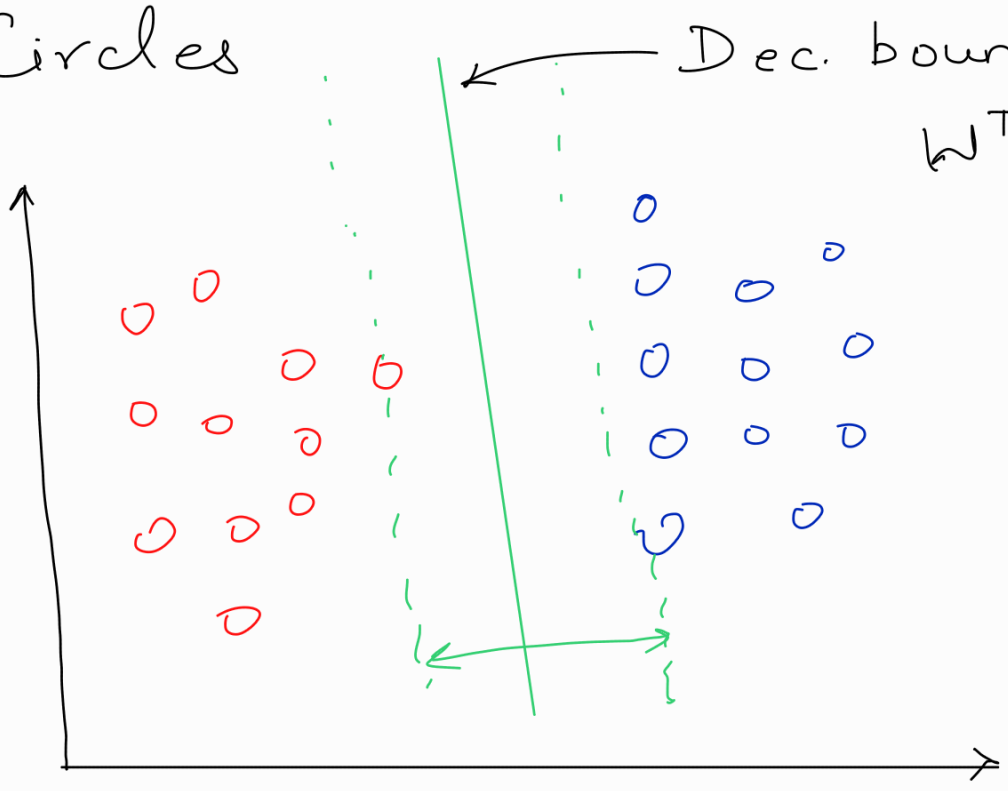
15/3/24

2 ways to visualize loss:

1] Bands

2] Circles



$W^T X + b = \pm k$

W is ⊥ to classifier.

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

We maximize k such that:

$w^T x_i + b \geq k$ for $d_i = 1$

$w^T x_i + b \leq -k$ for $d_i = -1$

Value of $g(x)$ depends on $\|w\|$ (unit vector)

Keep $\|w\| = 1$, maximize $g(x)$ or ap1

Let $g(x) \geq 1$, minimize $\|w\|$ ap2

$w^T x_{pos} + b = 1$

$w^T x_{neg} + b = -1$

$$\frac{w^T}{\|w\|}(x_{pos} - x_{neg}) = \boxed{\frac{2}{\|w\|}} \leftarrow \text{margin}$$

We use approach 2:

Minimize $\frac{1}{2} W^T W$

subject to $d_i (W^T X_i + b) \geq 1$

for $i = 1 \ldots N$.

Using Lagrange's multiplier:

$$\frac{1}{2} W^T W - \alpha_i \left[d_i (W^T X_i + b) - 1 \right]$$

$$J(w, b, \alpha) = \frac{1}{2} W^T W - \sum_{i=1}^{N} \alpha_i d_i (W^T X_i + b) + \sum_{i=1}^{N} \alpha_i$$

There are N constraints $\therefore$ 1 lagrange multiplier for each constraint.

Minimization problem wrt W

Maximization " " α_i

Differentiating wrt W,

$$W = \sum_{i=1}^{N} \alpha_i d_i X_i \quad \text{——— (1)}$$

W is the weighted sum of points

Differentiating wrt d,

$$\Sigma \alpha_i d_i = 0 \quad \text{——— (2)}$$

d_i is 1 or -1 (classes)

The product of α with positive points = negative points

$$\alpha_i \left[d_i (W_0^T X_i + b_0) - 1 \right] = 0 \quad \text{——— (3)}$$

KKT condition

α is non-zero for points which are support vectors

Dual Form

Simplifying original equation,

$$\frac{1}{2} w^T w - \sum_{i=1}^{N} \alpha_i d_i w^T x_i - \sum_{i=1}^{N} \alpha_i d_i b$$

$$+ \sum_{i=1}^{N} \alpha_i$$

$$Q(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j d_i d_j x_i^T x_j$$

Put ① ↗

↓ QP solver

Positives:

Quadratic

Only 1 variable

$$w_0 = \sum_{i=1}^{N} \alpha_i d_i x_i$$

↓

$$\alpha_i \left[d_i \left(w_0^T x_i + b_0 \right) - 1 \right] = 0$$

↓

$$b_0 = 1 - w_0^T x_{s+}$$

Converting 2-class to n-class

1] Do 1 vs. all for n classes

2] Do all combinations of 1 vs 1.