Answer of Assignment 3

第 5 章：线性判别函数

2020E8017782032_ 蒲尧

第一部分：计算与证明

1. 现有四个来自于两个类别的二维空间中的样本，其中第一类的两个样本为 $(1,4)^T$ 和 $(2,3)^T$ ，第二类的两个样本为 $(4,1)^T$ 和 $(3,2)^T$ 。这里，上标 T 表示向量转置。若采用规范化增广样本表示形式，并假设初始的权向量 $\boldsymbol{a}=(0,1,0)^T$ ，其中向量 $\boldsymbol{a}$ 的第三维对应于样本的齐次坐标。同时，假定梯度更新步长 η_k 固定为 1。试利用批处理感知器算法求解线性判别函数 $g(\boldsymbol{y})=\boldsymbol{a}^T\boldsymbol{y}$ 的权向量 $\boldsymbol{a}$。（注：“规范化增广样本表示”是指对齐次坐标表示的样本进行规范化处理）。

2. 对于多类分类情形，考虑 one-vs-all 技巧，即构建 c 个线性判别函数：

$$g_i(\boldsymbol{x})=\boldsymbol{w}_i^T\boldsymbol{x}+w_{i0}, i=1,2,...,c,$$

此时的决策规则为：对 $j\neq i$, 如果 $g_i(\boldsymbol{x})>g_j(\boldsymbol{x})$, x 则被分为 ω_i 类。现有三个二维空间内的模式分类器，其判别函数为：

$$g_1(\boldsymbol{x})=-x_1+x_2$$
$$g_2(\boldsymbol{x})=x_1+x_2-1$$
$$g_3(\boldsymbol{x})=-x_2$$

试画出决策面，指出为何此时不存在分类不确定性区域。

第二部分：计算机编程

本章所使用的数据：

	ω_1		ω_2		ω_3		ω_4	
sample	x_1	x_2	x_1	x_2	x_1	x_2	x_1	x_2
1	0.1	1.1	7.1	4.2	−3.0	−2.9	−2.0	−8.4
2	6.8	7.1	−1.4	−4.3	0.5	8.7	−8.9	0.2
3	−3.5	−4.1	4.5	0.0	2.9	2.1	−4.2	−7.7
4	2.0	2.7	6.3	1.6	−0.1	5.2	−8.5	−3.2
5	4.1	2.8	4.2	1.9	−4.0	2.2	−6.7	−4.0
6	3.1	5.0	1.4	−3.2	−1.3	3.7	−0.5	−9.2
7	−0.8	−1.3	2.4	−4.0	−3.4	6.2	−5.3	−6.7
8	0.9	1.2	2.5	−6.1	−4.1	3.4	−8.7	−6.4
9	5.0	6.4	8.4	3.7	−5.1	1.6	−7.1	−9.7
10	3.9	4.0	4.1	−2.2	1.9	5.1	−8.0	−6.3

1. Write a program to implement the "batch perceptron" algorithm.

   (a). Starting with $\boldsymbol{a} = 0$, apply your program to the training data from ω_1 and ω_2 . Note that the number of iterations required for convergence（即记录下收敛的步数）。

   (b). Apply your program to the training data from ω_3 and ω_2 . Again, note that the number of iterations required for convergence.

2. Implement the Ho-Kashyap algorithm and apply it to the training data from ω_1 and ω_3 . Repeat to apply it to the training data from ω_2 and ω_4 . Point out the training errors, and give some analyses.

3. 请写一个程序，实现MSE 多类扩展方法。每一类用前 8 个样本来构造分类器，用后两个样本作测试。请给出你的正确率。

第一部分回答

1. 由题可知：

$$a = \left(0, 1, 0\right)^T$$

$$\omega_1 : x_1 = \begin{pmatrix} 1 \\ 4 \end{pmatrix}, x_2 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}; \omega_2 : x_1 = \begin{pmatrix} 4 \\ 1 \end{pmatrix}, x_2 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

$$y_1 = \begin{pmatrix} 1 \\ 4 \\ 1 \end{pmatrix}, y_2 = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}, y_3 = -\begin{pmatrix} 4 \\ 1 \\ 1 \end{pmatrix}, y_4 = -\begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$$

所以：

$g(y_1) = a^T y_1 = 4 > 0$ (正确分类) , $g(y_2) = a^T y_2 = 3 > 0$ (正确分类) ;

$g(y_3) = a^T y_3 = -1 < 0$ (错误分类) , $g(y_4) = a^T y_4 = -2 < 0$ (错误分类)

第一次修正：

$$a = a + \eta_k \sum_{y \in Y(k)} y = a + y_3 + y_4 = -\left(7, 2, 2\right)^T$$

此时：

$g(y_1) = a^T y_1 = -17 < 0$(错误分类) , $g(y_2) = a^T y_2 = -22 > 0$ (错误分类) ;

$g(y_3) = a^T y_3 = 32 > 0$ (正确分类) , $g(y_4) = a^T y_4 = 27 > 0$ (正确分类)

第二次修正：

$$a = a + \eta_k \sum_{y \in Y(k)} y = a + y_1 + y_2 = (-4, 5, 0)^T$$

此时：

$g(y_1) = a^T y_1 = 16 > 0$(正确分类) , $g(y_2) = a^T y_2 = 7 > 0$ (正确分类) ;

$g(y_3) = a^T y_3 = 11 > 0$ (正确分类) , $g(y_4) = a^T y_4 = 2 > 0$ (正确分类)

$\therefore$ 不再修正，$a = (-4, 5, 0)^T$

2. 属于第 1 类的区域应该满足：$g_1(x) > g_2(x), g_1(x) > g_3(x)$
   故 ω_1 的决策面为:
   $g_1(x) - g_2(x) = -2x_1 + 1 = 0$
   $g_1(x) - g_3(x) = -x_1 + 2x_2 = 0$
   属于第 2 类的区域应该满足：$g_2(x) > g_1(x), g_2(x) > g_3(x)$
   故 ω_2 的决策面为:
   $g_2(x) - g_1(x) = 2x_1 - 1 = 0$
   $g_2(x) - g_3(x) = x_1 + 2x_2 - 1 = 0$
   属于第 3 类的区域应该满足：$g_3(x) > g_1(x), g_3(x) > g_2(x)$
   故 ω_3 的决策面为:
   $g_3(x) - g_1(x) = x_1 - 2x_2 = 0$
   $g_3(x) - g_2(x) = -x_1 - 2x_2 + 1 = 0$
   绘图如下：

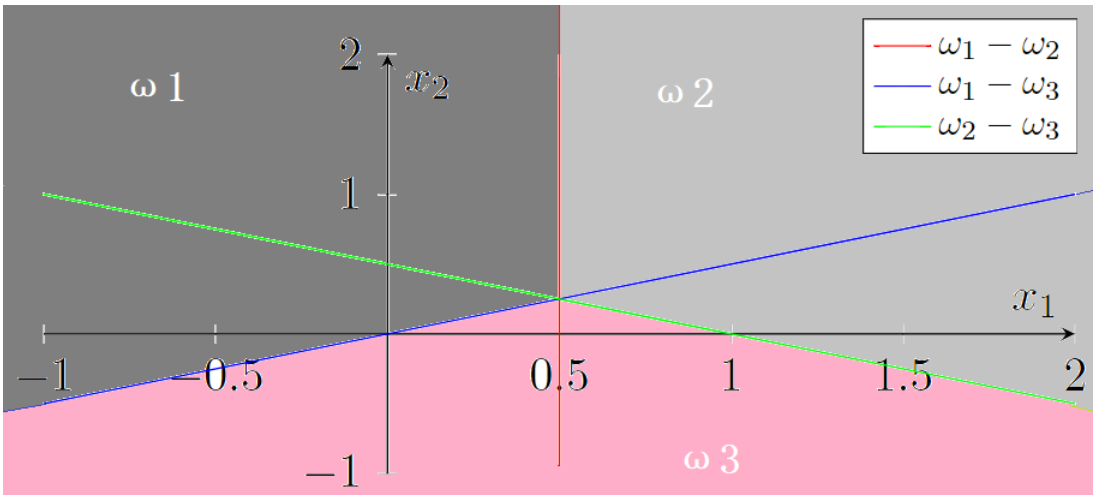

图 1: 3 分类决策面

由于三条分界面交于一点 (0.5,0.5)，不存在不确定性区域。

第二部分回答

1. 代码输出如下：

```
Batch_perceptron algorithm:
a = [ 3.4   -3.04   3.41]
Iterations required for convergence is 24
Batch_perceptron algorithm:
a = [ 1.9   -4.14   4.86]
```

```
Iterations required for convergence is 17
Ho–Kashyap algorithm：
No solution found!
Ho–Kashyap algorithm：
a = [[0.13381828] [0.03717747] [0.01669066]]
b = [[0.46787909] [0.01       ] [0.30111691]
[0.3947414 ] [0.32167591] [0.13245664]
[0.15628159] [0.12494896] [0.50786447]
[0.24952647] [0.08073817] [0.19372309]
[0.15084515] [0.23560033] [0.18203341]
[0.03832449] [0.17504972] [0.29644592]
[0.29204113] [0.26875263]]
Iterations required for convergence is 29132
MSE准则多分类：
target = [1. 1. 2. 2. 3. 3. 4. 4.]
output = [1. 1. 2. 2. 3. 3. 4. 4.]
accuracy = 1.0
```

2. 输出图像如下：

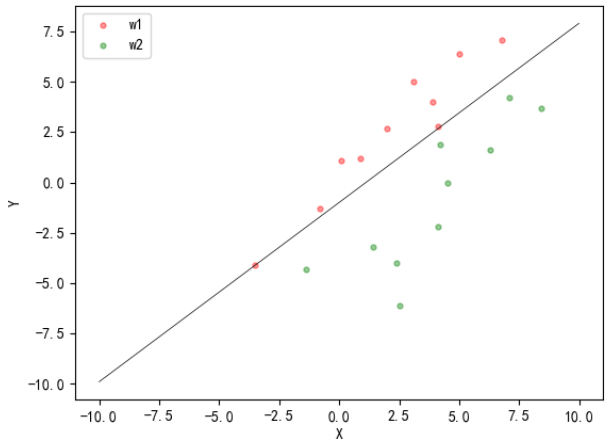

图 2: $\omega_1 \quad vs \quad \omega_2$

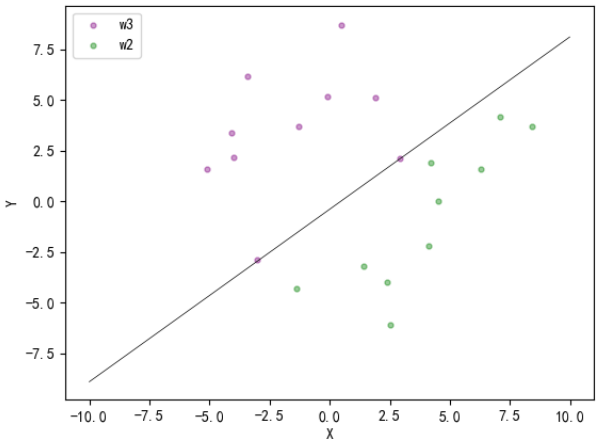

图 3: $\omega_3 \quad vs \quad \omega_2$

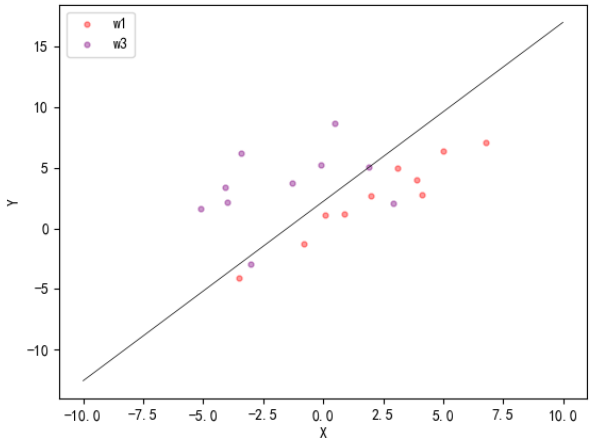

图 4: $\omega_1 \quad vs \quad \omega_3$ 线性不可分

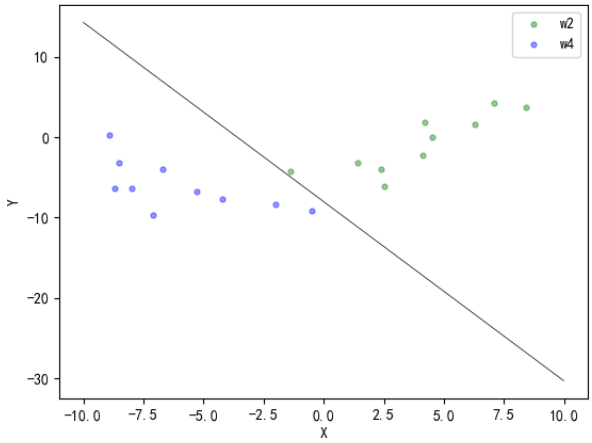

图 5: $\omega_2 \quad vs \quad \omega_4$

3. 代码见如下文件 Python 文件