$$P(e_1 \mid x) = \frac{1}{1 + e^{-w^T x + w_0}}$$

Discriminative Model

$$P(e_1 \mid x) = \frac{1}{1 + e^{-w_1 x + w_0}}$$

2 parameters (w_0, w_1)

Generative Model

4 parameters $(\mu_1, \mu_2, \sigma, \pi)$

image

cat

$$P(x \mid e_k)$$ Generative Model

$$P(e_k \mid x)$$ Discriminative Model

$$P(\mathcal{C}_1 \mid x) = \frac{1}{1 + e^{-w^T x}}$$

$w^T \phi(x)$

$$L(w) = P(t_1, \ldots, t_N \mid w)$$

$t_n = 1$ if "Class 1"

$t_n = 0$ o.w.

$$= \prod_{n=1}^{N} p(t_n \mid w)$$

$$= \prod_{n=1}^{N} \begin{cases} P(\mathcal{C}_1 \mid x_n) & \text{if } t_n \text{ is "Class 1"} \\ 1 - P(\mathcal{C}_1 \mid x_n) & \text{o.w.} \end{cases}$$

$$= \prod_{n=1}^{N} P(\mathcal{C}_1 \mid x_n)^{t_n} \left(1 - P(\mathcal{C}_1 \mid x_n)\right)^{1-t_n}$$

$$\sigma(a) = \frac{1}{1+e^{-a}}$$

$$\frac{u}{v} \qquad \frac{u'v - vu'}{v^2}$$

$$\frac{d}{da}\sigma(a) = \frac{+e^{-a}}{(1+e^{-a})^2}$$

$$= \sigma(a)\cdot\sigma(a)\cdot e^{-a} \qquad e^{-a} = \frac{(1-\sigma(a))}{\sigma(a)}$$

$$= \sigma(a)(1-\sigma(a))$$

$p(\vec{t}\,|\,\vec{w};\vec{x})$

$$p(\vec{t}\,|\,\vec{w}) = \prod_{n=1}^{N} p(t_n|w) = \prod \begin{cases} p(e_1|x_n) & t_n = 1 \\ 1 - p(e_1|x_n) & t_n = 0 \end{cases}$$

$$= \prod_{n=1}^{N} y_n^{t_n} (1 - y_n)^{1-t_n}$$

$$y_n \equiv p(e_1|x_n)$$

$$\ln p(\vec{t}\,|\,\vec{w}) = \sum_{n=1}^{N} t_n \ln y_n + (1-t_n) \ln(1-y_n)$$

$$\nabla \ln p(\vec{t}\,|\,\vec{w}) = \sum_{n=1}^{N} \frac{t_n}{y_n} \nabla y_n - \frac{(1-t_n)}{(1-y_n)} \nabla y_n$$

$$= \sum_{n=1}^{N} \frac{t_n}{\cancel{y_n}} \cancel{y_n}(1-y_n) x_n - \frac{(1-t_n)}{\cancel{(1-y_n)}} y_n \cancel{(1-y_n)} x_n$$

$$\sigma(a) = \frac{1}{1+e^{-a}}$$

$$= \sum_{n=1}^{N} t_n (1-y_n) x_n - (1-t_n) y_n x_n$$

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

$$= \sum_{n=1}^{N} t_n x_n - \cancel{t_n y_n x_n} - y_n x_n + \cancel{t_n y_n x_n}$$

$$= \sigma(a)(1 - \sigma(a))$$

$$L(w) = \prod_{n=1}^{N} P(e_1 \mid x_n)^{t_n} \left(1 - P(e_1 \mid x_n)\right)^{1-t_n}$$

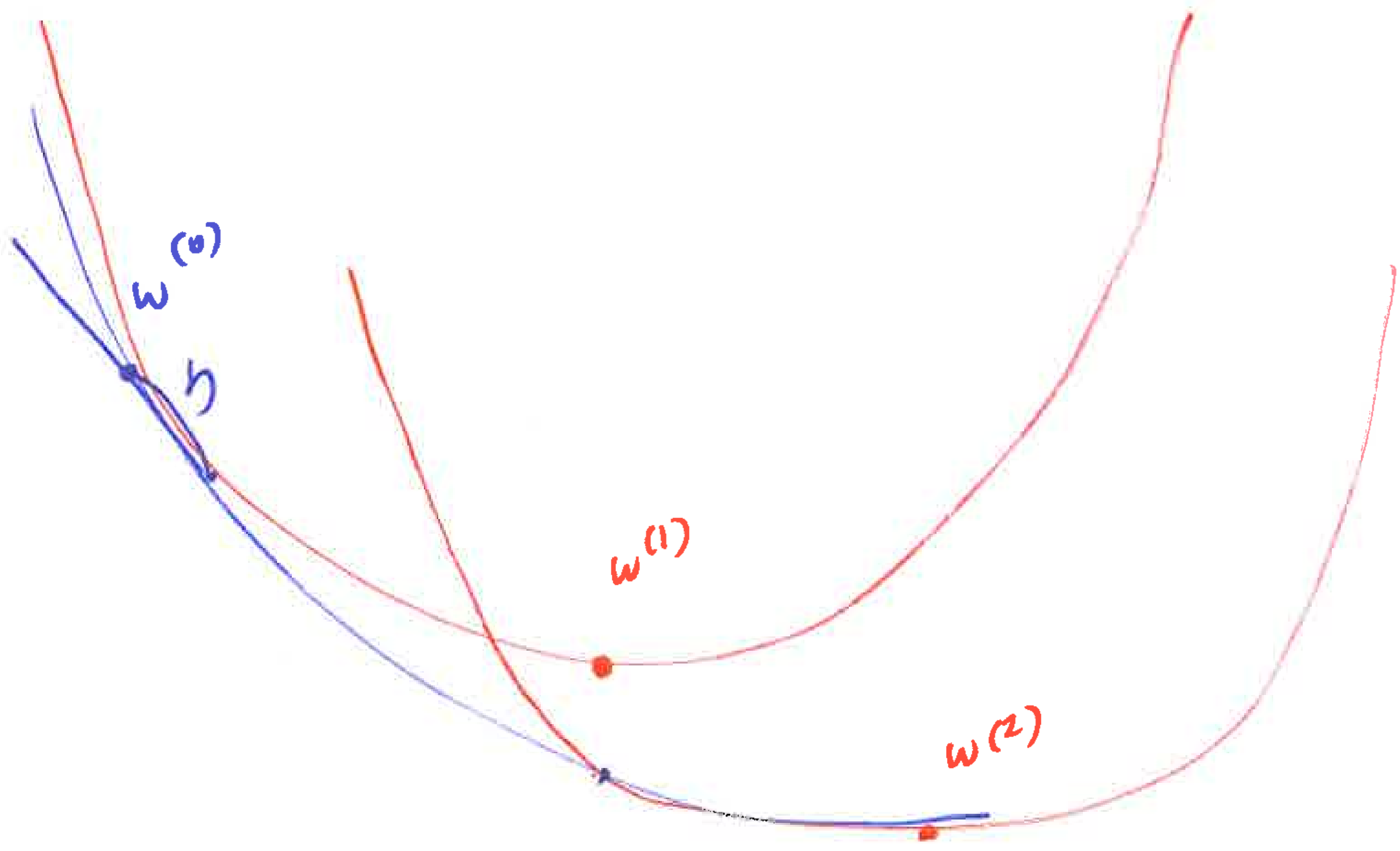
x_n 'image

$$t_n \text{ is } \begin{cases} 1 & \text{if Cat} \\ 0 & \text{o.w.} \end{cases}$$

$$l(w) = \sum_{n=1}^{N} t_n \ln P(e_1 \mid x_n) + (1-t_n) \ln\left[1 - P(e_1 \mid x_n)\right]$$

$$P(e_1 \mid x_n) = \frac{1}{1 + e^{-w^T x_n}}$$

$$\nabla l(w) = \sum_{n=1}^{N} \left(t_n - P(e_1 \mid x_n)\right) x_n$$



$w^{(0)}$
η
$w^{(1)}$
$w^{(2)}$