Lecture7-8编程作业

李星毅 U201712072 自实1701

训练集: 中国与日本的沿海城市的经纬度坐标向量，中国标签为+1, 日本为标签为-1.

测试集: 钓鱼岛的经纬度坐标向量

用支撑向量机设计分类器，（1）判断钓鱼岛属于哪一类；（2）增加几个非海边城市的经纬度坐标进行训练，判断这些城市是否影响分类结果，是否为支撑向量。

1. 训练集：

中国

[119.28, 26.08] 福州

[121.31, 25.03] 台北

[121.47, 31.23] 上海

[118.06, 24.27] 厦门

[113.53, 29.58] 武汉

[104.06, 30.67] 成都

[116.25, 39.54] 北京

[121.46, 39.04] 大连

[122.10, 37.50] 威海

[124.23, 40.07] 丹东

日本

[129.87, 32.75] 长崎

[130.33, 31.36] 鹿儿岛

[131.42, 31.91] 宫崎

[130.24, 33.35] 福冈

[136.54, 35.10] 名古屋

[132.27, 34.24] 广岛

[139.46, 35.42] 东京

[133.33, 15.43] 鸟取

[138.38, 34.98] 静冈

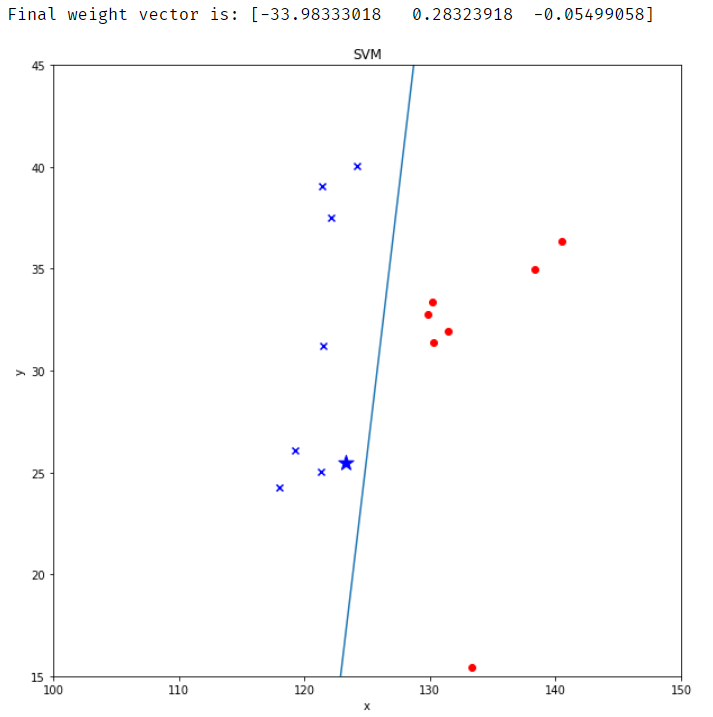
[140.47, 36.37] 水户

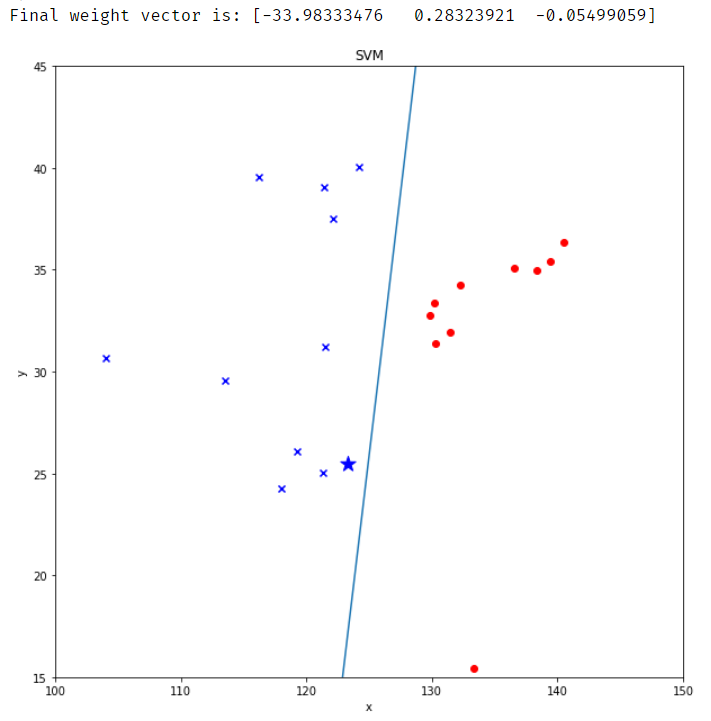
1. 测试集：

[123.28, 25.45] 钓鱼岛

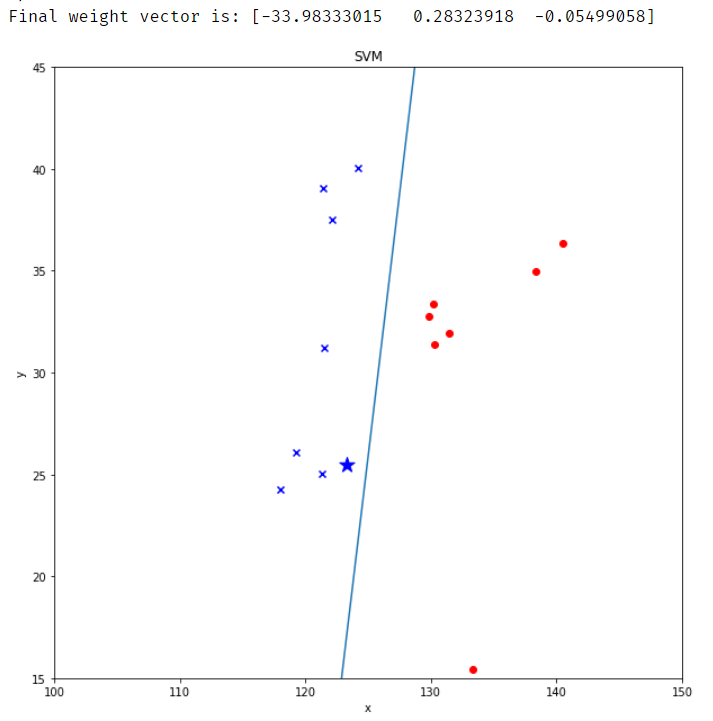
1. 结果展示（蓝色代表中国，红色代表日本，星星表示钓鱼岛）

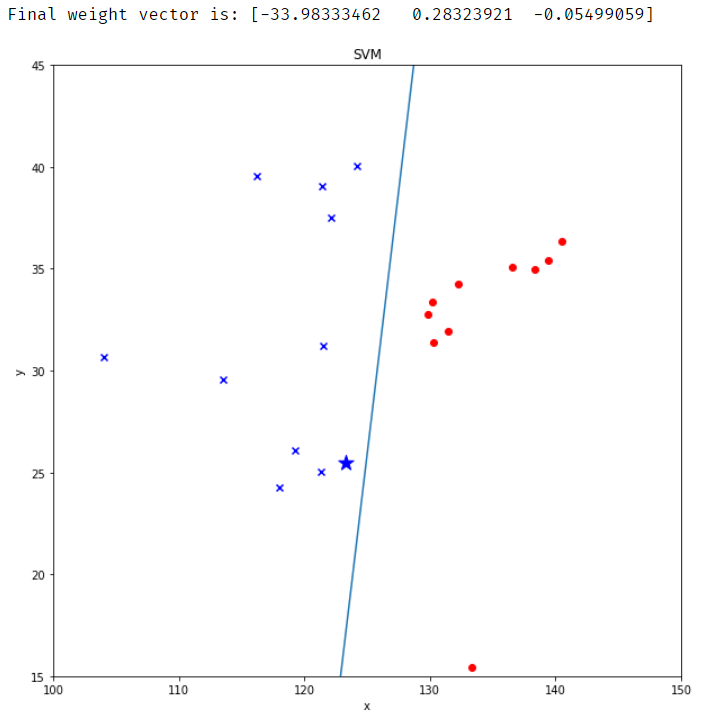
（1）Primal SVM



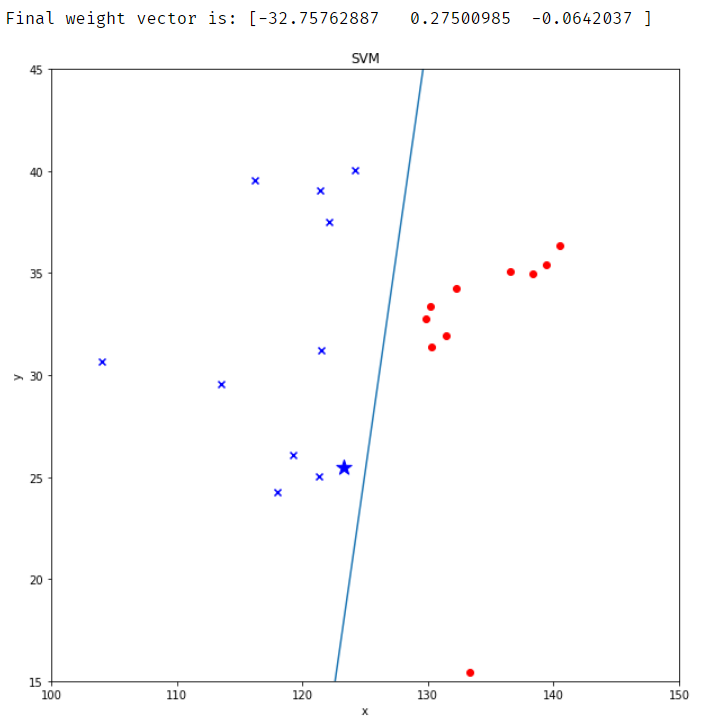
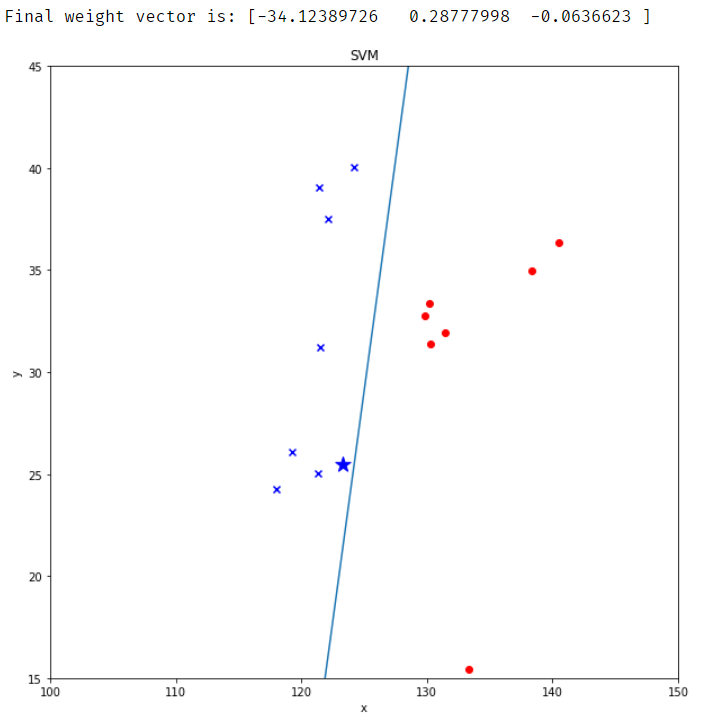


（2）Dual SVM





（3）SVM Pegasos次梯度下降



1. 实验分析
2. 通过Primal SVM、Dual SVM和SVM Pegasos都可以看出钓鱼岛属于中国，且内陆城市不影响分类结果，均不是支撑向量。
3. Primal SVM和Dual SVM的解几乎一致，稍有差别应该是Python CVXOPT计算的误差。
4. SVM Pegasos次梯度下降算法基于Hinge Loss，加入权重模长平方的正则项，使用迭代寻优的方法逼近局部最优解。可以看到，结果与Primal SVM和Dual SVM差不多。