

## 第 5 章：线性判别函数

### 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  $\mathbf{a}=(0,1,0)^T$ ，其中向量  $\mathbf{a}$  的第三维对应于样本的齐次坐标。同时，假定梯度更新步长 $\eta_k$  固定为 1。试利用批处理感知准则函数方法求解线性判别函数  $g(\mathbf{y})=\mathbf{a}^T\mathbf{y}$  的权向量  $\mathbf{a}$ 。（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：

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

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

$$g_1(\mathbf{x}) = -x_1 + x_2$$

$$g_2(\mathbf{x}) = x_1 + x_2 - 1$$

$$g_3(\mathbf{x}) = -x_2$$

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：

sample	$\omega_1$		$\omega_2$		$\omega_3$		$\omega_4$	
	$x_1$	$x_2$	$x_1$	$x_2$	$x_1$	$x_2$	$x_1$	$x_2$
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 perception” algorithm.
  - (a). Starting with  $\mathbf{a} = \mathbf{0}$ , apply your program to the training data from  $\omega_1$  and  $\omega_2$ . 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  $\omega_3$  and  $\omega_2$ . Again, note that the number of iterations required for convergence.

2. Implement the [Ho-Kashyap algorithm](#) and apply it to the training data from  $\omega_1$  and  $\omega_3$ . Repeat to apply it to the training data from  $\omega_2$  and  $\omega_4$ . 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。请写出主要计算步骤，并给出你的正确率。