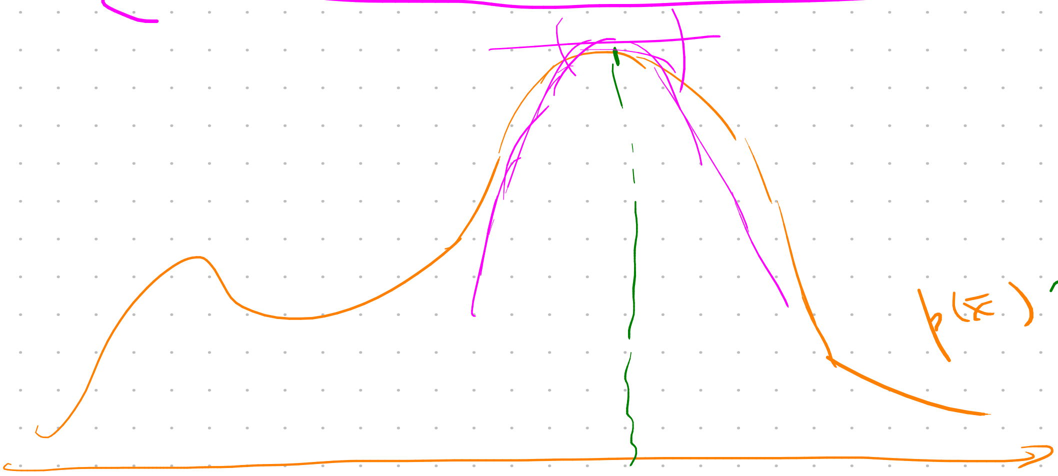
$$L = \sum_n (y - \bar{x}_n^T \bar{w})^2 + \underbrace{\alpha \cdot \sum_{i=1}^{d} |w_i|} \quad \Big\} \ \log p(\bar{w}|D)$$

$$p(y|\bar{x}_i, \bar{w}) = \mathcal{N}(y|\bar{w}^T\bar{x}, \sigma^2)$$

$$p(\bar{w}) \propto e^{-\beta \sum_i |w_i|}$$

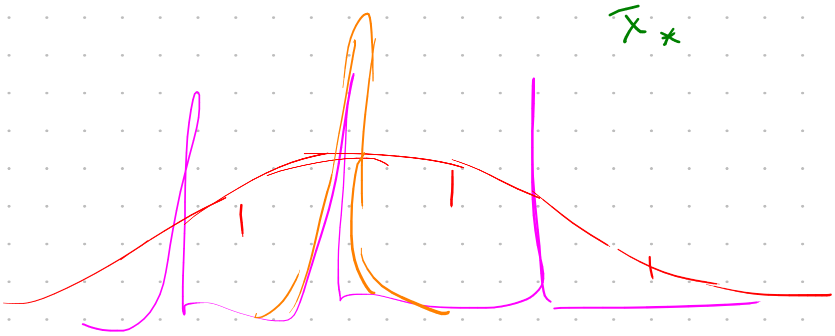
$$\boxed{p(y|\bar{x}, D) = \int p(y|\bar{x}, \bar{w}) \overbrace{p(\bar{w}|D)}^{\approx q(\bar{w})} d\bar{w}}$$

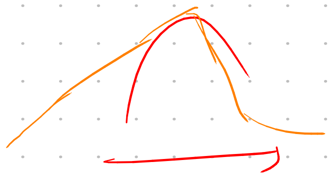


$\bar{x}_*$

$p(\bar{x}) \approx q(\bar{x})$

$\log p(\bar{x}) \approx$





$$\log p(\bar{x}) \approx \log p(\bar{x}_0)$$

$$\int \sigma(\bar{w}^T\bar{x})\, q(\bar{w})\, d\bar{w} = \int \sigma(a) \cdot \overbrace{\left[\int \delta(a - \bar{w}^T\bar{x})\, q(\bar{w})\, d\bar{w} \right]}^{p(a) = \mathcal{N}(a|\mu_a, \sigma_a^2)}$$

$$\mu_a = E[a] = \int a\, p(a)\, da = \iint \delta(a - \bar{w}^T\bar{x})\, a\, q(\bar{w})\, da\, d\bar{w} =$$

$$= \int (\bar{w}^T\bar{x}) \cdot q(\bar{w})\, d\bar{w} = E_{q(\bar{w})}[\bar{w}^T\bar{x}] = \bar{w}_{MAP}^T \bar{x}$$

$$q(\bar{w}) = \mathcal{N}(\bar{w}|\bar{w}_{MAP}, \Sigma_N)$$

$$\sigma_a^2 = \int (a-\mathbb{E}[a])^2 p(a)\,da = \int q(\bar w)\left(\int \delta(a-\bar w^T\bar x)\cdot(a-\mathbb{E}[a])^2\,da\right) d\bar w$$

$$= \int \left(\bar w^T \bar x - \bar w_{MAP}^T \bar x\right)^2 q(\bar w)\,d\bar w =$$

$$= \int (\bar w - \bar w_{MAP})^T \bar x\, \bar x^T (\bar w - \bar w_{MAP})\, q(\bar w)\,d\bar w$$

$$= \underline{\int \bar x^T (\bar w - \bar w_{MAP})(\bar w - \bar w_{MAP})^T \bar x\; q(\bar w)\,d\bar w} =$$

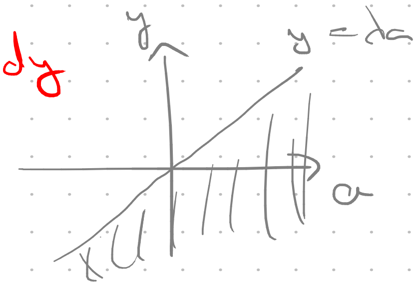
$$\int (\bar w - \bar\mu)(\bar w - \bar\mu)^T \mathcal{N}(\bar w|\bar\mu,\Sigma)\,d\bar w \;\; \text{"} \;\Sigma$$

$$= \bar x^T \Sigma_N \bar x$$

$$\bar w_{MAP}^T \bar x \qquad \bar x^T \Sigma_N \bar x$$

$$\int \sigma(\bar w^T \bar x)\, q(\bar w)\,d\bar w = \int \sigma(a)\, \mathcal{N}(a|\mu_a, \sigma_a^2)\,da =$$

$$\sigma(a) \approx \Phi(\lambda a) = \int_{-\infty}^{\lambda a} \mathcal{N}(y|0,1)\,dy$$



$$\approx \int_{-\infty}^{\infty}\int_{-\infty}^{\lambda a} \underbrace{\mathcal{N}(y|0,1)\,\mathcal{N}(a|\mu_a,\sigma_a^2)}_{-\frac{y^2}{2} - \frac{1}{2\sigma_a^2}(a-\mu_a)^2}\,dy\,da =$$

$$= \int \int \mathcal{N}(a|\mu_a,\sigma_a^2)\,\mathcal{N}(y|0,1)\,da\,dy$$

$$\frac{(a-\mu_a)}{\sigma_a} = z \qquad \lambda a = \lambda\frac{z}{\sigma_a} + \mu_a\lambda$$

$$\int \Phi(\lambda a)\, \mathcal{N}(a|\mu_a,\sigma_a^2)\,da$$

$$A \sim \mathcal{N}(a|\mu_a,\sigma_a^2)$$

$$= \int \Phi\left(\frac{\lambda}{\sigma_a}(z+\mu_a\sigma_a)\right) \mathcal{N}(z|0,1)\,dz \qquad B \sim \mathcal{N}(b|0,1)$$

$$\Pr[A \le b \mid B=b] = \underline{\Pr[A \le b]} = \int_{-\infty}^{b} \mathcal{N}(a|\mu_a,\sigma_a^2)\,da = \Phi\left(\frac{b-\mu_a}{\sigma_a}\right)$$

$$\Pr[A \le B] = \int_{-\infty}^{\infty} \Pr[A \le B \mid B = b]\, p(b)\, db = \qquad A \sim \mathcal{N}(a \mid \mu, \sigma^2) \quad B \sim \mathcal{N}(b \mid 0, 1)$$

$$= \int_{-\infty}^{\infty} \Phi\left(\frac{b-\mu}{\sigma}\right) \mathcal{N}(b \mid 0, 1)\, db$$

$$\Pr[A - B \le 0] = \int_{-\infty}^{0} \mathcal{N}(c \mid \mu_a, \sigma_a^2 + 1)\, dc = \Phi\left(\frac{-\mu_a}{\sqrt{\sigma_a^2 + 1}}\right)$$

$$A - B \sim \mathcal{N}(c \mid \mu, \sigma^2 + 1)$$

$$\boxed{\Phi\left(-\frac{\mu}{\sqrt{\sigma^2+1}}\right) = \int \Phi\left(\frac{b-\mu}{\sigma}\right) \mathcal{N}(b \mid 0, 1)\, db}$$

$$\int \Phi(\lambda a)\, \mathcal{N}(a \mid \mu_a, \sigma_a^2)\, da = \left[z = \frac{a - \mu_a}{\sigma_a}\right] =$$

$$= \int \Phi\big((\sigma_a \cdot z + \mu_a)\lambda\big)\, \mathcal{N}(z \mid 0, 1)\, dz$$

$$\frac{z - (-\mu_a/\sigma_a)}{1/\sigma_a\lambda} \qquad \mu = -\frac{\mu_a}{\sigma_a} \qquad \sigma = 1/\sigma_a\lambda$$

$$-\frac{\mu}{\sqrt{\sigma^2+1}} = \frac{\mu_a/\sigma_a}{\sqrt{1 + 1/\sigma_a^2\lambda^2}} = \frac{\lambda \mu_a}{\sqrt{\sigma_a^2\lambda^2 + 1}}$$

hyp.-perp.:

$$p(D|\bar{w}) = \prod_n p(y_n|\bar{w}^T\bar{x}_n, \sigma^2)$$

$$\beta = 1/\sigma^2$$

$$p(\bar{w}) = \mathcal{N}(\bar{w}|\bar{0}, \alpha I)$$

$$p(D|\alpha,\beta) \xrightarrow[\alpha,\beta]{} \max$$

$$\|$$

$$\int p(\bar{w}|\alpha)\, p(D|\bar{w},\beta)\, d\bar{w}$$

Optimal Bayes classifier

C_1 $\quad$ C_2

$p(\bar{x}|C_1)$ $\quad$ $p(\bar{x}|C_2)$

$p(\bar{x}|C_k)$ — $p(\bar{x}|C_k)$

$$p(C_k|\bar{x}) = \frac{p(C_k)\,p(\bar{x}|C_k)}{\sum_l p(C_l)\,p(\bar{x}|C_l)}$$

$$\approx$$

$$\text{softmax}(\bar{w}_k^T\bar{x})$$

Naive Bayes

$$p(\bar{x},y) = p(y)\cdot\prod_{i=1}^{d} p(x_i|y)$$

naive

$$y \sim w_0 + w_1x + w_2x^2$$

$$\dot{\bar{x}} \overset{?}{=} A\bar{x}$$

$$\dot{\bar{x}} = (A+\lambda I)\bar{x}$$

$$\bar{x} = A - \lambda \to 0$$