Hard-Margin SVM　　Support Vector Machine

支持向量机

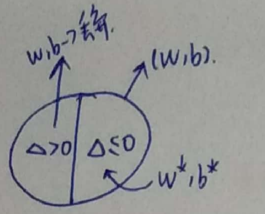
$Data=\{(x_i,y_i)\}_{i=1}^{N}$　$x_i\in\mathbb{R}^p$, $y_i\in\{+1,-1\}$.

带约束 $\begin{cases}\min\limits_{w,b}\ \frac{1}{2}w^Tw \\ s.t.\ y_i(w^Tx_i+b)\ge 1 \Leftrightarrow \underbrace{1-y_i(w^Tx_i+b)\le 0}_{\Delta} \\ \text{for } i=1,\dots,N\end{cases}$　⟶ primal problem 原问题

条件

dual problem 对偶问题.

拉格朗日函数: $\mathcal{L}(w,b,\lambda)=\frac{1}{2}w^Tw+\sum_{i=1}^{N}\underset{\lambda_i\ge 0}{\lambda_i}\underbrace{(1-y_i(w^Tx_i+b))}_{\le 0}$



弱对偶关系.

$\min\max\mathcal{L}\ \geq\ \max\min\mathcal{L}$

凤尾　鸡头

鸡头　凤尾

鸡 ≤ 凤

"=" 为强对偶关系.

无约束 $\begin{cases}\min\limits_{w,b}\max\limits_{\lambda}\mathcal{L}(w,b,\lambda) \\ s.t.\ \lambda_i\ge 0\end{cases}$

强对偶

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

(x_i,y_i)

如果 $1-y_i(w^Tx_i+b)>0$, $\max\limits_{\lambda}\mathcal{L}(\lambda,w,b)=\infty$　$=\frac{1}{2}w^Tw+\infty$

如果 $1-y_i(w^Tx_i+b)\le 0$, $\max\limits_{\lambda}\mathcal{L}(\lambda,w,b)$ 一定存在. $\frac{1}{2}w^Tw+0=\frac{1}{2}w^Tw$, $\min\limits_{w,b}\max\limits_{\lambda}\mathcal{L}=\min\limits_{w,b}\frac{1}{2}w^Tw$

$\min\limits_{w,b}\max\limits_{\lambda}\mathcal{L}(w,b,\lambda)=\min\limits_{w,b}(\infty,\frac{1}{2}w^Tw)=\min\limits_{w,b}\frac{1}{2}w^Tw$

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

$\frac{\partial\mathcal{L}}{\partial b}=\frac{\partial}{\partial b}\left[\sum_{i=1}^{N}\lambda_i-\sum_{i=1}^{N}\lambda_iy_i(w^Tx_i+b)\right]$

$=\frac{\partial}{\partial b}\left[-\sum_{i=1}^{N}\lambda_iy_ib\right]$

$=-\sum_{i=1}^{N}\lambda_iy_i=0$

$\lambda=\begin{pmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_n\end{pmatrix}$

$\frac{\partial\mathcal{L}}{\partial b}\triangleq 0\Rightarrow\sum_{i=1}^{N}\lambda_iy_i=0$, 将其代入 $\mathcal{L}(w,b,\lambda)$

$\mathcal{L}(w,b,\lambda)=\frac{1}{2}w^Tw+\sum_{i=1}^{N}\lambda_i-\sum_{i=1}^{N}\lambda_iy_i(w^Tx_i+b)$

$=\frac{1}{2}w^Tw+\sum_{i=1}^{N}\lambda_i-\sum_{i=1}^{N}\lambda_iy_iw^Tx_i+\sum_{i=1}^{N}\lambda_iy_ib.$

$=\frac{1}{2}w^Tw+\sum_{i=1}^{N}\lambda_i-\sum_{i=1}^{N}\lambda_iy_iw^Tx_i$

w^*

$\frac{\partial\mathcal{L}}{\partial w}=\frac{1}{2}\cdot 2w-\sum_{i=1}^{N}\lambda_iy_ix_i\triangleq 0\Rightarrow\boxed{w^*=\sum_{i=1}^{N}\lambda_iy_ix_i}$

$\mathcal{L}(w,b,\lambda)=\frac{1}{2}\left(\sum_{i=1}^{N}\lambda_iy_ix_i\right)^T\left(\sum_{j=1}^{N}\lambda_jy_jx_j\right)-\underbrace{\sum_{i=1}^{N}\lambda_iy_i\left(\sum_{j=1}^{N}\lambda_jy_jx_j\right)^Tx_i}_{\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_i\lambda_jy_iy_j\underbrace{x_j^Tx_i}_{x_i^Tx_j}}+\sum_{i=1}^{N}\lambda_i.$

$=-\frac{1}{2}$

$\sum_{i=1}^{N}\lambda_iy_ix_i^T\sum_{j=1}^{N}\lambda_jy_jx_j$

$\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_i\lambda_jy_iy_jx_i^Tx_j$

-1

$\frac{1}{2}$

$=-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_i\lambda_jy_iy_jx_i^Tx_j+\sum_{i=1}^{N}\lambda_i$

$\begin{cases}\max\limits_{\lambda}-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_i\lambda_jy_iy_jx_i^Tx_j+\sum_{i=1}^{N}\lambda_i \\ s.t.\ \lambda_i\ge 0\quad\sum_{i=1}^{N}\lambda_iy_i=0\end{cases}$