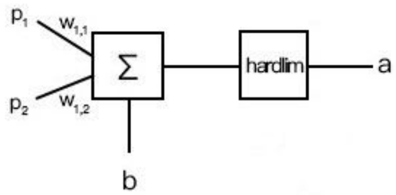
# Homework 1

5130309194    Shen Jiyuan

================================================================================

1.1 单神经元感知机图：（其中，$w_{1,1}$=0; $w_{1,2}$=-1; b=-0.5）




求解过程：

1、将 $P_1, P_2, P_3, P_4$ 四个点放在平面直角坐标系中观察，则易看出判定边界为直线 y=-0.5；

2、确定一个与判定边界垂直的权值向量 $w^T$=[0,-1];

3、选择判定边界上一点(0,-0.5)代入判定式 $w^Tp$+b=0 中，解得 b=-0.5.

1.2 由输入样例验证求解结果：

$a_1 = \text{hardlim}(w^Tp_1+b) = \text{hardlim}([0,-1]\cdot[1,-1]^T-0.5) = \text{hardlim}(0.5) = 1 = t_1$

$a_2 = \text{hardlim}(w^Tp_2+b) = \text{hardlim}([0,-1]\cdot[-1,-1]^T-0.5) = \text{hardlim}(0.5) = 1 = t_2$

$a_3 = \text{hardlim}(w^Tp_3+b) = \text{hardlim}([0,-1]\cdot[0,0]^T-0.5) = \text{hardlim}(-0.5) = 0 = t_3$

$a_4 = \text{hardlim}(w^Tp_4+b) = \text{hardlim}([0,-1]\cdot[1,0]^T-0.5) = \text{hardlim}(-0.5) = 0 = t_4$

由以上验证知求解结果是正确的。

1.3 分类样本 $p_5, p_6, p_7, p_8$ 如下：

$a_5 = \text{hardlim}(w^Tp_5+b) = \text{hardlim}([0,-1]\cdot[-2,0]^T-0.5) = \text{hardlim}(-0.5) = 0$

$a_6 = \text{hardlim}(w^Tp_6+b) = \text{hardlim}([0,-1]\cdot[1,1]^T-0.5) = \text{hardlim}(-1.5) = 0$

$a_7 = \text{hardlim}(w^Tp_7+b) = \text{hardlim}([0,-1]\cdot[0,1]^T-0.5) = \text{hardlim}(-1.5) = 0$

$a_8 = \text{hardlim}(w^Tp_8+b) = \text{hardlim}([0,-1]\cdot[-1,-2]^T-0.5) = \text{hardlim}(1.5) = 1$

则有 $p_5, p_6, p_7$ 为一类($p_3,p_4$)，$p_8$ 为另一类($p_1,p_2$)。

1.4 $p_6, p_7, p_8$ 的分类与 $w$，b 的选择无关，$p_5$ 的分类与 $w$，b 的选择有关。
因为 $w$，b 的选择所依赖的判定边界的斜率和偏移量都有一定的变化，而 $p_5$ 在此变化范围之内，$p_6, p_7, p_8$ 在此变化范围之外。

1.5 1、应用感知机学习规则修正 $w(0)^T$，b(0)值：

Iteration1: $a_1 = \text{hardlim}(w(0)^Tp_1+b(0)) = \text{hardlim}([0,0]\cdot[1,-1]^T+0) = \text{hardlim}(0) = 1,\ t_1=1$

$w(1)^T = w(0)^T = [0,0]$

$b(1) = b(0) = 0$

Iteration2: $a_2 = \text{hardlim}(w(1)^Tp_2+b(1)) = \text{hardlim}([0,0]\cdot[-1,-1]^T+0) = \text{hardlim}(0) = 1,\ t_2=1$

$w(2)^T = w(1)^T = [0,0]$

$b(2) = b(1) = 0$

Iteration3: $a_3 = \text{hardlim}(w(2)^Tp_3+b(2)) = \text{hardlim}([0,0]\cdot[0,0]^T+0) = \text{hardlim}(0) = 1,\ t_3=0$

$w(3)^T = w(2)^T - p_3 = [0,0]$

$b(3) = b(2) + t_3 - a_3 = -1$

Iteration4: $a_4 = \text{hardlim}(w(3)^T p_4 + b(3)) = \text{hardlim}([0,0]\cdot[1,0]^T - 1) = \text{hardlim}(-1) = 0,\ t_4 = 0$

$w(4)^T = w(3)^T = [0,0]$

$b(4) = b(3) = -1$

Iteration5: $a_1 = \text{hardlim}(w(4)^T p_1 + b(4)) = \text{hardlim}([0,0]\cdot[1,-1]^T - 1) = \text{hardlim}(-1) = 0,\ t_1 = 1$

$w(5)^T = w(4)^T + p_1 = [1,-1]$

$b(5) = b(4) + t_1 - a_1 = 0$

Iteration6: $a_2 = \text{hardlim}(w(5)^T p_2 + b(5)) = \text{hardlim}([1,-1]\cdot[-1,-1]^T + 0) = \text{hardlim}(0) = 1,\ t_2 = 1$

$w(6)^T = w(5)^T = [1,-1]$

$b(6) = b(5) = 0$

Iteration7: $a_3 = \text{hardlim}(w(6)^T p_3 + b(6)) = \text{hardlim}([1,-1]\cdot[0,0]^T + 0) = \text{hardlim}(0) = 1,\ t_3 = 0$

$w(7)^T = w(6)^T - p_3 = [1,-1]$

$b(7) = b(6) + t_3 - a_3 = -1$

Iteration8: $a_4 = \text{hardlim}(w(7)^T p_4 + b(7)) = \text{hardlim}([1,-1]\cdot[1,0]^T - 1) = \text{hardlim}(0) = 1,\ t_4 = 0$

$w(8)^T = w(7)^T - p_4 = [0,-1]$

$b(8) = b(7) - t_3 + a_3 = -2$

Iteration9: $a_1 = \text{hardlim}(w(8)^T p_1 + b(8)) = \text{hardlim}([0,-1]\cdot[1,-1]^T - 2) = \text{hardlim}(-1) = 0,\ t_1 = 1$

$w(9)^T = w(8)^T + p_1 = [1,-2]$

$b(9) = b(8) + t_1 - a_1 = -1$

Iteration10: $a_2 = \text{hardlim}(w(9)^T p_2 + b(9)) = \text{hardlim}([1,-2]\cdot[-1,-1]^T - 1) = \text{hardlim}(0) = 1,\ t_2 = 1$

$w(10)^T = w(9)^T = [1,-2]$

$b(10) = b(9) = -1$

Iteration11: $a_3 = \text{hardlim}(w(10)^T p_3 + b(10)) = \text{hardlim}([1,-2]\cdot[0,0]^T - 1) = \text{hardlim}(-1) = 0,\ t_3 = 0$

$w(11)^T = w(10)^T = [1,-2]$

$b(11) = b(10) = -1$

Iteration12: $a_4 = \text{hardlim}(w(11)^T p_4 + b(11)) = \text{hardlim}([1,-2]\cdot[1,0]^T - 1) = \text{hardlim}(0) = 1,\ t_4 = 0$

$w(12)^T = w(11)^T - p_4 = [0,-2]$

$b(12) = b(11) - t_3 + a_3 = -2$

Iteration13~Iteration16: 结果都是正确的，不需要修正。

即，经过修正得到 $w^T=[0,-2]$, $b= -2$.

2、验证(1.3)样本：

$a_5 = \text{hardlim}(w^T p_5 + b) = \text{hardlim}([0,-2]\cdot[-2,0]^T - 2) = \text{hardlim}(-2) = 0$

$a_6 = \text{hardlim}(w^T p_6 + b) = \text{hardlim}([0,-2]\cdot[1,1]^T - 2) = \text{hardlim}(-4) = 0$

$a_7 = \text{hardlim}(w^T p_7 + b) = \text{hardlim}([0,-2]\cdot[0,1]^T - 2) = \text{hardlim}(-4) = 0$

$a_8 = \text{hardlim}(w^T p_8 + b) = \text{hardlim}([0,-2]\cdot[-1,-2]^T - 2) = \text{hardlim}(2) = 1$

则有 $p_5, p_6, p_7$ 为一类($p_3,p_4$)，$p_8$ 为另一类($p_1,p_2$)。(同(1.3))

================================================================================

==========================================================================

2.1

LMS 解题详细过程：假设模型是线性的，$+h(x) = w_0 + w_1 x_1 + w_2 x_2$，

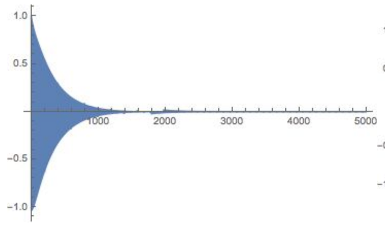
损失函数 $J(w) = \frac{1}{2}\sum_{i=1}^{m}\{h(x^{(i)}) - y^{(i)}\}^2$，考虑随机梯度下降算法：$w_i = w_i + \alpha(y^{(j)} - h(x^{(j)}))x_i^j$。

即：$W(k+1)=W(k)+2\alpha e(k)p(k)$，$b(k+1)=b(k)+2\alpha e(k)$。

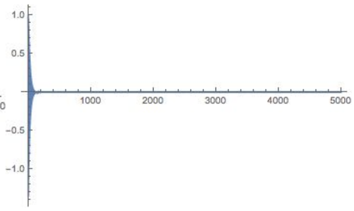
在 python 上实现 w,b 计算及 e-i 的文件输出，通过 matlab 读取 e-i 文件并作图。

```
import numpy as np
import time

p0 = np.matrix('1;-1')
p1 = np.matrix('-1;-1')
p2 = np.matrix('0;0')
p3 = np.matrix('1;0')

inputs = [p0, p1, p2, p3]
targets = [1, 1, -1, -1]
irritation = 5000 / 4
rates = [0.01, 0.1, 0.5]
deviation = []

def LMS(rate, index):
    file = open('deviations'+'%d'%index+'.txt', 'w')
    w = np.matrix('0,0')
    b = 0
    for i in range(irritation):
        for j in range(4):
            output = ((w * inputs[j]).tolist())[0][0] + b
            e = targets[j] - output
            if e != 0:
                w = w + 2*rate*e*inputs[j].transpose()
                b = b + 2*rate*e
            file.write('%.4f'%e+'\n')
    file.close()
    print w, b

for i in range(3):
    LMS(rates[i], i)
file = open('index', 'w')
for i in range(5000):
    file.write('%d'%(i+1)+'\n')
file.close()
```

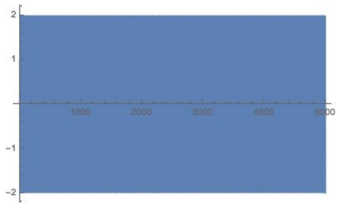
迭代 5000 次之后的误差(e)-迭代次数(i)曲线：




α = 0.01




α = 0.1




α = 0.5

2.2 最终的 $w^T$, b 结果：

α = 0.01：$w^T$ = [0,-2], b = -1

α = 0.1： $w^T$ = [0,-2], b = -1

α = 0.5： $w^T$ = [0,-3], b = -3

2.3 感知机和 ADALINE 都能处理题一中的分类问题。

感知机迭代到正确就会停止，而 ADALINE 会一直趋近于正确值。

当 ADALINE 的学习速率适当，且感知机与 ADALINE 都对 $w^T$ = [0, 0], b=0 的初始值进行修正时，它们都逐渐接近 $w^Tp$+b=0 分界面。其中，感知机基于硬极限函数逐渐接近分类为 1 的点，ADALINE 基于线性函数逐渐接近分类为 0 的点。

========================================================================