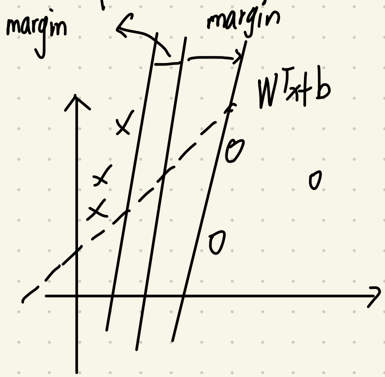
SVM 有三宝：间隔、对偶、核技巧

硬间隔 SVM 模型定义

SVM 有三宝：间隔，对偶，核技巧

类别
- 硬间隔 hard margin
- 软间隔 soft margin
- kernel svm



哪条线最好？

离样本点最远 (max margin(w,b))

$$margin(w,b) = \min_{x_i}\left(\frac{|w^Tx_i+b|}{\|w\|}\right)$$

超平面 w^Tx+b

模型 $f(w) = \underbrace{sign}_{符号函数}(w^Tx+b)$

无概率的判别模型

模型：
$$\begin{cases} \max\limits_{w,b}\ \min\limits_{x_i}\left(\frac{|w^Tx_i+b|}{\|w\|}\right) \\ 满足\ y_i(w^Tx_i+b)>0 \end{cases}$$

$$\begin{cases} \max\min\left(\frac{1}{\|w\|}\left[y_i(w^Tx_i+b)\right]\right) \\ y_i(w^Tx_i+b)>0 \Longrightarrow \exists r>0\ 使得\ \min y_i(w^Tx_i+b)=r\ (即\ y_i(w^Tx_i+b)>r) \end{cases}$$

$$\Longrightarrow \max\frac{1}{\|w\|}\min\left[y_i(w^Tx_i+b)\right]$$

代入

令 $r=1$ $\xrightarrow{\text{why?}}$

$w^T x+b$ 函数上
倾上 || 等价 不等价!
$2w^T x+2b$

$||w||$ 确定下来

$$M \begin{cases} \max\limits_{w,b} \dfrac{1}{||w||} \Longrightarrow \min \dfrac{1}{2} w^T w & \text{二次} \\ s.t.\ \min y_i(w^T x_i+b)=1 \quad (y_i(w^T x_i+b)\geq 1) & N\text{个约束} \end{cases}$$

$\begin{cases} \text{二次规划} \\ \text{凸优化} \end{cases}$ 问题

硬间隔　模型求解

$Data = \{(x_i, y_i)\}_{i=1}^{N}$　$x_i \in R$　$y_i = \{\pm 1\}$

$$\begin{cases} \min\limits_{w,b} \frac{1}{2} w^T w \\ s.t.\ y_i(w^T x_i + b) \geq 1 \Longleftrightarrow 1 - y_i(w^T x_i + b) \leq 0 \end{cases}$$

w,b带约束

定义拉格朗日函数

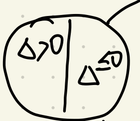
$$\mathcal{L}(w,b,\lambda) = \frac{1}{2} w^T w + \sum_{1}^{N} \underbrace{\lambda_i}_{\geq 0} [\underbrace{1 - y_i(w^T x_i + b)}_{\leq 0}]$$

$$\begin{cases} \max\limits_{w\,b} \max\limits_{\lambda} \mathcal{L}(w,b,\lambda) \\ s.t_i\ \lambda_i \geq 0 \end{cases}$$

w,b无约束

为什么等价

(w,b)的集合



$\Delta = 1 - y_i(w^T x_i + b)$

划成2部分

若 $\Delta > 0$　$\max\limits_{\lambda} \mathcal{L}(w,b,\lambda) = \infty$ (λ越大越好)

若 $\Delta \leq 0$　$\max\limits_{\lambda} \mathcal{L}(w,b,\lambda) = \mathcal{L}(w,b,0) = \frac{1}{2} w^T w$ (λ取0)

满足强对偶

对偶问题 (未证明)

对偶问题 $\begin{cases} \max\limits_{\lambda} \min\limits_{w,b} \mathcal{L}(w,b,\lambda) \\ \lambda_i \geq 0 \end{cases}$

$$\begin{cases} \min\max f \geq \max\min f \quad (鸡头 \leq 凤尾) & 弱对偶关系 \\ \min\max f = \max\min f & 强对偶关系 \end{cases}$$

求 $\min\limits_{w,b} \mathcal{L}(w,b,\lambda)$

① $\dfrac{\partial \mathcal{L}}{\partial b} = \dfrac{\partial}{\partial b}\left[\sum\limits_{i=1}^{N}\lambda_i - \sum\limits_{i=1}^{N}\lambda_i y_i (W^T x_i + b)\right]$

$= -\sum\limits_{i=1}^{N}\lambda_i y_i$

$\dfrac{\partial \mathcal{L}}{\partial b} = 0 \Longrightarrow \lambda_i y_i = 0 \Longrightarrow \mathcal{L}(w,b,\lambda) = \dfrac{1}{2}W^T W + \sum\limits_{i=1}^{N}\lambda_i - \sum\limits_{i=1}^{N}\lambda_i y_i W^T x_i$

② $\dfrac{\partial \mathcal{L}}{\partial W} = W - \sum\limits_{i=1}^{N}\lambda_i y_i x_i$

过程略

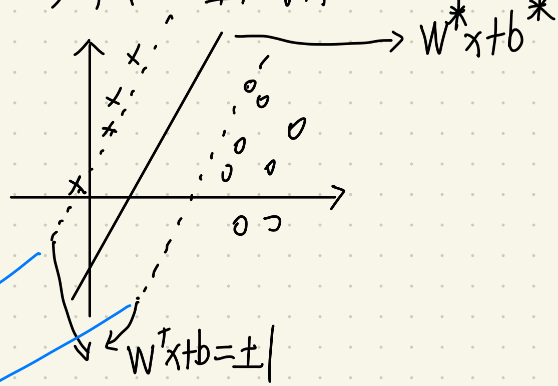
$\dfrac{\partial \mathcal{L}}{\partial W} = 0 \Longrightarrow W = \sum\limits_{i=1}^{N}\lambda_i y_i x_i \Longrightarrow \mathcal{L}(w,b,\lambda) = -\dfrac{1}{2}\sum\limits_{i=1}^{N}\sum\limits_{j=1}^{N}\lambda_i\lambda_j y_i y_j x_i^T x_j + \sum\limits_{i=1}^{N}\lambda_i$

$$\begin{cases} \max\limits_{\lambda} \ \underline{\quad \cdots\cdots \quad} \\ s.t.\ \lambda_i \geq 0,\ \sum\limits_{i=1}^{N}\lambda_i y_i = 0 \end{cases}$$

最终模型

KKT条件

$$\begin{cases} \frac{\partial L}{\partial w}=0 \quad \frac{\partial L}{\partial b}=0 \quad \frac{\partial L}{\partial \lambda}=0 \\ \lambda_i(1-y_i(w^Tx_i+b))=0 \longrightarrow \text{松弛互补条件} \\ \lambda_i \ge 0 \\ 1-y_i(w^Tx_i+b) \le 0 \end{cases}$$



$$\boxed{w^* = \sum_{i=1}^{N} \lambda_i y_i x_i} \qquad b^*$$

求 b^*: $\begin{cases} \text{当 } 1-y_i(w^Tx_i+b)=0 \text{ 时 } \lambda_i \text{ 取值无所谓 (在边界上)} \\ \text{当 } \neq 0 \text{ 时 } \lambda_i=0 \text{ (不在边界上)} \end{cases}$

$\exists k$ 使得 $1-y_k(w^Tx_k+b)=0$ (边界点)　　$b=y_k-w^Tx_k$　用w^*代替

$$\boxed{b=y_k-\sum_{i=1}^{N}\lambda_i y_i x_i^T x_k}$$

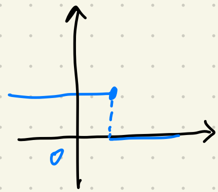
软间隔 SVM 定义

soft：允许一点点错误

$$\min \frac{1}{2}w^Tw + loss$$

令 $z = y_i(w^Tx_i+b)$

① loss = 点数：$\sum_{i=1}^{N} I\{y_i(w^Tx_i+b)<1\}$ 不连续



② loss = 距离：
(hinge loss)

$$\begin{cases} 0 & \text{当 } y_i(w^Tx_i+b) \geq 1 \\ 1-y_i(w^Tx_i+b) & <1 \end{cases}$$