CPSC 636-600 (Total 100 points)
See the course web page for the **due date** and **submission info**.

Instructor: Yoonsuck Choe

January 27, 2008

1 Learning Rules

Problem 1 (Written: 15 pts): Solve exercise 2.8 in the textbook.

Hint: For a random variable X and its mean μ, the variance is defined as $E[(X - \mu)^2]$. This is equivalent to $E[X^2] - E[X]^2$. The elementary techniques used in the derivation below will help you solve this exercise problem. Note: $E[X] = \mu$, by definition.

$$
\begin{aligned}
E[(X - \mu)^2] &= E[X^2 - 2\mu X + \mu^2] \\
&= E[X^2] - 2\mu E[X] + \mu^2 \\
&= E[X^2] - 2E[X]E[X] + E[X]^2 \\
&= E[X^2] - 2E[X]^2 + E[X]^2 \\
&= E[X^2] - E[X]^2.
\end{aligned}
$$

2 Learning Paradigms

Problem 2 (Written: 5 pts): Solve exercise 2.14.

3 Memory

Problem 3 (Written: 10 pts): Solve exercise 2.18.

Problem 4 (Written: 15 pts): Solve exercise 2.19.

Problem 5 (Written: 15 pts): Solve exercise 2.20.

4 Statistical Learning Theory

Problem 6 (Written: 15 pts): (1) List all possible dichotomies for the input set $\{x_1, x_2, x_3\}$ shown below (figure 1a). (2) Can classifiers defined by straight line decision boundaries shatter this set? Assume one side of the line is classified as 0, and the other side 1: figure 1b. (Note: if you rotate this line by 180 degrees, the classification of the points will be flipped.) (3) Can classifiers defined by rectangular decision boundaries (assume the rectangles have their sides parallel to the x and the y axis, i.e., you cannot rotate the rectangles: figure 1c)? Also assume the interior of the rectangle is classified as 1 and the exterior 0.
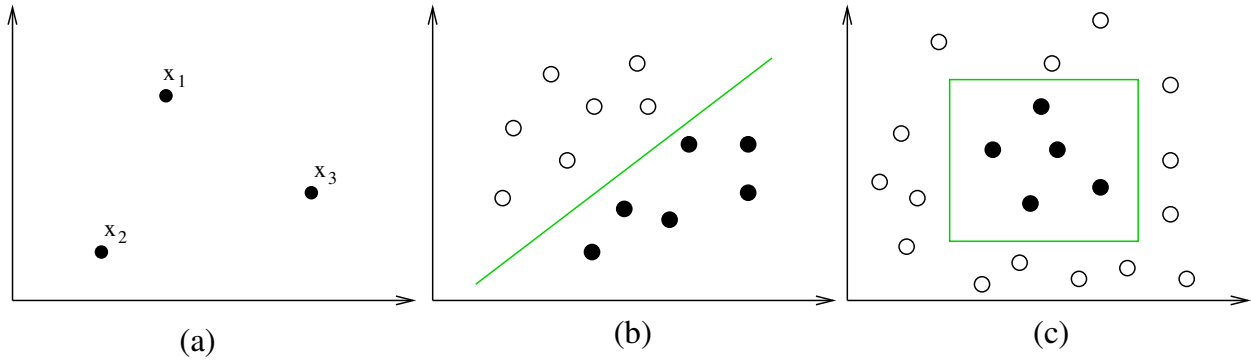
Figure 1: **VC-dimension Example.** Filled and open discs represent inputs of two different classes.

Problem 7 (Written: 5 pts):   What is the VC-dimension of the rectangular classifier in the above problem, for a 2D input space?

5   Programming exercise

Problem 8 (Written: 5 pts):   Show by example that equation 2.34 and 2.35 are equivalent, using the following $\mathbf{X}$ and $\mathbf{Y}$ (see equations 2.36 and 2.37: each column in the matrices represents one input or output vector). Follow the example in `src/partmat.m`. You can take out a particular vector from the matrix using the following trick: `k = 3; xk = X(:,k)`. This will give you $[8, 2, 4]^T$.

```
X =

   3   5   8   6   9
   6   6   2   5   3
   5   3   4   7   7

Y =

   6   9   7   6   7
   4   5   4   4   5
```

Problem 9 (Written: 15 pts):   In Matlab or Octave, follow the instructions below to generate the key matrix `X` and memory matrix `Y`. You don't need to understand what it does, at this point.

```
dat = rand(100,3); cv = cov(dat); [X,Xd] = eig(cv); X
Y = rand(5,3)
```

In Matlab or Octave, (1) Calculate the norm of each resulting column vector in `X`. You can easily do this by `X[:,1]'*X[:,1]`, etc. (2) Check if the column vectors in `X` are all *nearly* orthogonal with each other. Apply the technique in (1). (3) Construct $\hat{\mathbf{M}}$ as in equation 2.35, and verify that it has the property in equation 2.39 for all the key-vector/memory-vector pairs. Show all your results (print out).