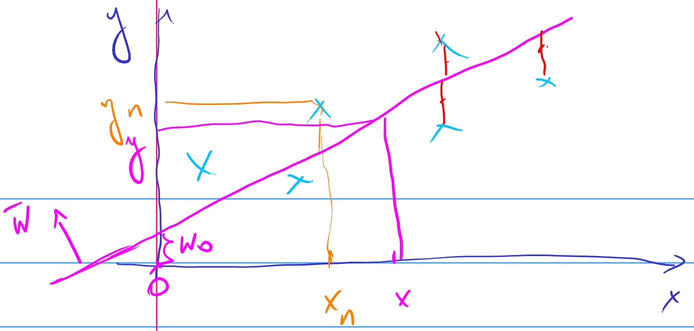


$$D = \{(\bar{x}_n, y_n)\}_{n=1}^{N}$$

$$\bar{w} \in \mathbb{R}^d \; ?$$

$$y \sim w_1 x + w_0$$

$$y \sim \bar{w}^T \bar{x}$$

$$L(\bar{w}) = \sum_{n=1}^{N} \left(y_n - \bar{w}^T \bar{x}_n \right)^2 \xrightarrow[\bar{w}]{} \min$$

$$\begin{pmatrix} | \\ y_n - \bar{w}^T \bar{x}_n \\ | \end{pmatrix}$$

$$L(\bar{w}) = (\bar{y} - X\bar{w})^T (\bar{y} - X\bar{w}) \xrightarrow[\bar{w}]{} \min$$

$$N\times 1 \quad N\times d \quad d\times 1$$

$$X = \begin{array}{c} \\ N \end{array}\begin{pmatrix} | \\ \bar{x}_n \\ | \end{pmatrix}_{d}$$

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

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

ynp $\rightarrow$ $\left[\nabla_{\bar{w}} (\bar{w}^T A \bar{w}) = (A + A^T)\bar{w} \right]$

$$L(\bar{w}) = \bar{y}^T \bar{y} - \bar{w}^T X^T y - y^T X \bar{w} + \bar{w}^T X^T X \bar{w} =$$

$$= \bar{y}^T \bar{y} - 2\bar{w}^T (X^T y) + \bar{w}^T X^T X \bar{w}$$

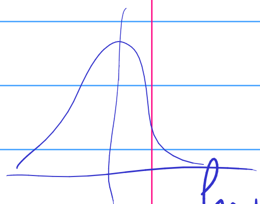
$$\nabla_{\bar{w}} L = -2X^T y + 2X^T X \bar{w} = 0$$

Moore-Penrose

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

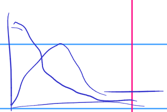
$$y = \bar{w}^T \bar{x} + \varepsilon \qquad \varepsilon \sim \mathcal{N}(0, \sigma^2)$$

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

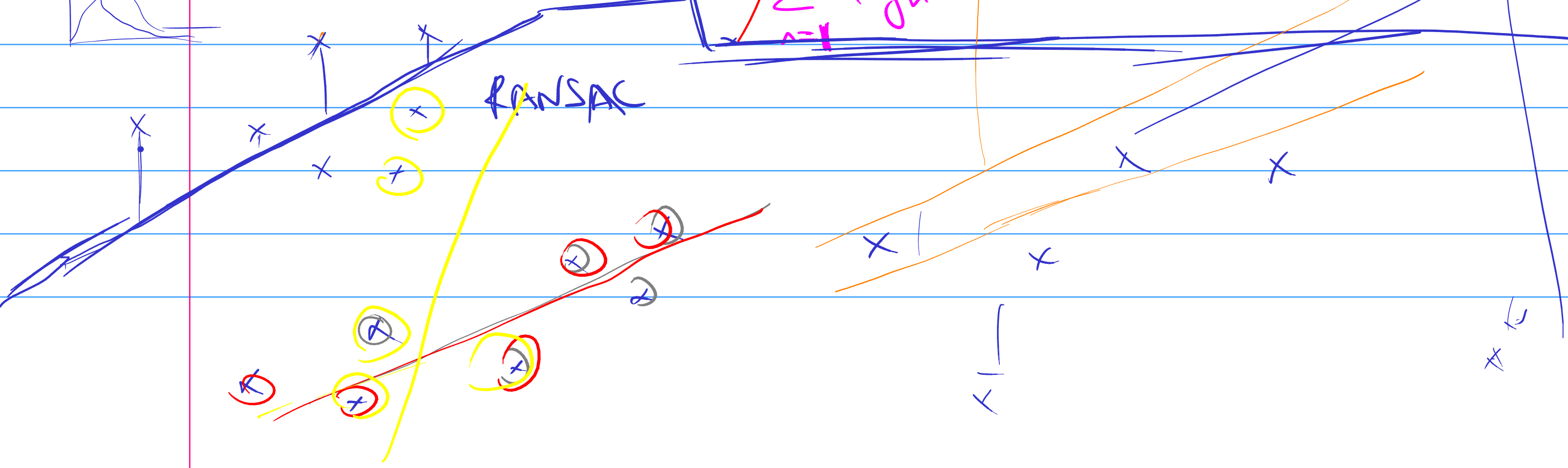


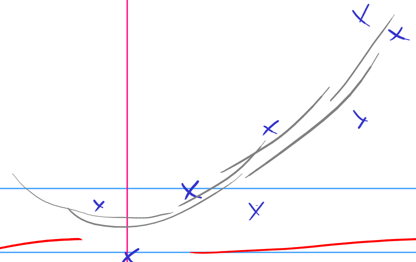
$$p(D \mid \bar{w}) = \prod_{n=1}^{N} \mathcal{N}(y_n \mid \bar{w}^T \bar{x}_n, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y_n - \bar{w}^T \bar{x}_n)^2} \rightarrow \max$$

$$\ln p(D \mid \bar{w}) = \sum_{n=1}^{N} \left(-\frac{1}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} (y_n - \bar{w}^T \bar{x}_n)^2 \right) \rightarrow \max$$

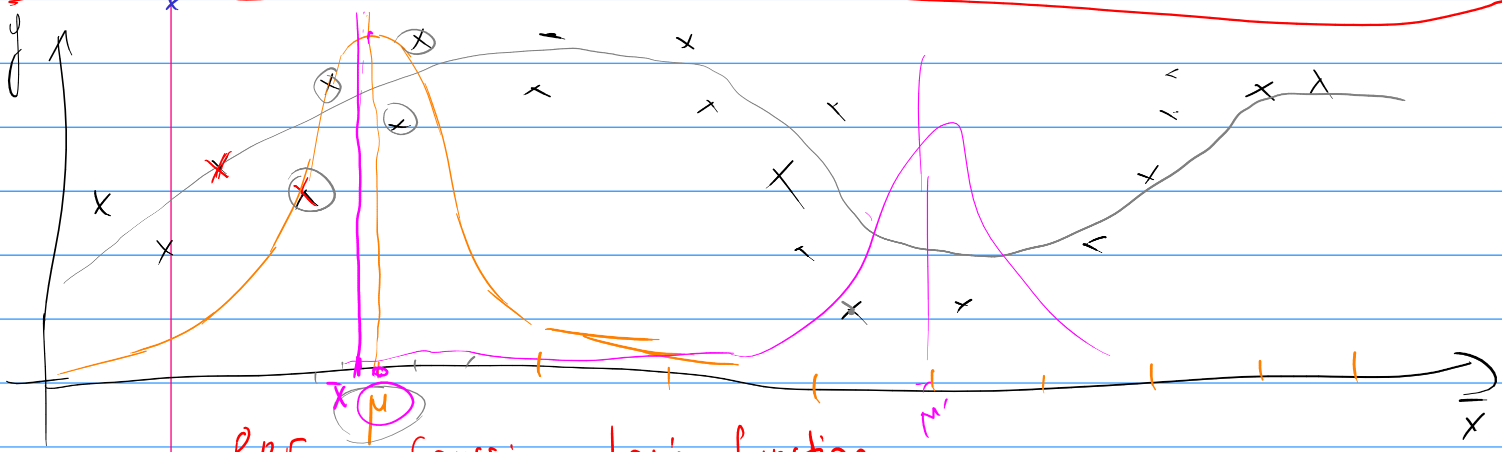


$$\sum_{n=1}^{N} (y_n - \bar{w}^T \bar{x}_n)^2 \rightarrow \min$$





$$\begin{pmatrix}x\\1\end{pmatrix} \mapsto \begin{pmatrix}x^2\\x\\1\end{pmatrix} \qquad y \sim w_2x^2+w_1x+w_0$$



RBF Gaussian basis function

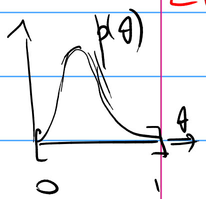
$$\varphi(\bar x,\bar\mu) = e^{-\frac{1}{s^2}\cdot\|\bar x-\bar\mu\|^2}$$

$$(\bar x) \mapsto \begin{pmatrix}\varphi_{\mu_1}(\bar x)\\ \varphi_{\mu_2}(\bar x)\\ \vdots \\ \varphi_{\mu_k}(\bar x)\end{pmatrix}$$

$$L(\bar w) = \sum_{n=1}^{N}\left(y_n-\bar w^T x_n\right)^2 + \Omega(\bar w)$$

$$\Omega(\bar w)=\sum|w|$$

$$\Omega(\bar w)=\sum w^2$$



$\ln p(D|\bar w)$

$$\boxed{p(\bar w|D)=\frac{p(\bar w)\,p(D|\bar w)}{p(D)}}$$

$\boxed{p(d|\bar w)}$

$\Sigma_0=(\sigma^2)$

$$\ln p(\bar w|D) = \ln p(D|\bar w)+\ln p(\bar w)+const \xrightarrow[\bar w]{} \max$$

$$p(\bar w|\bar\mu_0,\Sigma_0)=\frac{1}{(2\pi)^{d/2}\sqrt{\det\Sigma_0}}\cdot e^{-\frac12(\bar w-\bar\mu_0)^T\Sigma_0^{-1}(\bar w-\bar\mu_0)}$$

$\in \mathbb{R}^d$

$$\ln p(\bar w|D) = const + \ln p(\bar w|\bar\mu_0,\Sigma_0)+\ln p(D|\bar w)$$

$$= const - \cancel{\frac{d}{2}\ln 2\pi} - \cancel{\frac12\ln\det\Sigma_0} - \frac12(\bar w-\bar\mu_0)^T\Sigma_0^{-1}(\bar w-\bar\mu_0)$$

$$+\sum_{n=1}^{N}\left[-\cancel{\frac12\ln 2\pi}-\cancel{\ln\sigma}-\frac{1}{2\sigma^2}\left(y_n-\bar w^T\bar x_n\right)^2\right]$$

$$\boxed{p(\bar w|D)}\propto p(\bar w)\,p(D|\bar w)$$

$$p(y|\bar x,D)=\int p(y,\bar w|\bar x,D)\,d\bar w$$

$$=\int p(y|\bar w,\bar x)\cdot p(\bar w|D)\,d\bar w$$

$$p(\bar w|D)=\mathcal{N}(\bar w|\bar\mu',\Sigma')$$

$$\ln p(\bar w|D)=const-\frac12(\bar w-\bar\mu')^T\Sigma'^{-1}(\bar w-\bar\mu')=$$

$$=const-\frac12\bar w^T\Sigma'^{-1}\bar w+\bar w^T\left(\Sigma'^{-1}\bar\mu'\right)$$

$$\ln p(\bar w|D)=const$$

$$-\frac{1}{2\sigma^2}\sum_n\left(y_n-\bar w^T\bar x_n\right)^2$$

$$-\frac{1}{2\sigma_0^2}\cdot\bar w^T\bar w=$$

L_2-регуляризатор

$$\ln p(\bar{w}|D) = \text{const} - \frac{1}{2}\bar{w}^T\Sigma_0^{-1}\bar{w} + \bar{w}^T(\Sigma_0^{-1}\bar{\mu}_0) - \frac{1}{2\sigma^2}(\bar{y}-X\bar{w})^T(\bar{y}-X\bar{w})$$

$$= \sum_n (y_n - \bar{w}^T\bar{x}_n)^2 + \frac{\sigma^2}{\sigma_0^2}\|\bar{w}\|_2^2$$

$$= \text{const} - \frac{1}{2}\bar{w}^T\Sigma_0^{-1}\bar{w} + \bar{w}^T(\Sigma_0^{-1}\bar{\mu}_0) - \frac{1}{2\sigma^2}\bar{w}^T X^T X\bar{w} + \frac{1}{\sigma^2}\bar{w}^T X^T\bar{y}$$

$$= \text{const} - \frac{1}{2}\bar{w}^T\left(\Sigma_0^{-1} + \frac{1}{\sigma^2}X^TX\right)\bar{w} + \bar{w}^T\left(\Sigma_0^{-1}\bar{\mu}_0 + \frac{1}{\sigma^2}X^T\bar{y}\right)$$

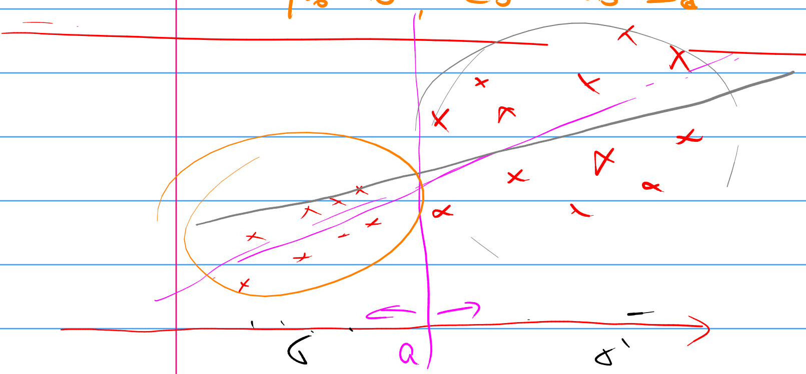
$$\Sigma' = \left(\Sigma_0^{-1} + \frac{1}{\sigma^2}X^TX\right)^{-1}$$

$$\bar{\mu}' = \Sigma' \cdot \left(\Sigma_0^{-1}\bar{\mu}_0 + \frac{1}{\sigma^2}X^T\bar{y}\right)$$

$$p(\bar{w}) \times p(D|\bar{w}) \propto p(\bar{w}|D)$$

$$\boxed{\mathcal{N}(\bar{w}|\bar{\mu}_0, \Sigma_0)} \times \prod \mathcal{N}(y_n|\bar{w}^T\bar{x}_n, \sigma^2) \propto \mathcal{N}(\bar{w}|\bar{\mu}', \Sigma')$$

$\bar{\mu}_0 = \bar{0}$, $\Sigma_0 = \sigma_0 \cdot I_d$



$$\ln p(D|\bar{w}) = \text{const} - \sum_n \frac{1}{2\sigma_n^2}(y_n - \bar{w}^T\bar{x}_n)^2$$

$$= \text{const} - \frac{1}{2\sigma^2}\sum_{\in a} \ldots - \frac{1}{2\sigma'^2}\sum_{a\to} \ldots$$

$$L(\bar{w}) = \sum_n \frac{1}{\sigma_n^2}(y_n - \bar{w}^T\bar{x}_n)^2$$