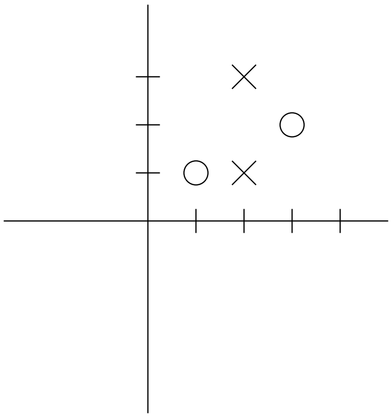
T-61.231 Principles of Pattern Recognition

Exercise 5: 29.10.2001

1. Formulate the SVM Dual problem with constraints for the not linearly separable case shown in the picture below, with vectors $x_1 = [1\ 1]^T \in \omega_1$, $x_2 = [2\ 1]^T \in \omega_2$, $x_3 = [3\ 2]^T \in \omega_1$, $x_4 = [2\ 3]^T \in \omega_2$.



2. Let us consider a 2-layer MLP network (1 hidden layer, 1 output layer), which is intended for classification task. Network has N_x input neurons, N_h hidden layer neurons, and N_y output neurons.

   a) How many weights, including offsets, does the network have?

   b) Let N_t be the number of training vectors and N_w number of weights, including offsets, calculated previously in a).

   According to Widrow's rule, the network has a good ability of generalization when

   $$N_t >= 10 \times N_w / N_y.$$

   If we would like to classify characters by using this network so that $N_y = 10$ (explain what this means) and the input images are raster images of size 16×16 (what is N_x ?), what is going to be the ratio between the size of the hidden layer and the size of the training set when we are at the Widrow's limit?

3. A two-class-classifying problem is illustrated in Figure 1. Determine by hand (without using any learning algorihm) a MLP network, which has 2 input neurons, 2 output neurons and a suitable number of neurons in hidden layer. Network has to give vector (1,0) or (0,1) as an output when input belongs to class 1 or 2, respectively.

   What is the smallest number of hidden layer neurons? Is there an upper limit to the number of hidden layer neurons after which the classifier does not work any more?

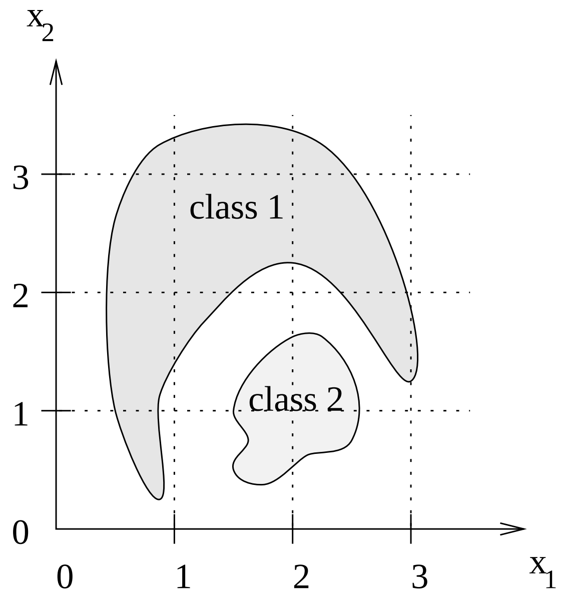


Figure 1:

4. Draw the three lines in the two-dimensional space

$$\begin{aligned} x_1 + x_2 &= 0 \\ x_2 &= \tfrac{1}{4} \\ x_1 - x_2 &= 0 \end{aligned}$$

   For each of the polyhedra formed by their intersections, determine the vertices of the cube into which they will be mapped by the first layer of a multilayer perceptron, realizing the preceeding lines. Combine the regions into two classes so that (a) a two-layer network is sufficient to classify them and (b) a three-layer network is necessary. For both cases compute analytically the corresponding synaptic weights.