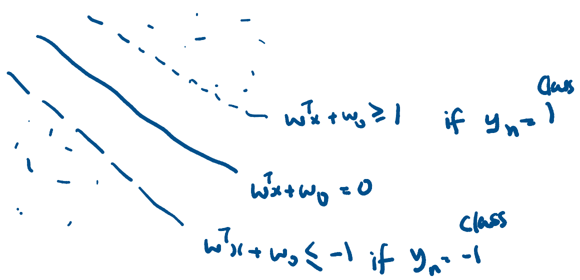
SVM

Friday, October 28, 2022 1:57 AM

https://towardsdatascience.com/implementing-svm-from-scratch-784e4ad0bc6a

2 Class $y_n = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$



$y_n(w^T x_n + w_0) \geq 1 \quad , \quad \forall n = 1, \ldots, N$

(min)

$\mathcal{L}(w) = \frac{1}{2}\|w\|_2^2$ s.t. $\boxed{\begin{array}{l} y_n(w^T x_n + w_0) \geq 1 \\ 0 \geq 1 - y_n(w^T x_n + w_0) \end{array}}$

$\therefore \mathcal{L}(w) = \frac{1}{2}\|w\|_2^2 + \frac{1}{N}\sum_{n=1}^{N} \max\left(0, 1 - y_n(w x_n + w_0)\right)$

$$\left.\begin{array}{l} \mathcal{L}(w) = \frac{1}{2}\|w\|_2^2 \\ \frac{\partial \mathcal{L}(w)}{\partial w} = w \\ \frac{\partial \mathcal{L}(w)}{\partial b} = 0 \end{array}\right\} \quad y_n(w^T x_n + w_0) \geq 1$$

$$\left.\begin{array}{l} \mathcal{L}(w) = \frac{1}{2}\|w\|_2^2 + \frac{1}{N}\sum_{n=1}^{N} 1 - y_n(w^T x_n + w_0) \\ \frac{\partial \mathcal{L}(w)}{\partial w} = w - y_n x_n^T \\ \frac{\partial \mathcal{L}(w)}{\partial w_0} = -y_n \end{array}\right\} \quad y_n(w^T x_n + w_0) \leq 1$$