Linear Classification

Machine Learning - Prof. Dr. Stephan Günnemann

Leonardo Freiherr von Lerchenfeld

November 24, 2017

Contents

1 Linear separability

1.1 Problem 1

The convex hull $co\mathcal{X}$ is the set of all points x given by

$$\begin{aligned} h(x) &= w^T x \\ &= w^T \sum_i \alpha_i x^i \\ &= \sum_i \alpha_i w^T x^i \\ with \quad & \sum_i \alpha_i = 1 \end{aligned}$$

Similarly the convex hull $co\mathcal{Y}$ can be described by

$$\begin{aligned} h(y) &= \sum_j \beta_j w^T y^j \\ with \quad & \sum_j \beta_j = 1 \end{aligned}$$

When the convex hulls intersect, they must have at least one point in common. This point belongs to both convex hulls, hence, there must be a set of common points, which is built by α_i and β_j. The linear discriminant of this common set, let‘s call it z, can be written by

$$h(z) = \sum_i \alpha_i w^T x^i = \sum_j \beta_j w^T y^j$$

For linear separability, the following conditions must be fulfilled

$$\begin{aligned} h(x^i) &= w^T x^i > 0 \\ h(y^j) &= w^T y^j < 0 \end{aligned}$$

It is a contradiction! The linear discriminant h(z) has to be simultaneously greater than zero and less than zero, which is impossible, since $\forall i: \; \alpha_i \geq 0, \beta_i \geq 0$.

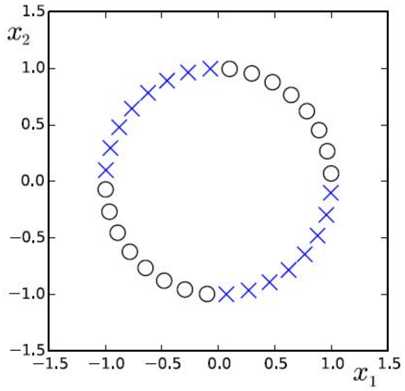
1.2 Problem 2

A hyperplane puts all points on the correct side. The probability of being in class 1 is given by

$$\begin{aligned} p(y=1|x) &= \frac{1}{1+exp(-(w^T x))} \\ &= \sigma(w^T x) \\ w^T x > 0 \quad & if\ x\ on\ normal\ side \\ w^T x < 0 \quad & else \end{aligned}$$

When the magnitude of w is infinity, the probability for each point is 100%. Thus, the probability function has not the shape of a sigmoid function, but a ramp function. To prevent this, one can add a regularization to penalize a large w.

1.3 Problem 3



I want to classify the circles. When you look at the values of the circles, they are either both positive or both negative. So an adequate basis function is:

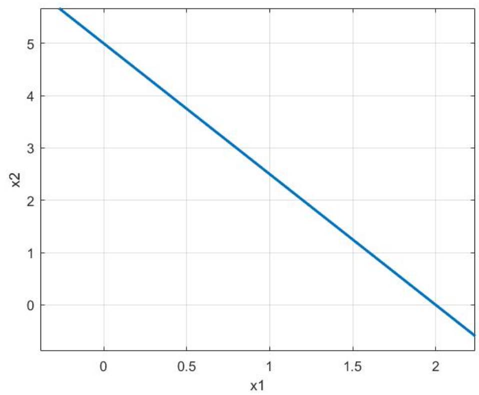
$$\Phi(x_1, x_2) = x_1 x_2$$

2 Basis functions

2.1 Problem 4

A hyperplane is defined by a normal vector $\mathbf{w}$ and an offset b $(= w_0)$

$$\mathbf{w}^T\mathbf{x} + b = 0 \text{ if x on the plane}$$



$$\begin{aligned}
w_0 + w_1 x_1 + w_1 x_2 &= 0 \\
w_0 + 2x_1 &= 0 \\
w_0 + 5x_2 &= 0 \\
-w_0 &= 5x_2 \\
-w_0 &= 2x_1 \\
w_0 &= -5 \\
w_2 &= 1 \\
w_1 &= \frac{5}{2}
\end{aligned}$$

Here I set $w_0 = -5$ but one can choose any other real value.