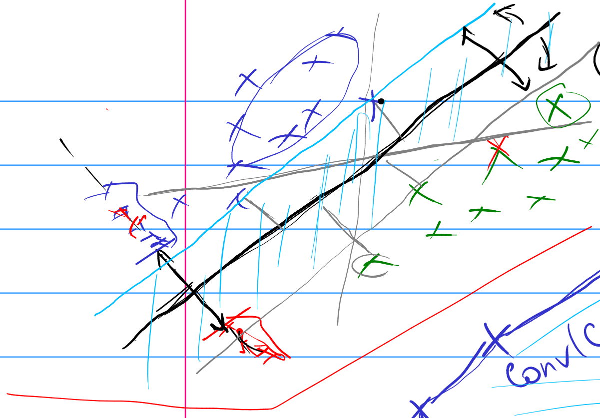
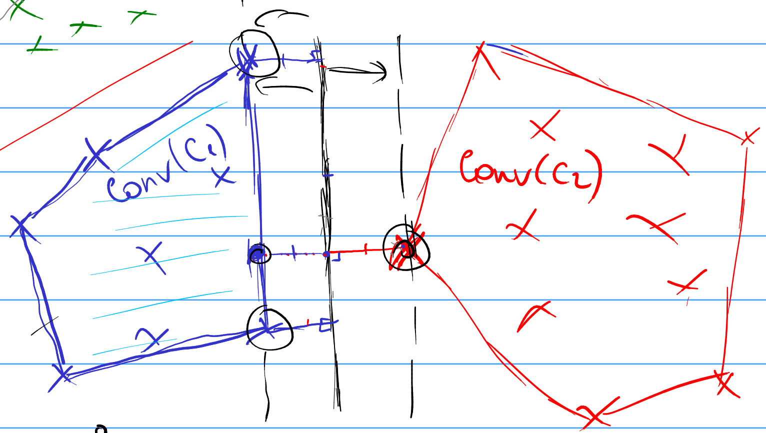


Generative: $p(\bar{x}|C_1),\ p(\bar{x}|C_2)$ — LDA, QDA

Discriminative: $p(C_1|\bar{x}),\ p(C_2|\bar{x})$ — Logistic regression



$$\min_{\bar{a},\bar{b}} \|\bar{a}-\bar{b}\|^2, \quad \bar{a}\in \mathrm{Conv}(C_1) = \Big\{\sum_{n:\bar{x}_n\in C_1} \alpha_n \bar{x}_n \ \Big|\ \alpha_n \geq 0,\ \sum \alpha_n = 1\Big\}$$

$$\bar{b}\in \mathrm{Conv}(C_2) = \Big\{\sum_{n:\bar{x}_n\in C_2} \alpha_n \bar{x}_n \ \Big|\ \alpha_n \geq 0,\ \sum \alpha_n = 1\Big\}$$

Вапник
Лернер
Червоненкис

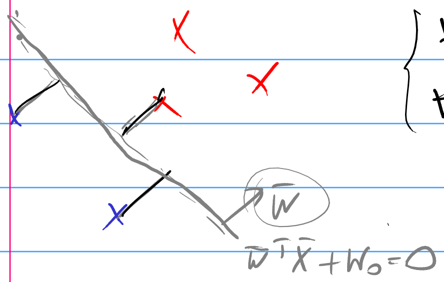
$$\min_{\alpha_n} \Big\|\sum_{\bar{x}_n\in C_1} \alpha_n \bar{x}_n - \sum_{\bar{x}_n\in C_2} \alpha_n \bar{x}_n\Big\|^2, \quad \forall n\ \alpha_n \geq 0, \quad \sum_{\bar{x}_n\in C_1}\alpha_n = \sum_{\bar{x}_n\in C_2}\alpha_n = 1$$

метод обобщенных портретов

quadratic programming

N переменных, N ограничений

VC-dimension



$$\begin{cases} \forall \bar{x}_n \in C_1 & \bar{w}^T\bar{x}_n + w_0 > 0 \\ \forall \bar{x}_n \in C_2 & \bar{w}^T\bar{x}_n + w_0 < 0 \end{cases} \qquad t_n \in \{\underset{C_2}{-1}, \underset{C_1}{+1}\}$$

$$\max_{\bar{w},w_0} \min_{n=1}^{N} \frac{|\bar{w}^T\bar{x}_n + w_0|}{\|\bar{w}\|} \qquad \boxed{\|\bar{w}\| = ?}$$

$$\forall n \quad t_n(\bar{w}^T\bar{x}_n + w_0) > 0$$

$$\max_{\bar{w},w_0} \min_n \frac{t_n(\bar{w}^T\bar{x}_n + w_0)}{\|\bar{w}\|} = \max_{\bar{w},w_0} \frac{1}{\|\bar{w}\|} \boxed{\min_n \underbrace{t_n(\bar{w}^T\bar{x}_n + w_0)}_{y(\bar{x}_n)}} \quad \text{"} 1$$

$$\|\bar{w}\| = 1$$

$$\max_{\bar{w},w_0} \min_n t_n(\bar{w}^T\bar{x}_n + w_0)$$

$$\forall n\ t_n(\bar{w}^T\bar{x}_n + w_0) > 0$$

$$\bar{w}^T\bar{w} = 1$$

$$\min_n y(\bar{x}_n) = 1$$

$$\forall n \quad y(\bar{x}_n) \geq 1$$

$$\min_n y(\bar{x}_n) = a > 1 \implies \bar{w}' = \frac{1}{a}\cdot\bar{w}, \quad w_0' = \frac{1}{a}\cdot w_0$$

$$\max_{\bar w, w_0} \frac{1}{\|\bar w\|} = \min_{\bar w, w_0} \|\bar w\|^2$$

$\cancel{\min_n t_n(\bar w^T \bar x_n + w_0) = 1}$

$\forall n \quad t_n(\bar w^T \bar x_n + w_0) \geq 1$

quadratic programming

D переменных, N ограничений

$$L(\bar w, w_0, \bar\alpha) = \frac{1}{2}\bar w^T \bar w - \sum_{n=1}^{N} \alpha_n \left(t_n(\bar w^T \bar x_n + w_0) - 1\right) \xrightarrow[\bar w, w_0, \bar\alpha]{} \min$$

$$\frac{\partial L}{\partial w_0} = 0 - \sum_n \alpha_n t_n = 0; \quad \boxed{\sum_n \alpha_n t_n = 0} \qquad \forall n \ \alpha_n \geq 0$$

$$\nabla_{\bar w} L = \bar w - \sum_n \alpha_n t_n \bar x_n = 0 \qquad \boxed{\bar w = \sum_n \alpha_n t_n \bar x_n}$$

$$L(\bar\alpha) = \frac{1}{2}\left(\sum_n \alpha_n t_n \bar x_n\right)^T \left(\sum_m \alpha_m t_m \bar x_m\right) - \sum_{n=1}^{N} \alpha_n \left(t_n \bar x_n^T \left(\sum_m \alpha_m t_m \bar x_m\right) - 1\right)$$

$$= \frac{1}{2}\sum_{n,m} \alpha_n \alpha_m t_n t_m \bar x_n^T \bar x_m - \sum_{n,m} \alpha_n \alpha_m t_n t_m \bar x_n^T \bar x_m + \sum_n \alpha_n$$

$$\boxed{L(\bar\alpha) = \sum_n \alpha_n - \frac{1}{2}\sum_n \sum_m \alpha_n \alpha_m t_n t_m \bar x_n^T \bar x_m \xrightarrow[\bar\alpha]{} \min,} \qquad \sum \alpha_n t_n = 0, \quad \forall n \ \alpha_n \geq 0$$

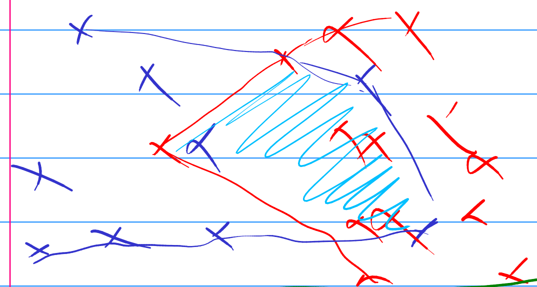
N переменных, N ограничений

$\bar w = \sum_n \alpha_n t_n \bar x_n \qquad w_0$

$$\hat t = \bar w^T \bar x + w_0 = \sum_n \alpha_n t_n \bar x_n^T \bar x + w_0$$

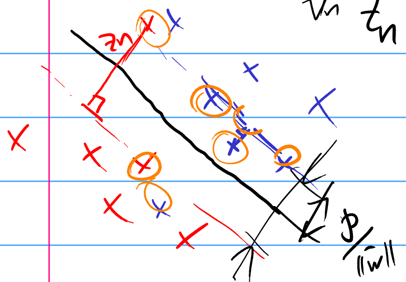
$$\approx \sum_n t_n k(\bar x_n, \bar x)$$

$$\alpha_n \bar x_n^T \bar x$$



$$\forall n \quad t_n(\bar w^T \bar x_n + w_0) + z_n \geq 1 \qquad \forall n \ z_n \geq 0$$

$$\min_{\bar z, \bar w, w_0} \left(\frac{1}{2}\bar w^T \bar w + C \cdot \sum z_n\right)$$



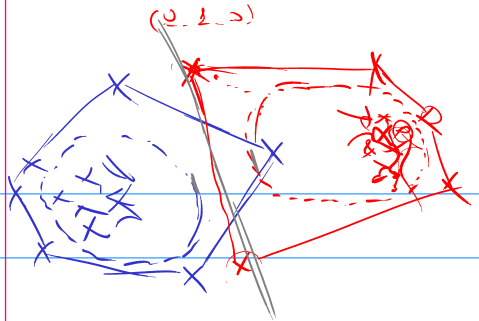
$$L(\bar\alpha) = \sum_n \alpha_n - \frac{1}{2}\sum_{n,m} \text{---} \qquad \forall n \quad C \geq \alpha_n \geq 0 \qquad \sum_n \alpha_n t_n = 0$$

ν-SVM

$$L_\nu(\bar w, \rho, \bar z) = \frac{1}{2}\bar w^T \bar w + \frac{1}{2}\sum z_n - \nu \rho$$

$$t_n(\bar w^T \bar x_n + w_0) + z_n \geq \rho \qquad z_n \geq 0 \quad \rho \geq 0$$

$$\boxed{\nu} \geq \frac{1}{N} \cdot \#\{n : z_n > 0\} \qquad \nu \leq \frac{1}{N} \# \{\text{support vectors}\}$$



$$\mathrm{Conv}(C) = \{ \sum \alpha_n \bar{x}_n \mid \bar{x}_n \in C,\ \sum \alpha_n = 1,\ \alpha_n \geq 0 \}$$

$$\mathrm{Conv}_D(C) = \{ \sum \alpha_n \bar{x}_n \mid 0 \leq \alpha_n \leq D \}$$

$1/2$

$$\bar{x}^T A \bar{x} + \bar{x}^T \bar{b} + c = 0$$

$$\bar{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$$

$\bar{\varphi} \downarrow$

$$\bar{\varphi}(\bar{x}) = \begin{pmatrix} x_1^2 \\ x_2^2 \\ x_1 x_2 \\ x_1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \end{pmatrix} = w_1 x_1^2 + w_2 x_2^2 + w_3 x_1 x_2 + w_4 x_1 + w_5 w_5$$

$\in \mathbb{R}^5$

$$\boxed{\bar{w}^T \bar{\varphi}(\bar{x}) + w_0 = 0}$$

$$\bar{x}_n \in \mathbb{R}^d \longmapsto \bar{\varphi}(\bar{x}_n) \in \mathbb{R}^D$$

$$\bar{w} \in \mathbb{R}^D$$

$$\boxed{\frac{d(d+1)}{2}} = O(d^2) \qquad \boxed{O(d^k)}$$

$d = 50 \qquad 50^5$

$$L(\bar{\alpha}) = \sum_n \alpha_n - \frac{1}{2} \sum_{n,m} \alpha_n \alpha_m t_n t_m \bar{x}_n^T \bar{x}_m \xrightarrow[\bar{\alpha}]{} \min$$

$$\sum \alpha_n t_n = 0, \quad C \geq \alpha_n \geq 0$$

$k(\bar{x}_n, \bar{x}_m)$

$$y(\bar{x}) = \sum_n \alpha_n t_n \bar{x}_n^T \bar{x} + w_0$$

$k(\bar{x}_n, \bar{x})$

$$\mathbb{R}^d \ni \bar{x} \xrightarrow{\varphi} \bar{\varphi}(\bar{x}) \in \mathbb{R}^D$$

kernel

$$\bar{y} \xrightarrow{\varphi} \bar{\varphi}(\bar{y}) \in \mathbb{R}^D$$

$$k(\bar{x}, \bar{y}) = \bar{\varphi}(\bar{x})^T \bar{\varphi}(\bar{y})$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \longmapsto \begin{pmatrix} x_1^2 \\ \sqrt{2} x_1 x_2 \\ x_2^2 \\ \sqrt{2} x_1 \\ \sqrt{2} x_2 \end{pmatrix} \qquad \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \longmapsto \begin{pmatrix} y_1^2 \\ \sqrt{2} y_1 y_2 \\ y_2^2 \\ \sqrt{2} y_1 \\ \sqrt{2} y_2 \end{pmatrix}$$

$$\bar{\varphi}(\bar{x})^T \bar{\varphi}(\bar{y}) = x_1^2 y_1^2 + 2 x_1 x_2 y_1 y_2 + x_2^2 y_2^2 = (x_1 y_1 + x_2 y_2)^2 = (\bar{x}^T \bar{y})^2$$

$$x_1^2 y_1^2 + 2 x_1 x_2 y_1 y_2 + x_2^2 y_2^2 + 2 x_1 y_1 + 2 x_2 y_2 = (\bar{x}^T \bar{y} + 1)^2 - 1$$

linear

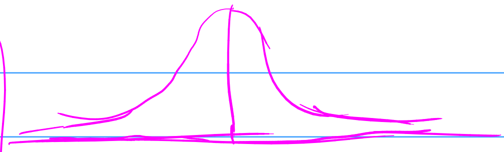
poly degree d

$$k(\bar{x}, \bar{y}) = (\bar{x}^T \bar{y} + 1)^d - 1$$

$$\forall \bar{x} \neq 0 \quad k(\bar{x}, \bar{x}) > 0 \qquad \forall \bar{x}\, \bar{y} \quad k(\bar{x}, \bar{y}) = k(\bar{y}, \bar{x})$$

$f(\bar{x}) = \bar{w}^T\bar{x} + w_0$

RBF - radial basis functions $\quad k(\bar{x}_n, \bar{x}) = e^{-s \cdot \|\bar{x} - \bar{x}_n\|^2}$



$\bar{x}_n$

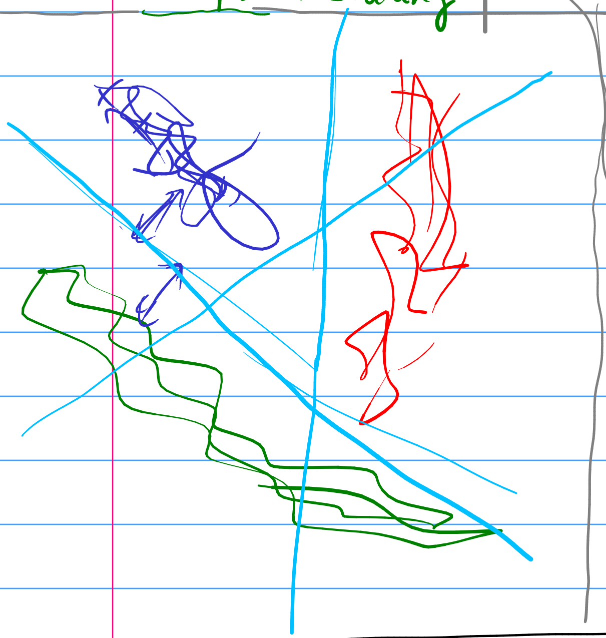
$$\boxed{y(\bar{x}) = \sum_n \alpha_n t_n k(\bar{x}_n, \bar{x}) + w_0}$$

Karush-Kuhn-Tucker: $\quad \alpha_n \left(t_n(\bar{w}^T\bar{x}_n + w_0) - 1 \right) = 0$

$\bar{x}_n$ - support vector: $\quad t_n(\bar{w}^T\bar{x}_n + w_0) = 1$

$$p(C_1|\bar{x}) = \frac{1}{1 + e^{af(\bar{x}) + b}}$$

Platt scaling



- one-vs-all $\quad y_k = \bar{w}_k^T\bar{x} + w_{0k} \quad C_k$ vs C_{-k}

$y(\bar{x}) = \arg\max_k y_k(\bar{x}) \quad k = 1, \dots, K$

LinearSVC. predict_proba

- one-vs-one $\quad \frac{K(K-1)}{2} \quad$ voting

$y_{kl}(\bar{x}) = \bar{w}_{kl}^T\bar{x} + w_{0,kl} \quad C_k$ vs C_l

$$\arg\max_{k=1}^{K} a_k = \arg\max_k \sum_l [a_k > a_l]$$

$f_k(\bar{x}) = p(C_k|\bar{x})$ $\quad$ Friedman

$$k^* = \arg\max_k p(C_k|\bar{x}) = \arg\max_k \left(\sum_{l=1}^{K} [p(C_k|\bar{x}) > p(C_l|\bar{x})] \right) =$$

$$= \arg\max_k \sum_l \left[\frac{p(C_k|x)}{p(C_k|x) + p(C_l|x)} > \frac{p(C_l|x)}{p(C_k|x) + p(C_l|x)} \right]$$

$$p(C_k|\bar{x} \in C_k \vee \bar{x} \in C_l) \overset{?}{>} p(C_l|\bar{x} \in C_k \vee \bar{x} \in C_l)$$

Hastie - Tibshirani

$$r_{kl} \approx p(\bar{x} \in C_k | \bar{x} \in C_k \vee \bar{x} \in C_l)$$

$$p(\bar{x} \in C_k) = p_k = ? \qquad r_{kl} \approx \frac{p_k}{p_k + p_l} = \mu_{kl}$$

$$L(\bar{p}) = \sum_{k<l} N_{kl} \left(r_{kl} \log\frac{r_{kl}}{\mu_{kl}} + (1 - r_{kl}) \log\frac{1 - r_{kl}}{1 - \mu_{kl}} \right) \xrightarrow[\bar{p}]{} \min$$