PRML 第三章作业

黄磊　　计 702　　2022E8013282156

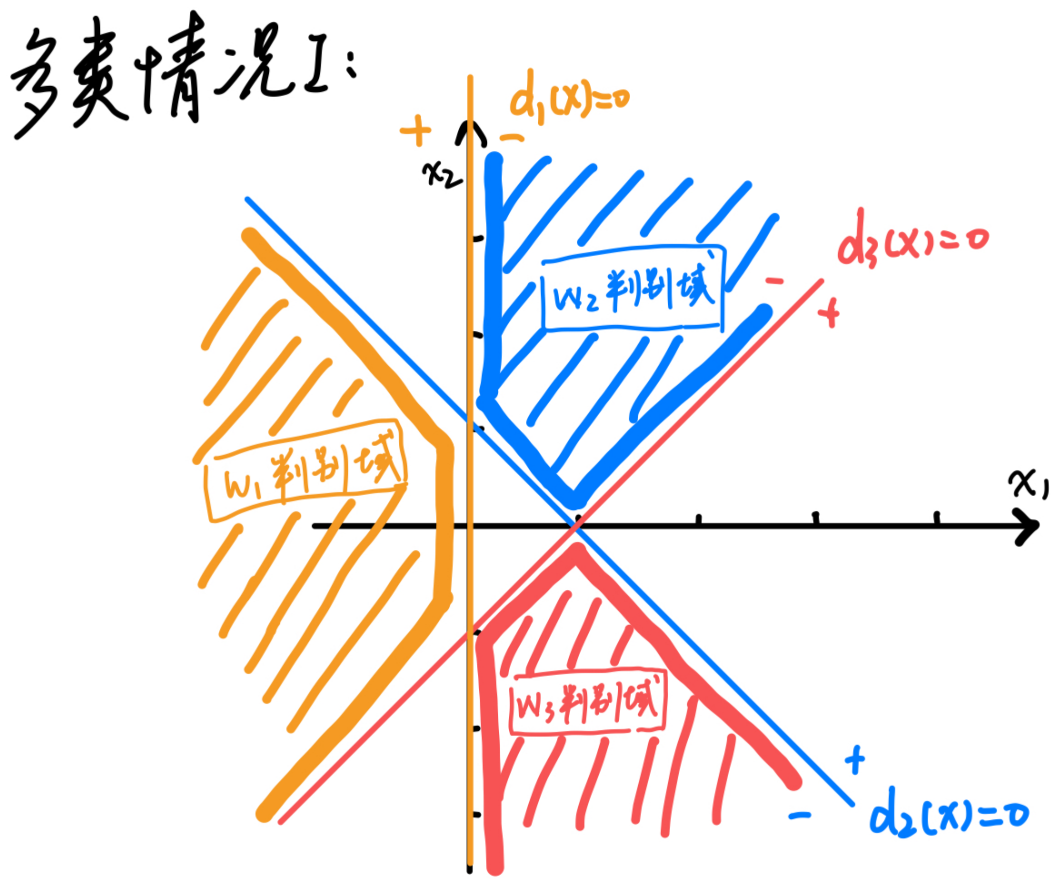
一、在一个 10 类的模式识别问题中，有 3 类单独满足多类情况 1，其余的类别满足多类情况 2。问该模式识别问题所需判别函数的最少数目是多少？

答：将 10 类问题可看作 4 类满足多类情况 1 的问题，即将剩下的满足多类情况 2 的 7 类单独划作一个子类 Θ，因此需要 4 个判别函数。在子类 Θ 中，运用多类情况 2 的判别法进行分类，因此需要 $7 \times (7-1)/2 = 21$ 个判别函数。因此一共需要 $4 + 21 = 25$ 个判别函数。

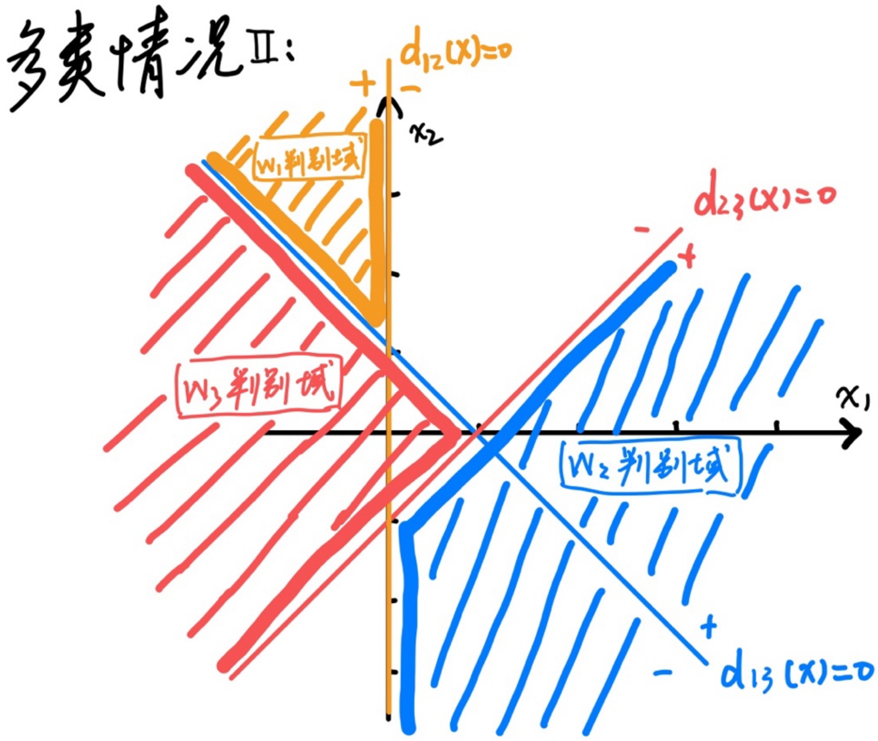
二、　一个三类问题，其判别函数如下：

$d_1(x)=-x_1$，　$d_2(x)=x1+x_2-1$，　$d_3(x)=x_1-x_2-1$

1. 设这些函数是在多类情况 1 条件下确定的，绘出其判别界面和每一个模式类别的区域。

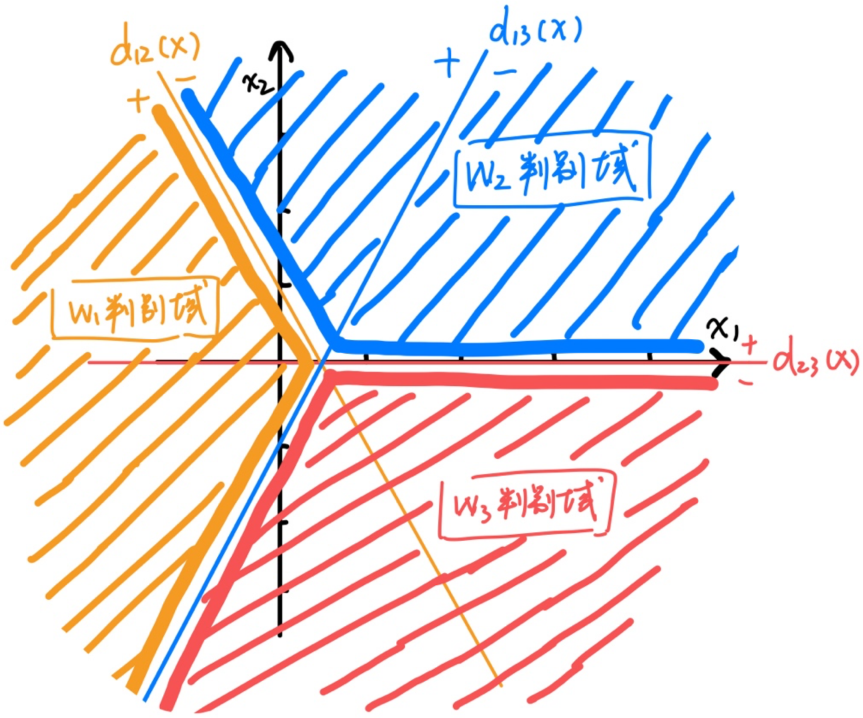


2. 设为多类情况 2，并使：$d_{12}(\mathbf{x})= d_1(\mathbf{x})$, $d_{13}(\mathbf{x})= d_2(\mathbf{x})$, $d_{23}(\mathbf{x})= d_3(\mathbf{x})$。绘出其判别界面和多类情况 2 的区域。



3. 设 $d_1(\boldsymbol{x})$, $d_2(\boldsymbol{x})$和 $d_3(\boldsymbol{x})$是在多类情况 3 的条件下确定的，绘出其判别界面和每类的区域。

多类情况 3 有：$d_{12}(x) = d_1(x) - d_2(x) = -x_1 - (x_1 + x_2 - 1) = -2x_1 - x_2 + 1$
同理有 $d_{13}(x) = -2x_1 + x_2 + 1$，$d_{23}(x) = 2x_2$ 得到判别界面和分类域如下图：



三、两类模式，每类包括 5 个 3 维不同的模式向量，且良好分布。如果它们是线性可分的，问权向量至少需要几个系数分量？假如要建立二次的多项式判别函数，又至少需要几个系数分量？（设模式的良好分布不因模式变化而改变。）

若为线性可分，则至少需要 $n+1=4$ 个系数分量；若建立二次多项式判别函数，则至少需要$\frac{(3+2)!}{3!2!}=10$ 个系数分量。

四、

1．用感知器算法求下列模式分类的解向量 *w*:

ω_1: $\{(0\ 0\ 0)^T, (1\ 0\ 0)^T, (1\ 0\ 1)^T, (1\ 1\ 0)^T\}$

ω_2: $\{(0\ 0\ 1)^T, (0\ 1\ 1)^T, (0\ 1\ 0)^T, (1\ 1\ 1)^T\}$

2. 编写求解上述问题的感知器算法程序（选做）。

具体程序见文件 Sensor.py ，初始化解向量 $\omega=[0,0,0]$,得到该分类需要 6 次迭代，得到迭代求解过程以及结果如下：

```
************************************************************
*********************** 测例二 *****************************
************************************************************
增广矩阵w1:   [[0, 0, 0, 1], [1, 0, 0, 1], [1, 0, 1, 1], [1, 1, 0, 1]]
负增广矩阵w2:   [[0.0, 0.0, -1.0, -1.0], [0.0, -1.0, -1.0, -1.0], [0.0,
 0.0, -1.0], [-1.0, -1.0, -1.0, -1.0]]
--------------------- the  1 _th epoch --------------------
w( 0 ) <= 0, w( 1 ) = w( 0 ) + C*x_ 0 = [0. 0. 0. 1.]
w( 1 ) > 0, w( 2 ) = w( 1 ) = [0. 0. 0. 1.]
w( 2 ) > 0, w( 3 ) = w( 2 ) = [0. 0. 0. 1.]
w( 3 ) > 0, w( 4 ) = w( 3 ) = [0. 0. 0. 1.]
w( 4 ) <= 0, w( 5 ) = w( 4 ) + C*x_ 4 = [ 0.  0. -1.  0.]
w( 5 ) > 0, w( 6 ) = w( 5 ) = [ 0.  0. -1.  0.]
w( 6 ) <= 0, w( 7 ) = w( 6 ) + C*x_ 6 = [ 0. -1. -1. -1.]
w( 7 ) > 0, w( 8 ) = w( 7 ) = [ 0. -1. -1. -1.]
--------------------- the  2 _th epoch --------------------
w( 0 ) <= 0, w( 1 ) = w( 0 ) + C*x_ 0 = [ 0. -1. -1.  0.]
w( 1 ) <= 0, w( 2 ) = w( 1 ) + C*x_ 1 = [ 1. -1. -1.  1.]
w( 2 ) > 0, w( 3 ) = w( 2 ) = [ 1. -1. -1.  1.]
w( 3 ) > 0, w( 4 ) = w( 3 ) = [ 1. -1. -1.  1.]
w( 4 ) <= 0, w( 5 ) = w( 4 ) + C*x_ 4 = [ 1. -1. -2.  0.]
w( 5 ) > 0, w( 6 ) = w( 5 ) = [ 1. -1. -2.  0.]
w( 6 ) > 0, w( 7 ) = w( 6 ) = [ 1. -1. -2.  0.]
w( 7 ) > 0, w( 8 ) = w( 7 ) = [ 1. -1. -2.  0.]
--------------------- the  3 _th epoch --------------------
w( 0 ) <= 0, w( 1 ) = w( 0 ) + C*x_ 0 = [ 1. -1. -2.  1.]
w( 1 ) > 0, w( 2 ) = w( 1 ) = [ 1. -1. -2.  1.]
w( 2 ) <= 0, w( 3 ) = w( 2 ) + C*x_ 2 = [ 2. -1. -1.  2.]
w( 3 ) > 0, w( 4 ) = w( 3 ) = [ 2. -1. -1.  2.]
w( 4 ) <= 0, w( 5 ) = w( 4 ) + C*x_ 4 = [ 2. -1. -2.  1.]
w( 5 ) > 0, w( 6 ) = w( 5 ) = [ 2. -1. -2.  1.]
```

```
w( 6 ) <= 0, w( 7 ) = w( 6 ) + C*x_ 6 = [ 2. -2. -2.  0.]
w( 7 ) > 0, w( 8 ) = w( 7 ) = [ 2. -2. -2.  0.]
-------------------- the  4 _th epoch --------------------
w( 0 ) <= 0, w( 1 ) = w( 0 ) + C*x_ 0 = [ 2. -2. -2.  1.]
w( 1 ) > 0, w( 2 ) = w( 1 ) = [ 2. -2. -2.  1.]
w( 2 ) > 0, w( 3 ) = w( 2 ) = [ 2. -2. -2.  1.]
w( 3 ) > 0, w( 4 ) = w( 3 ) = [ 2. -2. -2.  1.]
w( 4 ) > 0, w( 5 ) = w( 4 ) = [ 2. -2. -2.  1.]
w( 5 ) > 0, w( 6 ) = w( 5 ) = [ 2. -2. -2.  1.]
w( 6 ) > 0, w( 7 ) = w( 6 ) = [ 2. -2. -2.  1.]
w( 7 ) > 0, w( 8 ) = w( 7 ) = [ 2. -2. -2.  1.]
-------------------- the  5 _th epoch --------------------
w( 0 ) > 0, w( 1 ) = w( 0 ) = [ 2. -2. -2.  1.]
w( 1 ) > 0, w( 2 ) = w( 1 ) = [ 2. -2. -2.  1.]
w( 2 ) > 0, w( 3 ) = w( 2 ) = [ 2. -2. -2.  1.]
w( 3 ) > 0, w( 4 ) = w( 3 ) = [ 2. -2. -2.  1.]
w( 4 ) > 0, w( 5 ) = w( 4 ) = [ 2. -2. -2.  1.]
w( 5 ) > 0, w( 6 ) = w( 5 ) = [ 2. -2. -2.  1.]
w( 6 ) > 0, w( 7 ) = w( 6 ) = [ 2. -2. -2.  1.]
w( 7 ) > 0, w( 8 ) = w( 7 ) = [ 2. -2. -2.  1.]
最终该解向量为:  [ 2. -2. -2.  1.]
相应判别函数为: d(x) = 2.0*x1 - 2.0*x2 - 2.0*x3 + 1.0
```

最终得到解向量为 $\omega = [2, -2, -2, 1]^T$ ，相应的判别函数为：

$$d(x) = 2x_1 - 2x_2 - 2x_3 + 1 = 0$$

程序主要函数截图：

```
def Sensor(w1, w2, C):
    for i in range(len(w1)):
        w1[i] = w1[i] + [1]
    for i in range(len(w2)):
        w2[i] = w2[i] + [1]
    print("增广矩阵w1： ", w1)
    a = -1 * np.eye(len(w2[0]))
    oppo_w2 = np.dot(w2, a).tolist()
    print("负增广矩阵w2： ", oppo_w2)
    total_samples = w1
    total_samples.extend(oppo_w2)
    total_samples = np.array(total_samples)
    w = np.zeros(len(w2[0]))
    wrong_sum = len(total_samples)
    epoch = 0
    while (wrong_sum != 0):
        wrong_sum = len(total_samples)
        epoch += 1
        print("-------------------- the ", epoch, "_th epoch -------------------")
        for i in range(len(total_samples)):
            i_th_sum = sum(np.multiply(w, total_samples[i]))
            if i_th_sum > 0:
                print("w(", i, ") > 0, w(", i + 1, ") = w(", i, ") =", w)
                wrong_sum -= 1
                continue
            else:
                w = w + C * total_samples[i]
                print("w(", i, ") <= 0, w(", i + 1, ") = w(", i, ") + C*x_", i, "=", w)
                continue
    return w
```

五、用多类感知器算法求下列模式的判别函数：

ω_1: (-1 -1)^T，　ω_2: (0 0)^T，　ω_3: (1 1)^T

增广形式：$x^1 = [-1,-1,1]^T$,$x^2 = [0,\ 0\ ,1]^T$，$x^3 = [1,1,1\ ]^T$，初始化$\omega_1(1) = \omega_2(2) = \omega_3(3) = [0,0,0\ ]^T$,$C = 1$，进行迭代：

第一次迭代：以 x_1 为训练样本，$d_1(1) = d_2(1) = d_3(1) = 0$ ，故：

$$\omega_1(2) = \omega_1(1) + x^1 = [-1,-1,1]^T$$

$$\omega_2(2) = \omega_2(1) - x^1 = [1,1\ ,-1]^T$$

$$\omega_3(2) = \omega_3(1) - x^1 = [1,1\ ,-1]^T$$

第二次迭代：以 x_2 为训练样本，$d_1(2) = 1, d_2(2) = -1, d_3(2) = -1$, 因此：

$$\omega_1(3) = \omega_1(2) - x^2 = [-1,-1,0]^T$$

$$\omega_2(3) = \omega_2(2) + x^2 = [1,1\ ,0]^T$$

$$\omega_3(3) = \omega_3(2) - x^1 = [1,1\ ,-2]^T$$

第三次迭代：以 x_3 为训练样本，$d_1(3) = -2, d_2(3) = 2, d_3(3) = 0$, 因此：

$$\omega_1(4) = \omega_1(3) = [-1,-1,0]^T$$

$$\omega_2(4) = \omega_2(3) - x^3 = [0,0\ ,-1]^T$$

$$\omega_3(4) = \omega_3(3) + x^3 = [2,2,-1]^T$$

第四次迭代：以 x_1 为训练样本，$d_1(4) = 2, d_2(-1) = 2, d_3(4) = -5$, 因此：

$$\omega_1(5) = \omega_1(4) = [-1,-1,0]^T$$

$$\omega_2(5) = \omega_2(4) = [0,0\ ,-1]^T$$

$$\omega_3(5) = \omega_3(4) = [2,2,-1]^T$$

第五次迭代：以 x_2 为训练样本，$d_1(5) = 0, d_2(5) = -1, d_3(5) = -1$, 因此：

$$\omega_1(6) = \omega_1(5) - x^2 = [-1,-1,-1]^T$$

$$\omega_2(6) = \omega_2(5) + x^2 = [0,0\ ,0]^T$$

$$\omega_3(6) = \omega_3(5) - x^2 = [2,2,-2]^T$$

第六次迭代：以 x_3 为训练样本，$d_1(6) = -3, d_2(6) = 0, d_3(6) = 2,$ 因此：

$$\omega_1(7) = \omega_1(6) = [-1, -1, -1]^T$$

$$\omega_2(7) = \omega_2(6) = [0, 0, 0]^T$$

$$\omega_3(7) = \omega_3(6) = [2,2, -2]^T$$

第七次迭代：以 x_1 为训练样本，$d_1(7) = 1, d_2(7) = 0, d_3(7) = -6$ 因此：

$$\omega_1(8) = \omega_1(7) = [-1, -1, -1]^T$$

$$\omega_2(8) = \omega_2(7) = [0, 0, 0]^T$$

$$\omega_3(8) = \omega_3(7) = [2,2, -2]^T$$

第八次迭代：以 x_2 为训练样本，$d_1(8) = -1, d_2(8) = 0, d_3(8) = -2$ 因此：

$$\omega_1(9) = \omega_1(8) = [-1, -1, -1]^T$$

$$\omega_2(9) = \omega_2(8) = [0, 0, 0]^T$$

$$\omega_3(9) = \omega_3(8) = [2,2, -2]^T$$

最后三次的迭代结果相同，因此权向量的解为：

$$\omega_1 = [-1, -1, -1]^T$$

$$\omega_2 = [0, 0, 0]^T$$

$$\omega_3 = [2,2, -2]^T$$

对应的判别函数为：

$$d_1(x) = -x_1 - x_2 - 1$$

$$d_2(x) = 0$$

$$d_3(x) = 2x_1 + 2x_2 - 2$$

六、　采用梯度法和准则函数

$$J(w,x,b)=\frac{1}{8\|x\|^2}\left[\left(w^T x-b\right)-\left|w^T x-b\right|\right]^2$$

式中实数 b>0，试导出两类模式的分类算法。

有：

$$\frac{\partial J}{\partial w}=\frac{1}{4||x||^2}[(w^Tx-b)-|w^Tx-b|]*[x-x*sgn(w^Tx-b)]$$

其中，sgn 为示性函数，即：

$$sgn(w^Tx-b)=\begin{cases}1, & w^Tx-b>0\\ -1, & w^Tx-b\le 0\end{cases}$$

因此得到迭代式为：

$$w(k+1)=w(k)+\frac{C}{4||x||^2}[(w^Tx-b)-|w^Tx-b|]*[x-x*sgn(w^Tx-b)]$$

$$=w(k)+C*\begin{cases}0, & w^Tx-b>0\\ \dfrac{-(w^Tx-b)}{||x||^2}, & w^Tx-b\le 0\end{cases}$$

七、（选做）

1．用二次埃尔米特多项式的势函数算法求解以下模式的分类问题

ω_1: $\{(0\ 1)^T, (0\ -1)^T\}$

ω_2: $\{(1\ 0)^T, (-1\ 0)^T\}$

$$\varphi_1(x)=\varphi_1(x_1,x_2)=H_0(x_1)H_0(x_2)=1$$
$$\varphi_2(x)=\varphi_2(x_1,x_2)=H_0(x_1)H_1(x_2)=2x_2$$
$$\varphi_3(x)=\varphi_3(x_1,x_2)=H_0(x_1)H_2(x_2)=4x_2^2-2$$
$$\varphi_4(x)=\varphi_4(x_1,x_2)=H_1(x_1)H_0(x_2)=2x_1$$
$$\varphi_5(x)=\varphi_5(x_1,x_2)=H_1(x_1)H_1(x_2)=4x_1x_2$$
$$\varphi_6(x)=\varphi_6(x_1,x_2)=H_1(x_1)H_2(x_2)=2x_1(4x_2^2-2)$$
$$\varphi_7(x)=\varphi_7(x_1,x_2)=H_2(x_1)H_0(x_2)=4x_1^2-2$$
$$\varphi_8(x)=\varphi_8(x_1,x_2)=H_2(x_1)H_1(x_2)=2x_2(4x_1^2-2)$$
$$\varphi_9(x)=\varphi_9(x_1,x_2)=H_2(x_1)H_2(x_2)=(4x_1^2-2)(4x_2^2-2)$$

按照第一类势函数的定义，得到势函数：

$$K(x,x^k)=\sum_{i=1}^{9}\phi_i(x)\phi_i(x^k)$$

第一步：取 $x^1=[0,1]^T\in\omega_1$，因此：

$$K_1(x)=K(x,x^1)=-15+20x_2+40x_2^2+24x_1^2-32x_1^2x_2-64x_1^2x_2^2$$

第二步：取 $x^2=[0,-1]^T\in\omega_1$，因此 $K_1(x^2)=5>0$，则：

$$K_2(x)=K_1(x)$$

第三步：取 $x^3=[1,0]^T\in\omega_2$，因此 $K_2(x^3)=9>0$，则：

$$K_3(x)=K_2(x)-K(x,x^3)=20x_2+16x_2^2-20x_1-16x_1^2$$

第四步：取 $x^4=[-1,0]^T\in\omega_2$，因此 $K_3(x^4)=4>0$，则：

$$K_4(x)=K_3(x)-K(x,x^4)=15+20x_2-56x_1^2-8x_2^2-32x_1^2x_2+64x_1^2x_2^2$$

重复迭代一次：

第五步：取 $x^5=[0,1]^T\in\omega_1$，$K_4(x^5)=27>0$因此：

$$K_5(x)=K_4(x)$$

第六步：取 $x^6=[0,-1]^T\in\omega_1$，因此 $K_5(x^6)=-13<0$，则：

$$K_6(x)=K_5(x)+K(x,x^6)=-32x_1^2+32x_2^2$$

第七步：取 $x^7=[1,0]^T\in\omega_2$，因此 $K_6(x^7)=-32<0$，则：

$$K_7(x)=K_6(x)$$

第八步：取 $x^8=[-1,0]^T\in\omega_2$，因此 $K_7(x^8)=-32<0$，则：

$$K_8(x)=K_7(x)$$

检验 ω_1 类的情况，继续重复迭代：

第九步：取 $x^9=[0,1]^T\in\omega_1$，因此 $K_8(x^9)=32>0$，则:

$$K_9(x)=K_8(x)$$

第十步：取 $x^{10}=[0,-1]^T\in\omega_1$，因此 $K_9(x^{10})=32>0$，则:

$$K_{10}(x)=K_9(x)$$

由于第七、八、九、十步迭代结果都能正确分类样本，因此收敛于判别函数：

$$d(x)=-32x_1^2+32x_2^2$$

2. 用下列势函数

$$K(\boldsymbol{x},\boldsymbol{x}^k)=e^{-\alpha\left\|\boldsymbol{x}-\boldsymbol{x}^k\right\|^2}$$

求解以下模式的分类问题

ω_1: {(0 1)^T, (0 -1)^T}

ω_2: {(1 0)^T, (-1 0)^T}

取$\alpha=1$，在二维情况下，其势函数对应为：

$$K(x,x^k)=exp\left\{-\left||x-x^k|\right|^2\right\}=\exp\left\{-\left[\left(x-x_2^k\right)^2+\left(x-x_2^k\right)^2\right]\right\}$$

开始迭代：由于带有平方，因此改为采用下标表示各样本。

第一步：取 $x^1=[0,1]^T\in\omega_1$ ，因此：

$$K_1(x)=\exp\{-x_1^2-(x_2-1)^2\}$$

第二步：取 $x^2=[0,-1]^T\in\omega_1$ ，因此 $K_1(x^2)=e^{-4}>0$ ，则：

$$K_2(x)=K_1(x)$$

第三步：取 $x^3=[1,0]^T\in\omega_2$ ，因此 $K_2(x^3)=e^{-2}>0$ ，则：

$$K_3(x)=K_2(x)-K(x,x^3)$$

$$=\exp\{-x_1^2-(x_2-1)^2\}-\exp\{-(x_1-1)^2-x_2^2\}$$

第四步：取 $x^4=[-1,0]^T\in\omega_2$ ，因此 $K_3(x^4)=e^{-2}-e^{-4}>0$ ，则：

$$K_4(x)=K_3(x)-K(x,x^4)$$

$$=\exp\{-x_1^2-(x_2-1)^2\}-\exp\{-(x_1-1)^2-x_2^2\}-\exp\{-(x_1+1)^2-x_2^2\}$$

重复迭代一次：

第五步：取 $x^5=[0,1]^T\in\omega_1$ ，$K_4(x^5)=1-2e^{-2}>0$因此：

$$K_5(x)=K_4(x)$$

第六步：取 $x^6=[0,-1]^T\in\omega_1$ ，因此 $K_5(x^6)=e^{-4}-2e^{-2}<0$ ，则：

$$K_6(x)=K_5(x)+K(x,x^6)$$

$$=\exp\{-x_1^2-(x_2-1)^2\}-\exp\{-(x_1-1)^2-x_2^2\}-\exp\{-(x_1+1)^2-x_2^2\}$$

$$+\exp\{-x_1^2-(x_2+1)^2\}$$

第七步：取 $x^7=[1,0]^T\in\omega_2$ ，因此 $K_6(x^7)=2e^{-2}-1-e^{-4}<0$ ，则：

$$K_7(x)=K_6(x)$$

第八步：取 $x^8=[-1,0]^T\in\omega_2$ ，因此 $K_7(x^8)=2e^{-2}-e^{-4}-1<0$ ，则：

$$K_8(x)=K_7(x)$$

检验 ω_1 类的情况，继续重复迭代：

第九步：取 $x^9 = [0,1]^T \in \omega_1$，因此 $K_8(x^9) = e^{-4} + 1 - 2e^{-2} > 0$，则:

$$K_9(x) = K_8(x)$$

第十步：取 $x^{10} = [0,-1]^T \in \omega_1$，因此 $K_9(x^{10}) = 1 + e^{-4} - 2e^{-2} > 0$，则:

$$K_{10}(x) = K_9(x)$$

由于第七、八、九、十步迭代结果都能正确分类样本，因此收敛于判别函数：

$$\begin{aligned} d(x) = {} & \exp\{-x_1^2 - (x_2-1)^2\} - \exp\{-(x_1-1)^2 - x_2^2\} \\ & - \exp\{-(x_1+1)^2 - x_2^2\} + \exp\{-x_1^2 - (x_2+1)^2\} \end{aligned}$$