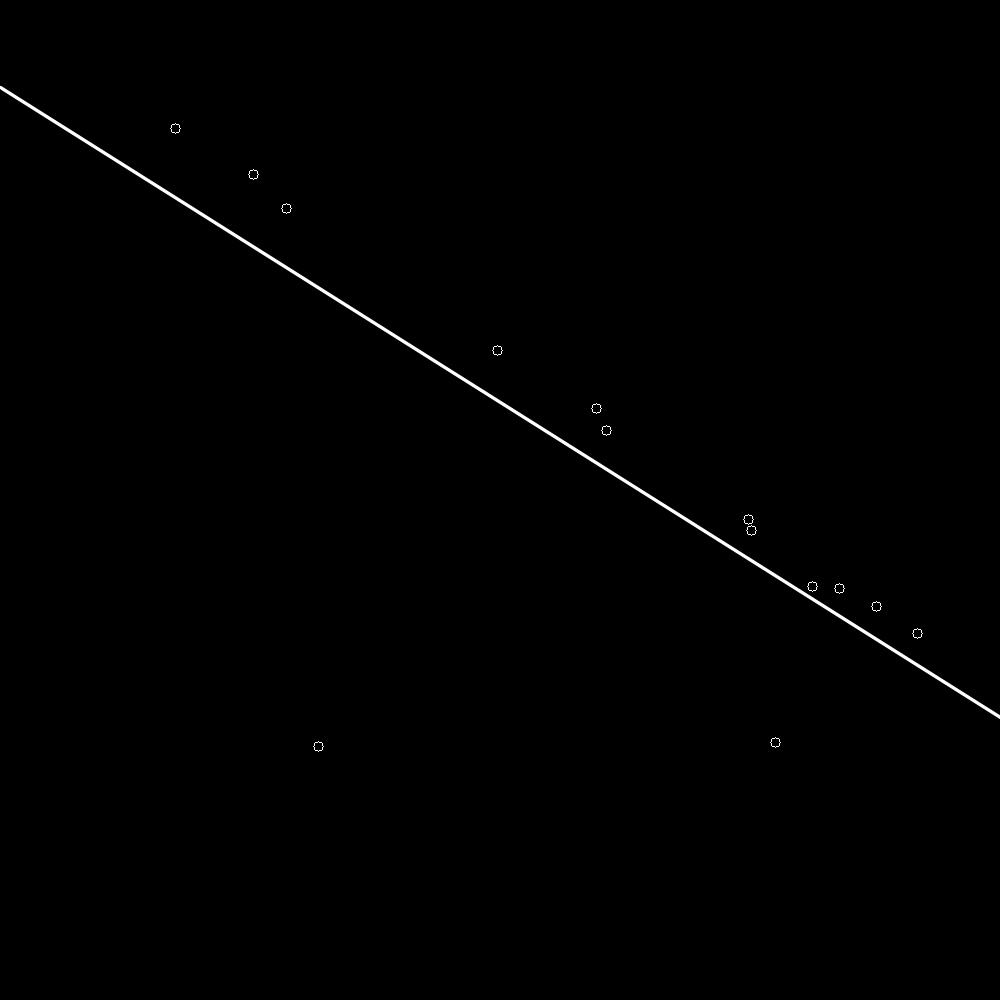
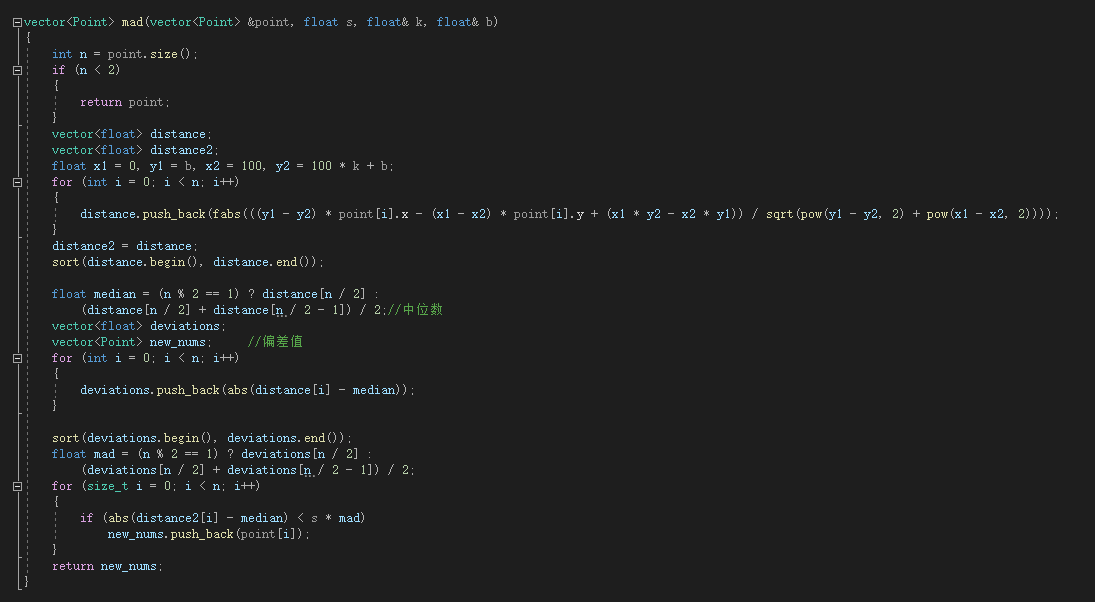


图1. 拟合直线未剔除离群点



这部分代码用于剔除离群点，主要方法是通过：

1. 计算所有点分别与所拟合直线的距离
2. 对距离进行排序，找出中间值
3. 计算每个距离与距离中位数的偏差值
4. 对偏差值进行排序，找出偏差值的中位数作为基准
5. 将所有偏差值与这个中位数缩放后的数值进行比较，其中s是缩放因子，用于调整可接受的离群值偏差范围，例如，s=1即为接受偏差最小的一半数据。如图2为s=1，图3为s=2。重新拟合直线即可获得剔除离群点的直线。

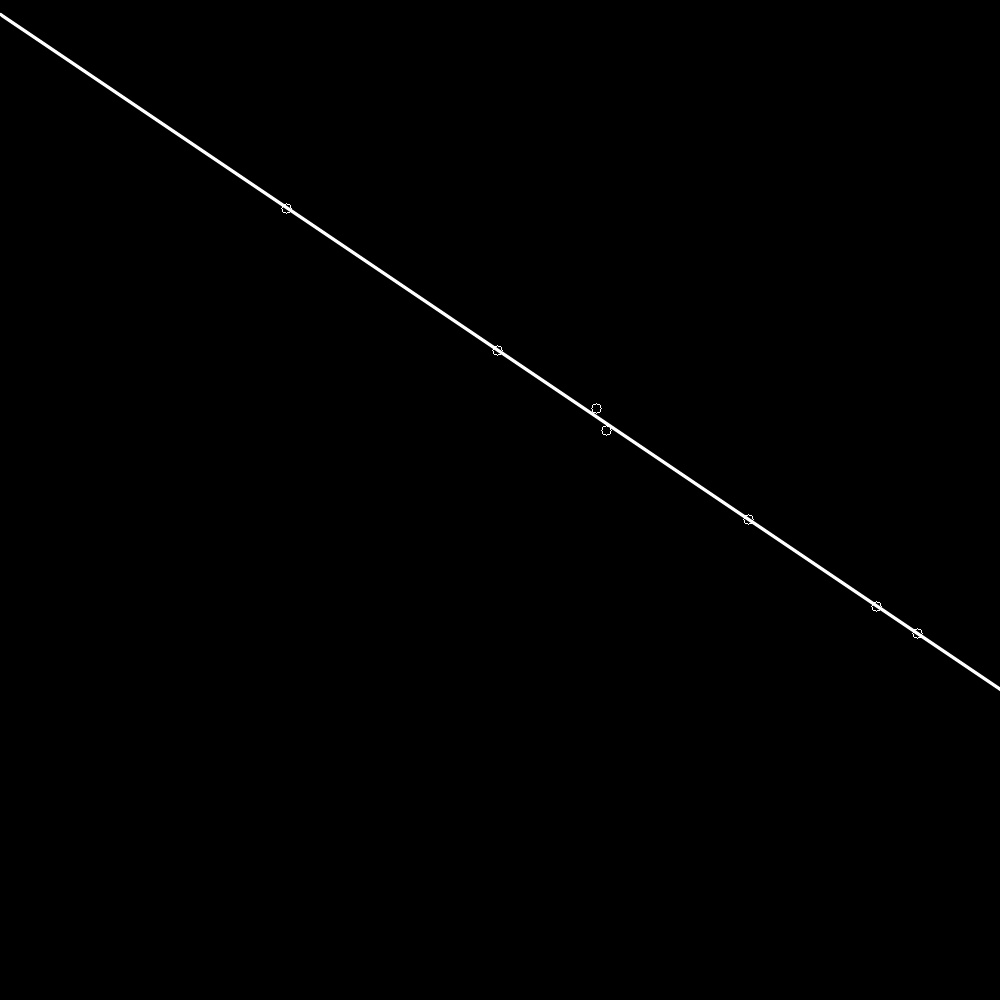


图2 剔除离群点拟合直线s=1

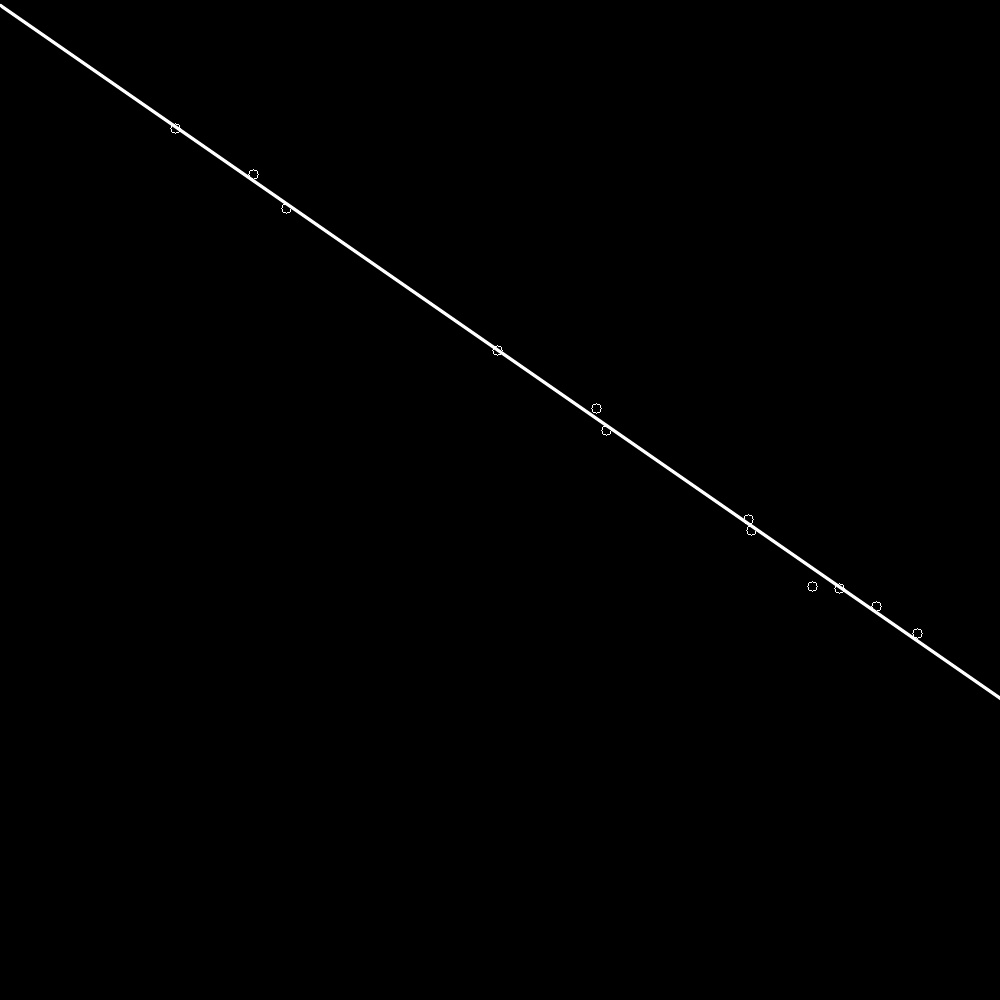


图3 剔除离群点拟合直线s=2