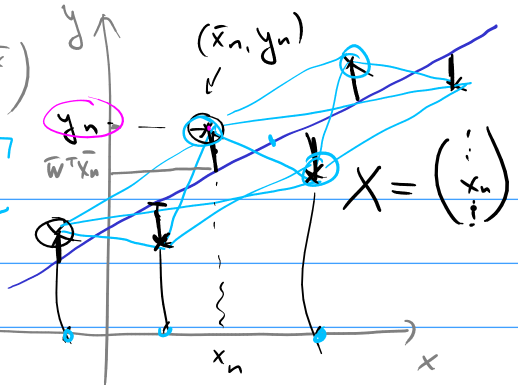
$n = 1..N$

$x_1, x_2, \dots, x_m \longrightarrow y$

$\bar{w}^T \bar{x} = \begin{pmatrix} \bar{x} \\ 1 \end{pmatrix}$

$\boxed{D = \{(\bar{x}_n, y_n)\}_{n=1}^{N}}$

$\boxed{y_n \approx w_1 x_{n1} + w_2 x_{n2} + \dots + w_m x_{nm} + w_0}$



$X = \begin{pmatrix} \vdots \\ x_n \\ \vdots \end{pmatrix}$

МНК / least squares

$$\min_{\bar{w}} L(D, \bar{w}) = \min_{\bar{w}} \sum_{n=1}^{N} (y_n - \bar{w}^T \bar{x}_n)^2$$

$\bar{y} = \begin{pmatrix} \vdots \\ y_n \\ \vdots \end{pmatrix} N\times 1$

$X = \begin{pmatrix} & \vdots & \\ - & \bar{x}_n & - \\ & \vdots & \end{pmatrix} N$, M — #точек

$$\nabla_{\bar{w}} \sum_{n=1}^{N} (y_n - \bar{w}^T \bar{x}_n)^2 =$$

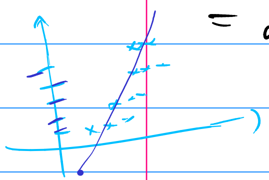
$$= \nabla_{\bar{w}} \left((\bar{y} - X\bar{w})^T (\bar{y} - X\bar{w}) \right) =$$

$$= \nabla_{\bar{w}} \left(\bar{y}^T \bar{y} - \underbrace{\bar{w}^T X^T \bar{y}}_{X^T y} - \underbrace{\bar{y}^T X \bar{w}}_{X^T y} + \underbrace{\bar{w}^T X^T X \bar{w}}_{2X^T X \bar{w}} \right)$$

$$= 2X^T X \bar{w} - 2X^T \bar{y} = 0$$

$$X^T X \bar{w} = X^T \bar{y}$$

$$\boxed{\bar{w}^* = (X^T X)^{-1} X^T \bar{y}}$$



$\nabla_{\bar{w}} (\bar{w}^T \bar{a}) = \bar{a}$ $\quad \nabla_{\bar{w}} f = \left(\frac{\partial f}{\partial w_i} \right)$

$\bar{w}^T \bar{a} = \sum w_i a_i$

$\nabla_{\bar{w}} (\bar{w}^T \bar{w}) = 2\bar{w} = \bar{w} + \bar{w}$

$\nabla_{\bar{w}} (\bar{w}^T A \bar{w}) = (A + A^T) \bar{w}$

$$\frac{\partial \left(\sum_{i,j} w_i w_j a_{ij} \right)}{\partial w_k} =$$

$$= \frac{\partial \left(\sum_{j \neq k} w_k w_j a_{kj} + \sum_{i \neq k} w_i w_k a_{ik} + w_k^2 a_{kk} \right)}{\partial w_k}$$

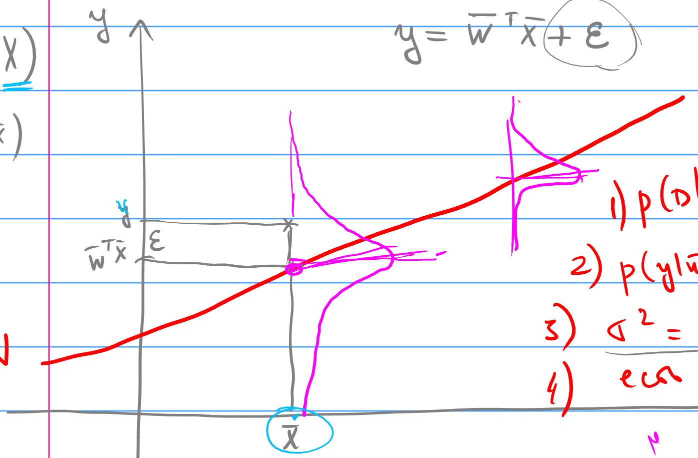
$$= \sum_{j \neq k} w_j a_{kj} + \sum_{i \neq k} w_i a_{ik} + 2 w_k a_{kk}$$

$$= \underbrace{\sum_j w_j a_{kj}}_{(A\bar{w})_k} + \underbrace{\sum_i w_i a_{ik}}_{(A^T \bar{w})_k}$$

$p(\bar{y} | \bar{w}, X)$

$p(y | \bar{w}, \bar{x})$

$y = \bar{w}^T \bar{x} + \varepsilon$



1) $p(D|\theta) = \prod_{n=1}^{N} p(d_n|\theta)$

2) $p(y|\bar{w}, \bar{x})$ нормально

3) $\sigma^2 = const$

4) если прямая $\bar{w}^T \bar{x}$

$\bar{y} = \begin{pmatrix} \vdots \\ y_n \\ \vdots \end{pmatrix} N$

$p(D|\theta)$

$$p(y | \bar{w}, \bar{x}) = \mathcal{N}(y \mid \bar{w}^T \bar{x}, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2} (y - \bar{w}^T \bar{x})^2}$$

$$p(\bar{y} | \bar{w}, X) = \prod_{n=1}^{N} p(y_n | \bar{w}, \bar{x}_n) = \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2} (y_n - \bar{w}^T \bar{x}_n)^2} \xrightarrow[\bar{w}]{} \max$$

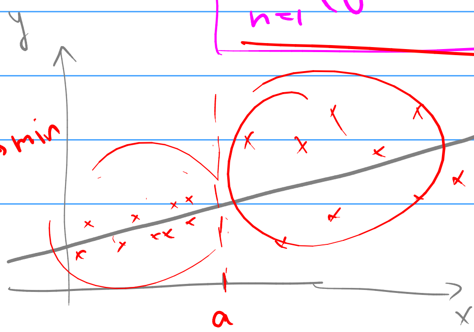
$$\ln p(\bar{y} | \bar{w}, X) = -\frac{N}{2} \ln(2\pi\sigma^2) + \sum_{n=1}^{N} \boxed{-\frac{1}{2\sigma^2}} (y_n - \bar{w}^T \bar{x}_n)^2 \to \max$$

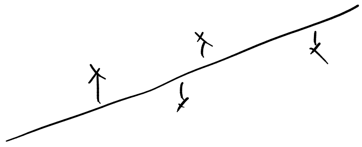
$$\boxed{\sum_{n=1}^{N} (y_n - \bar{w}^T \bar{x}_n)^2 \to \min}$$

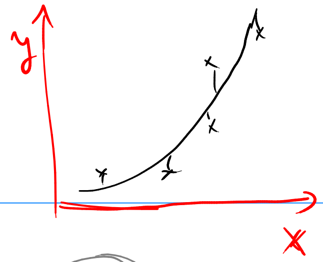
Generative $p(\bar{x}, y) = ?$

Discriminative $p(y | \bar{x}) = ?$

$$\sum_{n=1}^{N} \alpha_n (y_n - \bar{w}^T \bar{x}_n)^2 \to \min$$







$$y \sim w_0 + w_1 x + w_2 x^2$$

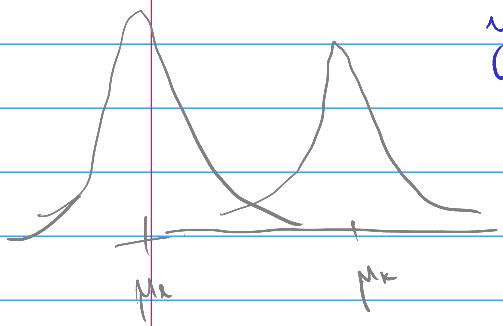
$$\sum_n \left(y_n - (w_0 + w_1 x_n + w_2 x_n^2)\right)^2 \to \min$$

$$\begin{pmatrix} x \\ 1 \end{pmatrix} \xmapsto{\bar{\varphi}} \begin{pmatrix} x^2 \\ x \\ 1 \end{pmatrix}$$

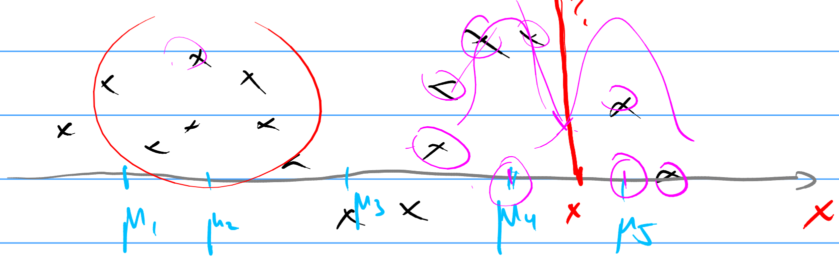
$$y \sim \bar{w}^T \bar{\varphi}(\bar{x})$$

↑ basis functions

$$\bar{\varphi}(x) = \begin{pmatrix} \vdots \\ \varphi_k(x) \\ \vdots \end{pmatrix}, \quad \varphi_k(x) = a \cdot e^{-b \cdot (x-\mu_k)^2}$$



RBF - radial basis functions



$$X = \begin{array}{c} \begin{pmatrix} \vdots \\ x_n \\ \vdots \end{pmatrix} N \\ 1 \end{array} \qquad \xrightarrow{\bar{\varphi}} \qquad \Phi = \begin{array}{c} \begin{pmatrix} & \vdots & \\ \ldots & \varphi_k(x_n) & \ldots \\ & \vdots & \end{pmatrix} N \\ k \end{array}$$

$$(x_n) \longmapsto \begin{pmatrix} \varphi_1(x_n) \\ \vdots \\ \varphi_k(x_n) \end{pmatrix}$$

$$p(\bar{y} \mid \bar{w}, \Phi) = \prod_{n=1}^{N} p(y_n \mid \bar{w}, \bar{\varphi}(\bar{x}_n))$$