1.(c) Does this mean we can turn any discriminative model (predicting class labels from inputs, $p(C_k|x)$ into a generative one (predictining inputs from class labels, $p(x|C_k)$ by "going backwards" through the model? Unfortunately, the answer is no, because continuous distributions for inputs are more complex than binary distributions.

To make this clear, consider the special case: Assume that the class distributions are both univariate Gaussians with different means and variances:

$$p(x|C_k) = \mathcal{N}(x; \mu_k, \sigma_k^2).$$

From your answer above, construct an explicit form of θ, θ_0 (that's two real numbers) as a function of $\mu_1, \mu_2, \sigma_1, \sigma_2, p(C_1)$.

Assume someone has performed logistic regression and given you the parameters θ, θ_0.

1. Can you recover the parameters $\mu_1, \mu_2, \sigma_1, \sigma_2, p(C_1)$?
2. Could you do so if we set $p(C_1) = p(C_2) = 1/2$ and $\sigma_1 = 1$ $\sigma_2 = 1$?

Answer:

① Start with the assumption

$$p(x|C_1) = \mathcal{N}(x; \mu_1, \sigma_1^2)$$

and $$p(x|C_2) = \mathcal{N}(x; \mu_2, \sigma_2^2)$$

From the previous answer we know that given exponential distributions.

$$p(x|C_1) = h(x) \exp\left[\phi^T(x)\, \omega_1 - \ln Z(\omega_1)\right]$$

$$p(x|C_2) = h(x) \exp\left[\phi^T(x)\, \omega - \ln Z(\omega_2)\right]$$

$$\frac{p(C_1|x)}{p(C_2|x)} = \exp\left[\phi^T(x)\,(\omega_1 - \omega_2) - \ln \frac{Z(\omega_1)}{Z(\omega_2)} + \ln \frac{p(C_1)}{p(C_2)} \right]$$

as $\quad a(x) = \ln \left[\frac{p(C_1|x)}{p(C_2|x)} \right]$

$$a(x) = \phi^T(x)\, \underbrace{(\omega_1 - \omega_2)}_{\theta} \underbrace{- \ln \frac{Z(\omega_1)}{Z(\omega_2)} + \ln \frac{p(C_1)}{p(C_2)}}_{\theta_0}$$

we know that the normal distribution also belongs to the exponential family. (Lecture-5, slide 13)

$$\mathcal{N}(x; \mu, \sigma^2) = \exp\left(\left[x \quad -\frac{x^2}{2}\right] \begin{bmatrix} \mu/\sigma^2 \\ 1/\sigma^2 \end{bmatrix} - \left(\frac{\mu^2}{2\sigma^2} + \log \sqrt{2\pi\sigma^2} \right) \right)$$

where $\phi^T(x) = [\phi_1(x) \quad \phi_2(x)] = \left[x \quad -\frac{x^2}{2}\right]$

$$\omega = \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} = \begin{bmatrix} \mu/\sigma^2 \\ 1/\sigma^2 \end{bmatrix}$$

$$\ln Z(\omega) = -\frac{1}{2}\left[\frac{\omega_1}{\omega_2} - \ln \omega_2 + \ln(2\pi) \right]$$

$$= -\left[\frac{\mu^2}{2\sigma^2} + \ln(\sqrt{2\pi\sigma^2}) \right]$$

Combining previous two equations we get

$$\frac{p(C_1|x)}{p(C_2|x)} = \exp\left\{ [x \quad -\tfrac{x^2}{2}] \begin{bmatrix} \mu_1/\sigma_1^2 - \mu_2/\sigma_1^2 \\ 1/\sigma_1^2 - 1/\sigma_2^2 \end{bmatrix} - \frac{1}{2}\left(\frac{\mu_1^2}{\sigma_1^2} - \frac{\mu_2^2}{\sigma_2^2} - \ln\frac{(2\pi\sigma_1^2)}{(2\pi\sigma_2^2)} \right) + \ln\frac{p(C_1)}{p(C_2)} \right\}$$

Let us assume $p(C_1) = p(C_2) = 1/2$ and $\sigma_1 = \sigma_2 = \sigma$

then

$$\frac{p(C_1|x)}{p(C_2|x)} = \exp\left\{ [x \quad -\tfrac{x^2}{2}] \begin{bmatrix} \mu_1/\sigma^2 - \mu_2/\sigma^2 \\ 0 \end{bmatrix} - \frac{1}{2}\left(\frac{\mu_1^2}{\sigma^2} - \frac{\mu_2^2}{\sigma^2} \right) \right\}$$

$$a(x) = \frac{\mu_1 - \mu_2}{\sigma^2} x - \frac{\mu_1^2 - \mu_2^2}{2\sigma^2}$$

$$a(x) = \left(\frac{\mu_1 - \mu_2}{\sigma^2}\right) \left[x - \frac{(\mu_1 + \mu_2)}{2} \right]$$

Clearly $\mu_1 = 3$, $\mu_2 = 0$, $\sigma^2 = 3$ will give $a(x) = x - \frac{3}{2}$

and $\mu_1 = 2$, $\mu_2 = 1$, $\sigma^2 = 1$ will result also result in $a(x) = x - 3/2$

Thus there does not exist a unique inverse mapping from parameters θ, θ_0 to $\mu_1, \mu_2, \sigma_1, \sigma_2$

② Now consider the case $p(C_1) = p(C_2) = 1/2$

$\sigma_1 = \sigma_2 = \sigma = 1$

$$\frac{p(C_1|x)}{p(C_2|x)} = \exp\left\{ \begin{bmatrix} x & -\frac{x^2}{2} \end{bmatrix} \begin{bmatrix} \mu_1 - \mu_2 \\ 0 \end{bmatrix} - \frac{1}{2}\left(\mu_1^2 - \mu_2^2\right) \right\}$$

$$a(x) = (\mu_1 - \mu_2)\, x - \frac{1}{2}(\mu_1^2 - \mu_2^2)$$

Here given $\theta_1 = \mu_1 - \mu_2$

$\theta_0 = -\frac{1}{2}\, \mu_1^2 - \mu_2^2 = -\frac{1}{2}\, \theta_1 (\mu_1 + \mu_2)$

or $\mu_1 - \mu_2 = \theta_1$

$\mu_1 + \mu_2 = -2\theta_0/\theta_1$

$$2\mu_1 = \theta_1 - \frac{2\theta_0}{\theta_1} = \frac{\theta_1^2 - 2\theta_0}{\theta_1}$$

$$\mu_1 = \frac{\theta_1^2 - 2\theta_0}{2\theta_1} \qquad \mu_2 = \frac{\theta_1^2 - 2\theta_0}{2\theta_1} - \theta_1 = -\frac{2\theta_0 - \theta_1^2}{2\theta_1}$$

Thus assuming $\theta_1 \neq 0$ μ_1 & μ_2 are recoverable from $a(x)$ given θ_1 and θ_0 from the result of logistic regression.