Question 1

$$L(\mathbf{w},\lambda)=\mathbf{w}^{\mathrm{T}}(\mathbf{m}_2-\mathbf{m}_1)+\lambda(\mathbf{w}^{\mathrm{T}}\mathbf{w}-1)$$

$$\frac{\partial}{\partial \mathbf{w}}L(\mathbf{w},\lambda)=\mathbf{m}_2-\mathbf{m}_1+2\lambda\mathbf{w}=0$$

Hence,

$$\mathbf{w}=-\frac{1}{2\lambda}(\mathbf{m}_2-\mathbf{m}_1)\propto(\mathbf{m}_2-\mathbf{m}_1)$$

Question 2

$$\begin{aligned}
J(\boldsymbol{w}) &= \frac{(m_2-m_1)^2}{s_1^2+s_2^2}\\
&= \frac{(\boldsymbol{w}^T\boldsymbol{m}_2-\boldsymbol{w}^T\boldsymbol{m}_1)^2}{\sum_{n\in\mathcal{C}_1}(\boldsymbol{w}^T\boldsymbol{x}_1-\boldsymbol{w}^T\boldsymbol{m}_1)+\sum_{n\in\mathcal{C}_2}(\boldsymbol{w}^T\boldsymbol{x}_2-\boldsymbol{w}^T\boldsymbol{m}_2)}\\
&= \frac{\boldsymbol{w}^T(\boldsymbol{m}_2-\boldsymbol{m}_1)(\boldsymbol{m}_2-\boldsymbol{m}_1)^T\boldsymbol{w}}{\boldsymbol{w}^T\left(\sum_{n\in\mathcal{C}_1}(\boldsymbol{x}_1-\boldsymbol{m}_1)+\sum_{n\in\mathcal{C}_2}(\boldsymbol{x}_2-\boldsymbol{m}_2)\right)\boldsymbol{w}}\\
&= \frac{\boldsymbol{w}^T\boldsymbol{S}_B\boldsymbol{w}}{\boldsymbol{w}^T\boldsymbol{S}_W\boldsymbol{w}}
\end{aligned}$$

Question 3

For a point ϕ_n from $\mathcal{C}_k$,

$$p(\phi_n,\mathcal{C}_k)=p(\mathcal{C}_k)p(\phi_n|\mathcal{C}_k)=\pi_k p(\phi_n|\mathcal{C}_k)$$

so the likelihood function is

$$p(\{\phi_n,\mathbf{t}_n\}|\boldsymbol{\pi},\boldsymbol{\mu},\Sigma)=\prod_{n=1}^{N}\prod_{k=1}^{K}[\pi_k p(\phi_n|\mathcal{C}_k)]^{t_{nk}}$$

Take the log-likelihood

$$\sum_{n=1}^{N}\sum_{k=1}^{K}[t_{nk}\ln\pi_k+t_{nk}\ln p(\phi_n|\mathcal{C}_k)]$$

As $\sum_{k=1}^{K}\pi_k=1$, we use Lagrange Multiplier to maximize

$$L(\pi_i,\lambda)=\sum_{n=1}^{N}\sum_{k=1}^{K}[t_{nk}\ln\pi_k+t_{nk}\ln p(\phi_n|\mathcal{C}_k)]+\lambda(\sum_{k=1}^{K}\pi_k-1)$$

$$\frac{\partial}{\partial\lambda}L(\pi_i,\lambda)=\sum_{k=1}^{K}\pi_k-1=0$$

$$\frac{\partial}{\partial\pi_i}L(\pi_i,\lambda)=\frac{\sum_{n=1}^{N}t_{ni}}{\pi_i}+\lambda=0$$

$$\pi_i = -\frac{\sum_{n=1}^{N} t_{ni}}{\lambda} = -\frac{N_i}{\lambda}$$

Since

$$\sum_{k=1}^{K} \pi_k = \sum_{k=1}^{K} -\frac{N_i}{\lambda} = -\frac{N}{\lambda} = 1$$

then $\lambda = -N$

Hence,

$$\pi_i = \frac{N_i}{N}$$

Question 4

$$\begin{aligned} \frac{\partial}{\partial a}\sigma(a) &= \frac{e^{-a}}{(1+e^{-a})^2} \\ &= \frac{e^{-a}}{1+e^{-a}} \frac{1}{1+e^{-a}} \\ &= \frac{e^{-a}}{1+e^{-a}} (1 - \frac{e^{-a}}{1+e^{-a}}) \\ &= \sigma(1-\sigma) \end{aligned}$$

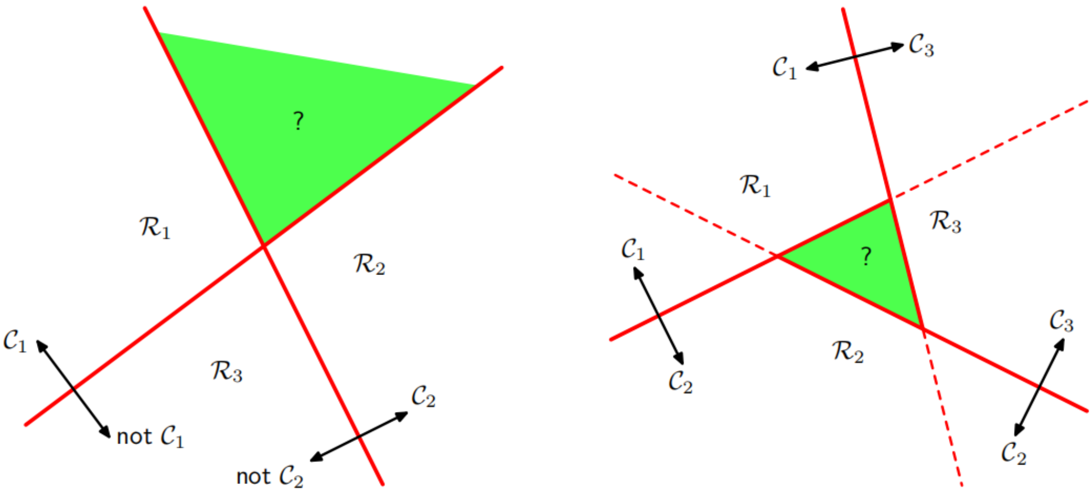
Question 5

$$\begin{aligned} y_n &= \sigma(a_n) \\ a_n &= \boldsymbol{w}^T \boldsymbol{\phi}_n \end{aligned}$$

$$\frac{\partial y_n}{\partial \boldsymbol{w}} = \frac{\partial y_n}{\partial \boldsymbol{a}_n} \frac{\partial a_n}{\partial \boldsymbol{w}} = y_n(1-y_n)\phi_n$$

$$\begin{aligned} \nabla \mathbb{E}(\boldsymbol{w}) &= -\sum_{n=1}^{N} \left(\frac{t_n}{y_n} \frac{\partial y_n}{\partial \boldsymbol{w}} - \frac{1-t_n}{1-y_n} \frac{\partial y_n}{\partial \boldsymbol{w}} \right) \\ &= -\sum_{n=1}^{N} (t_n(1-y_n) - (1-t_n)y_n)\phi_n \\ &= \sum_{n=1}^{N} (y_n - t_n)\phi_n \end{aligned}$$

Question 6



Question 7

Proof by contradiction

If their convex hulls intersect, then

$$\exists \boldsymbol{y}, \quad s.t. \, \boldsymbol{y} = \sum_n \alpha_n \boldsymbol{x}^n = \sum_m \beta_n \boldsymbol{z}^n$$

where $\alpha, \beta > 0$ and $\sum_n \alpha_n = \sum_m \beta_n = 1$.

Assume $\{\boldsymbol{x}^n\}$ and $\{\boldsymbol{z}^m\}$ are linearly separable, there exists a vector $\hat{\boldsymbol{w}}$ and a scalar w_0 such that $\hat{\boldsymbol{w}}^T \boldsymbol{x}^n + w_0 > 0$ for all $\boldsymbol{x}^n$, and $\hat{\boldsymbol{w}} \boldsymbol{z}^m + w_0 < 0$ for all $\boldsymbol{z}^m$.

We figure out

$$\sum_n \hat{\boldsymbol{w}}^T \alpha_n \boldsymbol{x}^n + \sum_n \alpha_n w_0 = \sum_n \hat{\boldsymbol{w}}^T \boldsymbol{y} + w_0 > 0$$
$$\sum_m \hat{\boldsymbol{w}}^T \beta_m \boldsymbol{z}^m + \sum_m \beta_m w_0 = \sum_m \hat{\boldsymbol{w}}^T \boldsymbol{y} + w_0 < 0$$

It leads a contradiction.

Hence, if their convex hulls intersect, the two sets of points cannot be linearly separable, and conversely that, if they are linearly separable, their convex hulls do not intersect.