$$p(\text{COVID} \mid \text{test}=1) = \frac{\overbrace{p(\text{test}=1 \mid \text{COVID})}^{1-\text{false\_negative}} \overbrace{p(\text{COVID})}^{0.01}}{\underbrace{p(t=1 \mid \text{COVID})}_{1-\text{false\_negative}} \overbrace{p(\text{COVID})}^{0.01} + \underbrace{p(t=1 \mid \overline{\text{COVID}})}_{\text{false\_positive}} \overbrace{p(\overline{\text{COVID}})}^{0.99}}$$

$$L(\bar{w}) = \frac{1}{2}\|\bar{w}\|^2 + C \cdot \sum z_n$$

$z_n \geq 0$

$t_n(\bar{w}^T \bar{x}_n + w_0) \geq 1 - z_n$

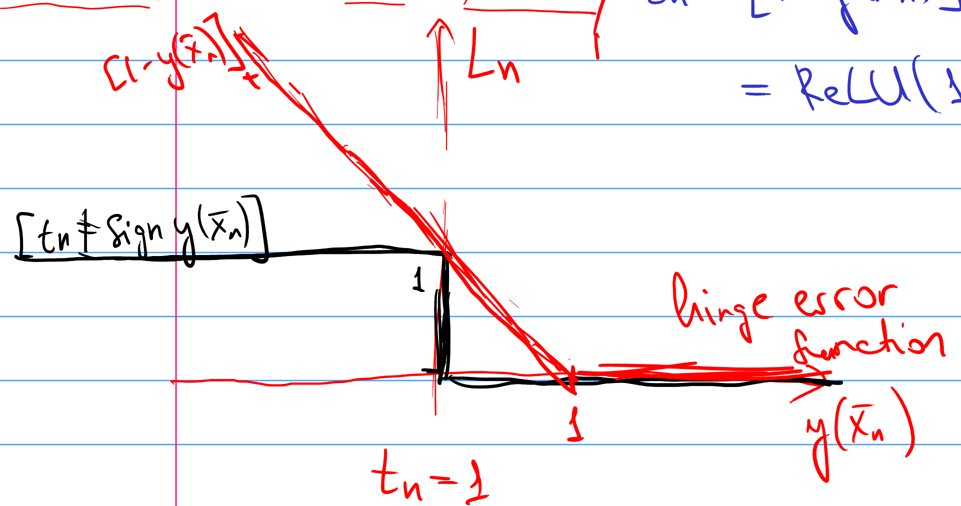
$$= \frac{1}{2}\|\bar{w}\|^2 + C \cdot \sum [1 - y(\bar{x}_n)]_+$$

$t_n = 1$ $\quad y(\bar{x}_n) = \bar{w}^T \bar{x}_n + w_0$

$$z_n = \begin{cases} 1 - y(\bar{x}_n), & y(\bar{x}_n) < 1 \\ 0, & y(\bar{x}_n) \geq 0 \end{cases}$$

$$L(\bar{w}) = \sum [1 - y(\bar{x}_n)]_+ + \alpha \cdot \|\bar{w}\|^2$$

$$z_n = [1 - y(\bar{x}_n)]_+ = \text{ReLU}(1 - y(\bar{x}_n))$$



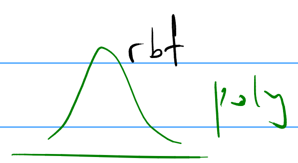
$$y(\bar{x}_n) = \sigma(\bar{w}^T \bar{x}_n)$$

$$-\ln \sigma(\bar{w}^T \bar{x}_n)$$

$$e^{-y(\bar{x}_n)}$$

Relevance Vector Machines (RVM)

$$p(y \mid \bar{x}, \bar{w}, \beta) = \mathcal{N}(y \mid \hat{y}(\bar{x}), \beta^{-1})$$



$\bar{w}^T \varphi(\bar{x}) \quad \begin{pmatrix} 1 \\ k(\bar{x}, \bar{x}_n) \\ 1 \end{pmatrix}$

$$\hat{y}(\bar{x}) = \sum_{n=1}^{N} w_n \cdot k(\bar{x}, \bar{x}_n) + w_0$$

Ant.-persp.:

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

$$p(y \mid \bar{x}, \bar{w}, \beta) = \mathcal{N}(y \mid \bar{w}^T \varphi(\bar{x}), \beta^{-1})$$

$$p(\bar{w} \mid \bar{\alpha}) = \mathcal{N}\left(\bar{w} \mid \bar{0}, \underbrace{\begin{pmatrix} \alpha_1^{-1} & 0 \\ 0 & \alpha_N^{-1} \end{pmatrix}}_{\Sigma_0}\right) = \prod_{n=1}^{N} \mathcal{N}(w_n \mid 0, \alpha_n^{-1})$$

$$\bar{\mu}_N = \beta \Sigma_N \Phi^T \bar{y}$$

$$\Sigma_N = (\Sigma_0^{-1} + \beta \underbrace{\Phi^T}_{h \times n} \Phi)^{-1}$$

$$p(D \mid \bar{\alpha}, \beta)$$

$$p(\bar{y} \mid \Phi, \bar{\alpha}, \beta) = \int \overbrace{p(\bar{y} \mid \Phi, \bar{w}, \beta)}^{\prod_n N(y_n \mid \hat{y}(\bar{x}_n), \beta)} \overbrace{p(\bar{w} \mid \bar{\alpha}, \Phi)}\, d\bar{w} \xrightarrow[\bar{\alpha}, \beta]{} \max$$

$$\int \prod N(y_n \mid \ldots) \cdot N(\bar{w} \mid \bar{\mu}_N, \Sigma_N)\, d\bar{w} =$$

$$= N(\bar{y} \mid \ldots, \ldots)$$

$$\ln p(\bar{y} \mid \Phi, \bar{\alpha}, \beta) = \ldots \bar{\alpha} \ldots \beta \underset{fix}{\bar{w}} \xrightarrow[\bar{\alpha}, \beta]{} \max$$

$$\alpha, \beta \longrightarrow \mu_N, \Sigma_N$$

$$\bar{\alpha}, \beta \longrightarrow \mu_N, \Sigma_N$$

$$\alpha_n \to \infty$$

$$N(w_n \mid 0, \alpha_n^{-1}) \to 0$$

$$\hat{y}(\bar{x}) = \sum_n w_n k(\bar{x}, \bar{x}_n)$$

$$p(\bar{w}_n \mid D)$$

$\underset{\bar{x}_n}{\alpha_n \neq \infty}$ – relevance vectors

---

$$\hat{y}(\bar{x}, \bar{w}) = \sigma(\bar{w}^T \varphi(\bar{x}))$$

$$\varphi(\bar{x}) = \begin{pmatrix} \vdots \\ k(\bar{x}, \bar{x}_n) \\ \vdots \end{pmatrix}$$

$$p(\bar{w} \mid \bar{\alpha}) = \prod N(w_n \mid 0, \alpha_n^{-1})$$

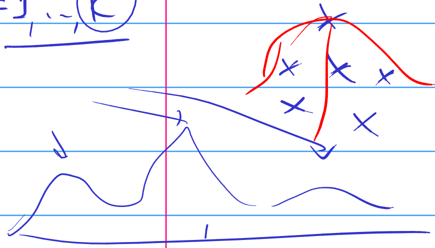
$$\ln p(\bar{w} \mid D, \bar{\alpha}, \beta) = \sum_n \left[t_n \ln \sigma(\bar{w}^T \bar{\varphi}(\bar{x}_n)) + (1 - t_n) \ln(1 - \sigma(\bar{w}^T \bar{\varphi}(\bar{x}_n))) \right]$$

$$- \frac{1}{2} \underbrace{\sum_n \alpha_n w_n^2}_{= \bar{w}^T \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_W \end{pmatrix} \bar{w}} + \text{const}$$

$$\mu_N, \Sigma_N$$

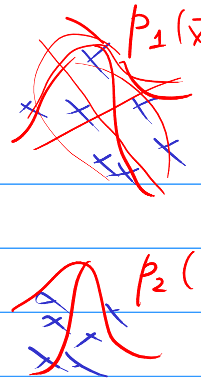
$$p(\bar{y} \mid \bar{\alpha}) \xrightarrow[\bar{\alpha}]{} \max$$

$k = 1, \dots, K$



$p_3(\bar{x}) = \mathcal{N}(\bar{x} | \bar{\mu}_3, \Sigma_3)$

$p_1(\bar{x}) = \mathcal{N}(\bar{x} | \bar{\mu}_1, \Sigma_1)$



$p_2(\bar{x}) = \mathcal{N}(\bar{x} | \bar{\mu}_2, \Sigma_2)$

1) $z_n \sim p(z_n = k) = \boxed{\pi_k}$, $z_n \in \{1, \dots, k\}$

2) $\bar{x}_n \sim \mathcal{N}(\bar{x}_n | \bar{\mu}_{z_n}, \Sigma_{z_n})$

$$p(\bar{x}_n | \theta) = \sum_{z_n} p(\bar{x}_n, z_n | \theta) = \sum_k p(z_n | \theta) \cdot p(\bar{x}_n | z_n, \theta) = \sum_k \pi_k \mathcal{N}(\bar{x}_n | \bar{\mu}_k, \Sigma_k)$$

$\sum \pi_k = 1$

$$p(\bar{x} | \underbrace{\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_3, \Sigma_1, \Sigma_2, \Sigma_3, \bar{\pi}}_{\theta}) = \pi_1 p_1(\bar{x}) + \pi_2 p_2(\bar{x}) + \pi_3 p_3(\bar{x})$$

$z_n \sim \pi_k$

$\bar{x}_n \sim p_{z_n}(\bar{x}_n)$

$$p(\bar{x} | \theta) = \sum_k p(C_k) \cdot p(\bar{x} | C_k)$$

$\bar{\pi}, \bar{\mu}_k, \Sigma_k \to \max$

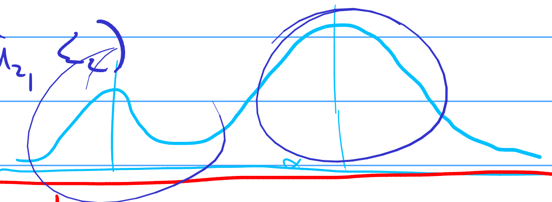
$$p(X | \theta) = \prod_{n=1}^{N} p(\bar{x}_n | \theta) = \prod_{n=1}^{N} \left(\pi_1 \mathcal{N}(\bar{x}_n | \bar{\mu}_1, \Sigma_1) + \pi_2 \mathcal{N}(\bar{x}_n | \bar{\mu}_2, \Sigma_2) + \pi_3 \mathcal{N}(\bar{x}_n | \bar{\mu}_3, \Sigma_3) \right)$$

$$\pi_k = \frac{\#[x_n \in C_k]}{N} \qquad \bar{\mu}_1 = \frac{1}{\#[\bar{x}_n \in C_1]} \sum_{\bar{x}_n \in C_1} \bar{x}_n \qquad \Sigma_1$$

$\pi_1 = \pi$ $\quad \pi_2 = 1 - \pi$

$$p(\bar{x} | \theta) = \pi_1 \mathcal{N}(\bar{x} | \bar{\mu}_1, \Sigma_1) + \pi_2 \mathcal{N}(\bar{x} | \bar{\mu}_2, \Sigma_2)$$

$$z_n = [\bar{x}_n \in C_1] \in \{0, 1\}$$



$p(X | \theta) \xrightarrow[\theta]{} \max$ — трудно

$p(X, Z | \theta) \xrightarrow[\theta]{} \max$ — легко

$$p(X, Z | \theta) = \prod_{n=1}^{N} p(\bar{x}_n, z_n | \theta) = \prod_{n=1}^{N} \pi_{z_n} \cdot \mathcal{N}(\bar{x}_n | \bar{\mu}_{z_n}, \Sigma_{z_n}) = \prod_{n=1}^{N} \prod_{k=1}^{K} \pi_k^{z_{nk}} \mathcal{N}(\bar{x}_n | \bar{\mu}_k, \Sigma_k)^{z_{nk}}$$

$\bar{z}_n = (0 \dots 1 \dots 0)$

EM-алгоритм для смеси гауссианов:

E-шаг: Expectation

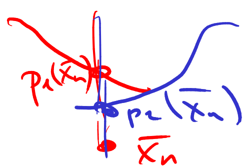
$$\mathbb{E}[z_n] = p(C_1 | \bar{x}_n) = \frac{p(\bar{x}_n | C_1) p(C_1)}{p(\bar{x}_n | C_1) p(C_1) + p(\bar{x}_n | C_2) p(C_2)} = \frac{\pi \cdot \mathcal{N}(\bar{x}_n | \bar{\mu}_1, \Sigma_1)}{\pi \cdot \mathcal{N}(\bar{x}_n | \bar{\mu}_1, \Sigma_1) + (1 - \pi) \mathcal{N}(\bar{x}_n | \bar{\mu}_2, \Sigma_2)}$$

fix $\theta = (\mu_k, \Sigma_k, \pi_k, \dots)$

M-шаг: Maximization

fix $\mathbb{E} z_n$

$$p(X, Z | \theta) = \prod_n \prod_k \pi_k^{\mathbb{E} z_{nk}} \cdot \mathcal{N}(\bar{x}_n | \bar{\mu}_k, \Sigma_k)^{\mathbb{E} z_{nk}} \qquad \mathbb{E} z_{nk} \in [0, 1]$$



$$\ln p(X, Z|\theta) = \sum_n \sum_k \left[\mathbb{E} z_{nk} \ln \pi_k + \mathbb{E} z_{nk} \ln \mathcal{N}_k(\bar{x}_n) \right] \xrightarrow[\theta]{} \max$$

$$\pi_k = \frac{\sum_n \mathbb{E} z_{nk}}{N}$$

$$\sum_n \sum_k \mathbb{E}_{z_{nk}} (\bar{x}_n - \bar{\mu}_k)^T \Sigma_k^{-1} (\bar{x}_n - \bar{\mu}_k)$$

$$\sum_n \mathbb{E} z_{nk} (\ldots)^T \Sigma_k^{-1} (\ldots) \longrightarrow \max$$

$$\mathbb{E} z_{nk} \in [0,1]$$

$$p(X|\theta) \longrightarrow \max$$

$$\ell(\theta) = \log p(X|\theta) \longrightarrow \max$$

$$\theta^{(0)} \to \theta^{(1)} \to \ldots \to \boxed{\theta^{(m)}} \to \theta^{(m+1)} \to \ldots$$

$$p(X, Z|\theta)$$

$$\ell(\theta) - \ell(\theta^{(m)}) = \log p(X|\theta) - \log p(X|\theta^{(m)}) =$$

$$= \log \int p(X, Z|\theta)\, dZ - \log p(X|\theta^{(m)}) =$$

$\log \int \ldots$

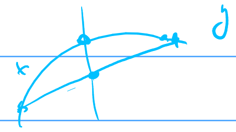
$\log \mathbb{E}_Z[\ldots]$

$\log \int \ldots q(Z)\, dZ$

$$= \log \int p(X, Z|\theta) \cdot \frac{p(Z|X, \theta^{(m)})}{p(Z|X, \theta^{(m)})} dZ - \log p(X|\theta^{(m)})$$

$$= \log \int \underbrace{\frac{p(X, Z|\theta)}{p(Z|X, \theta^{(m)})}}_{f(Z)} p(Z|X, \theta^{(m)})\, dZ - \log p(X|\theta^{(m)}) \geq$$

$$\log \mathbb{E}_{p(Z|X, \theta^{(m)})} [f(Z)]$$



$$f(\alpha x + (1-\alpha) y) \geq$$
$$\geq \alpha f(x) + (1-\alpha) f(y)$$

$$\geq \int \log \frac{p(X, Z|\theta)}{p(Z|X, \theta^{(m)})} p(Z|X, \theta^{(m)})\, dZ - \log p(X|\theta^{(m)})$$

$$f\left(\sum_k \alpha_k x_k\right) \geq$$
$$\geq \sum_k \alpha_k f(x_k)$$

$$= \int \log \frac{p(X, Z|\theta)}{\underbrace{p(Z|X, \theta^{(m)})\, p(X|\theta^{(m)})}_{p(X, Z|\theta^{(m)})}} p(Z|X, \theta^{(m)})\, dZ$$

$$\ell(\theta) - \ell(\theta^{(m)}) \geq \int \log \frac{p(X, Z|\theta)}{\ldots} p(Z|X, \theta^{(m)})\, dZ$$

$$f(\mathbb{E}[x]) \geq \mathbb{E}[f(x)]$$

Jensen's inequality

"$\mathcal{L}(\theta)$"

$$\ell(\theta) \geq \mathcal{L}(\theta) + \ell(\theta^{(m)})$$

$$\theta = \theta^{(m)}$$

$$\mathcal{L}(\theta) \xrightarrow[\theta]{} \max$$

$$Q(\theta, \theta^{(m)}) = \int \log p(X, Z|\theta)\, p(Z|X, \theta^{(m)})\, dZ$$

$$\theta^{(m+1)} := \arg\max_\theta Q(\theta, \theta^{(m)})$$

$$= \int \log p(X, Z|\theta)\, p(Z|X, \theta^{(m)})\, dZ$$

$$- \cancel{\int \log p(X, Z|\theta^{(m)})\, p(Z|X, \theta^{(m)})\, dZ}$$