第 9 章　支持向量机习题解答

9.1　解　（1）分类方法

记x_1和x_2分别表示蠓虫的触角和翅膀长度，已知观测样本为$[a_i, y_i]$（$i=1,\cdots,9$），其中$a_i \in R^2$，$y_i=1$表示 Af，$y_i=-1$表示 Apf。

首先进行线性分类，即要找一个最优分类面$(\omega \cdot x)+b=0$，其中$x=[x_1, x_2]$，$\omega \in R^2$，$b \in R$，w, b待定，满足如下条件

$$\begin{cases}(\omega \cdot a_i)+b \geq 1, & y_i=1\text{时}, \\ (\omega \cdot a_i)+b \leq -1, & y_i=-1\text{时},\end{cases}$$

即有$y_i((\omega \cdot a_i)-b) \geq 1$，$i=1,\cdots,n$，其中，满足方程$(\omega \cdot a_i)+b=\pm 1$的样本为支持向量。

要使两类总体到分类面的距离最大，则有

$$\max \frac{2}{\|\omega\|} \Rightarrow \min \frac{1}{2}\|\omega\|^2,$$

于是建立 SVM 的如下数学模型。

模型 1

$$\min \quad \frac{1}{2}\|\omega\|^2,$$

$$\text{s.t.} \quad y_i((\omega \cdot a_i)+b) \geq 1，\quad i=1,2,\cdots,n.$$

求得最优值对应的ω^*, b^*，可得分类函数

$$g(x)=\operatorname{sgn}((\omega^* \cdot x)+b^*).$$

当$g(x)=1$时，把样本归于 Af 类，当$g(x)=-1$时，把样本归于 Apf 类。

模型 1 是一个二次规划模型，为了利用 Matlab 求解模型 1，下面把模型 1 化为其对偶问题。

定义广义拉格朗日函数

$$L(\omega, \alpha)=\frac{1}{2}\|\omega\|^2+\sum_{i=1}^{n} \alpha_i[1-y_i((\omega \cdot a_i)+b)],$$

其中$\alpha=[\alpha_1,\cdots,\alpha_n]^T \in R^{n+}$。

由 KKT 互补条件，通过对ω和b求偏导可得

$$\frac{\partial L}{\partial \omega}=\omega-\sum_{i=1}^{n} \alpha_i y_i a_i=0,$$

$$\frac{\partial L}{\partial b}=\sum_{i=1}^{n} \alpha_i y_i=0,$$

得$\omega=\sum_{i=1}^{n} \alpha_i y_i a_i$，$\sum_{i=1}^{n} \alpha_i y_i=0$，代入原始拉格朗日函数得

$$L=\sum_{i=1}^{n} \alpha_i-\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j (a_i \cdot a_j).$$

于是模型 1 可以化为

模型 2

$$\max \ \sum_{i=1}^{n}\alpha_i-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j (a_i\cdot a_j),$$

$$\text{s.t.}\ \begin{cases}\sum_{i=1}^{n}\alpha_i y_i=0,\\ 0\le\alpha_i, i=1,2,\cdots,n.\end{cases}$$

解此二次规划得到最优解α^*，从而得权重向量$\omega^*=\sum_{i=1}^{n}\alpha_i^* y_i a_i$。

由 KKT 互补条件知

$$\alpha_i^*[1-y_i((\omega^*\cdot a_i)+b^*)]=0,$$

这意味着仅仅是支持向量a_i，使得α_i^*为正，所有其它样本对应的α_i^*均为零。选择α^*的一个正分量α_j^*，并以此计算

$$b^*=y_j-\sum_{i=1}^{n}y_i\alpha_i^*(a_i\cdot a_j).$$

最终的分类函数表达式如下

$$g(x)=\mathrm{sgn}(\sum_{i=1}^{n}\alpha_i^* g_i(t_i\cdot x)+b^*). \tag{9.1}$$

实际上，模型 2 中的$(a_i\cdot a_j)$是核函数的线性形式。非线性核函数可以将原样本空间线性不可分的向量转化到高维特征空间中线性可分的向量。

将模型 2 换成一般的核函数$K(x,y)$，可得一般的模型。

模型 3

$$\max \ \sum_{i=1}^{n}\alpha_i-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j K(a_i,a_j),$$

$$\text{s.t.}\ \begin{cases}\sum_{i=1}^{n}\alpha_i y_i=0,\\ 0\le\alpha_i, i=1,2,\cdots,n.\end{cases}$$

分类函数表达式

$$g(x)=\mathrm{sgn}(\sum_{i=1}^{n}\alpha_i^* y_i K(a_i,x)+b^*). \tag{9.2}$$

（2）未知样本的分类

使用模型 1 或模型 2，利用 Matlab 软件，把 3 个待判定的样本点全部判为 Apf 类，且该方法对已知样本点的误判率为 0。

计算的 Matlab 程序如下

```
clc, clear
x0=[1.24,1.27; 1.36,1.74; 1.38,1.64; 1.38,1.82; 1.38,1.90; 1.40,1.70
    1.48,1.82; 1.54,1.82; 1.56,2.08; 1.14,1.82; 1.18,1.96; 1.20,1.86
    1.26,2.00; 1.28,2.00; 1.30,1.96]; %输入已知样本数据
x=[1.24,1.80; 1.28,1.84; 1.40,2.04]; %输入待判样本点数据
```

```
group=[ones(9,1); -ones(6,1)]; %输入已知样本标志
s=svmtrain(x0,group); %使用线性核函数训练支持向量机的分类器
check=svmclassify(s,x0)    %对已知样本点进行检验
solution=svmclassify(s,x) %对未知样本点进行判别
```

9.2 解 问题

$$\min \ \|\omega\|^2 + c_1\sum_{i=1}^{n}\xi_i\,,$$

$$\text{s.t.}\quad g_i((\omega\cdot x_i)+b)\ge 1-\xi_i\,,\quad i=1,2,\cdots,n\,,$$

$$\xi_i\ge 0\,,\quad i=1,2,\cdots,n\,.$$

与问题

$$\min\|\omega\|^2 + c_1\sum_{i=1}^{n}\xi_i + c_2\sum_{i=1}^{n}\xi_i^2\,,$$

$$\text{s.t.}\quad g_i((\omega\cdot x_i)+b)\ge 1-\xi_i\,,\quad i=1,2,\cdots,n\,,$$

$$\xi_i\ge 0\,,\quad i=1,2,\cdots,n\,.$$

是等价的。我们只需考虑参数c_1变化产生的影响即可。

当参数$c_1\to 0$时，即目标函数的惩罚因子较小，允许ξ_i取较大的值。当$c_1\to+\infty$，不允许ξ_i取正值，问题等价于

$$\min \ \|\omega\|^2\,,$$

$$\text{s.t.}\quad g_i((\omega\cdot x_i)+b)\ge 1\,,\quad i=1,2,\cdots,n\,.$$

下面给出对偶表示形式。首先引入 Lagrange 函数

$$L(\omega,b,\xi,\alpha,\beta)=\|\omega\|^2+c_1\sum_{i=1}^{n}\xi_i+c_2\sum_{i=1}^{n}\xi_i^2-\sum_{i=1}^{n}\alpha_i(g_i[(\omega\cdot x_i)+b]-1+\xi_i)-\sum_{i=1}^{n}\beta_i\xi_i\,,$$

其中$\alpha_i\ge 0$和$\beta_i\ge 0$，根据 Wolf 对偶定义，对L关于w,b,ξ求极小，即

$$\nabla_w L(\omega,b,\xi,\alpha,\beta)=0\,,\quad \nabla_b L(\omega,b,\xi,\alpha,\beta)=0\,,\quad \nabla_\xi L(\omega,b,\xi,\alpha,\beta)=0\,,$$

得到

$$w=\frac{1}{2}\sum_{i=1}^{n}\alpha_i g_i x_i\,,$$

$$\sum_{i=1}^{n}\alpha_i g_i=0\,,$$

$$c_1+2c_2\xi_i-\alpha_i-\beta_i=0\,.$$

然后将上述极值条件代入 Lagrange 函数，对α求极大，得到对偶问题

$$\max_{\alpha}\ -\frac{1}{4}\sum_{i=1}^{n}\sum_{j=1}^{n}g_i g_j\alpha_i\alpha_j(x_i,x_j)+\sum_{i=1}^{n}\alpha_i\,,$$

$$\text{s.t.}\quad \sum_{i=1}^{n}g_i\alpha_i=0\,,$$

$$0\le\alpha_i\le c_1\,,\quad i=1,\cdots,n\,.$$