ECS 171 HW1 Weiran Guo (912916431)

1. (a) $w^T x = w_0 + w_1 x_1 + w_2 x_2$

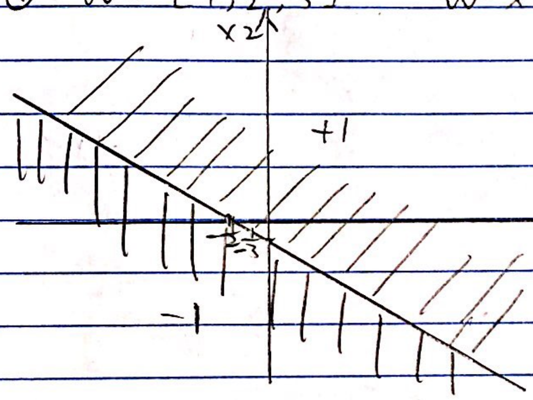
$sign(w^T x)$ seperate by line $w_0 + w_1 x_1 + w_2 x_2 = 0$

which is known as $a x_1 - x_2 + b = 0$

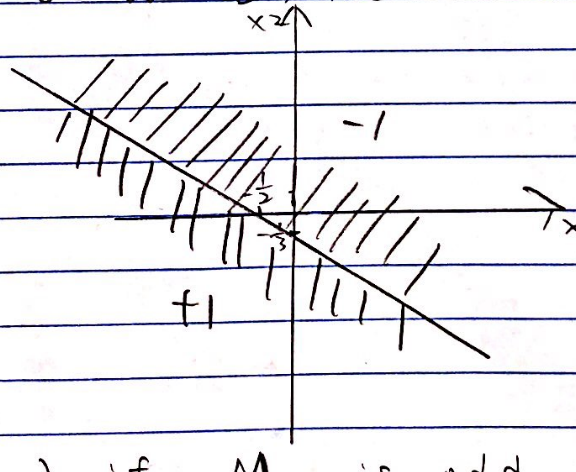
$\therefore -\frac{w_0}{w_2} - \frac{w_1}{w_2} x_1 - x_2 = 0$

$a = -\frac{w_1}{w_2}$ $b = -\frac{w_0}{w_2}$

(b). ① $w = [1, 2, 3]^T$ $w^T x = 1 + 2x_1 + 3x_2$



②. $w = -[1, 2, 3]^T$ $w^T x = -1 - 2x_1 - 3x_2$



2. (a) if M is odd, N is odd

$E_{off}(h, f) = \frac{M+1}{2} \cdot \frac{1}{M}$

if M is odd, N is even

$E_{off}(h, f) = \frac{M-1}{2} \cdot \frac{1}{M}$

if M is even

$E_{off}(h, f) = \frac{M}{2} \cdot \frac{1}{M} = \frac{1}{2}$

(b) for M out-set case, there are $2^M - 1$ combinations, hence total is $(2^M - 1 + 1) = \boxed{2^M}$ cases

(c) k mismatches $\Rightarrow$ $M-k$ matches

hence, total $\binom{M}{k} = \frac{M!}{k!(M-k)!}$ target f.

(d)

$$E_f(E_{off}(h,f)) = \sum_{k=0}^{M} \frac{k}{M} \frac{M!}{k!(M-k)!} \cdot \frac{1}{2^M}$$

$$= \frac{1}{2^M} \sum_{k=0}^{M} \frac{(M-1)!}{(k-1)!(M-k)!} \qquad k' = k-1$$

$$= \frac{1}{2^M} \cdot \sum_{k=0}^{M-1} \frac{(M-1)!}{k!(M-1-k)} = \frac{1}{2^M} \cdot 2^{M-1}$$

$$= \frac{1}{2}$$

(e) By (d), known that all f are equally likely, no matter what h is,

$$E_f(E_{off}(A_1(D), f)) = E_f(E_{off}(A_2(D), f)) = \frac{1}{2}$$

3. (a) when stepsize ≥ 0.001, objective function increases (diverge) RuntimeWarning: overflow in exp occurs.

when stepsize ≤ 0.0001, it gets good results.

when stepsize is smaller, it takes more iterations to converge.

the objective function is keep decreasing starting from 0.69

(b) accuracy is 0.986

(c) By SGD, with fixed step size, and batch size 1,

the accuracy table:

step size	accuracy
0.1	diverge
0.01	diverge
0.001	diverge
0.0001	diverge
0.00001	0.982
0.000009	0.985

(d) by using $\eta^t \sim t^{-a}$ $a = 0.4$
the accuracy becomes 0.984, which increases

4. (a) if $[\text{sign}(w^T x_n) \neq y_n] = 1$

$y_i w^T x_n < 0 \Rightarrow 1 - y_i w^T x_n > 1$

hence $e_n(w) = \max(0, 1 - y_i w^T x_n) > 1$ is an upper bound of $[\text{sign}(w^T x_n) \neq y_n] = 1$

else if $[\text{sign}(w^T x_n) \neq y_n] = 0$

$y_i w^T x_n > 0 \Rightarrow 1 - y_i w^T x_n < 1$

if $(1 - y_i w^T x_n) > 0$, $\max(0, 1 - y_i w^T x_n) = 1 - y_i w^T x_n > 0$

hence it is an upper bound of $[\text{sign}(w^T x_n \neq y_n] = 0$

else if $(1 - y_i w^T x_n) < 0$, $\max(0, 1 - y_i w^T x_n) = 0$ $= [\text{sign}(w^T x_n \neq y_n] = 0$, it satisfies that it is an upper bound.

hence, in whatever cases, $e_n(w)$ is an upper bound for the $[\text{sign}(w^T x_n) \neq y_n]$.

hence $\frac{1}{N} \sum_{n=1}^{N} e_n(w)$ is an upper bound for the in-sample classification error.

(b) the accuracy is 0.9825