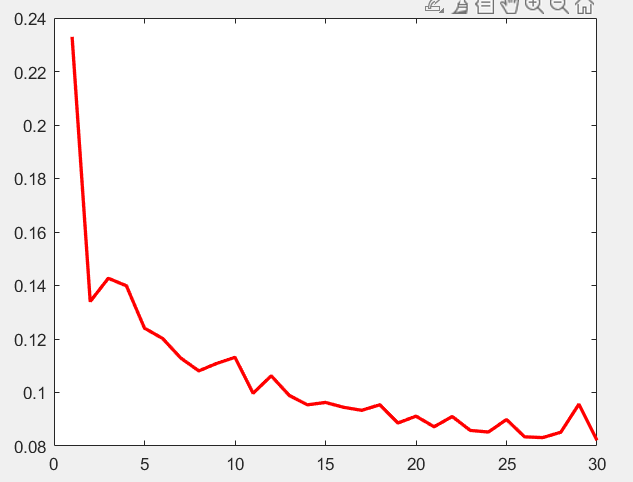


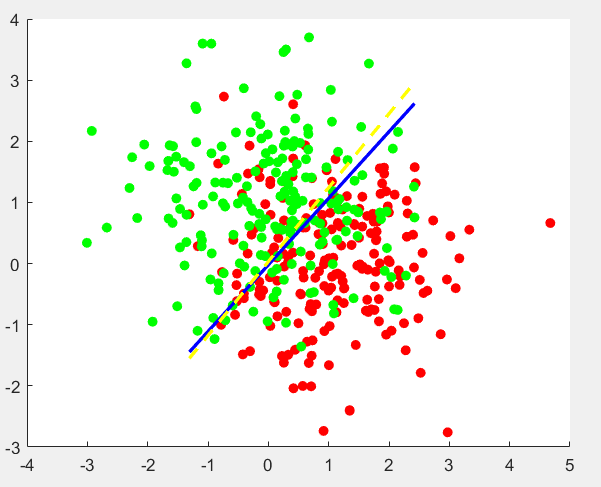
蓝色为广义逆计算分类面，黄色为随机梯度下降得到分类面，其中batchsize=20，epoch=30，步长为定长0.05。

分类正确率均为100%。



损失函数随epoch变化曲线如上图所示。

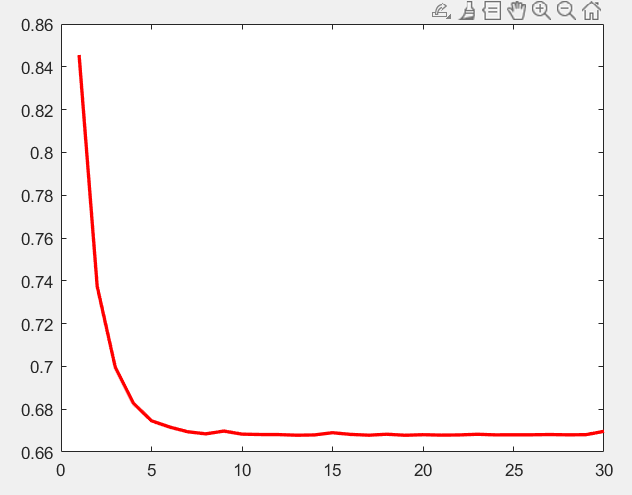
改变样本后分类面如下：



广义逆训练集错误个数为70，测试集错误个数为26；

随机梯度下降训练集错误个数为72，测试集错误个数为25；

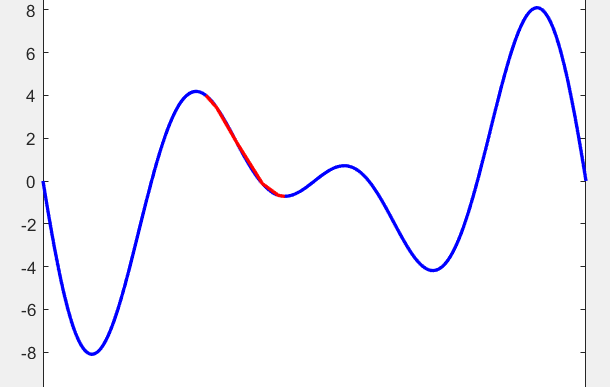
损失函数随epoch变化曲线如下：



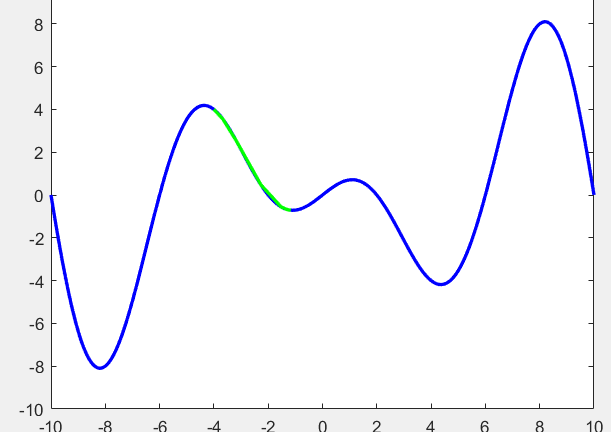
改变各类超参后发现，对于样本是线性可分的情况，广义逆和SGD均能取得较好成绩。SGD在batch很小时，损失函数随epoch变化曲线不平滑，可能突然增大。在batch很大时，分类结果可能有少数错误，损失函数随epoch整体呈平滑下降趋势。步长太大时不收敛，步长太小时w更新太慢，最终分类面无法较好分类。

线性不可分时，两种方法均无法较好分类。

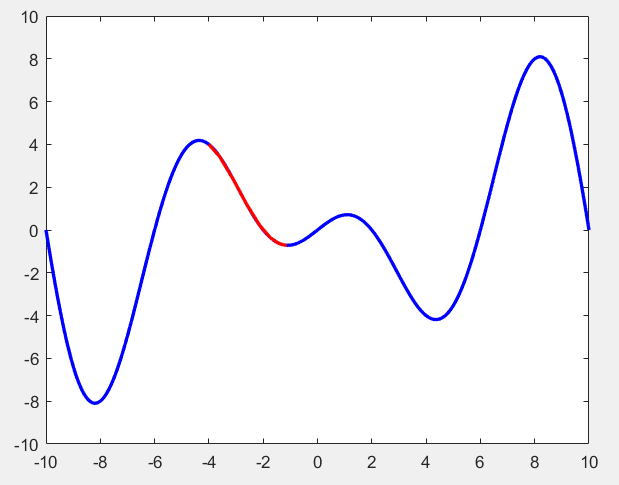
迭代十次：



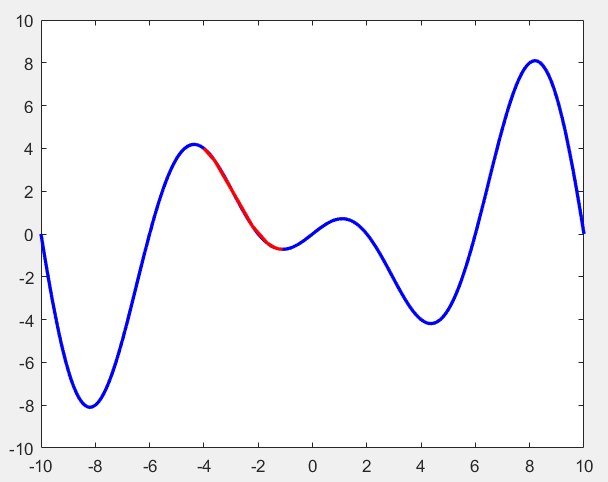
梯度下降法



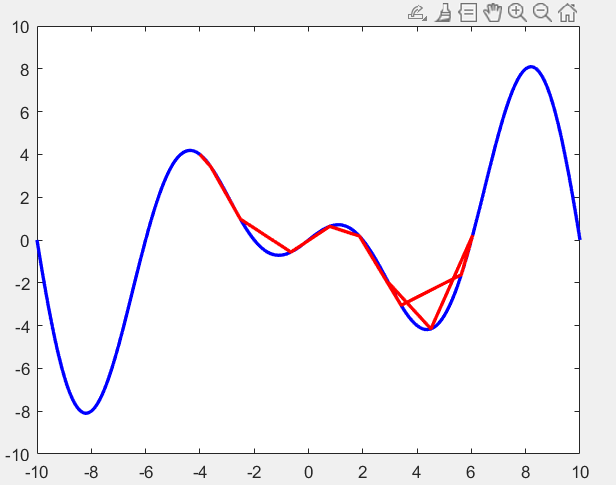
随机梯度下降



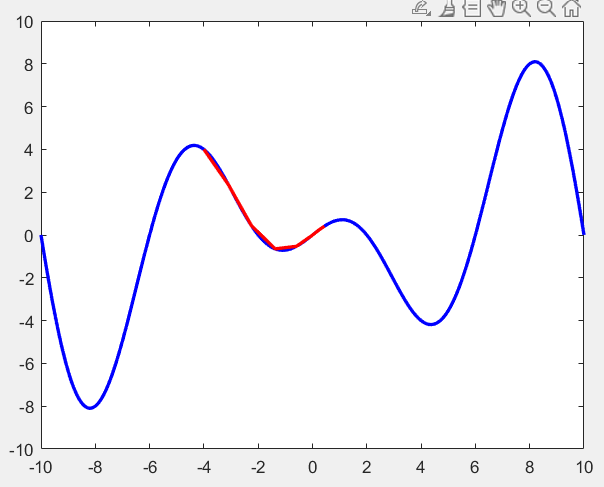
Adagrad



RMSProp

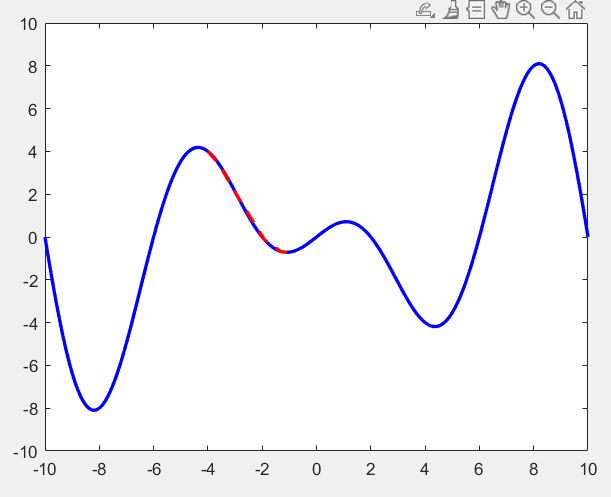


动量法

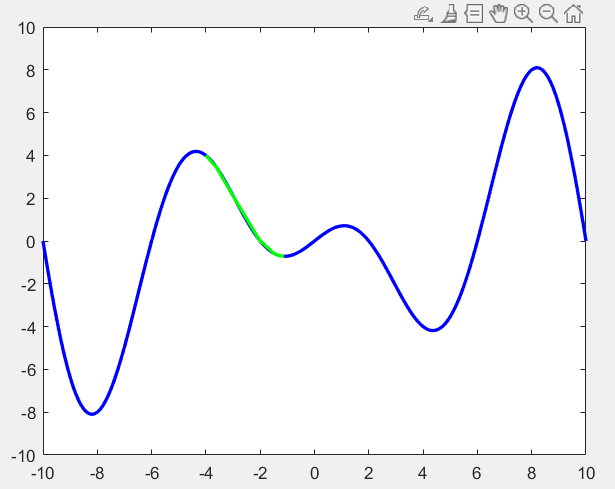


Adam

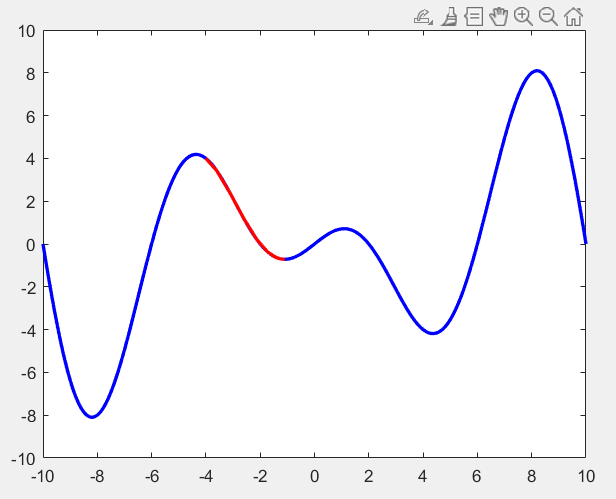
迭代50次时：



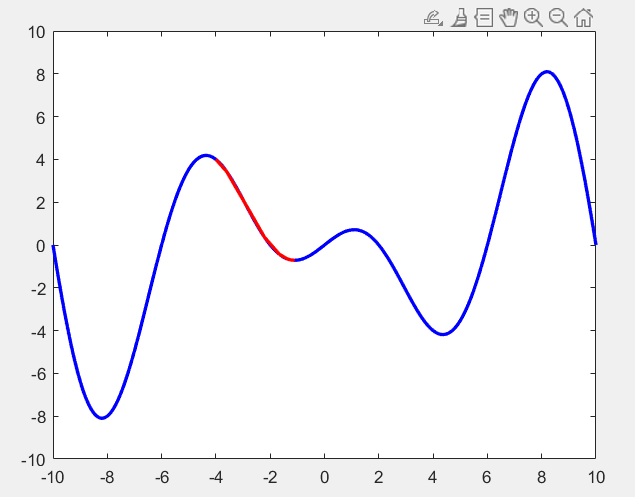
梯度下降法



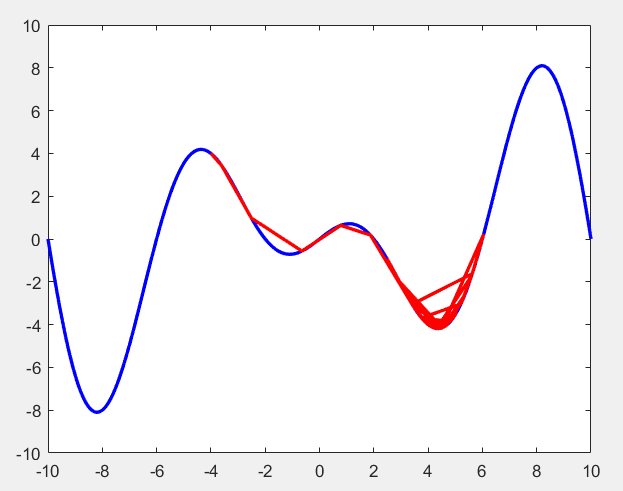
随机梯度下降



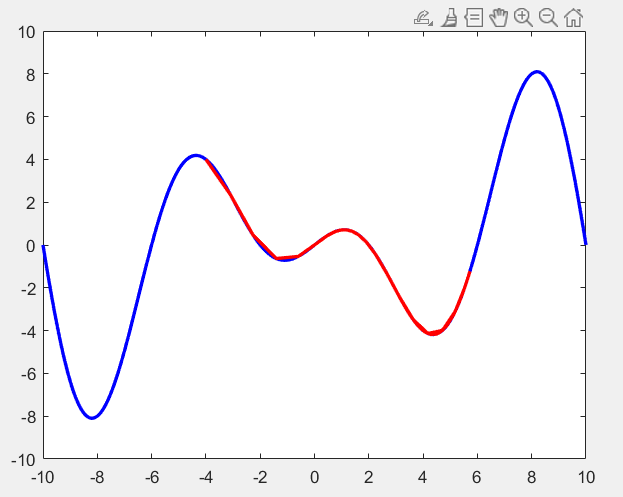
Adagrad



RMSProp



动量法（最终落在4.16）



Adam(最终落在4.67)