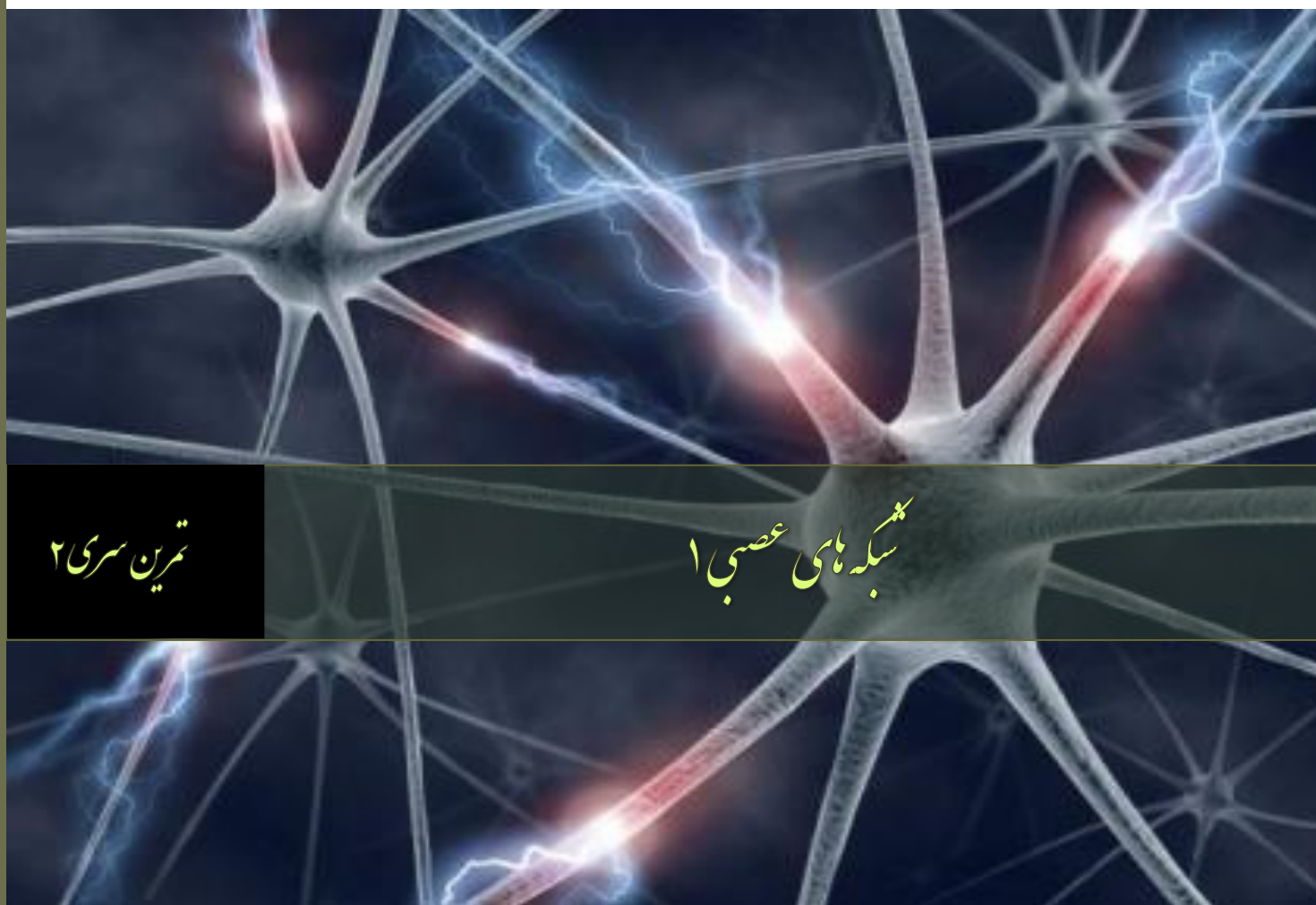


۱۳۹۱/۱۲/۱۳



تمرین سری ۲

شبکه های عصبی ۱

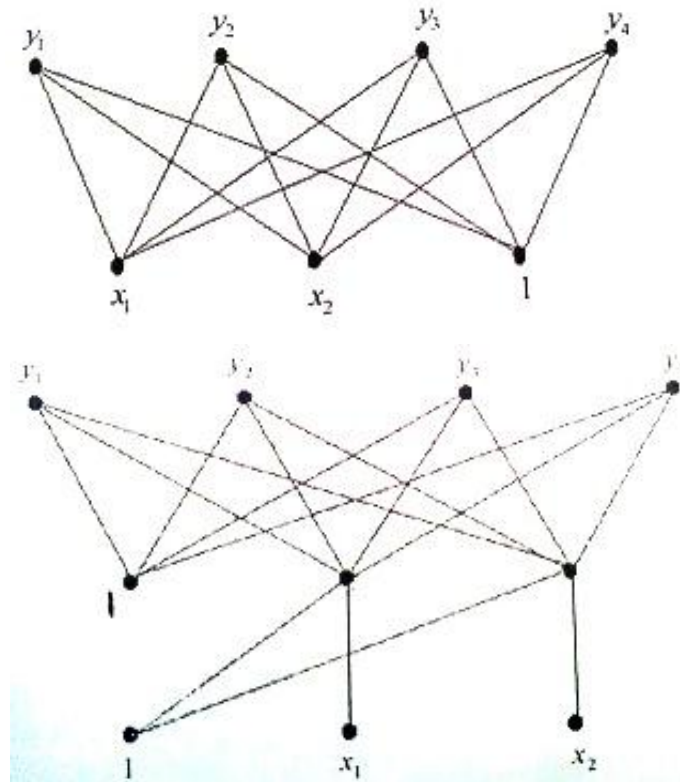
استاد درس:

جناب آقای دکتر سید صالحی

دانشجو: بهنام عادلی | ۹۰۱۳۳۰۲۲

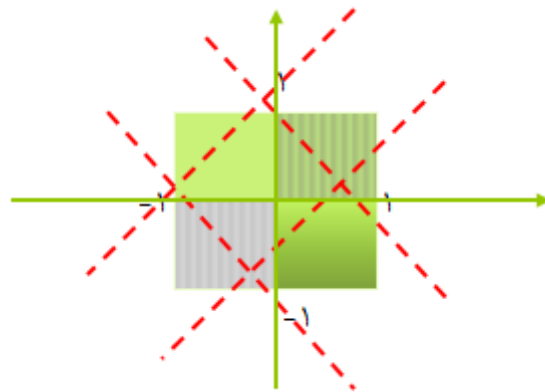
## پاسخ پرسش ۱:

فرق اصلی این دو شبکه در وجود یک پنهان در شبکه پایینی است. در شبکه شکل ۱-۲ الف، در نهایت چهار ابر صفحه ( $2-1=1$  بعدی) که در واقع یک خط است خواهیم داشت. این چهار خط فضای دو بعدی ورودی را می‌توانند بر روی ۱۶ راس یک مکعب تصور کنند. در نتیجه ما تصویر این چهار بخش را در ۱۶ نقطه خواهیم داشت که این خروجی شبکه اول خواهد بود یعنی در نهایت به چهار خط مورب در فضای ورودی خواهیم رسید پس می‌بینیم که نیاز به لایه ای دیگر هست تا این فضا را به چهار بخش تقسیم نماید تا به خروجی نهایی دست پیدا کنیم پس شبکه ۱-۲ الف) برای این کار مناسب نخواهد بود اما در شبکه ۱-۲ ب) از آنجایی که می‌توان با سطح آستانه ۱ که به عنوان ورودی دیگری برای لایه پنهان تعریف شده است  $x_1$  یا  $x_2$  را خاموش یا روشن نمود. بدین معنی که با دو خط که در فضای لایه پنهان ایجاد می‌شود می‌توان در بخشی از فضا خروجی  $x_1$  را صفر و در بخشی دیگر ۱ کرد در نتیجه می‌توان در لایه بعدی با تفکیک دو بخش منفی و مثبت هر کدام از  $x_1$  و  $x_2$  ها فضا را به چهار بخش دلخواه صورت مساله تقسیم نمود.



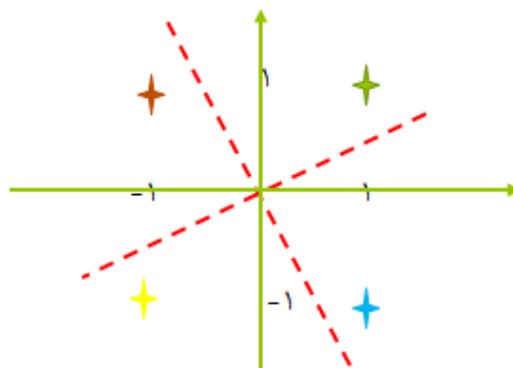
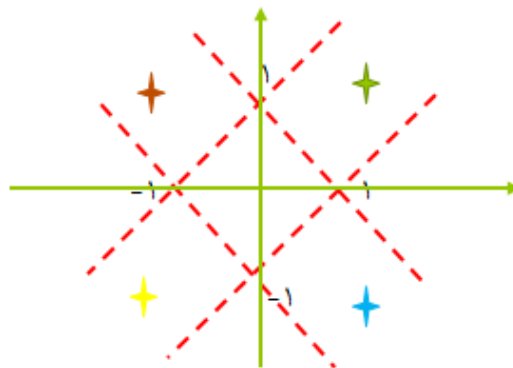
شکل ۱-۲ الف: شبکه جلوسوی تک لایه ب: شبکه جلوسوی با یک لایه پنهان

در حالتی که تنها خروجی داشته باشیم به شکل زیر می‌رسیم:



شکل ۲-۲

اما در حالتی که یک لایه پنهان داشته باشیم می‌توانیم به شکل ۲-۳ الف) و سپس با تنظیم وزن‌ها به شکل ۲-۳ ب) نیز دست پیدا کنیم:



شکل ۲-۳ الف) بالا و ب) پایین

## جواب ۲:

## (الف)

برای انجام کار به روش محاسباتی ابتدا باید ابرصفحات تصمیم را که در اینجا خط هستند بدست آوریم برای این کار از خطوط متصل کننده‌ی ورودی‌ها استفاده نموده و عمودمنصف آن‌ها را به عنوان مرز تصمیم در نظر می‌گیریم. ( با این انتخاب سیستم تصمیم گیری نسبت به نویز قوی بوده و احتمال خطای آن کمتر خواهد بود ) برای این کار:

$$x_i = \frac{\theta_j}{w_{ij}} \Rightarrow$$

$$y_3 : x_1 = \frac{\theta_3}{w_{13}} = \infty \Rightarrow w_{13} = 0, x_2 = \frac{\theta_3}{w_{23}} = -1 \Rightarrow w_{23} = -\theta_3 = -2$$

$$y_2 : x_1 = \frac{\theta_2}{w_{12}} = \frac{-2}{\sqrt{3}} \Rightarrow w_{12} = -\sqrt{3}, x_2 = \frac{\theta_2}{w_{22}} = 2 \Rightarrow w_{22} = \frac{\theta_2}{2} = 1$$

$$y_1 : x_1 = \frac{\theta_1}{w_{11}} = \frac{2}{\sqrt{3}} \Rightarrow w_{11} = \sqrt{3}, x_2 = \frac{\theta_2}{w_{22}} = 2 \Rightarrow w_{21} = \frac{\theta_2}{2} = 1$$

$$y_1: \sqrt{3}x_1 + x_2 - 2 = 0$$

$$y_2: -\sqrt{3}x_1 + x_2 - 2 = 0$$

$$y_3: -x_2 - 1 = 0$$

در نتیجه ماتریس وزن‌ها به شکل زیر در خواهد آمد:

$$w = \begin{bmatrix} \sqrt{3} & -\sqrt{3} & 0 \\ 1 & 1 & -1 \\ -2 & -2 & -1 \end{bmatrix}$$

حال Y را محاسبه می‌کنیم:

$$Y = \bar{x}\bar{w} = \begin{bmatrix} 0 & 0 & 1 \\ \sqrt{3} & 1 & 1 \\ -\sqrt{3} & 1 & 1 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{3} & -\sqrt{3} & 0 \\ 1 & 1 & -1 \\ -2 & -2 & -1 \end{bmatrix} = \begin{bmatrix} -2 & -2 & -1 \\ 2 & -4 & -2 \\ -4 & 2 & -2 \\ -4 & -4 & -2 \end{bmatrix}$$

حال با استفاده از تابع غیرخطی سخته پله واحد نتیجه خروجی را محاسبه می‌کنیم:

$$y = f(Y) = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

در صورت سوال داشتیم:

$$\bar{d} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

درنتیجه اختلاف این دو:

$$d - y = 0$$

پس می‌بینیم که طبقه بندی کننده بدون خطا کار می‌کند ( که این ایده آل است و فقط در طراحی‌های

دستی به آن می‌توان رسید).

(ب)

با استفاده از الگوریتم تعلیم پرسپترون که در فایل‌های درسی موجود بود کدهای زیر را در MATLAB

پیداسازی نمودیم:

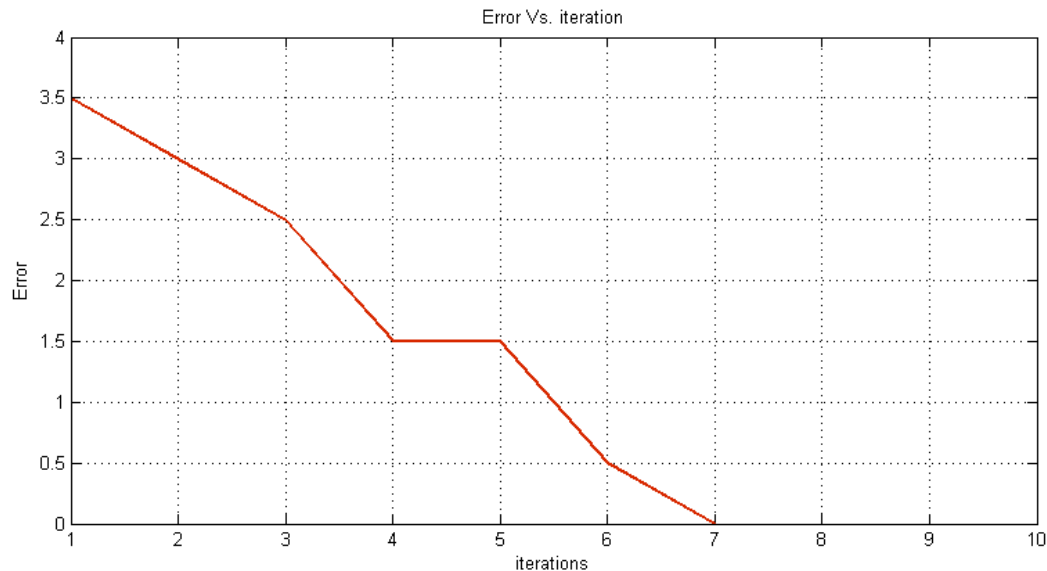
```
clear all
clc
flag = 1;
P = 4; % Number of training sample
K = 1; % Counter of training iterations
xp = [ 0 0 1;
       sqrt(3) 1 1;
       -sqrt(3) 1 1;
       0 -2 1]; % Training samples
dp = [0 0 0; 1 0 0; 0 1 0; 0 0 1]; % Desired output(s)
C = 0.1;
W = rand(3,3); % initialize W cells
```

```
f=inline('(1+hardlims(x))/2');
while(flag==1)
```

```
% f is normal step function
```

```
    E(K) = 0; % Total Square Error
    for p=1:P
        y= f(xp(p,:)*W);
        W = W + C*xp(p,:)'.*(dp(p,:)-y);
        E(K) = E(K) + sum(0.5*(dp(p,:)-y).^2);
    end
    if (E(K) < 1e-6) || (K > 50)
        flag=0;
    end
    K = K + 1;
end
k=1:length(E);
plot(k,E)
grid on
hold on
```

نمودار خطا بر حسب تعداد دفعات تکرار در شکل زیر دیده می‌شود:



شکل ۲-۴ نمودار خطا بر حسب تعداد تکرار تعلیم

ماتریس به‌دست آمده از طریق تعلیم با MATLAB به صورت زیر است:

W =

0.6020 -0.1768 -0.0894

0.4630 0.2482 -0.2710

-0.0459 -0.3495 -0.0867

مدل ساده شده ماتریس وزن‌ها که در قسمت الف) بدست آمد در زیر آمده است:

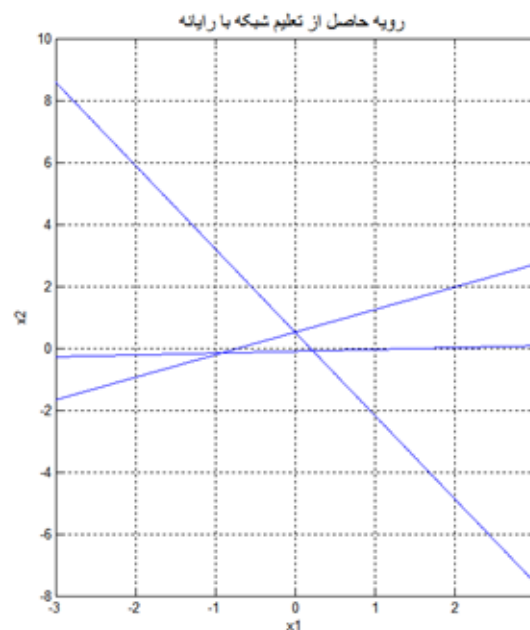
$w =$

1.7321   -1.7321   0

1.0000   1.0000   -1.0000

-2.0000   -2.0000   -1.0000

بدیهی است که شبکه ای که خودمان طراحی کرده‌ایم قابلیت بهتری دارد زیرا حین تعلیم ممکن است ابرصفحات از حالت ایده آل خود جابجا شده و تحمل اعوجاج آن‌ها به علت برهم خوردن تقارن کم شود زیرا ماتریس وزن‌ها را از یک مقدار تصادفی کوچک آغاز کردیم، پس ممکن است در هر بار اجرا نقاط مجاور در تقسیم بندی مقداری فرق کنند.



شکل ۲-۵ رویه‌های تصمیم محاسبه شده با رایانه

(ج)

برای این کار با استفاده از تابع rand مقدار اتفاقی کوچکی به ورودی شبکه‌ها در کدهای بالا اضافه کرده و نتیجه را مشاهده می‌کنیم:

x1 =

0.5962 0.3290 1.4782

ans =

0 0 0

ans =

0 0 0

x2 =

2.3292 1.1614 1.8295

ans =

1 0 0

ans =

1 0 0

x3 =

-0.7759 1.5955 1.0287

ans =

0 1 0

ans =

1 1 0

x4 =

0.8121 -1.3899 1.7015

ans =

0 0 0

ans =

0 0 1



به ازای ورودی نویزی ۳ شبکه اول و به ازای ورودی نویزی ۴ شبکه دوم جواب صحیح داده است. بنابراین به طور کلی نمی توان گفت که پاسخ کدام شبکه حساسیت کمتری به نویز داشته و این مهم بسته به شرایط مسئله می تواند متفاوت باشد.