Homework Set 10

1. Consider the function:

$$f(x_1, x_2) = 5x_1^2 + x_2^2 + 4x_1x_2 - 14x_1 - 6x_2 + 20$$

   - What is the maximum or minimum of this function? Use the theory that we learned earlier.
   - Suppose we begin at the point $(0, 10)$ and set the step size to be a fixed small number, say 0.001.

     By hand, perform gradient descent twice using the step size in place of the Golden Section search.

2. Code up the solution to the pattern classification problem (Exercise 1). Be sure to use the examples in the text to help you! For your convenience, here was that problem:

   > Let's put all of this together to solve another pattern classification problem using Hebb's rule. Suppose we are given the following associations:
   >
   > | Point | Class |
   > |---|---|
   > | $(1, 1)$ | 1 |
   > | $(1, 2)$ | 1 |
   > | $(2, -1)$ | 2 |
   > | $(2, 0)$ | 2 |
   > | $(-1, 2)$ | 3 |
   > | $(-2, 1)$ | 3 |
   > | $(-1, -1)$ | 4 |
   > | $(-2, -2)$ | 4 |
   >
   > In this example, take Class 1 to be the vector $[-1, -1]^T$, Class 2 as vector $[-1, 1]^T$, Class 3 as $[1, -1]^T$, and Class 4 as $[1, 1]^T$- this puts the 4 classes are on the vertices of a square.
   >
   > Now for the details of the program. First write the inputs as an 2×8 matrix, with a corresponding output matrix that is also 2×8. Parameters that can be placed first will be the maximum number of times through the data N and the learning rate, a, which we will set to 0.04. We can also set an error bound so that we might stop early. Set the initial weights to the 2×2 identity, and the bias vector b to $[1, 1]^T$.
   >
   > Be sure you have trained long enough to get a good error, and plot the decision boundaries as well.

3. To be sure you understood what was going on in the previous problem, do the first iteration by hand. That is, take W to be the identity matrix and $\mathbf{b}$ to be the vector $[1, 1]^T$. Take $\alpha = 0.04$ and use the first data point to update the matrix W and $\mathbf{b}$ using the Widrow-Hoff update.

4. Solve the previous problem again using batch training (again, be sure to use the examples from the handout).

5. For the nonlinear feedforward neural network, use $\sigma(r) = \frac{1}{1+e^{-r}}$, and suppose you have a $1-2-2$ network, with

$$W^{(1)} = \begin{bmatrix} 1 \\ -0.5 \end{bmatrix}, \mathbf{b}^{(1)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad W^{(2)} = \begin{bmatrix} 0.1 & 0.2 \\ -0.2 & 0.1 \end{bmatrix}, \mathbf{b}^{(2)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Let $x = 1$, and find the output of the neural network (by hand). Be sure to denote the prestate and state vectors for each layer.

6. **Practice with Matlab and photos:** Download the sample data set from the class website, `FacesTrial01.mat`. In Matlab, to open the file (navigate to where you saved the file first), type:

```
load FacesTrial01
```

You'll see a matrix X that is 1567500×32. Each column represents a photo that is 1425×1100 (that is, $1425 \times 1100 = 1567500$).

To see what the 31st photo is, for example, you would type:

```
Temp=reshape(X(:,31),1425,1100);   %Creates the photo
imagesc(Temp);    %Display as a scaled image (the sc is for scaling)
colormap(gray); axis equal; axis off;  % Squares up the image
```

You'll also see a vector G and a vector M- This is an index telling you which photos are for women and men (respectively).

Assignment: Look up the "mean" command, and see if you can figure out how to produce the mean photo for all the men, the mean photo for all the women, and the mean photo for everyone. Display these images as well.