Lecture 19 (Support Vector Machines)

1 Mathematical Setup

1. $y \in \{-1, 1\}$
2. $w^T x + b = 0$ is the equation of the hyperplane
3. Not all examples contribute to the model, we try to find the "support vectors" from the training set
4. SVMs find hyperplanes which maximise the minimum margin (margin defined below)

1.1 Margin of Point

(normalised) Signed distance of point from hyperplane, it is given by:

$$\frac{w^T x^i + b}{||w||_2}$$

1.2 Max-Margin Based Classifier (SVMs property)

$$\max_{\gamma,w,b}(\gamma : \forall i, \gamma^i \geq \gamma)$$

$$\gamma^i = y^i \times \left(\frac{w^T x^i + b}{||w||_2}\right)$$

$$\hat{\gamma} = ||w||\gamma$$

$$\max_{\gamma,w,b}(\frac{\hat{\gamma}}{||w||} : \forall i, y^i \times (w^T x^i + b) \geq \hat{\gamma})$$

Suppose $\hat{\gamma}^*, w^*, b^*$ is an optimal solution, then any multiple of these is also an optimal solution.

$$\implies 1, \frac{w^*}{\hat{\gamma}^*}, \frac{b^*}{\hat{\gamma}^*} \text{ is also an optimal solution}$$

$$\implies 1, w^{'*}, b^{'*}$$

$$\implies \max_{w,b}(\frac{1}{||w||} : \forall i, y^i \times (w^T x^i + b) \geq 1)$$

$$\implies \max_{w,b}(\frac{1}{||w||^2} : \forall i, y^i \times (w^T x^i + b) \geq 1)$$

$$\implies \min_{w,b}(\frac{1}{2}w^T w : \forall i, y^i \times (w^T x^i + b) \geq 1)$$

1. This is the optimisation problem for SVMs 1. It is a subproblem of convex (constrained) optimisation problem

2 Constrained Optimisation Problem

$$\min_{w}(f(w) : (g_i(w) \leq 0, \forall i \in \{1, \ldots, m\}) \wedge (h_l(w) = 0, \forall l \in \{1, \ldots, k\}))$$

2.1 Convex Constrained Optimisation

1. f is convex function
2. g_i are convex functions
3. h_l are affine, i.e., $h_l(w) = w^T x + b$ (almost linear function, it allows for intercept term too)