SMAI Class 8 - SVM

August 29, 2022

Contents

Maximum Margin Classification

- **Margin**: Width of a band around decision boundary without any training samples. Also the Radius of region around each training sample through which the decision boundary cannot pass.

Margin varies with position and origin of the separating hyperplane. As margin increases, the feasible region (for the decision boundary) decreases.

The circles touching the margin are the corresponding samples that *control* the boundary. Such samples are called ***Support Vectors*** and the resulting linear classifier is called the **SVM**.

Radius $= R$ Margin $= \rho = 2R$

$w^T x + b = 0$: Decision Boundary $w^T x + b = \pm\epsilon$: Margins

We want to maximize ϵ.

Some constraints

- Normalize w ($\|w\| = 1$).

Problem

$g(x) = w^T x + b$

Minimize $1/2 w^T w$ subject to $y_i(w^T x_i + b) - 1 \geq 0 \forall i$.

Pegasos Algorithm

Pegasos paper ref.

```
INPUT: Training Data: S, lambda, Max iterations: T
```

```
INITIALIZE: w_1

FOR t = 1, 2, 3 ... T
 - Choose i = {1, 2, ... |S|} at random and pick sample (x_i, y_i)
 - Set eta_t = 1 / lambda_t
 - IF( y_i (w_t, x_i) < 1):
  - Update w_{t+1} <- w_t - eta_t lambda w_t + eta_t y_i x_i
 - ELSE
  - Update w_{t+1} <- w_t - eta_t lambda w_t

OUTPUT W_{T+1}
```

SVM Theory (A high level view)

Concept called VC-dimension(h) that measures the complexity of classifiers.

R test (chance of classifier making error during testing). Sort of expected loss.

The R-test is R-train + some function.

$$R_{test}(\alpha) = R_{train}(\alpha) + f(h, N)$$

$$h \leq min\{d, \lceil 1/m^2 \rceil\} + 1$$

where d is dimension of the feature space. m is defined as the relative margin.

$$m = \rho / D$$

ρ: Margin, D: Data diameter