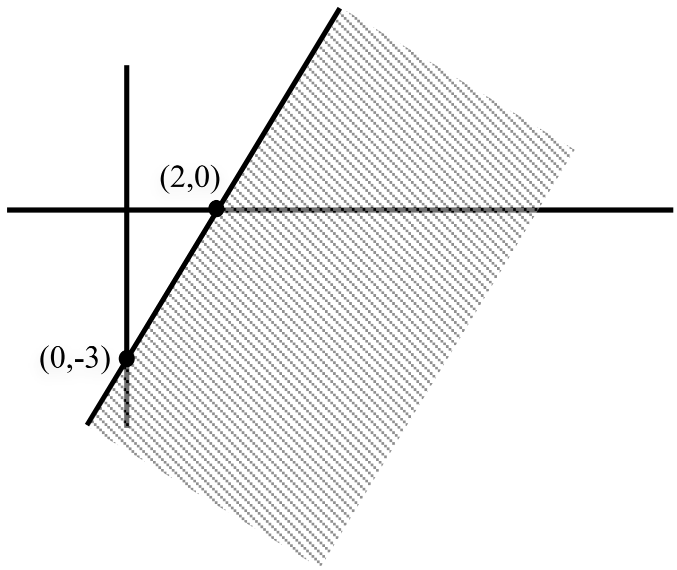
The Johns Hopkins University
JHU ENGINEERING FOR PROFESSIONALS PROGRAM
NEURAL NETWORKS: 625-438/605-447

Solutions to Problem Set #1

1.1 Sketch the region $3x - 2y \geq 6$.

Ans:



1.2 Consider the x-y plane. Given that the points (0,0), (0,1), and (1,0) belong to the set **R** and that the point (1,1) belongs to set **S**. Write the equation of a straight line (i.e., $ax + by = k$) which separates sets **R** and **S** (*i.e.,* all points of **R** are on one side of the boundary line and all points of **S** are on the other side.)

Ans. $x + y = 1.5$

1.3 Now let (1, 1) and (-1,-1) belong to set R and (1, -1) and (-1, 1) belong to set S. Can sets R and S be separated with a straight line? How about with a general curve of the form $ax^2 + bx + cy^2 + dy + exy = k$?

Ans. No, the two sets cannot be separated by a linear function. This problem is similar to the XOR problem. A function of the form noted above however can segregate the two sets **R** and **S** because of the second order terms.

1.4 The sigmoid activation function below has been slightly modified by employing a coefficient c that is multiplied to the activity function value A. Answer the following questions:

$$v = \frac{1}{1 + e^{-cA}} = \frac{1}{1 + e^{-c\left(\sum_{i=1}^{n} w_i x_i + \theta\right)}}$$

a) What effect does changing c to be different from 1 have on the function?

Ans. It affects the steepness of the curve as it transitions from values close to zero to close to one. For example, if $c = 1$, the slope of the function when $x = 0$ is ¼. Increasing c increases the slope at $x = 0$. Another way to look at it is that it stretches the graph horizontally.

b) What effect does changing the bias value θ have on the function?

Ans. If $\theta > 0$ in the equation above, the sigmoid curve is shifted to the left. Similarly for $\theta < 0$ which shifts the curve to the right.

1.5 A Perceptron's output value y is often based on the *activity function.* This output may then be modified by another function or functions such as the *activation function* and/or a *threshold logic* function. Please explain the utility of using these modifications and why they are used in some situation and not in others.

Ans. These activation functions perform modifications of the activity function that are necessary and convenient to further, useful processing. For example, the threshold logic function could be used to either signal or not signal some class membership in a classification problem; or implement some logic function. Different activations functions can segregate different parts of the input space in different ways.

1.6 Sometimes we want to model output as a binary value; that is, a function that produces only one of two possible values such as 0 and 1. Sometimes we want this property to have additional mathematical characteristics that involve combining two outputs. In the binary example, two different outputs (0 and 1) can be combined using certain mathematical operations to produce a 0. How would you design an output function using just two output values so a combined output using the same mathematical operation as alluded to above, produces a 1?

Ans. For the AND function, the connective operation could simply be multiplication (x). For the OR function, it could be addition (+). For XOR, we could combine addition, subtraction and multiplication thusly: A XOR B = A+B – AxB.

1.7 As mentioned in the videos, the action potential is best described by an electro-chemical 'pulse' and that these pulses can occur with a frequency up to several hundred per second. In electrical engineering, one modeling technique employs the impulse function, a single value of infinite height and zero width --- a mathematical trick to capture the effects of a sudden 'pulse'. Consider a single pulse of zero width and unit height and a train of two such pulses and their corresponding outputs. Notice that the outputs of each pulse add together. Now consider a train of such pulses.

What is the effect on the output of increasing the frequency (shortening the time between pulses) of the input pulses? Describe in plain English, the long run behavior of the output with varying frequency of the input.

Ans. The closer the pulses are, the higher the resultant amplitude or magnitude. Increasing the frequency therefore increases the resultant amplitude. This is akin to a duty cycle in electronic systems where the average actuation depends on the frequency of pulses.