$x_1 \ldots x_k$ (x_i, y_i)

$x_i \in \mathbb{R}^n$ $y_i \in \mathbb{R}$

$$\hat{f}(x_i) = y_i + \xi$$

$$\theta^T x_i = \overbrace{y_i + \xi}^{\hat{y}_i}$$

1) $L(y_i, \hat{y}_i)$

$\|y_i - \hat{y}_i\|_2$

$\|y_i - \theta^T x_i\|_2$

2) $\theta_0 \sim N(0, 1)$

3) $\theta_{n+1} = \theta_n - \alpha \cdot \nabla_\theta L(y_i, \hat{y}_i)$

Частотный подход:

$x_1, x_2 \ldots x_n$ $P(X|\theta)$

$$\theta = \arg\max_\theta P(x_1, x_2 \ldots x_n | \theta)$$

i.i.d.

$$\theta \arg\max_\theta \prod_{i=1}^{n} \overbrace{P(x_i|\theta)}^{\text{likelihood}} =$$

$$= \arg\max_{\Theta} \sum_{i=1}^{n} \log P(x_i | \theta)$$

H T
5 3

$$P(X|\theta) = \theta^{x}(1-\theta)^{(1-x)}$$

$$\hat{\theta} = \arg\max_{\theta} \underbrace{\sum_{i=1}^{n} \log\left(\theta^{x}(1-\theta)^{(1-x)}\right)}_{} =$$

$$= \sum_{i=1}^{n} x_i \log\theta + (1-x_i)\log(1-\theta) =$$

$$= \sum_{i=1}^{n} \left[x_i \log\theta - x_i \log(1-\theta)\right] + n\log(1-\theta) =$$

$$\frac{\partial}{\partial\theta}$$

$$\sum_{i=1}^{n}\left[\frac{x_i}{\theta} + \frac{x_i}{1-\theta}\right] - \frac{n}{1-\theta} = 0$$

$$\sum_{i=1}^{n} \frac{x_i}{\theta-\theta^2} - \frac{\theta n}{\theta-\theta^2}$$

$$\hat{\Theta} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \frac{1}{8}\cdot 5 = \frac{5}{8}$$

H T

8 0.

$\hat{\theta} = 1.$

likelihood.

prior.

$$P(\theta|x) = \frac{P(x|\theta)\,P(\theta)}{\underbrace{P(x)}}$$

evidence

aposteriori.

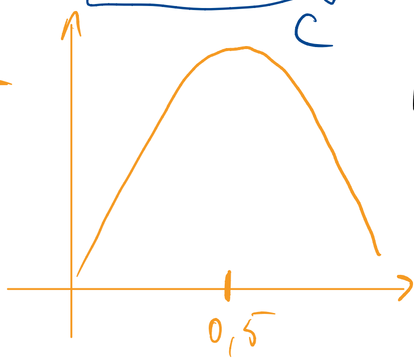
x_i i.i.d.

$$\hat{\theta} = \arg\max_{\theta} \prod_{i=1}^{n} P(x_i|\theta)\,P(\theta).$$

MAP.

$$P(\theta|\alpha,\beta) = \underbrace{\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)}}_{C}\,\theta^{\alpha-1}(1-\theta)^{\beta-1}$$

$\alpha = \beta = 3$



$$P(\theta) = C\cdot\theta^{2}(1-\theta)^{2}.$$

$$\hat{\theta} = \arg\max_{\theta}\left[\prod_{i=1}^{n} P(x_i|\theta)\right]\cdot P(\theta) =$$

$$= \sum_{i=1}^{n} \log\left(\theta^{x}(1-\theta)^{(1-x)}\right) + \log \theta^{2}(1-\theta)^{2} =$$

$$= \left[\sum_{i=1}^{n} x \log \theta + (1-x)\log(1-\theta)\right] + 2\log\theta +$$

$$+ 2\log(1-\theta)$$

$$\frac{\partial}{\partial \theta}$$

$$\|$$

$$\left[\sum_{i=1}^{n} \frac{x_i}{\theta-\theta^2}\right] - \frac{\theta n}{\theta-\theta^2} + \frac{2}{\theta}^{(1-\theta} - \frac{2}{1-\theta}^{(\theta} =$$

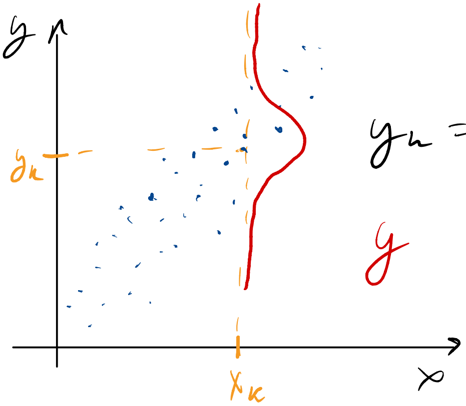
$$= \sum_{i=1}^{n} \frac{x_i}{\theta-\theta^2} - \frac{\theta n - 2 + 4\theta}{\theta-\theta^2}$$

$$\sum_{i=1}^{n} x_i - \theta(n+4) + 2$$

$$\theta = \frac{1}{n+4}\left[\left(\sum_{i=1}^{n} x_i\right) + 2\right]$$

$$\boxed{\frac{10}{12}}$$

$x_i \quad y_i$



$$y_k = \hat{f}(x_k, \theta) + \xi \qquad \xi \sim N(0, \sigma)$$

$$y \sim N(\hat{f}(x, \theta), \sigma)$$

$$P(y \mid \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$$

$$\hat{\theta} = \arg\max \sum_{i=1}^{n} \left[\log \frac{1}{\sqrt{2\pi\sigma^2}} - \frac{1}{2\sigma^2} \left(y_i - \hat{f}(x_i, \theta)\right)^2 \right]$$

$$= \arg\max_{\theta} -\frac{1}{2\sigma^2} \sum_{i=1}^{n} \left[y_i - \hat{f}(x_i, \theta) \right]^2 + n \log \frac{1}{\sqrt{2\pi\sigma^2}}$$

MSE

$$\hat{\theta} = \arg\max_{\theta} \prod_{i=1}^{n} P(x_i \mid \theta) \cdot P(\theta)$$

$$\theta \sim N(0, 1 \cdot \tfrac{1}{2})$$

$$\frac{\sqrt{\alpha}}{\sqrt{(2\pi)^m \sigma^2}} \cdot e^{-\frac{1}{2\sigma^2}(y_i - \theta^T x_i)^2 - \frac{\alpha}{2}\theta^T\theta}$$

$$= \arg\max_\theta -\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \theta^T x_i)^2 - \frac{\alpha}{2}\theta^T\theta$$

MSE L2 Reg.

$$P(x|\eta) = h(x)g(\eta)\cdot e^{\eta^T u(x)}$$

$$P(x|\mu) = \mu^x (1-\mu)^{(1-x)}$$

$$P(x|\mu) = e^{x \ln\mu + (1-x)\ln(1-\mu)} =$$

$$= (1-\mu)\, e^{\ln\left(\frac{\mu}{1-\mu}\right)x}$$

$\sigma(x)$

$$\eta = \ln\frac{\mu}{1-\mu} \qquad \mu = \frac{1}{1+\exp(-\eta)}$$

$$P(x|\eta) = \sigma(-\eta)\, e^{\eta x}$$

$u(x) = x \qquad g(\eta) = \sigma(-\eta)$

$h(x) = 1.$

GLM.

$$y = \psi(\theta^T \varphi(x)).$$

activation. basis function.

ψ^{-1} – link function.

$f(x, \theta)$

$$h_1 = \psi_1(\theta_1^T \varphi_1(x)).$$

$$h_2 = \psi_2(\theta_2^T \varphi_2(h_1))$$

$\vdots$

$$y = \psi_k(\theta_k^T \phi_k(h_{k-1}))$$

MLP.

1) Локальные минимумы.
2) Седловые точки
3) Много параметров.