3.24 《机器学习基础》第3次作业

王宇哲 1800011828

T_1

证：存在性显然，根据线性可分的定义，

$\exists H: w\cdot x+b=0$ s.t.

$w\cdot x_i+b>0,\ y_i=+1$

$w\cdot x_i+b<0,\ y_i=-1$

必存在 H s.t.

间隔 $\frac{2}{\|w\|}$ 最大，此时 $w=\arg\min_H \|w\|$；

下证唯一性.

假设存在2个不同的硬间隔最大超平面

$H_1(w_1,b_1)$

$H_2(w_2,b_2)$

满足 $\|w_1\|=\|w_2\|=c$

考虑 $w=\frac{w_1+w_2}{2}$，$b=\frac{b_1+b_2}{2}$

由于 H_1. H_2 均满足 $y_i(w\cdot x_i+b)-1\geq 0$

故分别有 $y_i(w_1\cdot x_i+b)\geq 1$ ①

$y_i(w_2\cdot x_i+b)\geq 1$ ②，$\forall i$

(①+②)÷2，得

$$y_i\left(\frac{w_1+w_2}{2}x_i+\frac{b_1+b_2}{2}\right)\geq 1$$

即 $y_i(w\cdot x_i+b)\geq 1$

故 (w,b) 也是符合要求的超平面

故 $\|w\|\geq c$

而 $\|w\|=\left\|\frac{w_1+w_2}{2}\right\|\leq\frac{1}{2}(\|w_1\|+\|w_2\|)$

$=\frac{1}{2}(c+c)$

$=c$

由此可见：不等式

$\|w_1+w_2\|\leq\|w_1\|+\|w_2\|$ 等号可以取到

而取等条件即

$w_1=\lambda w_2$，$\lambda>0$ $\in\mathbb{R}$

又由 $\|w_1\|=\|w_2\|$ 知 $\lambda=1$

故 $w_1=w_2=w$；

$H_1(w,b_1)$，$H_2(w,b_2)$，下证 $b_1=b_2$

考虑硬间隔分离超平面的定义，对于

$H_1(w,b_1)$，$\exists x_1',x_1''\in\{x_i\}$，s.t.

$w\cdot x_1'+b_1=1$ ③

$w\cdot x_1''+b_1=-1$ ④

同理对 $H_2(w,b_2)$，$\exists x_2',x_2''\in\{x_i\}$，s.t.

$w\cdot x_2'+b_2=1$ ⑤

$w\cdot x_2''+b_2=-1$ ⑥

并且满足（"正确分类"&边界）

$w\cdot x_1'+b_2\geq 1$ ⑦

$w\cdot x_1''+b_2\leq -1$ ⑧

$w\cdot x_2'+b_1\geq 1$ ⑨

$w\cdot x_2''+b_1\leq -1$ ⑩

~~由③④⑤⑥，解出~~

~~$b_1-b_2=\frac{w(x_2'+x_2''-x_1'-x_1'')}{2}$~~

~~⑦⑧即~~

~~$w\cdot x_1'+b_2\geq$~~

由③⑦，有 $w\cdot x_1'+b_1\leq w\cdot x_1'+b_2$

故 $b_1\leq b_2$

由⑤⑨，有 $w\cdot x_2'+b_2\leq w\cdot x_2'+b_1$

故 $b_2\leq b_1$

因此可推知 $b_1=b_2$；

综上，可知 $H_1(w_1,b_1)$ 与 $H_2(w_2,b_2)$

是同一个超平面

故硬间隔最大超平面唯一 证毕！

T_2

解：考虑约束 $y_i(w\cdot x_i+b)\geqslant 1-\xi_i$

引入 Lagrange 乘子 $\alpha_i\geqslant 0$；

考虑约束 $\xi_i\geqslant 0$

引入 Lagrange 乘子 $\beta_i\geqslant 0$；

构造 Lagrange 函数

$$L(w,b,\xi,\alpha,\beta)=\frac{1}{2}\|w\|^2+C\sum_{i=1}^{N}\xi_i^2$$

$$-\sum_{i=1}^{N}\alpha_i\left[y_i(w\cdot x_i+b)-1+\xi_i\right]$$

$$-\sum_{i=1}^{N}\beta_i\xi_i$$

考虑

$$\nabla_w L=w-\sum_{i=1}^{N}\alpha_i y_i x_i=0\Rightarrow w=\sum_{i=1}^{N}\alpha_i y_i x_i$$

$$\nabla_b L=-\sum_{i=1}^{N}\alpha_i y_i=0\Rightarrow\sum_{i=1}^{N}\alpha_i y_i=0$$

$$\nabla_{\xi_i}L=2C\xi_i-\alpha_i-\beta_i=0\Rightarrow\alpha_i+\beta_i=2C\xi_i$$

得到对偶问题

$$\max_{\alpha}\sum_{i=1}^{N}\alpha_i-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j x_i x_j-\sum_{i=1}^{N}C\xi_i^2$$

即

$$\max_{\alpha}\left[\sum_{i=1}^{N}(\alpha_i-C\xi_i^2)-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j x_i x_j\right]$$

s.t. $\sum_{i=1}^{N}\alpha_i y_i=0$,

$0\leqslant\alpha_i\leqslant 2C\xi_i$