Lecture 35

niceguy

April 10, 2023

1 Support Vector Machines

Regression

$$y = f(x), x \in \mathbb{R}^m, y \in \mathbb{R}$$

Classification

$$x \in \mathbb{R}^n, y \in \{-1, 1\}$$

2 Hyperplane

Let $w \in \mathbb{R}^n$ be a vector, and $b \in \mathbb{R}$. Then consider

$$w^T x - b = 0$$

Example 2.1.

$$w = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, b = 1$$

Then the solution is

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, x_2 = 1 - 2x_1$$

2.1 Data

Consider the data (x_i, y_i), with $x_i \in \mathbb{R}^m$ and $y_i = \pm 1$. The problem is given a new x, predict $y = ax + b$. Now the support vector machine predicts that $y = -1$ if $w^T x - b < 0$, vice versa.

2.2 Optimisation

We want to do this in a way that maximises the degree of ”parallel shift” the hyperplane can make while still making the same predictions. Note that the distance is given by

$$w^T x^* = 1$$

where x^* is a multiple of w for it to be normal to the plane. Then letting $x^* = kw$ gives

$$k||w||^2 = 1 \Rightarrow ||x|| = \frac{1}{||w||}$$

Maximising x is equivalent to minimising $||w||^2$ with the constraint

$$y_i(^T x_i - b) \geq 1$$