Lab 6

SVM Problem Set

Raymond Lin and Kenny Guan

1) $\vec{p} = \vec{x}_i - \gamma_i y_i \cdot \frac{\vec{w}}{\|\vec{w}\|}$

$0 = \vec{w} \cdot \vec{p} + \vec{b}$

$0 = \vec{w}(\vec{x}_i - \gamma_i y_i \cdot \frac{\vec{w}}{\|\vec{w}\|}) + \vec{b}$

$0 = \vec{w}\vec{x}_i - \gamma_i y_i \cdot \frac{\vec{w} \cdot \vec{w}}{\|\vec{w}\|} + b$

$0 = \vec{w}\vec{x}_i - \gamma_i y_i \cdot \frac{\|\vec{w}\|\|\vec{w}\| \cos\theta}{\|\vec{w}\|} + b$    * $\theta = 0$ because the vector is dotted with itself so the angle between them is 0

$0 = \vec{w}\vec{x}_i - \gamma_i y_i \|\vec{w}\| + b$

$\gamma_i y_i \|\vec{w}\| = \vec{w} \cdot \vec{x}_i + b$

$\gamma_i = \left(\frac{\vec{w} \cdot x_i}{\|\vec{w}\|} + \frac{b}{\|\vec{w}\|}\right) \frac{1}{y_i}$    * since $y_i$ is just a sign, $\frac{1}{y_i} = y_i$

$\gamma_i = y_i \left(\frac{\vec{w} \cdot \vec{x}_i}{\|\vec{w}\|} + \frac{b}{\|\vec{w}\|}\right)$

2) Corollary 1

We want to show $\|\vec{w}\|^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} y_i y_j \alpha_i \alpha_j \vec{x}_i \vec{x}_j$

$\|\vec{w}\|^2 = w_1^2 + w_2^2 + \dots + w_n^2$

$w_1^2 = (\sum_{i=1}^{n} \alpha_i y_i x_{i1})^2 = (\alpha_i y_i x_{11} + \alpha_i y_i x_{12} + \dots)^2$

$\sum_{i=1}^{n} \sum_{j=1}^{n} y_1 y_j \alpha_i \alpha_j \vec{x}_i \vec{x}_j$

and we do this for n weights so we have

$\|\vec{w}\|^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} y_i y_j \alpha_i \alpha_j \vec{x}_i \vec{x}_j$

We want to show $L(\vec{w}, b, \vec{\alpha}) = \frac{1}{2}\|\vec{w}\|^2 - \sum \alpha_i [y_i(\vec{w} \cdot \vec{x}_i + b) - 1] = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i y_j \alpha_i \alpha_j \vec{x}_i \cdot \vec{x}_j$

$L(\vec{w}, b, \vec{\alpha}) = \frac{1}{2}\|\vec{w}\|^2 - \sum \alpha_i [y_i(\vec{w} \cdot \vec{x}_i + b) - 1]$

$= \frac{1}{2}\|\vec{w}\|^2 - \sum(\alpha_i y_i \vec{w} \cdot \vec{x}_i + \alpha_i y_i b - \alpha_i)$

$= \frac{1}{2}\|\vec{w}\|^2 - \sum \alpha_i y_i \vec{w} \cdot \vec{x}_i - \sum \alpha_i y_i b + \sum \alpha_i$    * $\sum \alpha_i y_i b = 0$

$= \frac{1}{2}\|\vec{w}\|^2 - \sum \alpha_i y_i \vec{w} \cdot \vec{x}_i + \sum \alpha_i$

$= \frac{1}{2}\|\vec{w}\|^2 - \vec{w} \sum \alpha_i y_i \vec{x}_i + \sum \alpha_i$    * $\sum \alpha_i y_i \vec{x}_i = \vec{w}$

$= \frac{1}{2}\|\vec{w}\|^2 - \vec{w}\cdot\vec{w} + \sum \alpha_i$     * $\vec{w}\cdot\vec{w} = \|w\|^2\cos(0) = \|\vec{w}\|^2$

$= -\frac{1}{2}\|\vec{w}\|^2 + \sum \alpha_i$

$= -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} y_i y_j \alpha_i \alpha_j \vec{x}_i \cdot \vec{x}_j + \sum_{i=1}^{n}\alpha_i$ * by Corollary 1

3) We want to show $\max_x \min_y f(x,y) \leq \min_y \max_x f(x,y)$

Let $(x^*, y^*) = \max_x \min_y f(x,y)$

Then we know $(x^*, y^*) \leq (x^*, y)$ for any $y$.

Thus we know every value in $\max_x f(x,y) \geq (x^*, y)$

by the transitive property we know every value in $\max_x f(x,y) \geq (x^*, y) \geq (x^*, y^*)$

Thus $\min_y \max_x f(x,y) \geq (x^*, y^*)$ since all max values are greater than or equal to $(x^*, y^*)$.

So we have $(x^*, y^*) \leq \min_y \max_x f(x,y)$ or $\max_x \min_y f(x,y) \leq \min_y \max_x f(x,y)$

4) We know that $\hat{y}_i = y_i(\vec{w}\cdot\vec{x}_i + b)$

and on the margin $\hat{y}_i = 1$

So $1 \leq y_i(\vec{w}\cdot\vec{x} + b)$

$-y_i(\vec{w}\cdot\vec{x} + b) \leq 0$

When $y_i = -1$

$-1(-1)(\vec{w}\cdot\vec{x} + b) \leq 0$

$\vec{w}\cdot\vec{x} + b \leq 0$

$b \leq -\vec{w}\cdot\vec{x}$     * and we want to be close to margin so we want to maximize $-\vec{w}\cdot\vec{x}$

When $y_i = 1$

$-1(\vec{w}\cdot\vec{x} + b) \leq 0$

$-b \leq \vec{w}\cdot\vec{x}$     * so we want to minimize $-\vec{w}\cdot\vec{x}$

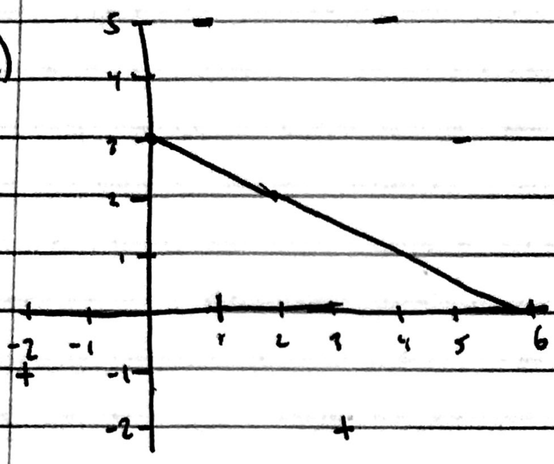
$b \geq -\vec{w}\cdot\vec{x}$

In order to find the optimal bias, we want to find the midpoint of the support vectors.

$b^* = \frac{1}{2}\left(\max_{i:y_i=-1} -\vec{w}^*\cdot\vec{x}_i + \min_{i:y_i=1} -\vec{w}^*\cdot\vec{x}_i\right)$ * distribute the negative

$b^* = -\frac{1}{2}\left(\max_{i: y_i=-1} \vec{w}^* \cdot \vec{x}_i + \min_{i: y_i=1} \vec{w}^* \cdot x_i\right)$

5) a)




b) $x_2 = -\frac{1}{2}x_1 + 3$

c) $\vec{x}_1, \vec{x}_3, \vec{x}_5$

d) $\vec{w} = a\begin{bmatrix}-1\\-2\end{bmatrix}$

e) $1 = -a + b$

$1 = -1(-a - 10a) + b$

$1 = 11a + b$

$11a - b = -a + b$

$12a = 2b$

$b = 6a$

$1 = 5a$

$a = 0.2 \qquad b^* = 1.2$

$\vec{w}^* = \begin{bmatrix}-0.2\\-0.4\end{bmatrix}$

f) $\begin{bmatrix}-1/5\\-2/5\end{bmatrix} = -\alpha_1\begin{bmatrix}1\\5\end{bmatrix} - \alpha_3\begin{bmatrix}5\\3\end{bmatrix} + \alpha_5\begin{bmatrix}1\\0\end{bmatrix}$

$\alpha_1 + \alpha_3 = \alpha_5$

$2(-\alpha_1 - 5\alpha_3 + \alpha_5) = -5\alpha_1 - 3\alpha_3$

$-2\alpha_1 - 10\alpha_3 + 2\alpha_5 = -5\alpha_1 - 3\alpha_3$

$3\alpha_1 - 7\alpha_3 + 2\alpha_5 = 0$

$3\alpha_1 - 7\alpha_3 + 2\alpha_1 + 2\alpha_3 = 0$

$5\alpha_1 - 5\alpha_3 = 0$

$\alpha_1 = \alpha_3$

$\alpha_1 = \alpha_3$

$\alpha_5 = 2\alpha_1$

$-\frac{1}{5} = \alpha_1 - 5\alpha_1 + 2\alpha_1$

$-\frac{1}{5} = -4\alpha_1$

$\alpha_1 = \frac{1}{20} \qquad \alpha_2 = 0$

$\alpha_3 = \frac{1}{20} \qquad \alpha_4 = 0$

$\alpha_5 = \frac{1}{10} \qquad \alpha_6 = 0$