(where $\|w\| = \sqrt{w^\top w}$ is the Euclidean norm of w), it is convenient to temporarily assume that $\|w\| = 1$, so that

$$d(x, H) = |w^\top x - b|.$$

See Figure 50.13. Then with our sign convention, we have

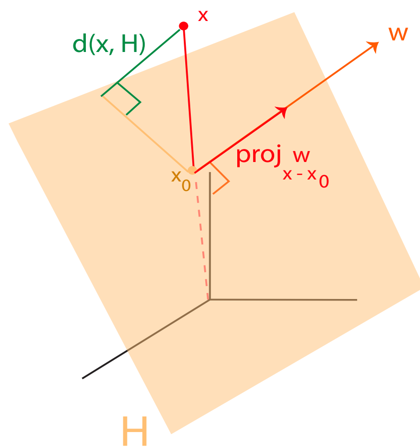

Figure 50.13: In $\mathbb{R}^3$, the distance from a point to the plane $w^\top x - b = 0$ is given by the projection onto the normal w.

$$\begin{aligned} d(u_i, H) &= w^\top u_i - b & i &= 1, \ldots, p \\ d(v_j, H) &= -w^\top v_j + b & j &= 1, \ldots, q. \end{aligned}$$

If we let

$$\delta = \min\{d(u_i, H),\ d(v_j, H) \mid 1 \leq i \leq p,\ 1 \leq j \leq q\},$$

then the hyperplane H should chosen so that

$$\begin{aligned} w^\top u_i - b &\geq \delta & i &= 1, \ldots, p \\ -w^\top v_j + b &\geq \delta & j &= 1, \ldots, q, \end{aligned}$$

and such that $\delta > 0$ is *maximal*. The distance δ is called the *margin* associated with the hyperplane H. This is indeed one way of formulating the two-class separation problem as an optimization problem with a linear objective function $J(\delta, w, b) = \delta$, and affine and quadratic constraints (SVM$_{h1}$):

$$\begin{aligned} &\text{maximize} \quad \delta \\ &\text{subject to} \\ &\qquad w^\top u_i - b \geq \delta \qquad i = 1, \ldots, p \\ &\qquad -w^\top v_j + b \geq \delta \qquad j = 1, \ldots, q \\ &\qquad \|w\| \leq 1. \end{aligned}$$