Lecture 8 Generative Models

$P(y|x)$

$x \rightarrow$ video

$P(x)$

0.5

0.5

$P(x|y)$

(x, y)

$$P(x,y) = \underbrace{P(y)}_{\text{class prior}} P(x|y)$$

$P(y|x)$? ← $P(y)\ P(x|y)$

$$\boxed{P(y|x) \propto P(x|y)\,P(y)}$$

$P(x)$ difficult to compute

$P(y|x)$

1: x is given

2. distribution on y

$\propto$ $\left\{ \begin{array}{l} P(x|y)\,P(y) \\ \boxed{P(x|y)\,P(y)} \end{array} \right.$

$$= \frac{\boxed{P(x|y)\,P(y)}}{Z = P(x)}$$

1. data augmentation $\rightarrow$ more data

$\downarrow$ train

disriminati

$$\boxed{\begin{array}{l} P(cat) = 0.8 \\ P(dog) = 0.2 \end{array}} \qquad \begin{array}{l} 0.5 \\ 0.5 \end{array}$$

human knowledge

logistic + $\boxed{\text{regularizer}}$

$P(y)$

y cat, dog - -

latent variable models

$P(x|y)$

image x

data (x, y)

disrim..

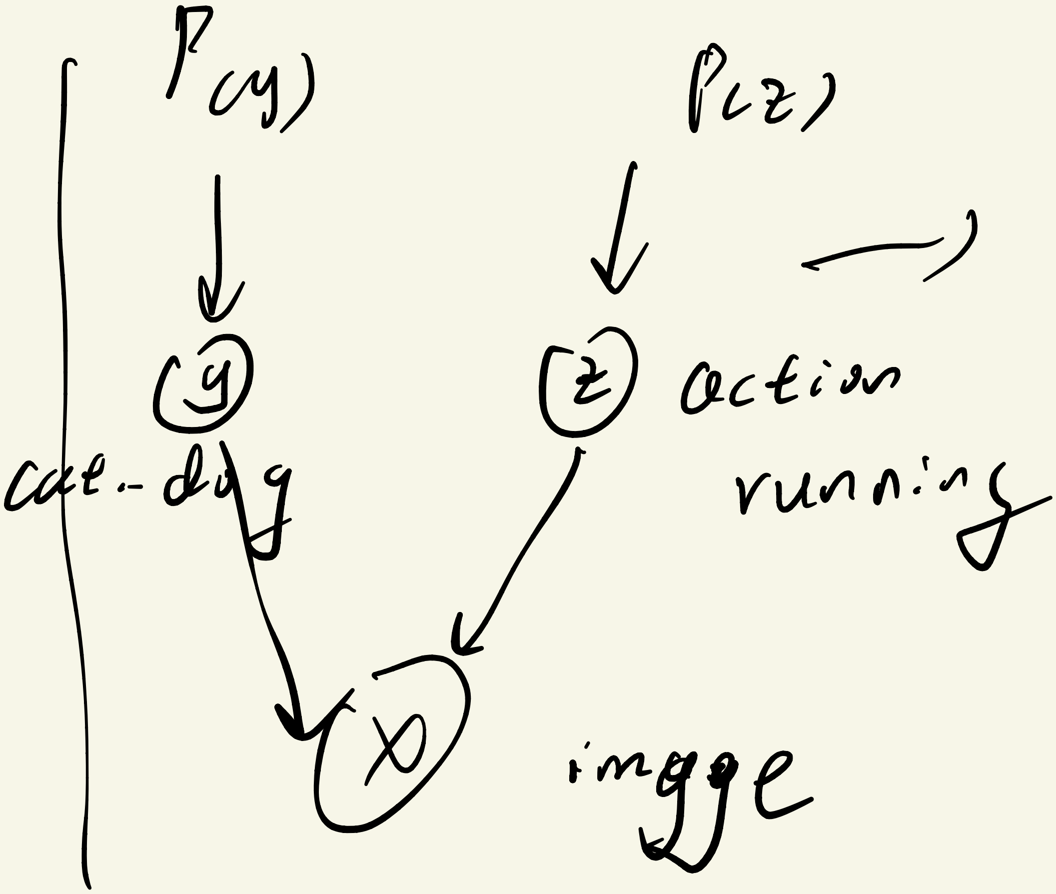
$P(y|x)$

data (x)

$$P(x) = \sum_y P(x, y)$$

$$= \sum_y P(y) P(x|y)$$

$\arg\max \log P(x)$



Probabilistic Graphical models

eating

standing

(x, y, z)

(x, y)

$\underline{\underline{(x)}}$

compression is all you need

X

generative model

$$\underbrace{p^*(x)}_{\text{true distribution}} \longleftrightarrow q(x)$$

$$KL(p \| q) = E_p\left[\log \frac{p}{q}\right]$$

$$\arg\max_\theta \sum_{i=1}^{n} q_\theta(x^{(i)}) \Rightarrow$$

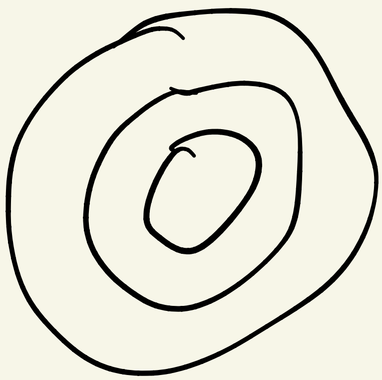
$$\arg\max_\theta E_{x^{(i)} \sim p^*(x)} q_\theta(x^{(i)})$$

$$|||$$

$$\arg\min_\theta KL\left(p^*(x) \,\|\, q_\theta(x^{(i)})\right)$$

$$I = \begin{bmatrix} 1 & \ddots & 0 & 0 \\ & \ddots & 1 & 0 \\ 0 & & \ddots & \\ 0 & 0 & & 1 \end{bmatrix}$$

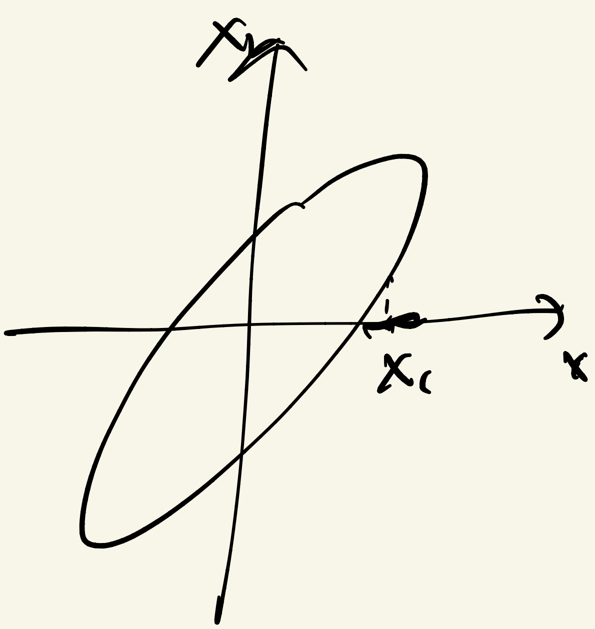
$$\Sigma = 0.6\, I$$



$\underline{x_1 > 0}$

$\underline{x_1 \;,\; x_2}$ $\qquad x_2$

$$\Sigma = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$$



$\mu_0 \quad \mu_1 \quad \phi \quad \Sigma$

$$\mu_0 = \frac{\sum_{i=1}^{n} \mathbb{1}(y^{(i)}=0)\, x^{(i)}}{\sum_{i=1}^{n} \mathbb{1}(y^{(i)}=0)}$$

$$\arg\max_y P(y|x) = \arg\max_y \underbrace{P(y)}_{\text{Ber}}\ \underbrace{P(x|y)}_{\text{Gau.}}$$

$$= \arg\max_y \log P(y) + \log P(x|y)$$

$y=0 \qquad L_0$

$L_0 > L_1$

$y=1 \qquad L_1$

$$y=1 \Rightarrow \log\phi - \frac{1}{2}(x-\mu_1)^T\Sigma^{-1}(x-\mu_1)$$

$x^T\Sigma^{-1}x$

$$y=0 \Rightarrow \log\phi - \frac{1}{2}(x-\mu_0)^T\Sigma^{-1}(x-\mu_0)$$

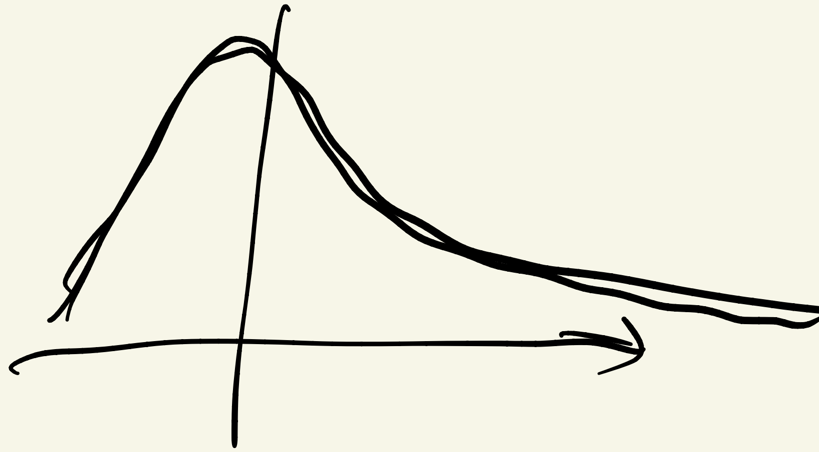
$x^T\Sigma^{-1}x$

linear boundary

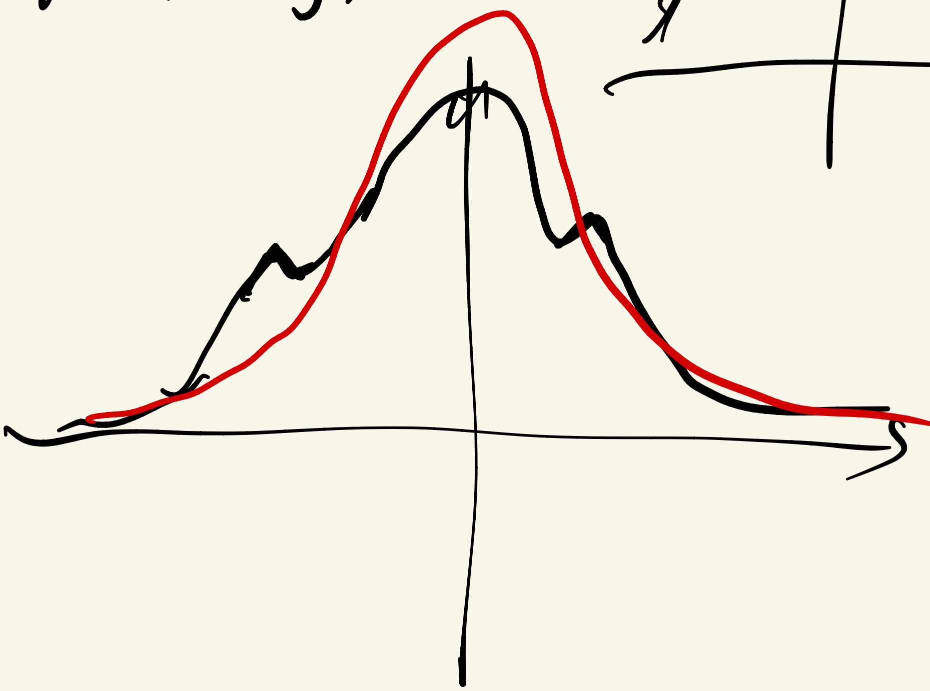
$$P(y|x) = \frac{1}{1+\exp(-\theta^T x)}$$

$P(x|y) \sim \text{Poisson}(\lambda)$

$P(y|x) \rightarrow \text{logistic}$



$P(x|y)$



a boy computer scien

lose

1. count

$\boxed{P(x|y)}$

2. order

$\boxed{2^{50000}}$

$$\underbrace{P(x_i \mid y)} = \underbrace{P(x_i \mid y, x_j \cdots x_k \ldots)}$$

y is given

computer $\rightarrow$ electrical