Lec 2: Friday, 02 August

02 August 2013
13:35

1) Noise Continued - Chapter 1 of DHS
2) ML vs MAP vs Bayesian approaches - Chapter 3 of DHS
3) Discriminative vs Generative methods - Chapter from Tom Mitchell's book

---

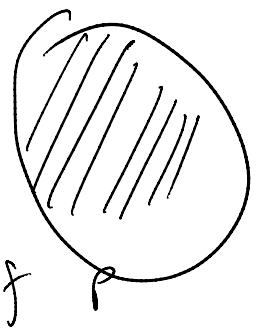
x_1 x_2 x_3 --- x_n

y_1 y_2 y_3 y_n

$y = f(x)$ ✓

$x = g(y)$

$p(x, y)$



(1) FIND

(2) PARAMETERS

(3) $x, \hat{x}$

(4) $y, \hat{y}$

X , Y , x , y

$N = 100$, $n = 45$, $y = 1$

$p(y = 1 \mid N, n)$

$p(y = -1 \mid N, n)$

$p(y = +1 \mid N, n) > p(y = -1 \mid N, n)$

$1 - p(y = +1 \mid \cdots)$

$p(y = +1 \mid N, n) \geq 1/2$

$p(y = +1 \mid N, n) = \theta$

$\underset{\theta}{\text{ARGMAX}}\ p(n, N \mid \theta)$

① $p(n, N \mid \theta) = {}^{N}C_{n}\, \theta^{n} (1-\theta)^{N-n}$

$\underset{\theta}{\text{ARGMAX}}\ C_n\, \theta^{n} (1-\theta)^{N-n}$

$\theta_{ML} = \frac{n}{N} = \frac{45}{100}$

MAP

$\theta_{ML} = \underset{\theta}{\text{ARGMAX}}\ P(D|\theta)$

$P(n, N|\theta)$

$\theta_{MAP} \equiv \underset{\theta}{\text{ARGMAX}}\ P(\theta|D)$

$P(\theta|n, N)$

$$P(\theta|n,N) = \frac{P(n,N|\theta)\,P(\theta)}{P(n,N)}$$

$$\Rightarrow \underset{\theta}{\text{ARGMAX}}\ P(\theta|N,n) = \underset{\theta}{\text{ARGMAX}}\ P(n,N|\theta)P(\theta)$$

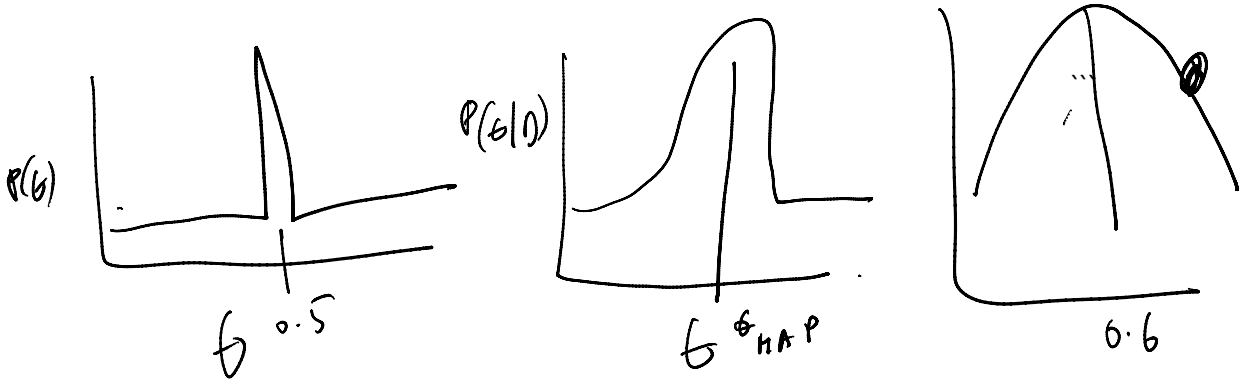
$$P(n,N|\theta) = {}^{N}C_{n}\,\theta^{n}(1-\theta)^{N-n}$$

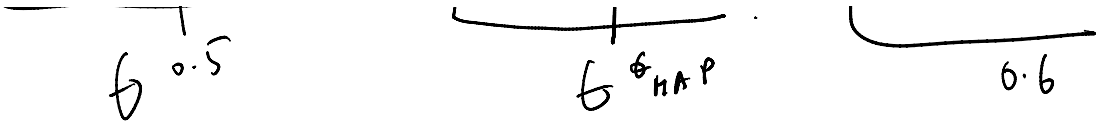
$$P(\theta|a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\,\theta^{a-1}(1-\theta)^{b-1}$$

$$\underset{\theta}{\text{ARGMAX}}\ \theta^{n+a-1}(1-\theta)^{N-n+b-1}$$

$$\theta_{MAP} = \frac{n+a-1}{N+a+b-2} > 1/2$$

$$\theta_{ML} = \frac{n}{N} > 1/2$$





$$p(y=+1 \mid n, N)$$

$$= \int_\theta p(y=+1, \theta \mid n, N)\, d\theta \quad \checkmark$$

$$= \int_\theta p(y=+1 \mid \theta, n, N)\, p(\theta \mid n, N)\, d\theta \quad \checkmark$$

$$\int_\theta p(y=+1 \mid \theta)\, p(\theta \mid n, N)\, d\theta \quad \checkmark$$

$$\int_\theta \theta\, \beta(\theta \mid n+a,\ N-n+b)\, d\theta$$

$$= \frac{n+a}{N+a+b}$$

$$p(y=+1 \mid x, X, Y)$$

$$= \int_\theta p(y=+1, \theta \mid x, X, Y)\, d\theta$$

$$= \int_\theta P(y=+1 \mid \theta, x, X, Y) P(\theta \mid x, X, Y) d\theta$$

$$= \int_\theta P(y=+1 \mid \theta, x) P(\theta \mid X, Y) d\theta$$

$\theta \rightarrow p(y=+1 \mid x, \theta_{MAP})$

$\theta_{MAP} = \underset{\theta}{\text{ARGMAX}}\ p(\theta \mid X, Y)$

$\equiv$ ML ASSUMPTION $= \dfrac{p(X, Y \mid \theta)\, p(\theta)}{p(X, Y \mid \theta)}$

$$p(Y \mid X) = \frac{p(X \mid Y)\, p(Y)}{p(X)}$$

$\underset{Y}{\text{ARGMAX}}\ p(Y \mid X) \equiv \underset{Y}{\text{ARGMAX}}\ p(X \mid Y) p(Y)$

DISC GEN

SVM