CPSC 636-600 Homework 3 (Total 100 points)

See the course web page for the **due date** and **submission info**.

Instructor: Yoonsuck Choe

February 27, 2008

Problem 1 (Written: 20 pts): Solve Exercise 5.3.

Problem 2 (Written: 20 pts): Repeat Exercise 5.3 with only two hidden units, centered at $\mathbf{t}_1 = [0, 1]^T$ and $\mathbf{t}_2 = [1, 0]^T$.

Note: Instead of ϕ^{-1}, you need to find ϕ^{+}, the pseudo-inverse.

Problem 3 (Program: 40 pts): Write two Matlab/Octave programs to conduct (1) RBF learning and (2) Generalized RBF learning for 2D input and 1D output. That is, there are two input units and one output unit. For the GRBF, compare the two approaches that use fixed centers selected at random vs. self-organized selection of centers.

Programming hints: You can quickly calculate the ϕ matrix or the $\mathbf{G}$ matrix using matrix operations.

```
# The input: each row is one input vector
octave:1> x = [0 0 ; 0 1; 1 0 ; 1 1]
x =

  0  0
  0  1
  1  0
  1  1

# The target: each row is one target
octave:2> d = [0 1 1 0]'
d =

  0
  1
  1
  0

# Extract 1st vector component of the input vector
octave:3> x1 = x(:,1)
x1 =

  0
```

```
  0
  1
  1

# Extract 2nd vector component of the input vector
octave:4> x2 = x(:,2)
x2 =

  0
  1
  0
  1

# Center vectors: each row is one center vector
octave:5> t = [0 1; 1 0]
t =

  0  1
  1  0

# Extract 1st vector component of the center vector
octave:6> t1 = t(:,1)
t1 =

  0
  1

# Extract 2nd vector component of the center vector
octave:7> t2 = t(:,2)
t2 =

  1
  0


# Calculate input count and hidden layer count
octave:8> [m,junk] = size(t)
m = 2

octave:9> [N,junk] = size(x)
N = 4

# Make t1 matrix (N x m)
octave:10> T1=ones(N,1)*t1'
T1 =

  0  1
  0  1
```

```
   0  1
   0  1

# Make t2 matrix (N x m)
octave:11> T2=ones(N,1)*t2'
T2 =

   1  0
   1  0
   1  0
   1  0

# Make x1 matrix (N x m)
octave:12> X1= x1*ones(1,m)
X1 =

   0  0
   0  0
   1  1
   1  1

# Make x2 matrix (N x m)
octave:13> X2=x2*ones(1,m)
X2 =

   0  0
   1  1
   0  0
   1  1

# Set width of RBF
octave:14> sigma=0.4
sigma = 0.40000

# In one line, get the G matrix using matrix operations.
octave:105> G = exp(-((X1-T1).^2 + (X2-T2).^2)/(2*sigma^2))
G =

   0.0439369  0.0439369
   1.0000000  0.0019305
   0.0019305  1.0000000
   0.0439369  0.0439369

# Find pseudoinverse
octave:15> w = inv(G'*G)*G'*d
w =

   0.99045
```

```
  0.99045

# Verify answer (each row in G corresponds to one input,
# so the resulting column vector contains all outputs
# for all the inputs.
octave:16> G*w
ans =

  0.087035
  0.992367
  0.992367
  0.087035
```

Problem 4 (Written: 20 pts): Test RBF and GRBF on a classification task similar to section 4.8 in the textbook.

First, construct the positive (class 1) and negative inputs (class 2):

```
pos=[randn(100,1),randn(100,1)];
neg=[randn(100,1)*sqrt(5)+5,randn(100,1)*sqrt(5)];
plot(pos(:,1),pos(:,2),'r+',neg(:,1),neg(:,2),'bx');

x=[pos;neg];
d=[ones(100,1);zeros(100,1)];
```

You can see that the two point clouds are scattered near $[0, 0]^T$ and $[5, 0]^T$, with a Gaussian distribution. For classification tasks using RBF, usually multiple output units are used, but here, let us just use 1 output unit for similicity (class 1 if output > 0.5, and class 2 otherwise).

You can also construct a separate *test set* using the same commands and saving the results in a different variable (running `randn()` multiple times will give you a new set of points). Do not use the test set during training.

1. Test performance (classification accuracy) using RBF on the training set and the test set. Experiment with various standard deviations (spread).

2. Test performance using GRBF with bias, with two random centers picked from the training input set, and std 2 for both centers, on both the training set and the test set. Experiment with various standard deviations (spread).

3. Test performance using GRBF with bias, with two self-organized centers (see Haykin p. 301), and std 2 for both centers, on both the training set and the test set. Experiment with various standard deviations (spread). Report the learned centers and compare those to the ideal centers.

4. Test performance using GRBF with bias, with centers $[0, 0]^T$ and $[5, 0]^T$, and std 2 for both centers, on both the training set and the test set. Compare the performance with (3). Experiment with various standard deviations (spread).

5. Based on your experiments, summarize the strengths and weaknesses of the various approaches above.