

One only needs to separate the all-zero vector from the rest by a linear hyperplane.

This may be done by a classifier of the form

$$x^T \underline{1} \geq c \text{ for any } c \in (0, 1)$$

where $\underline{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ is the all-ones vector.

The distance of $\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ to this plane is $\frac{c}{\sqrt{d}}$

(since the nearest point in the plane to the origin is

$(\frac{c}{d}, \dots, \frac{c}{d})$). The distance to both classes is maximized for $c = \frac{1}{2}$ (halfway between the two extremes) which gives $2/\sqrt{d}$ as the margin.

In the second case the optimal separating plane is

$$x^T \underline{1} \geq d/2$$

The margin is half the distance between the planes

$$x^T \underline{1} = \frac{d+1}{2} \text{ and } x^T \underline{1} = \frac{d-1}{2}$$

which is again $2/\sqrt{d}$