$$\ln p(\bar{z}) \approx \ln p(\bar{z}_0) - \frac{1}{2}(\bar{z}-\bar{z}_0)^T A (\bar{z}-\bar{z}_0)$$

$$\int \delta(\bar{w}^T\bar{x}-a)\,\sigma(a)\,da \qquad A = -\nabla\nabla \ln p$$

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

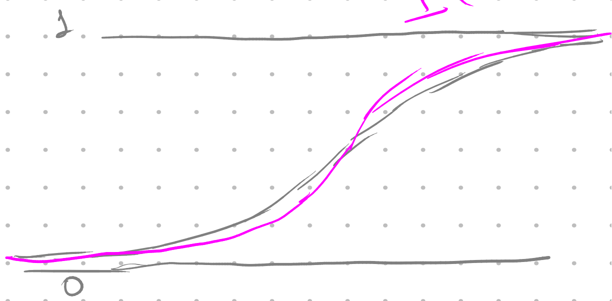
$a = \bar{w}^T\bar{x}$

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

$$\approx \int \sigma(a) \left(\int \delta(\bar{w}^T\bar{x}-a)\,\mathcal{N}(\bar{w}|\bar{w}_{MAP}, \Sigma_N)\, d\bar{w} \right) da =$$

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

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



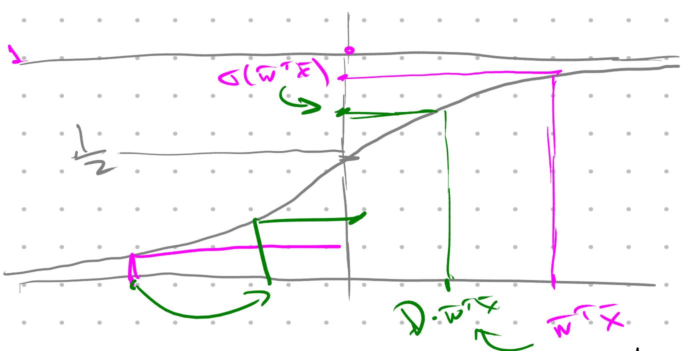
$$= \int \sigma(a)\,\mathcal{N}(a|\mu_a, \sigma_a^2)\, da \approx$$

$$\approx \int \int_{-\infty}^{a} \mathcal{N}(\lambda x)\, dx \cdot \mathcal{N}(a|\mu_a, \sigma_a^2)\, da$$

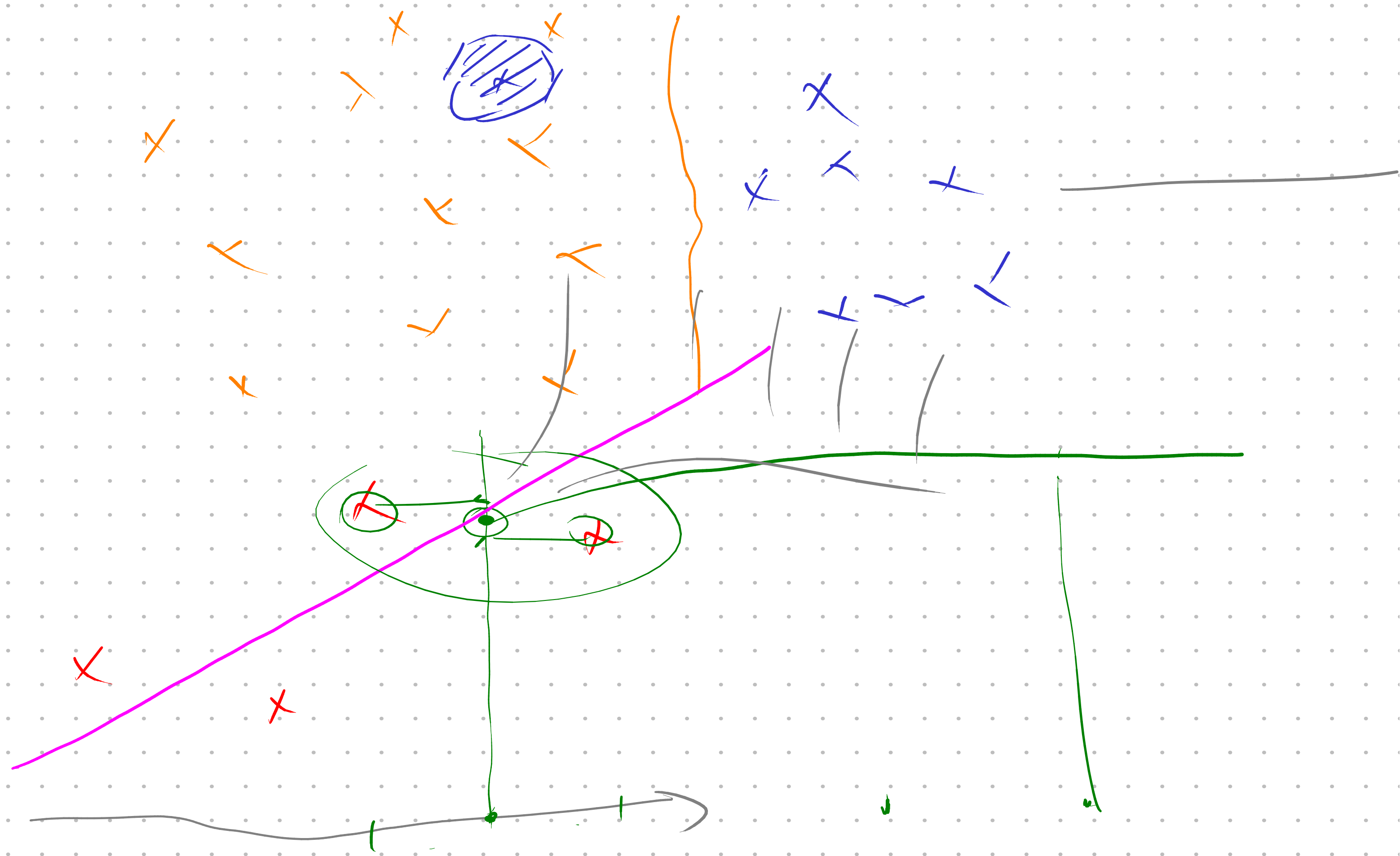
$\bar{w}_{MAP}^T \bar{x}$ =

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

$$\kappa = \frac{1}{\sqrt{1 + \frac{\pi}{8}\sigma^2}}$$



$\boxed{k}$-NN

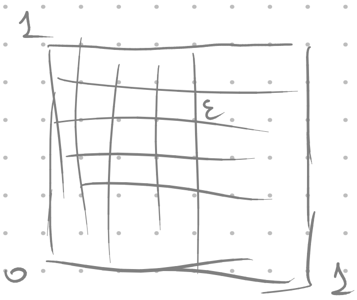


Curse of dimensionality

(1)

$1/\varepsilon$

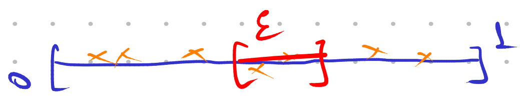
$$\int_0^1 f(x)\,dx \approx \sum$$



$1/\varepsilon^2$

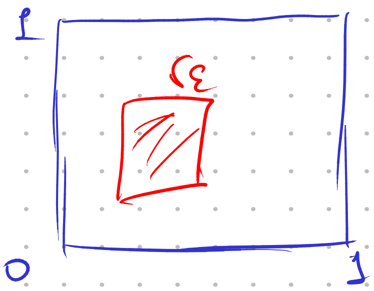
$1/\varepsilon^d$

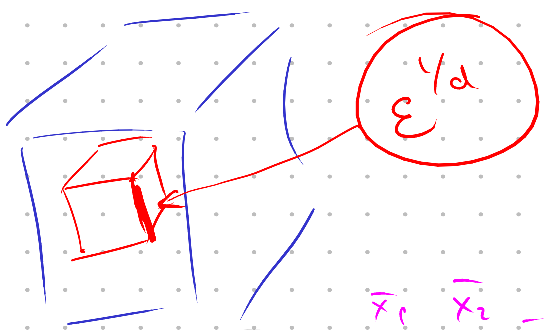
(2)



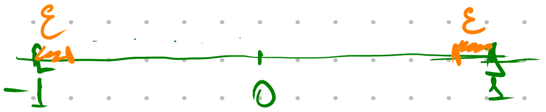
Δοια ε στ [illegible]

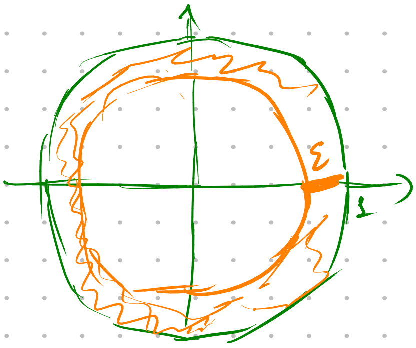
$$\sum_{i=1}^{d} (x_i - y_i)^2$$



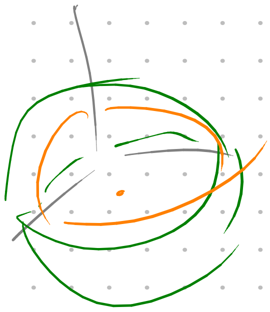


③ ε



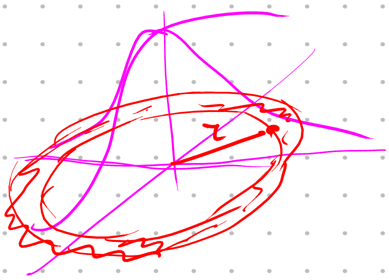


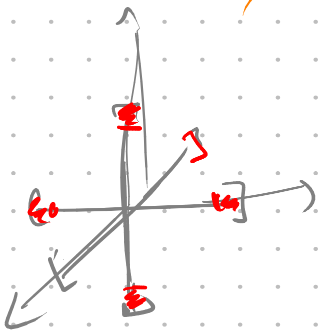
$$\frac{\pi r^2 - \pi((1-\varepsilon)r)^2}{\pi r^2} = 1-(1-\varepsilon)^2 = 2\varepsilon-\varepsilon^2$$

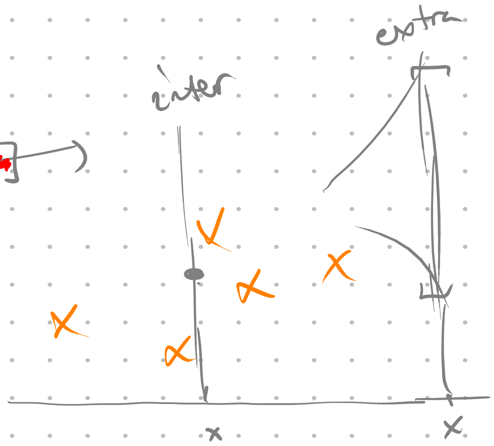


$$\frac{\cancel{\tfrac{4}{3}\pi}\, \cancel{r^3} - \cancel{\tfrac{4}{3}\pi}((1-\varepsilon)\cancel{r})^3}{\cancel{\tfrac{4}{3}\pi r^3}}$$

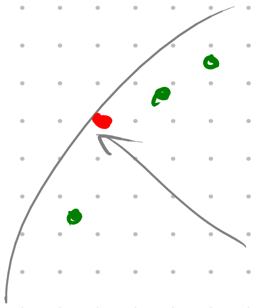
$$\boxed{1-(1-\varepsilon)^d} \xrightarrow[d\to\infty]{} 1$$







$$\bar{x} \sim \mathcal{N}(\bar{x} \mid \bar{0}, I)$$



d pag

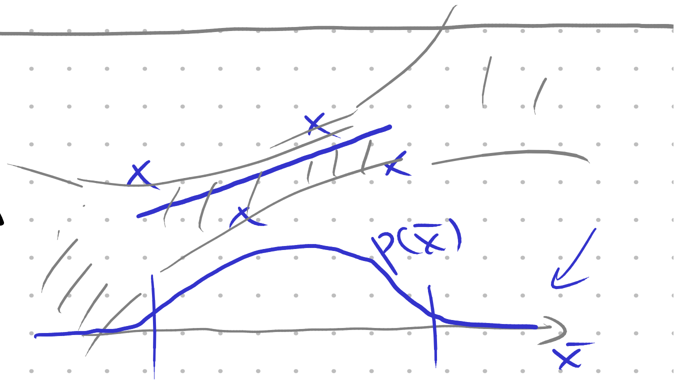
$$x_i \sim \mathcal{N}(x_i \mid 0, 1)$$

$$r^2(\bar{x}) = \sum^{d} x_i^2 \approx \mathcal{N}(r^2 \mid \_\_, \_\_)$$

Statistical decision theory

$\bar{x} \in \mathbb{R}^d, \; y \in \mathbb{R}$ $\qquad D = \{(\bar{x}, y)\}$

$p(y \mid \bar{x})$ $\qquad \boxed{p(\bar{x}, y)}$



$$L(y, f(\bar{x})) = (y - f(\bar{x}))^2$$

$$EPE[f] = \mathbb{E}_{p(\bar{x},y)}\left[L(y, f(\bar{x}))\right] = \underbrace{\iint (y - f(\bar{x}))^2 \, p(\bar{x}, y)\, d\bar{x}\, dy}$$

$f \to \min$

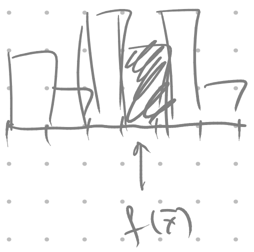
$$p(\bar{x}) \int (y - f(\bar{x}))^2 \, p(y|\bar{x}) \, dy \qquad \cancel{\int (x-a)^2 p(x) dx}$$

$$\boxed{\hat{f}(\bar{x}) = E_{p(y|\bar{x})}[y] = E[y|\bar{x}]}$$ ⇐ regression function

$$L(y, f(\bar{x})) = [y \neq f(\bar{x})] = \begin{cases} 0, & y = f(\bar{x}) \\ 1, & y \neq f(\bar{x}) \end{cases} \qquad y \in \{C_1, \dots, C_K\}$$

$$EPE[f] = \iint L(y, f(\bar{x})) \, p(\bar{x}, y) \, d\bar{x} \, dy =$$

$$= \int \left[\sum_{k=1}^{K} L(C_k, f(\bar{x})) \, p(C_k|\bar{x}) \right] p(\bar{x}) \, d\bar{x} \xrightarrow[f(\bar{x})]{} \min$$

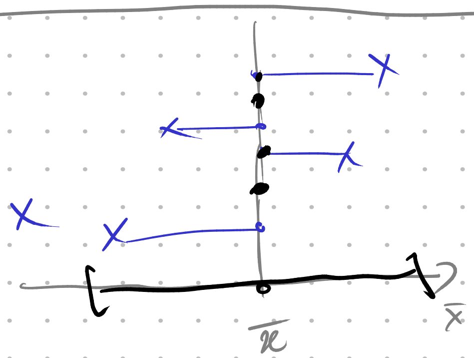


$$\boxed{\hat{f}(\bar{x}) = \arg\max_k \, p(C_k|\bar{x})}$$ optimal Bayes classifier

Test

Disease	pos	neg
pos	0	1000
neg	1	0

$$\hat{f}(\bar{x}) = \arg\max_k \left[L(C_k, f(x)) \, p(C_k|\bar{x}) \right]$$



$$\hat{f}(\bar{x}) = E[y|\bar{x}] \approx \frac{1}{R} \sum_{r=1}^{R} y_r \approx$$

где $y_r \sim p(y|\bar{x})$

kNN $\approx \frac{1}{R} \sum_r y_r$, где y_r [illegible] $\in NN(\bar{x})$

$$EPE[f] = \mathbb{E}[(y - f(\bar{x}))^2] = \iint (y - f(\bar{x}))^2 \, p(\bar{x}, y) \, dy \, d\bar{x} =$$

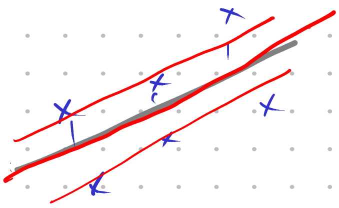
$$= \iint \left((y - \hat{f}(\bar{x})) + (\hat{f}(\bar{x}) - f(\bar{x})) \right)^2 p(\bar{x}, y) \, dy \, d\bar{x} =$$

$$= \int (\hat{f} - f) \, p(\bar{x}) \int (y - \hat{f}) \, p(y|\bar{x}) \, dy$$

$$= \mathbb{E}[(y - f(\bar{x}))^2] + \cancel{2\mathbb{E}[(y - \hat{f}(\bar{x}))(\hat{f}(\bar{x}) - f(\bar{x}))]}$$

$$+ \mathbb{E}[(\hat{f}(\bar{x}) - f(\bar{x}))^2]$$

$$EPE[f] = \underbrace{\mathbb{E}[(y - \hat{f}(x))^2]}_{\text{Noise}} + \underbrace{\mathbb{E}[(\hat{f}(\bar{x}) - f(\bar{x}))^2]}_{\to \min}$$



$$f(\bar{x}; D) \qquad D \sim p(\bar{x}, y)$$

$$\mathbb{E}[(\hat{f} - f)^2] = \mathbb{E}[(\hat{f} - \mathbb{E}_D f + \mathbb{E}_D f - f)^2] =$$

$$\mathbb{E}_D f - f(\bar{x})$$

$$= \mathbb{E}[(\hat{f} - \mathbb{E}_D f)^2] + \mathbb{E}[(f - \mathbb{E}_D f)^2] + \underbrace{\cancel{2\mathbb{E}[(\hat{f} - \mathbb{E}_D f)(\mathbb{E}_D f - f)]}}_{= 0}$$

$$EPE[f] = \mathbb{E}[(\hat{f}(\bar{x}) - \mathbb{E}_D f(\bar{x}; D))^2] \qquad \text{Bias}$$

$$+ \mathbb{E}[(f(\bar{x}; D) - \mathbb{E}_D f(\bar{x}; D))^2] \qquad \text{Variance}$$

$$+ \mathbb{E}[(y - \hat{f}(\bar{x}))^2] \qquad \text{Noise}$$