第 7 章 支持向量机

7.1 对于样本空间中的一划分超平面 $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}+b=0$，有 $\boldsymbol{w}=(-1, 3, 2)$，$b=1$。则判断如下向量是否为支持向量，并求出间隔。

(1) $\boldsymbol{x}_1=(4, -2, 2)$

(2) $\boldsymbol{x}_2=(2, 5, -6.5)$

(3) $\boldsymbol{x}_3=(4, -2, 4)$

解答：使下列公式满足等号的样本成为“支持向量”

$$\begin{cases} \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i+b\geqslant +1, & y_i=+1\ ; \\ \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i+b\leqslant -1, & y_i=-1\ . \end{cases}$$

（1）$w^Tx_i+b=-1*4+(-2)*3+2*2+1=-5$，显然不是

（2）$w^Tx_i+b=-1*2+3*5+2*(-6.5)+1=1$，满足条件，是支持向量

（3）$w^Tx_i+b=-1*4+3*(-2)+2*4+1=-1$，满足条件，是支持向量

间隔$\gamma=\dfrac{2}{||w||}=\dfrac{2}{\sqrt{1+9+4}}=\dfrac{\sqrt{14}}{7}$

7.2 假设输入空间是 $\mathbf{R}^2$，核函数是$\kappa(\boldsymbol{x},\boldsymbol{z})=(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{z})^2$，试找出其相关的特征空间$\mathcal{H}$和映射 $\Phi(\boldsymbol{x})$：

1) 取特征空间$H=R^3$，记$x=\left(x^{(1)},x^{(2)}\right)^T,z=\left(z^{(1)},z^{(2)}\right)^T$，而

$$(x\cdot z)^2=\left(x^{(1)}z^{(1)}+x^{(2)}z^{(2)}\right)^2=\left(x^{(1)}z^{(1)}\right)^2+2x^{(1)}z^{(1)}x^{(2)}z^{(2)}+\left(x^{(2)}z^{(2)}\right)^2,$$

所以我们得到映射$\phi(x)=\left(\left(x^{(1)}\right)^2,\sqrt{2}\,x^{(1)}x^{(2)},\left(x^{(2)}\right)^2\right)^T$，易得核函数

$$\phi(x)\cdot\phi(z)=(xz)^2=K(x,z)$$

2) 取特征空间$H=R^3$，映射还可以为

$$\phi(x)=\frac{1}{\sqrt{2}}\left(\left(x^{(1)}\right)^2-\left(x^{(2)}\right)^2,2x^{(1)}x^{(2)},\left(x^{(1)}\right)^2=\left(x^{(2)}\right)^2\right)^T$$

3) 取特征空间$H=R^4$，映射此时为$\phi(x)=\left(\left(x^{(1)}\right)^2,x^{(1)}x^{(2)},x^{(1)}x^{(2)},\left(x^{(2)}\right)^2\right)^T$

7.3 设计几种 SVM 实现多分类的方案。

SVM 多分类的实现主要从两种思路：

1) 直接法，也即直接修改目标函数，将多分类的参数通过计算合并到一个最优化的问题，通过对这个问题的求解得到多分类的结果

2) 间接法，也即组合多个二分类器实现多分类。通过查阅资料可得间接法主要有 one-against-one 和 one-against-all 两种方法

- **一对多法（one-versus-rest,简称 OVR SVMs）**

  训练时依次把某个类别的样本归为一类,其他剩余的样本归为另一类，这样 k 个类别的样本就构造出了 k 个 SVM。分类时将未知样本分类为具有最大分类函数值的那类。

  假如我有四类要划分（也就是 4 个 Label），他们是 A、B、C、D。

  于是我在抽取训练集的时候，分别抽取

  （1）A 所对应的向量作为正集，B，C，D 所对应的向量作为负集；

  （2）B 所对应的向量作为正集，A，C，D 所对应的向量作为负集；

  （3）C 所对应的向量作为正集，A，B，D 所对应的向量作为负集；

  （4）D 所对应的向量作为正集，A，B，C 所对应的向量作为负集；

  使用这四个训练集分别进行训练，然后的得到四个训练结果文件。

  在测试的时候，把对应的测试向量分别利用这四个训练结果文件进行测试。最后每个测试都有一个结果 f1(x),f2(x),f3(x),f4(x)。于是最终的结果便是这四个值中最大的一个作为分类结果。

- 一对一法（**one-versus-one,**简称 **OVO SVMs** 或者 **pairwise**）

其做法是在任意两类样本之间设计一个 SVM，因此 k 个类别的样本就需要设计 k(k-1)/2 个 SVM。当对一个未知样本进行分类时，最后得票最多的类别即为该未知样本的类别。

假设有四类 A,B,C,D 四类。在训练的时候我选择 A,B; A,C; A,D; B,C; B,D;C,D 所对应的向量作为训练集，然后得到六个训练结果，在测试的时候，把对应的向量分别对六个结果进行测试，然后采取投票形式，最后得到一组结果。

投票是这样的：

A=B=C=D=0;

(A,B)-classifier 如果是 A win,则 A=A+1;otherwise,B=B+1;

(A,C)-classifier 如果是 A win,则 A=A+1;otherwise, C=C+1;

...

(C,D)-classifier 如果是 C win,则 C=C+1;otherwise,D=D+1;

The decision is the Max(A,B,C,D)

分别选取两个不同类别构成一个 SVM 子分类器，这样对于 K 个类别来说，共有（k*(k-1)/2）个分类器。在构造 i 和 j 的分类器时，可以将类别 i 的训练样本置为 1，j 的样本置为-1 来进行训练。

在进行测试的时候，使用最多的就是 Friedman 提出的投票策略：将测试数据 x 对所有的分类器分别进行测试，若由得到 x 属于第 i 类，则第 i 类加 1，属于 j 类，则第 j 类投票加 1.累计各类别的得分，选择得分最高者所对应的类别为测试数据的类别。