2 Neural Network Representation

We first need to work in the x_1 and x_2 plane in order to delimit the region of class 'triangle' ($f(x) = 1$) and the class 'square' ($f(x) = 0$) (see Fig. 1).

$$\begin{aligned}
&d1 : 5\mathrm{x}_1 + 4\mathrm{x}_2 + 20 = 0 \\
&d2 : -5\mathrm{x}_1 + 4\mathrm{x}_2 + 20 = 0 \\
&d3 : -\mathrm{x}_1 - \mathrm{x}_2 + 4 = 0
\end{aligned} \tag{2.1}$$

We can then define that the pre-activation layer directly from the above equations, that is

$$\begin{aligned}
\mathrm{y}_1 &= 5\mathrm{x}_1 + 4\mathrm{x}_2 + 20 \\
\mathrm{y}_2 &= -5\mathrm{x}_1 + 4\mathrm{x}_2 + 20 \\
\mathrm{y}_3 &= -\mathrm{x}_1 - \mathrm{x}_2 + 4
\end{aligned} \tag{2.2}$$

Here, we have defined $\mathrm{w}_{11} = 5$, $\mathrm{w}_{12} = 4$, $\mathrm{w}_{21} = -5$, $\mathrm{w}_{22} = 4$, $\mathrm{w}_{31} = -1$, $\mathrm{w}_{32} = -1$, $\mathrm{b}_1 = \mathrm{b}_2 = 20$ and, finally, $\mathrm{b}_2 = 4$

The region of class 'triangle' is defined in the 2-dimensional space x_1 and x_2 as

$$H(\mathrm{y}_1) = H(\mathrm{y}_2) = H(\mathrm{y}_3) = 1. \tag{2.3}$$

This solution in the 3-dimensional space formed by y_1, y_2 and y_3 correspond to the point $(1, 1, 1)$, as presented in Fig. 2. There are three regions associated to the class 'square', namely $(0, 1, 1)$, $(1, 0, 1)$, and $(1, 1, 0)$. Thus, we need to find a set of parameters u_1, u_2, u_3 and c such that the classifier will able to discriminate the two classes.

The solution correspond a plane with normal vector pointing along the point $(1, 1, 1)$ and goes through any point that satisfies the classification. We chose to the point $(1, 0.5, 1)$. Again, see 2 for more details.

The equation of a plane in 3-dimension is given by

$$a\mathrm{y}_1 + b\mathrm{y}_2 + c\mathrm{y}_3 + d = 0 \tag{2.4}$$

where the vector (a, b, c) is the normal to the plane. We choose the vector $(1, 1, 1)$ which each component corresponds to the values of u_1, u_2 and u_3 respectively. We can find the parameter d by solving the above equation for the point $(1, 0.5, 1)$. We find that $d = -2.5$.

In summary, we have that

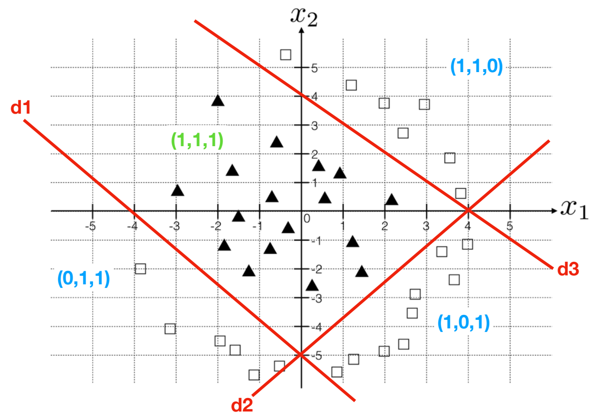

Figure 1: 2-dimensional space formed by the vectors x_1, x_2. The coordinates in blue and green correspond to the different classes in the 3-dimensional space formed by the vectors y_1, y_2 and y_3. See 2 for more details.

.

$$\begin{pmatrix} w_{1,1} & w_{1,2} & b_1 \\ w_{2,1} & w_{2,2} & b_2 \\ w_{3,1} & w_{3,2} & b_3 \end{pmatrix} = \begin{pmatrix} 5 & 4 & 20 \\ -5 & 4 & 20 \\ -1 & 1 & 4 \end{pmatrix}$$

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ -2.5 \end{pmatrix}. \tag{2.5}$$

It is worth mentioning that we could have defined the lines d_1 and d_2 such that $f(0) = -4$ instead of $f(0) = -5$. In that case, the class 'triangle' would fall in the two regions $(0, 0, 1)$ and $(1, 0, 0)$ in the 3-dimensional space formed by y_1, y_2 and y_3. We decided otherwise.

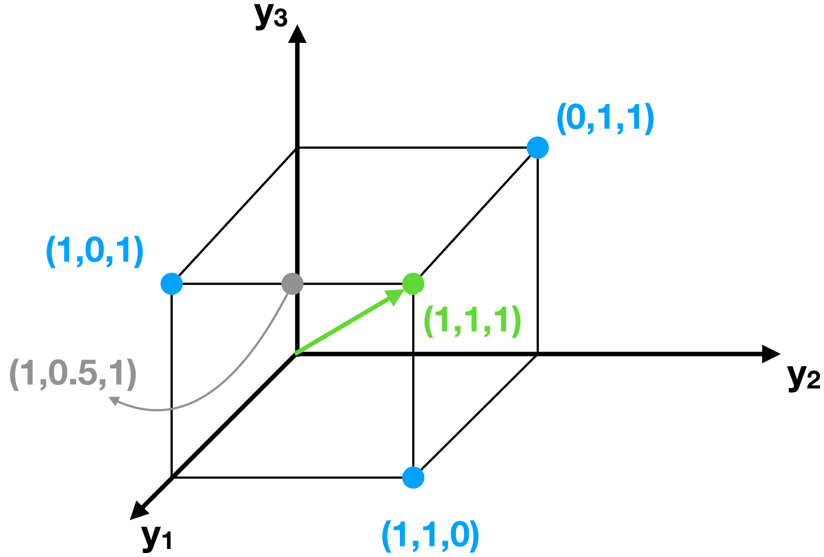

Figure 2: 3-dimensional space formed by the vectors y_1, y_2 and y_3. These components correspond to the the activated layer.

.