Deep Learning - Foundations and Concepts

Chapter 5. Single-layer Networks: Classification

nonlineark@github

February 12, 2025

Outline

Discriminant functions

- The goal in classification is to take an input vector $x \in \mathbb{R}^D$ and assign it to one of K discrete classes $\mathcal{C}_k$.
- A discriminant is a function that takes an input vector x and assigns it to one of K classes, denoted $\mathcal{C}_k$.
- We will restrict attention to linear discriminants, for which the decision surfaces are hyperplaines.

Two classes

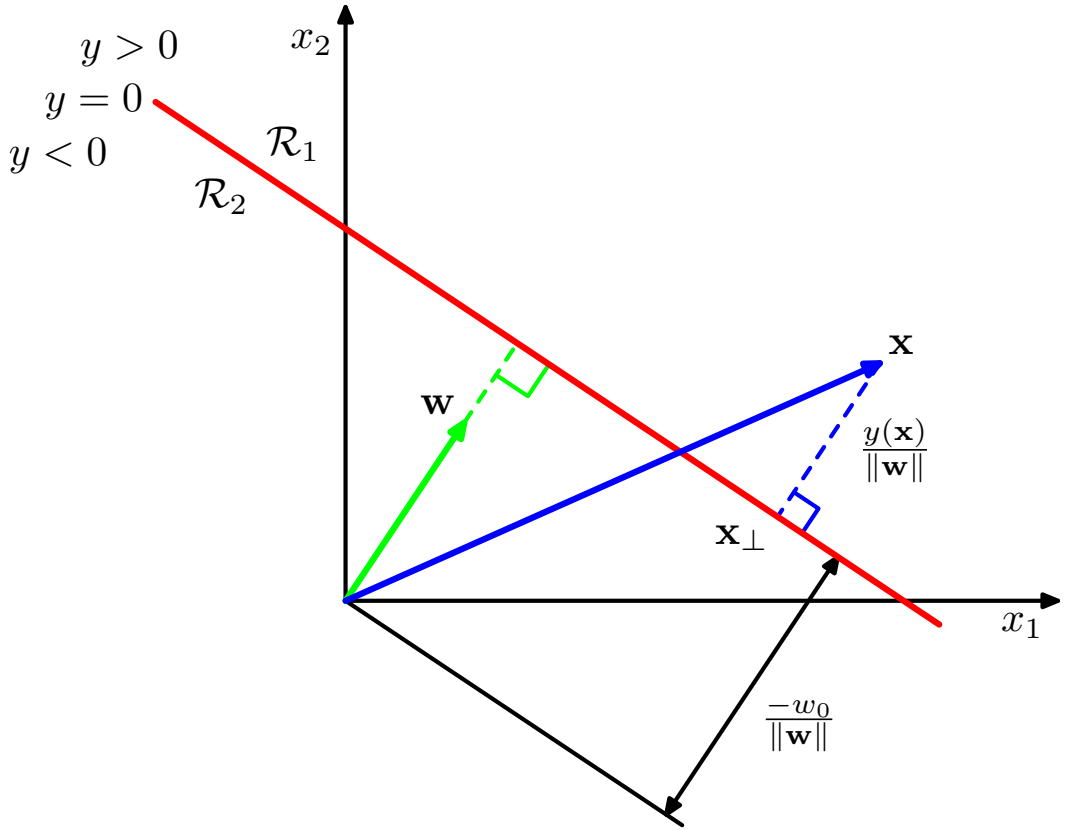
Taking a linear function of the input vector:

$$y(x) = w^T x + w_0$$

- An input vector is assigned to class $\mathcal{C}_1$ if $y(x) \geq 0$ and to class $\mathcal{C}_2$ otherwise.
- The decision boundary is a $(D-1)$-dimensional hyperplane.

Two classes

Figure: The geometry of a linear discriminant function in two dimensions



Two classes

It's easy to see that:

- w is orthogonal to the decision surface.
- w points to the direction of the increase of y.

Also the value of $y(x)$ gives a signed measure of the perpendicular distance r of the point x from the decision surface:

$$x = x_\perp + r\frac{w}{||w||}$$

$$y(x) = w^T x + w_0 = w^T x_\perp + w_0 + r||w|| = r||w||$$

$$r = \frac{y(x)}{||w||}$$

In particular, the signed distance of the origin from the decision surface is given by $\frac{w_0}{||w||}$.