1

(1)

recall that

$$\begin{aligned}
\bar{p}_n(x) &= \boldsymbol{E}(p_n(x)) \\
&= \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{E}[\frac{1}{V_n}\varphi(\frac{x-x_i}{h_n})] \\
&= \int \frac{1}{V_n}\varphi(\frac{x-v}{h_n})p(v)dv \\
&= \int \delta_n(x-v)p(v)dv
\end{aligned}$$

what we need to verify is $\lim\limits_{n\to\infty}\delta_n(x-v)=\delta(x-v)$ where $\delta(x-v)$ refers to Dirac function. Dirac function is defined as below

$$\delta(x)=\begin{cases}\infty, & x=0\\ 0, & otherwise\end{cases} \tag{1}$$

$$\int_R \delta(x)dx = 1 \tag{2}$$

Thus

$$\begin{aligned}
\delta_n(x-v) &= \frac{1}{V_n}\varphi(\frac{x-v}{h_n}) \\
&= \frac{1}{\prod_{i=1}^{d} h_{ni}}\prod_{i=1}^{d}\frac{x_i-v_i}{h_{ni}}\varphi(\frac{x-v}{h_n})\prod_{i=1}^{d}\frac{h_{ni}}{x_i-v_i} \\
&= \prod_{i=1}^{d}\frac{1}{x_i-v_i}\ \varphi(u)\prod_{j=1}^{d}u_j
\end{aligned}$$

$$\because \lim_{n\to\infty}V_n=0$$

$$\therefore \lim_{n\to\infty}h_{ni}=0,\ \ i=1,2,\ldots,d$$

$$\therefore u_i=\lim_{n\to\infty}\frac{x-v}{h_{ni}}=\begin{cases}0, & x=v\\ \infty, & x\neq v\end{cases}$$

$$\therefore \lim_{||u||\to\infty}\varphi(u)\prod_{i=1}^{d}u_i=0,\ \ x\neq v$$

Since

$$\int_R \delta_n(x-v)dx=\int_R\frac{1}{V_n}\varphi(\frac{x-v}{h_n})dx=\int_R\varphi(u)du=1$$

We can conclude that equations (19)~(22) are sufficient to the convergence of equation (17)

(2)

Recall that

$$
\begin{aligned}
\sigma_n^2(x) &= \sum_{i=1}^{n} \boldsymbol{E}[(\frac{1}{nV_n}\varphi(\frac{x-x_i}{h_n}) - \frac{1}{n}\bar{p}_n(x))^2] \\
&= n\boldsymbol{E}[\frac{1}{n^2V_n^2}\varphi^2(\frac{x-x_i}{h_n})]p(v)dv - \frac{1}{n}\bar{p}_n^2(x) \\
&= \frac{1}{nV_n}\int \frac{1}{V_n}\varphi^2(\frac{x-v}{h_n})p(v)dv - \frac{1}{n}\bar{p}_n^2(x)
\end{aligned}
$$

$$
\because \quad \sup_u \varphi(u) < \infty
$$

$$
\sigma_n^2(x) \leq \frac{\sup_u \varphi(u)\bar{p}_n(x)}{nV_n}
$$

$$
\because \lim_{n\to\infty} \bar{p}_n(x) = p(x), \quad \lim_{n\to\infty} nV_n = \infty
$$

$$
\therefore \lim_{n\to\infty} \sigma_n^2(x) = 0
$$

4

(a)

$$
g_i(x) = \sum_{j=1}^{n}(a_{ji})(e^{\frac{-(x-w_j)^t(x-w_j)}{2\sigma^2}}) = nP(w_i)P(x|w_i)
$$

(b)

$$
Risk(\alpha_i|x) = \sum_{j=1}^{c}\lambda_{ji}P(w_j|x) = \sum_{j=1}^{c}\lambda_{ji}\frac{P(x|w_j)P(w_j)}{P(x)}
$$

Since priori probability of each class is identical, we can define new $g_i\hat(x)$

$$
g_i\hat(x) = -\sum_{j=1}^{c}\lambda_{ji}P(x|w_j)
$$

(c)

Without the condition that the priori probability of each class is identical, $g_i\hat(x)$ can be written as

$$
g_i\hat(x) = -\sum_{j=1}^{c}\lambda_{ji}g_j(x)
$$

In classification algorithm, after we update g_i totally, we calculate $g_i\hat(x)$ with the equation above, and we select one of the class k that satisfies $k = \arg\max_i g_i(x)$

6

(a)

$$
\begin{aligned}
P(error) &= P(w_1)P(n_{w_1} < n_{w_2}) + P(w_2)P(n_{w_2} < n_{w_1}) \\
&= P(n_{w_1} < n_{w_2}) \\
&= \sum_{j=0}^{(k-1)/2} \binom{n}{j}(\frac{1}{2})^j(\frac{1}{2})^{n-j} \\
&= \frac{1}{2^n}\sum_{j=0}^{(k-1)/2} \binom{n}{j}
\end{aligned}
$$

(b)

$$
\because \binom{n}{j} > 0, \quad j = 0, 1, 2..., n
$$
$$
\therefore \sum_{j=0}^{(k-1)/2} \binom{n}{j} > \sum_{j=0}^{0}\binom{n}{j}, \quad k > 1
$$

The right hand side represents the nearest neighbor rule. Therefore, it has lower error rate.

(c)

$$
P_n(e) < \frac{1}{2^n}\sum_{j=0}^{c\sqrt{n}}\binom{n}{j}
$$

Since

$$
(1+1)^n = \sum_{j=0}^{n}\binom{n}{j}
$$

can be devided into $m = \lfloor\frac{n}{c\sqrt{n}}\rfloor$ pieces in order. (Elements in interval $[\lfloor\frac{n}{c\sqrt{n}}\rfloor \times c\sqrt{n}, n]$ are merged into the last piece)

Note that $\binom{n}{j}$ is a symmetric convex function. We can conclude that

$$
\sum_{j\in Piece_1} \binom{n}{j} = \min \sum_{j\in Piece_k} \binom{n}{j}, \quad k = 1, 2, \ldots, m
$$

$$
\therefore \sum_{j=0}^{c\sqrt{n}}\binom{n}{j} < \frac{2^n}{m}
$$

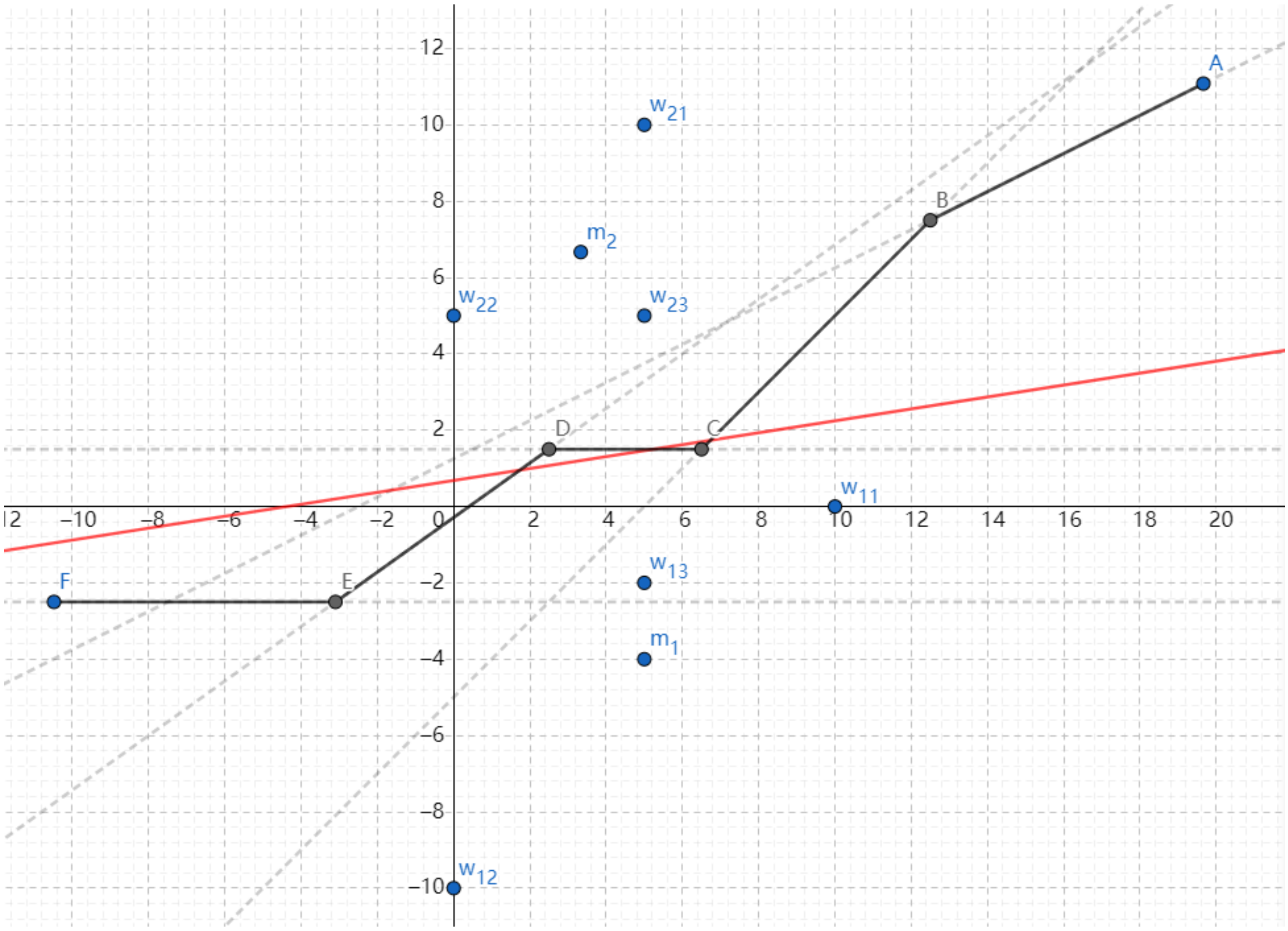
$$
\frac{1}{2^n}\sum_{j=0}^{c\sqrt{n}}\binom{n}{j} < \frac{1}{m} = O(\frac{1}{\sqrt{n}})
$$

$$
\therefore \lim_{n\to\infty} P_n(e) = 0
$$

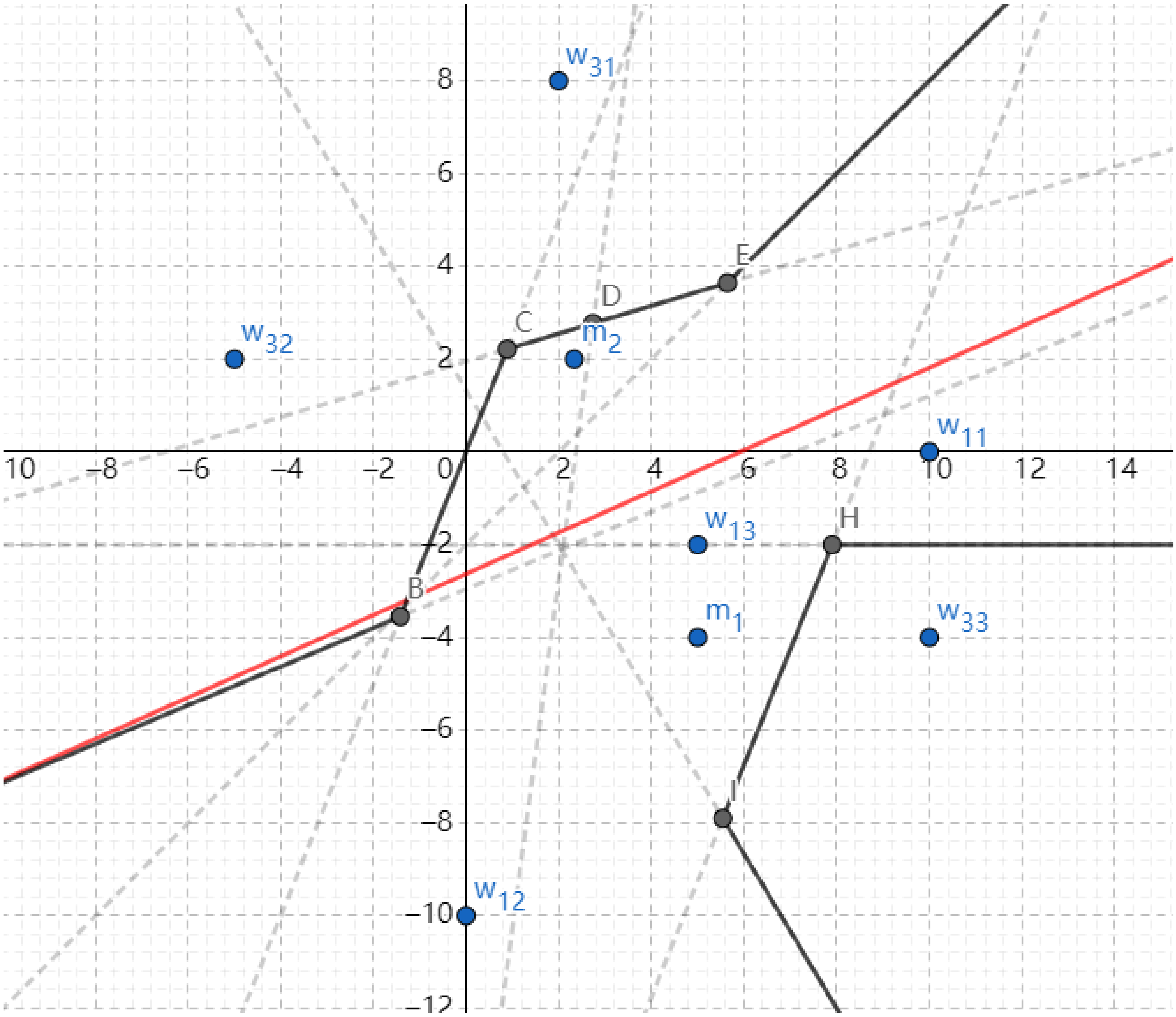
9

$m_1 = (5, -4), \quad m_2 = (\frac{10}{3}, \frac{20}{3}), \quad m_3(\frac{7}{3}, 2)$

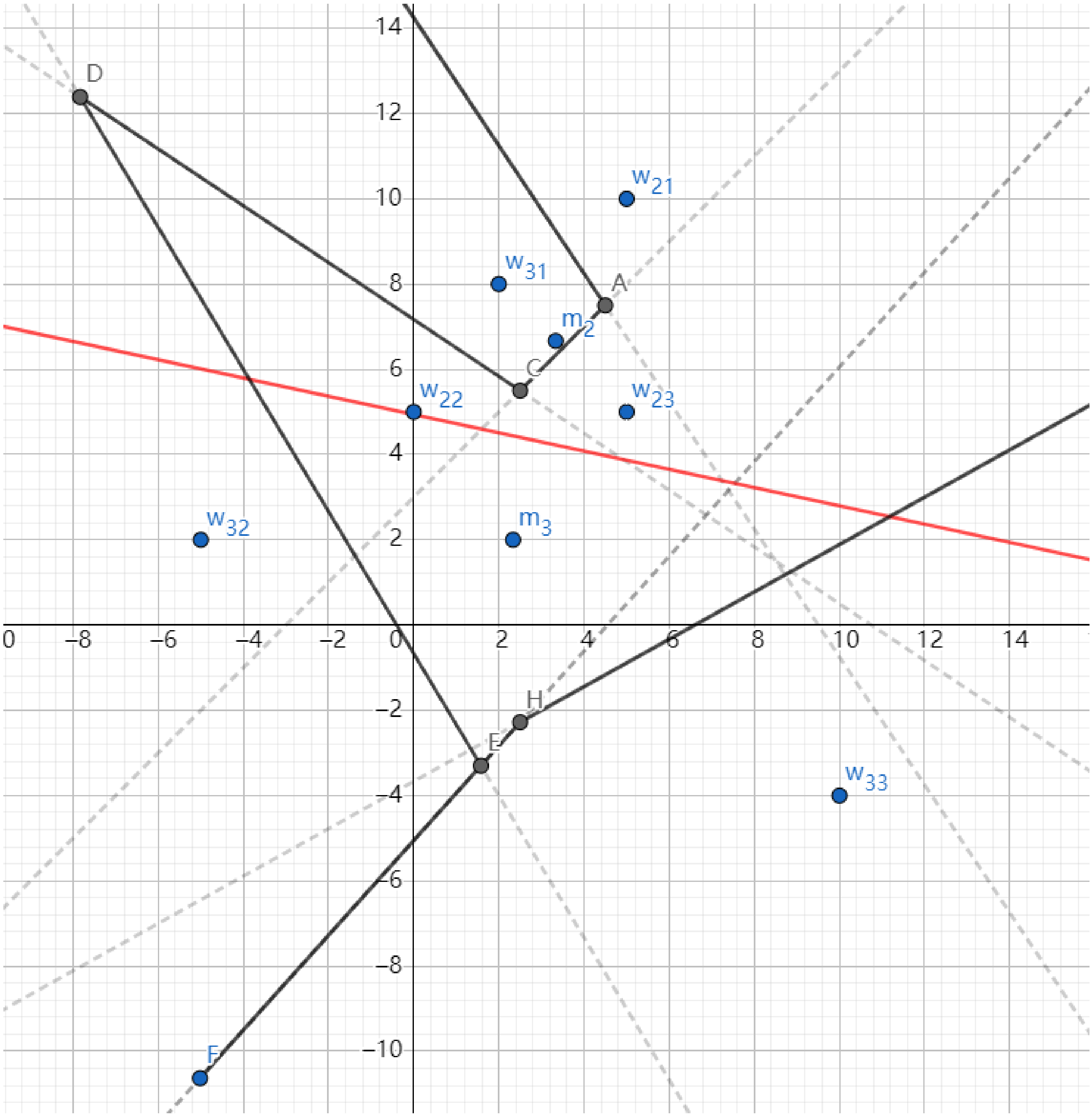
(a)



(b)

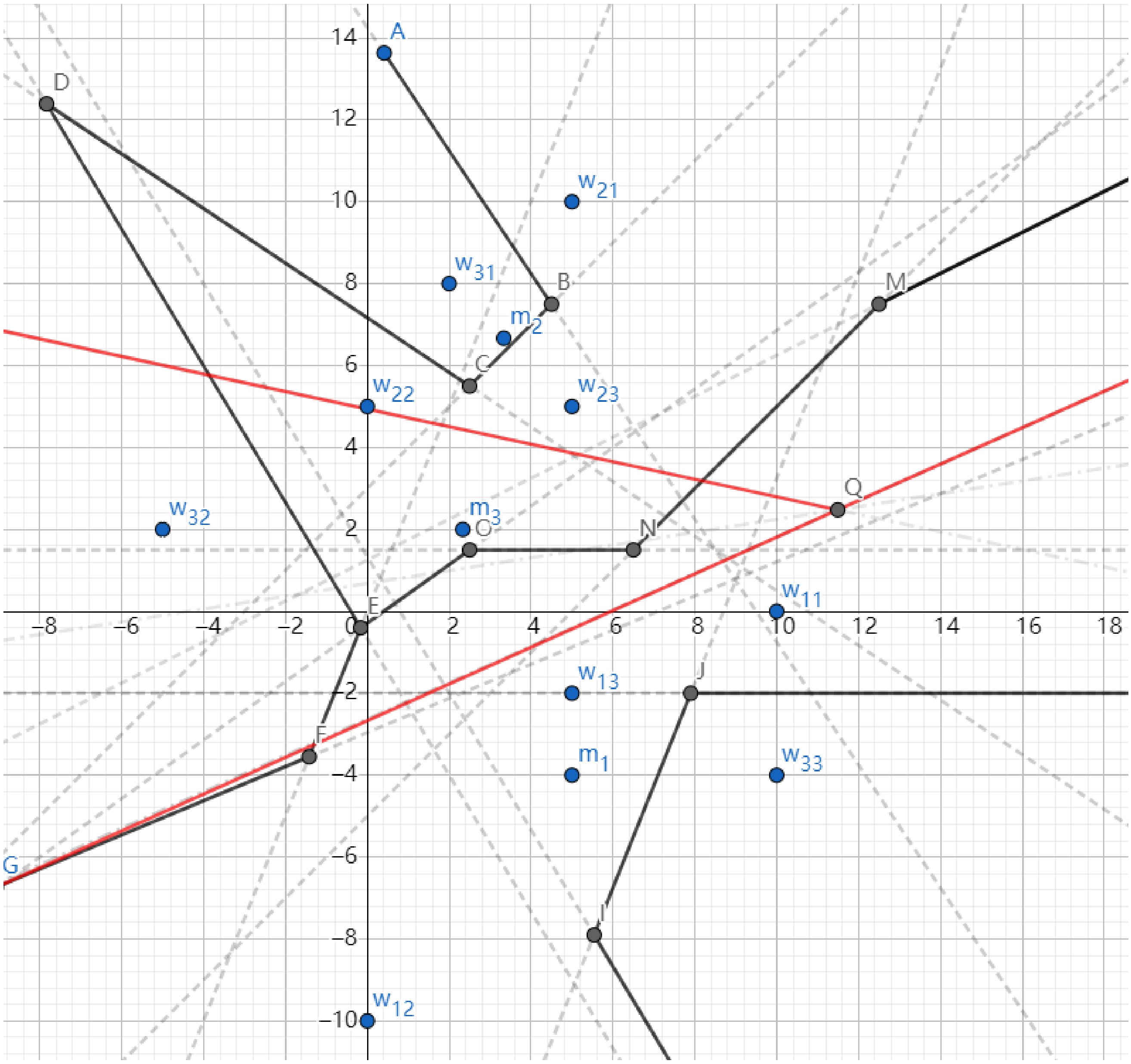


(c)



D
A
C
H
E
F
w_{21}
w_{31}
m_2
w_{22}
w_{23}
w_{32}
m_3
w_{33}
14
12
10
8
6
4
2
−2
−4
−6
−8
−10
−8
−6
−4
−2
0
2
4
6
8
10
12
14

(d)



17

Since the probability distribution of each class is identical, we have $P(x|w_i) = P(x|w),\ i = 1, 2, \ldots, c$.

$$P(x) = \sum_{i=1}^{c} P(x|w_i)P(w_i) = \frac{1}{c}\sum_{i=1}^{c} P(x|w_i) = P(x|w)$$

$$\therefore\ \ P(w_i|x) = \frac{P(x|w_i)P(w_i)}{P(x)} = P(w_i) = \frac{1}{c}$$

$$\begin{aligned} P = \lim_{n\to\infty} P_n(e) = &\lim_{n\to\infty} \int P_n(e|x)p(x)dx \\ &= \int [1 - \sum_{i=1}^{c} P^2(w_i|x)]p(x)dx \\ &= 1 - \frac{1}{c} \end{aligned}$$

$$P^* = \int P^*(error|x)p(x)dx = \int (1 - P(w_j|x))p(x)dx = 1 - \frac{1}{c}$$

$$\therefore \ \ P = P^*(2 - \frac{c}{c-1}P^*) = 1 - \frac{1}{c}$$