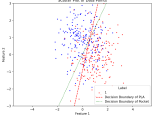


L9T4实验记录

	epoch_num=3			epoch_num=1		
	lr=0.1	lr=0.25	lr=0.005	lr=0.1	lr=0.25	lr=0.005
test_acc_PLA	0.7	0.7	0.7	0.7	0.7	0.7
test_acc_Pocket	0.775	0.775	0.775	0.7625	0.7625	0.7625
runtime_PLA	0.00252	0.00299	0.00300	0.00099	0.00199	0.00199
runtime_Pocket	0.01604	0.01501	0.01300	0.00300	0.00410	0.00799
plot						

数据配置说明

以上数据集配置与第三题同，即均值向量为[0, 1]与[1, 0],协方差为单位矩阵，两类样本各200个，且设置随机种子以保证数据的随机性不变。

超参数说明

- epoch轮数
- lr学习率

结果分析

1. 在epoch轮次一定时，如图三种学习率下，结果的准确度总是保持不变

可能的原因：由于两类别数据分布过于接近，已属于线性不可分数据，感知机的效果有上限，而经过 $3 \times 200 \times 2 = 1200$ 次迭代（或 $1 \times 200 \times 2 = 400$ ），即使是0.005的学习率，也能达到模型学习的上限。

2. 在相同条件下，pocket算法总比一般pla算法在测试集上总有更高的准确性与更长的运行时间

可能的原因：对于线性不可分数据，pocket算法能得到pla算法的最优化结果，而由于pocket算法流程更复杂，涉及到更多的指令，运行时间会略长。

3. 在学习率一定时，epoch数为3时，pocket算法模型的效果优于epoch数为1时的模型效果，而pla算法效果一致

可能的原因：对于一轮迭代，pocket算法接近但仍未收敛到其模型上限，因此加强训练轮数，在一定程度上有助于模型效果的提升；而pla算法更simple，相较于pocket算法更易收敛，因此一轮迭代已收敛，增加训练轮数对模型效果无影响。