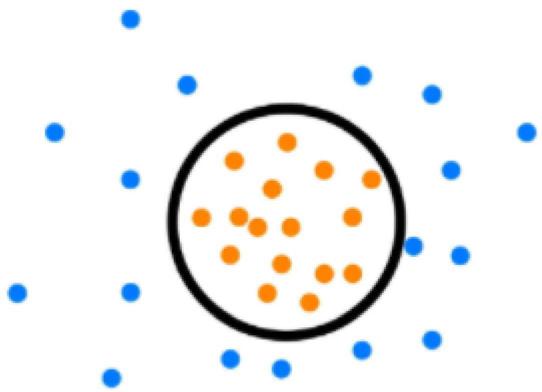
OneClassSVM



- 算法思想
    - 找到一个超球面，球面中心为a, 半径为R 使得所有正常样本在球面内，同时球面尽可能的小

  此时目标函数与约束条件为：

$$
\begin{aligned}
&\min_{R,a} R^2 \\
&s.t.\ \|\mathbf{x}_i - \mathbf{a}\|^2 \le R^2
\end{aligned}
$$

    - 由于一般情况下不能通过球面将正常样本和非正常样本完全分开，因此如果球面内包含所有正常样本，则可能包含更多的非正常样本。因此给样本点 x_i引入松弛条件 ξ_i ,允许部分正常样本点在球面之外，同时对松弛距离在目标函数加入惩罚项，C为惩罚系数,C>0 ,根据需要人为设定。

      此时目标函数与约束条件为：

$$
\begin{aligned}
&\min_{R,a} R^2 + C\sum_i \xi_i \\
&s.t.\ \|\mathbf{x}_i - \mathbf{a}\|^2 \le R^2 + \xi_i, \quad \xi_i \ge 0 \quad \forall i
\end{aligned}
$$

- 求解步骤
    1. 通过拉格朗日乘子法转化将条件极值转化为无约束极值

       拉格朗日函数如下：

$$
\begin{aligned}
L(R, \mathbf{a}, \alpha_i, \gamma_i, \xi_i) =& R^2 + C\sum_i \xi_i \\
&- \sum_i \alpha_i \left\{ R^2 + \xi_i - \left(\|\mathbf{x}_i\|^2 - 2\mathbf{a}\cdot\mathbf{x}_i + \|\mathbf{a}\|^2 \right) \right\} - \sum \gamma_i \xi_i
\end{aligned}
$$

$\alpha_i >= 0$，$\gamma_i >= 0$,对拉格朗日函数求最大值的条件等价于原条件极值的条件。此时原问题转化为先对拉格朗 日函数求最大值再求最小值即：

$$\min_{R,a,\xi} \max_{\alpha,\gamma} L \tag{1}$$

2. 转化为对偶问题求解

$$\min_{R,a,\xi} \max_{\alpha,\gamma} L = \max_{\alpha,\gamma} \min_{R,a,\xi} L \tag{2}$$

根据 $\min_{R,a,\xi} L$ 求偏导

$$\begin{aligned} \frac{\partial L}{\partial R} = 0: &\quad \sum_i \alpha_i = 1 \\ \frac{\partial L}{\partial \mathbf{a}} = 0: &\quad \mathbf{a} = \frac{\sum_i \alpha_i \mathbf{x}_i}{\sum_i \alpha_i} = \sum_i \alpha_i \mathbf{x}_i \\ \frac{\partial L}{\partial \xi_i} = 0: &\quad C - \alpha_i - \gamma_i = 0 \end{aligned}$$

将等式代入L得到:

$$L = \sum_i \alpha_i (\mathbf{x}_i \cdot \mathbf{x}_i) - \sum_{i,j} \alpha_i \alpha_j (\mathbf{x}_i \cdot \mathbf{x}_j)$$

此时目标函数和约束条件如下：约束条件由求偏导的第三条等式可得

$$\max_{\alpha} \sum_i \alpha_i (\mathbf{x}_i \cdot \mathbf{x}_i) - \sum_{i,j} \alpha_i \alpha_j (\mathbf{x}_i \cdot \mathbf{x}_j) \text{ subject to } 0 \le \alpha_i \le C$$

此时仅有 α 未知 可根据二次规划求解 α，然后根据 α 和求偏导第三条等式得到 γ ,根据 α 和第二条求偏导等式得到 a

根据KTT条件即原来的约束条件乘以拉格朗日乘数等于零即

$$\begin{aligned} &\gamma_i \xi_i = 0 \\ &||x_i - a||^2 - R^2 - \xi_i = 0 \end{aligned} \tag{3}$$

分别得到 ξ_i 和 R

- 最终结果即为超球面$(x-a)^2 <= R^2$

分类器对样本 x 的判断为

$$f(x) = sign(R^2 - (x-a)^2) \tag{4}$$

- 核函数定义

  定义两个向量的内积函数 $K(x,z)$

  此时存在函数 ϕ 使得 $\phi(x)$ 和 $\phi(z)$ 的内积 $\phi(x)\cdot\phi(z) = K(x,z)$, 同一个核函数 K可能对应多个映射函数 ϕ

  ϕ 一般将向量 x 映射到高维空间，或对向量 x 的每个维度进行非线性变换，由于目标函数可以转化为关于测试数据 $\bar{x}$ 的项都是和已知样本 x_i 的内积项，因此仅考虑给出核函数 K 无需寻找对于的映射函数 ϕ

- 定义核函数之后的超平面球面参数求解过程

  1. 求解 α_i

     根据最终化简的二次优化问题

     $$\max_{\alpha}\sum_i \alpha_i\left(\mathbf{x}_i\cdot\mathbf{x}_i\right)-\sum_{i,j}\alpha_i\alpha_j\left(\mathbf{x}_i\cdot\mathbf{x}_j\right)\ \text{subject to}\ 0\le\alpha_i\le C$$

     将目标函数中的 $x_i\cdot x_i$ 和 $x_i\cdot x_j$ 替换为 $K(x_i,x_i)$ 和 $K(x_i,x_j)$ 求解 α_i ，进而根据偏导第三条方程求得求得 γ

  2. 根第二条方程替换 a 为 $\sum_i \alpha_i x_i$

  3. 因此判别函数为

$$\begin{aligned} f(x) &= sign(R^2-(x-\sum_i \alpha_i x_i)^2) \\ f(x) &= sign(R^2-x^2+2\sum_i x x_i+(\sum \alpha_i x_i)^2) \\ f(x) &= sign(R^2-K(x,x)+2\sum K(x,x_i)+\sum_i\sum_j \alpha_i\alpha_j K(x_i,x_j)) \end{aligned} \tag{6}$$

- 常见的核函数

  线性核（Linear Kernel)

  $$k(x,y)=x^Ty+c$$

  多项式核（Polynomial Kernel)

  $$k(x,y)=\left(ax^Ty+c\right)^d$$

  径向基核函数（Radial Basis Function)

  $$k(x,y)=\exp\left(-\gamma\|x-y\|^2\right)$$

  也叫**高斯核（Gaussian Kernel)**，因为可以看成如下核函数的领一个种形式：

  $$k(x,y)=\exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right)$$

对于RBF核函数 $\gamma\rightarrow+\infty$ 任意两点的距离越趋于0，越不可分。$\gamma\rightarrow-\infty$ 任意两点距离趋于无穷，太稀疏，

$\gamma\rightarrow 0$，任意两点距离趋于1。