ENEE434 HW#2 solution YJ 02/22/2006

1). Matlab function ternsig.m:

$$f(u) = ternsig(u) = \begin{cases} \tanh(u-1) & for \quad u > 1 \\ 0 & for \quad -1 \le u \le 1 \\ -\tanh(u+1) & for \quad u < -1 \end{cases}$$

modify tansig.m as following to get ternsig.m

```
nsize=size(n);
s_r=nsize(1,1);
s_c=nsize(1,2);
for i=1:s_r
   for j=1:s_c
      if(n(i,j)>1)
         a(i,j)=tanh(n(i,j)-1);
      elseif(n(i,j)<-1)
         a(i,j)=-tanh(n(i,j)+1);
      else
         a(i,j)=0;
      end
   end
end
```

The derivative of ternsig, dternsig.m

```
nsize=size(n);
s_r=nsize(1,1);
s_c=nsize(1,2);
for i=1:s_r
   for j=1:s_c
      if(n(i,j)>1)
         a(i,j)=1/[cosh(n(i,j)-1)]^2;
      elseif(n(i,j)<-1)
         a(i,j)=-1/[cosh(n(i,j)+1)]^2;
      else
         a(i,j)=0;
      end
   end
end
```

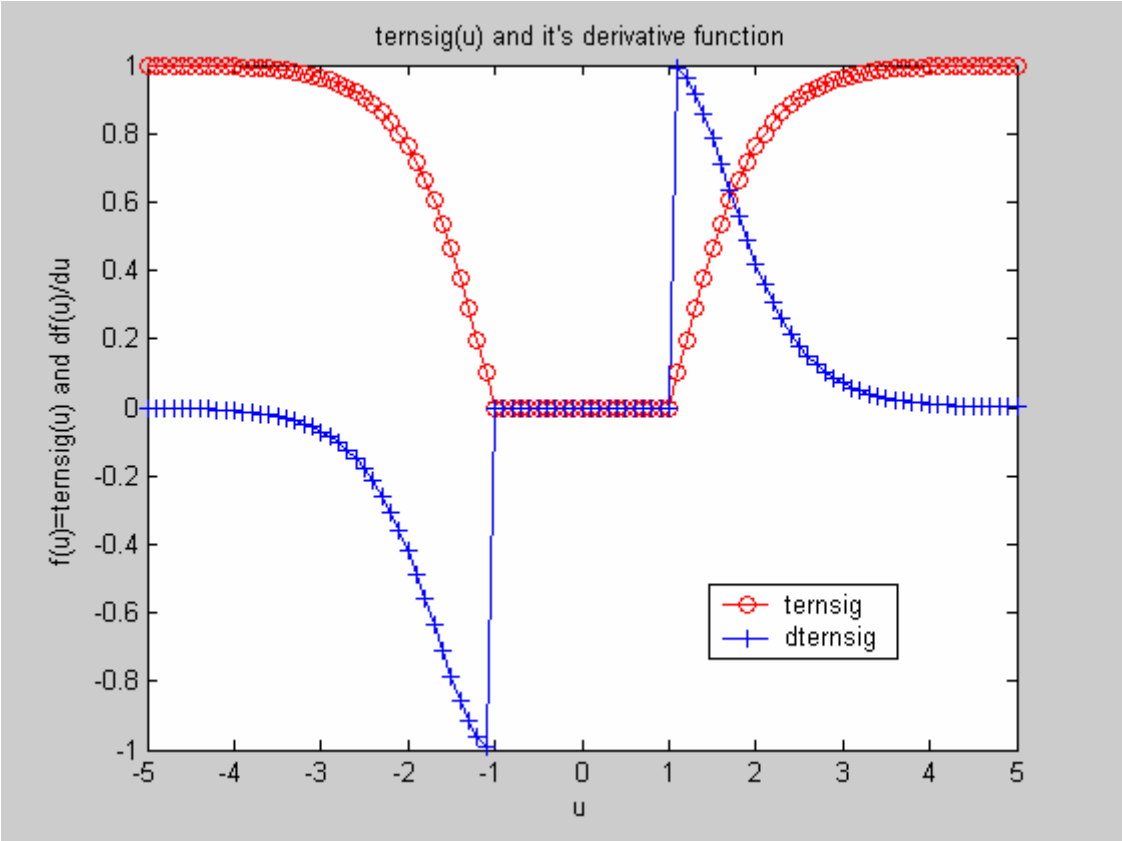
The NN: important code lines:

```
netHw2_1 = newff([-20 20; -20 20],[2 2 2],{'ternsig' 'ternsig' 'purelin'});

netHw2_1.trainParam.epochs = 100;
netHw2_1 = train(netHw2_1,P,T);
```

2. For the NN from problem 1, there are 3 layers, each layer has 2 neorons. The weights for all three layers are 2 by 2 matrices, and the bias are 2 by 1 matrices.
The activation functions for the first two layers are ternsig (from HW1), the activation function for the last layer is purelin. To find the sensitivity, the dirivation function of ternsig is need, a function named dternsig is defined as the derivative function of ternsig. Obviously, the derivative of purelin is just 1.

$$\frac{df(u)}{du} = dternsig(u) = \begin{cases} \dfrac{1}{\cosh^2(u-1)} & for \quad u > 1 \\ 0 & for \quad -1 \le u \le 1 \\ \dfrac{-1}{\cosh^2(u+1)} & for \quad u < -1 \end{cases}$$



$$a^3 = f^3\left(W^3 \cdot f^2\left(W^2 \cdot f^1\left(W^1 \cdot P + b^1\right) + b^2\right) + b^3\right)$$

a) First pass:

i) calculating the output and error from input:

W^1	$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$	W^2	$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$	W^3	$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$
b^1	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	b^2	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	b^3	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$
$n^1 = W^1 \cdot P + b^1$	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	$n^2 = W^2 \cdot a^1 + b^2$	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	$n^3 = W^3 \cdot a^2 + b^3$	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$
$a^1 = ternsig\left(n^1\right)$	$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$	$a^2 = ternsig\left(n^2\right)$	$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$	$a^3 = purelin\left(n^3\right)$	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$$e = T - a^3 = \begin{bmatrix} 1 \\ -5 \end{bmatrix}$$

$$F = e^T \cdot e = 26$$

ii) Backpropogating the sensitivity:

$\dot{F}^3(n^3)$ pureline	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\dot{F}^2(n^2)$ ternsig	$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$	$\dot{F}^1(n^1)$ ternsig	$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$
$S^3 = -2 \cdot \dot{F}^3(n^3) \cdot e$	$\begin{bmatrix} -2 \\ 10 \end{bmatrix}$	$S^2 = \dot{F}^2(n^2) \cdot (W^3)^T \cdot S^3$	$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$	$S^1 = \dot{F}^1(n^1) \cdot (W^2)^T \cdot S^2$	$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$

iii)Updating weights and biases:

$W^3 = W^3 - \alpha \cdot S^3 \cdot (a^2)^T$	$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$	$W^2 = W^2 - \alpha \cdot S^2 \cdot (a^1)^T$	$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$	$W^1 = W^1 - \alpha \cdot S^1 \cdot (P)^T$	$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$
$b^3 = b^3 - \alpha \cdot S^3$	$\begin{bmatrix} 1.4 \\ -1 \end{bmatrix}$	$b^2 = b^2 - \alpha \cdot S^2$	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	$b^1 = b^1 - \alpha \cdot S^1$	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

Only the bias for the third layer b3 has been changed.

b) 2nd pass:

W^1	$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$	W^2	$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$	W^3	$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$
b^1	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	b^2	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	b^3	$\begin{bmatrix} 1.4 \\ -1 \end{bmatrix}$
n^1	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	n^2	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	n^3	$\begin{bmatrix} 1.4 \\ -1 \end{bmatrix}$
a^1	$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$	a^2	$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$	a^3	$\begin{bmatrix} 1.4 \\ -1 \end{bmatrix}$

$$e = T - a^3 = \begin{bmatrix} 0.6 \\ -3 \end{bmatrix}$$

$$F = e^T \cdot e = 9.36$$

$\dot{F}^3(n^3)$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\dot{F}^2(n^2)$	$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$	$\dot{F}^1(n^1)$	$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$
S^3	$\begin{bmatrix} -1.2 \\ 6 \end{bmatrix}$	S^2	$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$	S^1	$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Update weights and bias

W^1	$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$	W^2	$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$	W^3	$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$
b^1	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	b^2	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	b^3	$\begin{bmatrix} 1.64 \\ -2.2 \end{bmatrix}$