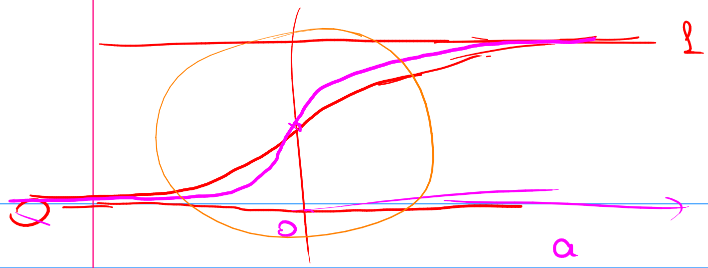
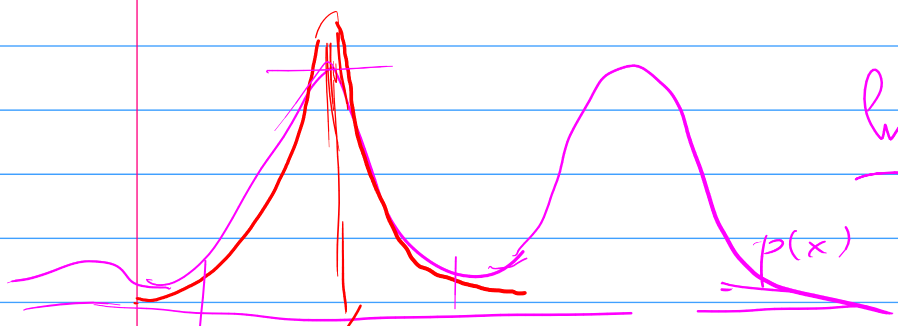


$p(x)$

$F(x) = \int_{-\infty}^{x} p(y)\,dy$

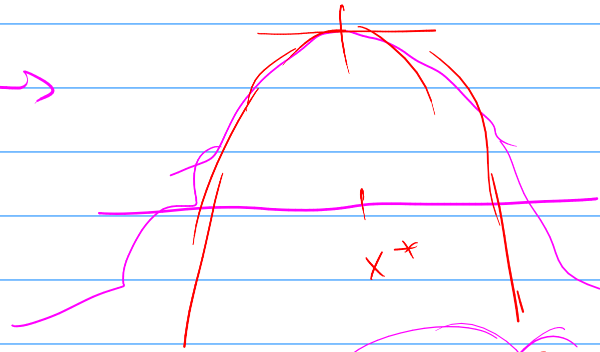
$$\Phi(a) = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,dx$$

probit



$\mu = x^*$

$\ln \longrightarrow$



$$\ln p(x) \approx \ln p(x^*) + (x-x^*) \ln p(x^*)' + \frac{1}{2}(x-x^*)^2 \frac{\partial^2 \ln}{\partial x^2}$$

Laplace approximation

$$-\frac{a}{2}(x-x^*)^2 + c$$

$$p(x) \approx c \cdot e^{-\frac{a}{2}(x-x^*)^2}$$

$$-\nabla\nabla \ln p(\bar{x})$$

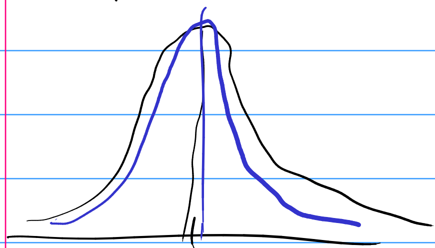
$$\ln p(\bar{x}) \approx \ln p(\bar{x}^*) - \frac{1}{2}(\bar{x}-\bar{x}^*)^T A (\bar{x}-\bar{x}^*)$$

$$p(\bar{x}) \approx p(\bar{x}^*) \cdot e^{-\frac{1}{2}(\bar{x}-\bar{x}^*)^T A (\bar{x}-\bar{x}^*)}$$

$M_1 \ldots M_k \qquad D$

$$p(M_i|D) \propto p(M_i)\, p(D|M_i) = \int p(D|\theta, M_i)\, p(\theta|M_i)\, d\theta$$

$$\approx p(D|\theta_{MAP}, M_i) \cdot p(\theta_{MAP}|M_i) \cdot \frac{(\ )}{\sqrt{\det}} \quad \frac{(2\pi)^{M/2}}{\sqrt{\det A}} \cdot e^{-\frac{1}{2}(\ldots)^T A (\ldots)}$$



θ_{MAP}

$$p(M_i|D) \propto p(M_i)\cdot p(D|\theta_{MAP}, M_i)\cdot p(\theta_{MAP}|M_i)\cdot \frac{(2\pi)^{M/2}}{\sqrt{\det A}}$$

$$\ln p(M_i|D) = \text{Const} + \ln p(M_i) + \ln p(D|\theta_{MAP}, M_i) +$$

Occam's factor

$$\boxed{+ \ln p(\theta_{MAP}|M_i) + \frac{M}{2}\ln 2\pi - \frac{1}{2}\ln\det A \quad \text{где } A = -\nabla\nabla \ln p(\theta_{MAP}|D)}$$

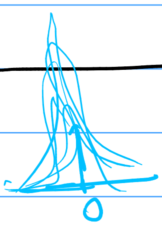
$$p(C_1|\bar x) \approx \sigma(\bar w^T \bar x) \qquad \begin{cases} p(\bar t|\bar w, X) \to \max \\ p(\bar w) = \mathcal{N}(\bar w|\bar\mu_0, \Sigma_0) \end{cases}$$

$$\ln p(\bar w|D) \overset{C_T}{=} \sum_n \left[t_n \ln\sigma_n + (1-t_n)\ln(1-\sigma_n)\right] - \frac{1}{2}(\bar w-\bar\mu_0)^T\Sigma_0^{-1}(\bar w-\bar\mu_0)$$

$$\nabla \ln p(\bar w|D) = \sum_n \left[t_n-\sigma_n\right]\bar x_n - \Sigma_0^{-1}(\bar w-\bar\mu_0)$$

$$\Sigma_N = -\nabla\nabla \ln p = \_\_\_ - \Sigma_0^{-1}$$

$$\approx \mathcal{N}(\bar w|\bar w_{MAP}, \Sigma_N)$$



$$p(C_1|\bar x, D) = \int \underbrace{p(C_1|\bar x, \bar w)}_{= \sigma(\bar w^T \bar x)} p(\bar w|D)\, d\bar w \approx$$

$$p(\bar w|D) \propto \mathcal{N}(\bar w|\mu_0, \Sigma_0)\cdot \prod_n \sigma(\bar w^T \bar x)^{t_n}(1-\sigma)^{1-t_n}$$

$$\boxed{\int \delta(x-a) f(x)\, dx = f(a)}$$

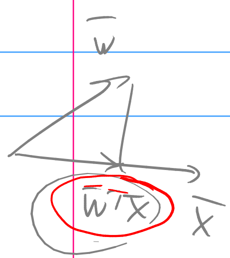
$$\int \delta(x)\, dx = 1$$

$$\Sigma_N = \Sigma_0^{-1} + \sum_{n=1}^{N} \sigma_n(1-\sigma_n)\bar x_n \bar x_n^T$$

$$\sigma(\bar w^T\bar x) = \int \delta(a-\bar w^T\bar x)\sigma(a)\, da$$

$$a = \bar x^T \bar w$$

$$\approx \int \underbrace{\sigma(\bar x^T \bar w)}_{a} \mathcal{N}(\bar w|\bar w_{MAP}, \Sigma_N)\, d\bar w =$$

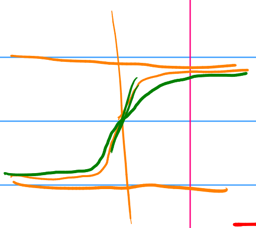


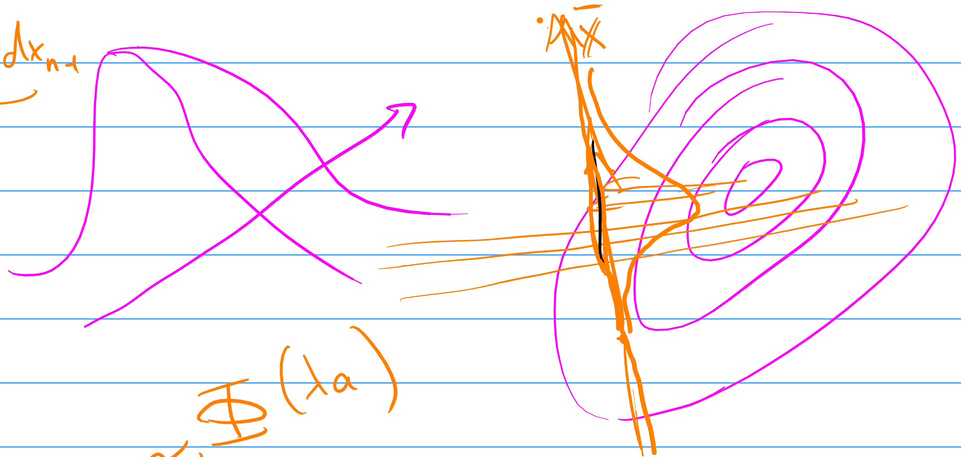
$$= \iint \underbrace{\delta(a-\bar w^T\bar x)}_{\mathcal{N}(a|\bar w^T\bar x, \varepsilon^2)}\sigma(a)\, \mathcal{N}(\bar w|\bar w_{MAP}, \Sigma_N)\, da\, d\bar w$$

$$= \int \sigma(a) \underbrace{\left[\int \overbrace{\delta(a - \bar{w}^T \bar{x})}^{N(a|\_)} \, N(\bar{w} | \bar{w}_{MAP}, \Sigma_N) \, d\bar{w} \right]}_{= N(a|\mu_a, \sigma_a^2)} da$$

$$\int p(\bar{x}) \, dx_1 dx_2 \ldots dx_{n+1} = \boxed{p(x_n)}$$

Σ





$$\approx \Phi(\lambda a)$$

$$= \int \sigma(a) \, N(a|\mu_a, \sigma_a^2) \, da \approx$$

$$\sigma(a) \approx \Phi\left(\underbrace{\sqrt{\tfrac{\pi}{8}}}_{= \lambda} \cdot a\right)$$

$$\approx \int \Phi(\lambda a) \, N(a|\mu_a, \sigma_a^2) \, da =$$

$$= \int \int_{-\infty}^{\lambda a} N(z|0,1) \, dz \cdot N(a|\mu_a, \sigma_a^2) \, da$$

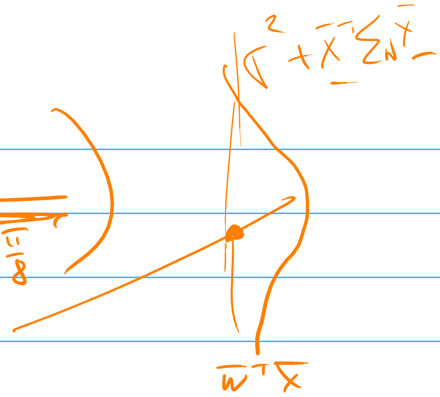
$$= \int_{-\infty}^{\ldots} \left(\underbrace{\int N(z|\_) \, N(a|\_) \, da}_{N(z|\_)} \right) dz$$

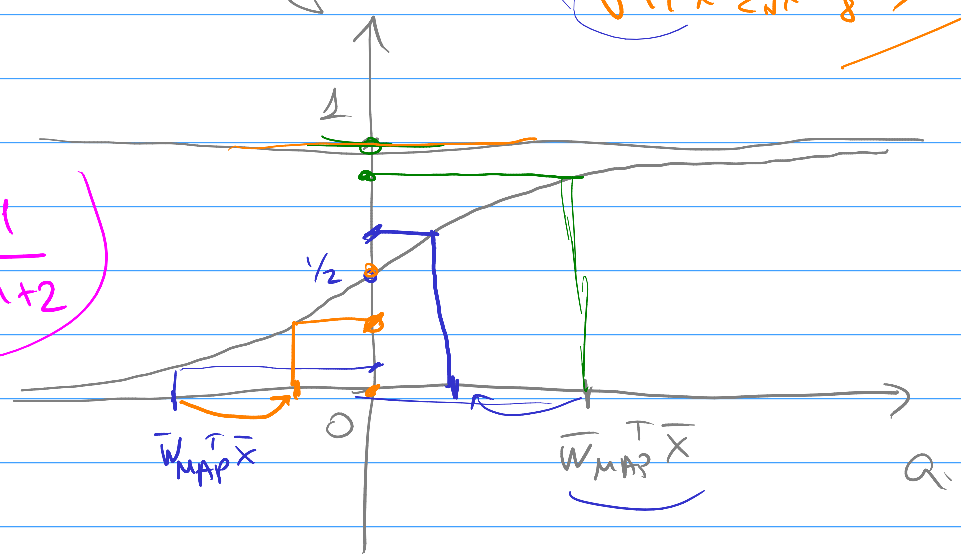
$$= \Phi\left(\frac{\mu_a}{\sqrt{\frac{1}{\lambda^2} + \sigma_a^2}} \right) \approx \sigma\left(\frac{1}{\lambda} \ldots \right)$$

$$\boxed{p(C_1 | \bar{x}, D) \approx \sigma\left(\underbrace{\mu_a}_{= \bar{w}_{MAP}^T \bar{x}} \cdot \frac{1}{\sqrt{1 + \frac{\pi}{8} \underbrace{\sigma_a^2}_{= \bar{x}^T \Sigma_N \bar{x}}}} \right)}$$

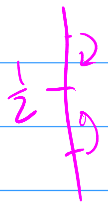
$$p(C_1|\bar{x}, \bar{w}_{MAP}) = \sigma(\bar{w}_{MAP}^T \bar{x})$$

$$p(C_1|\bar{x}, D) = \sigma\left(\bar{w}_{MAP}^T \bar{x} \cdot \frac{1}{\sqrt{1+\bar{x}^T \Sigma_N \bar{x} \cdot \frac{\pi}{8}}}\right)$$



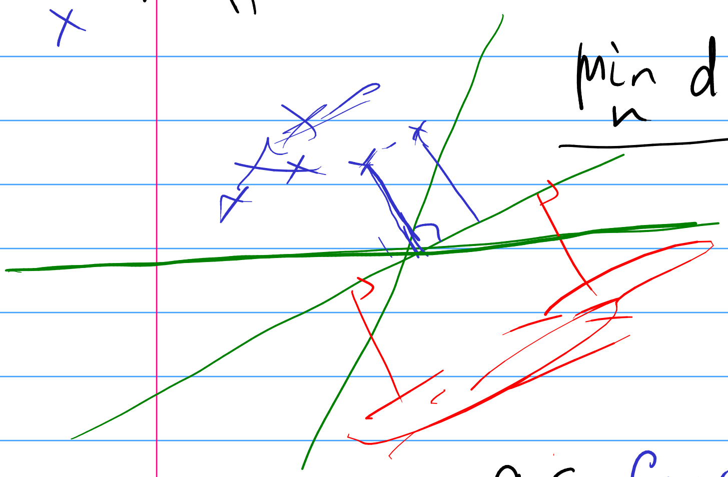


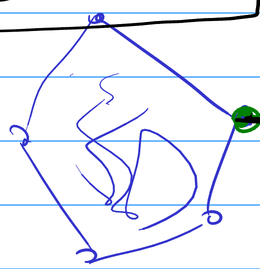
$$p(x|D) = \frac{n+1}{n+m+2}$$

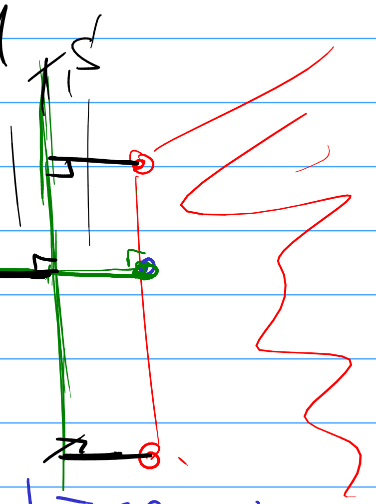


Support vector machines SVM

$$\min_n d(\bar{x}_n, S) \longrightarrow \max$$







$$a \in Conv(C_1) = \left\{ \sum \alpha_n \bar{x}_n \,\middle|\, \bar{x}_n \in C_1, \sum \alpha_n = 1, \alpha_n \geq 0 \right\}$$

$$b \in Conv(C_2) = \left\{ \sum \alpha_n \bar{x}_n \,\middle|\, \bar{x}_n \in C_2, \sum \alpha_n = 1, \alpha_n \geq 0 \right\}$$

$$\|a-b\|^2 \longrightarrow \min$$

$$\left\| \sum_{\bar{x}_n \in C_1} \alpha_n \bar{x}_n - \sum_{\bar{x}_m \in C_2} \beta_m \bar{x}_m \right\|^2 \longrightarrow \min$$

$$\sum \alpha_n = 1, \quad \sum \beta_m = 1$$

$$\alpha_n \geq 0 \qquad \beta_m \geq 0$$

Linear programming

$$\bar{c}^T \bar{x} \xrightarrow[\bar{x}]{} \min$$

$$A\bar{x} = \bar{b}$$

$$D\bar{x} \geqslant \bar{e}$$

quadratic programming

$$\bar{c}^T \bar{x} + \bar{x}^T H \bar{x} \xrightarrow[\bar{x}]{} \min$$

$$A\bar{x} = \bar{b}, \quad D\bar{x} \geqslant \bar{e}$$