II 核函数：Kernel function

解决非线性可分.

将样本从原始空间映射到更高维空间，使其变成线性可分问题.

如果原始空间是有限维，那么一定存在一个高维特征空间使样本可分.

令$\phi(x)$表示x映射后的特征向量，于是在特征空间中划分超平面所对应的模型可表示为：

$$f(x)=w^T\phi(x)+b.$$

$\Downarrow$

$$\begin{cases} \min\limits_{w,b} \frac{1}{2}\|w\|^2 \\ s.t.\ y_i(w^T\phi(x_i)+b)\geq 1,\quad i=1,2,\cdots,m. \end{cases}$$

$\Downarrow$

对偶问题：$\begin{cases} \max\limits_{\alpha} \sum\limits_{i=1}^{m}\alpha_i-\frac{1}{2}\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{m}\alpha_i\alpha_j y_i y_j \phi(x_i)^T\phi(x_j) \\ s.t.\ \sum\limits_{i=1}^{m}\alpha_i y_i=0 \quad (\alpha_i\geq 0) \end{cases}$

$\phi(x_i)^T\phi(x_j)$是x_i与x_j映射到特征空间之后的内积

直接计算$\phi(x_i)^T\phi(x_j)$通常是困难的

Vazyme

$$K(x_i,x_j)=\langle\phi(x_i),\phi(x_j)\rangle=\phi(x_i)^T\phi(x_j)$$

对偶式重写为：

$$\max_{\alpha}\sum_{i=1}^{m}\alpha_i-\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_i\alpha_j y_i y_j K(x_i,x_j)$$

$$s.t.\ \sum_{i=1}^{m}\alpha_i y_i=0 \quad (\alpha_i\geq 0,\ i=1,2,\cdots,m)$$

求解后即可得到：

$$\begin{aligned} f(x)&=w^T\phi(x)+b \\ &=\sum_{i=1}^{m}\alpha_i y_i\phi(x_i)^T\phi(x)+b \\ &=\sum_{i=1}^{m}\alpha_i y_i K(x,x_i)+b. \end{aligned}$$

定理：$K(\cdot,\cdot)$是定义在$\mathcal{X}\times\mathcal{X}$上的对称函数，当且仅当对于任意数据$D=\{x_1,x_2,\cdots,x_m\}$“核矩阵”$K$总是半正定：

$$K=\begin{bmatrix} K(x_1,x_1) & \cdots & K(x_1,x_n) \\ \vdots & \ddots & \vdots \end{bmatrix}$$