Title. Date. Page.

Hyun-Joon Yang

HW 1

1) $E(\vec{x}_i, \vec{x}_j) = \| \vec{x}_i - \vec{x}_j \|_2^2 = (\vec{x}_i - \vec{x}_j)^T (\vec{x}_i - \vec{x}_j)$

$$= \vec{x}_i^T \vec{x}_i - 2\vec{x}_i^T \vec{x}_j + \vec{x}_j^T \vec{x}_j$$

$\| x_i \|_2 = \vec{x}_i^T \vec{x}_i = 1 = \| x_j \|_2$

$$= 2 - 2\vec{x}_i^T \vec{x}_j$$

$$E(\vec{x}_i, \vec{x}_j) = 2(1 - \vec{x}_i^T \vec{x}_j)$$

$$C(\vec{x}_i, \vec{x}_j) = 1 - \frac{\vec{x}_i^T \vec{x}_j}{\|\vec{x}_i\|_2 \|\vec{x}_j\|_2} = 1 - \vec{x}_i^T \vec{x}_j$$

$$E(\vec{x}_i, \vec{x}_j) = 2C(\vec{x}_i, \vec{x}_j)$$

thus

if $C(\vec{x}_i, \vec{x}_j) \le C(\vec{x}_i, \vec{x}_o)$ then $E(\vec{x}_i, \vec{x}_j) \le E(\vec{x}_i, \vec{x}_o)$

Title. Date. Page.

2.2) a) $w_0^* = \underset{w_0}{\operatorname{argmin}} \| \vec{y} - (w_0 + X\vec{w}) \|$ $*$ $X \in \mathbb{R}^{N \times D}, \vec{w} \in \mathbb{R}^D$

b) $\frac{d}{dw_0} \Rightarrow -\cancel{2}\vec{1}_N^T(\vec{y} - w_0^*\vec{1}_N - X\vec{w}) = 0$

$= \vec{1}_N^T(\vec{y} - w_0^*\vec{1}_N - X\vec{w}) = 0$

since $\frac{1}{N}\sum_n X_{nd} = 0$ for $\forall d = 1, 2, \ldots D$, $X\vec{w} = 0$

$= \vec{1}_N^T(\vec{y} - w_0^*\vec{1}_N) = 0$

c) $w_0^* = \frac{1}{N}\vec{1}_N^T\vec{y}$

Title. Date. Page.

3.1) $\vec{W}_{k+1} = \vec{W}_k + y_i \vec{x}_i$

$\vec{W}_{k+1}^T \vec{W}_{opt} = \vec{W}_k^T \vec{W}_{opt} + y_i \vec{W}_{opt}^T \vec{x}_i$

$\vec{W}_{k+1}^T \vec{W}_{opt} - \vec{W}_k^T \vec{W}_{opt} = y_i \vec{W}_{opt}^T \vec{x}_i$

$$\gamma = \min \frac{|\vec{W}_{opt}^T \vec{x}_i|}{\|\vec{W}_{opt}\|}$$

$$\gamma \|\vec{W}_{opt}\| = \min |\vec{W}_{opt}^T \vec{x}_i|$$

since $y_i \in \{-1, 1\}$ and $y_i \vec{W}_{opt}^T \vec{x}_i \geq 0$, $y_i \vec{W}_{opt}^T \vec{x}_i \geq \min |\vec{W}_{opt}^T \vec{x}_i|$

$\vec{W}_{k+1}^T \vec{W}_{opt} - \vec{W}_k^T \vec{W}_{opt} \geq \gamma \|\vec{W}_{opt}\|$

$= \vec{W}_{k+1}^T \vec{W}_{opt} \geq \vec{W}_k^T \vec{W}_{opt} + \gamma \|\vec{W}_{opt}\|$

Title. Date. Page.

3.2) $\vec{W}_{k+1} = \vec{W}_k + y_i\vec{x}_i$

$$\|\vec{W}_{k+1}\|^2 = \|\vec{W}_k + y_i\vec{x}_i\|^2$$
$$= (\vec{W}_k + y_i\vec{x}_i)^T(\vec{W}_k + y_i\vec{x}_i)$$
$$= \vec{W}_k^T\vec{W}_k + 2y_i\vec{W}_k^T\vec{x}_i + y_i^2\vec{x}_i^T\vec{x}_i$$

$$\vec{x}_i^T\vec{x}_i = \|x_i\|^2 = 1^2 = 1$$

$$= \vec{W}_k^T\vec{W}_k + 2y_i\vec{W}_k^T\vec{x}_i + y_i^2$$

since $y \in \{-1,1\}$ and $y_i\vec{W}_k^T\vec{x}_i < 0$,
$\Rightarrow y_i^2 = 1$ and $2y_i\vec{W}_k^T\vec{x}_i < 0$

thus

$$\|\vec{W}_{k+1}\|^2 \leq \vec{W}_k^T\vec{W}_k + 1$$

$$\|\vec{W}_{k+1}\|^2 \leq \|\vec{W}_k\|^2 + 1$$

Title. Date. Page.

3.3) a) $\vec{W_1}^T \vec{W_{opt}} \geq \vec{W_0}^T \vec{W_{opt}} + \gamma \|\vec{W_{opt}}\|$ * $k=1$

since $\vec{W_0} = \vec{0}$, $\vec{W_1}^T \vec{W_{opt}} \geq \gamma \|\vec{W_{opt}}\|$

$\vec{W_2}^T \vec{W_{opt}} \geq \vec{W_1}^T \vec{W_{opt}} + \gamma \|\vec{W_{opt}}\| \geq 2\gamma \|\vec{W_{opt}}\|$ * $k=2$

after M mistakes, $\vec{W_{k+1}}^T \vec{W_{opt}} \geq \gamma M \|\vec{W_{opt}}\|$

$\gamma M \|\vec{W_{opt}}\| \leq \vec{W_{k+1}}^T \vec{W_{opt}} \leq \|\vec{W_{k+1}}\| \|\vec{W_{opt}}\|$

$$\gamma M \leq \|\vec{W_{k+1}}\|$$

b) $\|\vec{W_1}\|^2 \leq \|\vec{W_0}\|^2 + 1 = 1$ * $k=1$
$\|\vec{W_2}\|^2 \leq \|\vec{W_1}\|^2 + 1 = 1+1 = 2$ * $k=2$

after M mistakes, $\|\vec{W_{k+1}}\|^2 \leq M$

$\sqrt{\|\vec{W_{k+1}}\|^2} \leq \sqrt{M}$

$\|\vec{W_{k+1}}\| \leq \sqrt{M}$

therefore

$$\gamma M \leq \|\vec{W_{k+1}}\| \leq \sqrt{M}$$

3.4) $\gamma M \leq \sqrt{M}$
$\gamma^2 M^{\cancel{2}} \leq \cancel{M}$
$M \leq \gamma^{-2}$