支持向量机预测原理

支持向量机（Support Vector Machine，SVM）预测算法根据给定的训练样本 $S=\{(x_1,y_1),...,(x_l,y_l)\}\subset R^n\times R$，可以构建回归函数：

$$y_i=w^T\varphi(x_i)+b+e_i \tag{1}$$

其中，e_i 为误差。定义损失函数为：

$$I(y,f(x))=(y-f(x))^2=e^2 \tag{2}$$

其中，y 为实际值，$f(x)$为预测值。

最小二乘支持向量机（LSSVM）根据结构最小化原理，定义其风险函数为：

$$J_I=\frac{1}{2}\|\omega\|^2+\gamma\sum_{i=1}^{l}e_i^2 \tag{3}$$

由此推得原始最优化问题：

$$\min J_I(\omega,e)=\frac{1}{2}\|\omega\|^2+\gamma\sum_{i=1}^{l}e_i^2 \tag{4}$$

其中，γ为正实数，表示惩罚系数。它能够调节模型精度与模型复杂度的关系，在两者之间寻求一种折中，使所训练出来的模型具有比较强的泛化能力。

构造新的拉格朗日乘子并置零：

$$\begin{cases}\dfrac{\partial L}{\partial\omega}=0\rightarrow\omega=\sum_{i=1}^{l}\alpha_i\phi(x_i)\\ \dfrac{\partial L}{\partial b}=0\rightarrow\sum_{i=1}^{l}\alpha_i=0\\ \dfrac{\partial L}{\partial e_i}=0\rightarrow\alpha_i=\gamma e_i\\ \dfrac{\partial L}{\partial\alpha_i}=0\rightarrow b=y_i-\omega^T\phi(x_i)-e_i\\ i=1,\ldots\ldots l\end{cases} \tag{5}$$

消去 ω 和 e_i，得到以下行方程：

$$\begin{bmatrix}0 & {I_v}^T\\ I_v & \Omega+{D_\gamma}^{-1}\end{bmatrix}\begin{bmatrix}b\\ \alpha\end{bmatrix}=\begin{bmatrix}0\\ y\end{bmatrix} \tag{6}$$

其中，$I_v=[1,1,\ldots\ldots1]^T$，对角阵 $D_\gamma=diag[\gamma,\gamma,\ldots\ldots,\gamma],y=[y_1,y_2,\ldots\ldots,y_l]^T$，

$\alpha=[\alpha_1,\alpha_2,\ldots\ldots,\alpha_l]^T$，矩阵$\Omega$的各元素为

$$\Omega_{ij}=\phi(x_i)^T\phi(x_j)=K(x_i,x_j) \tag{7}$$

其中，$K(x,x_i)$是满足 Mercer 条件的核函数，它把原始空间中的特征量映射到高维空间中去，将原始空间中的二次规划问题转到高维空间的线性问题中去解决。径向基核函数的泛化能力强，结构相对简单，因此选用径向基函数作为预测问题的核函数，其表达式为：

$$k(x,x^*)=\exp(-\frac{\left\|x-x^*\right\|}{2\sigma^2}) \tag{8}$$

对式（6）进一步求解得到变量α与b，则预测模型为：

$$y(x)=\sum_{i=1}^{l}\alpha_i K(x,x_i)+b \tag{9}$$

SVM 预测流程如图 1 所示。

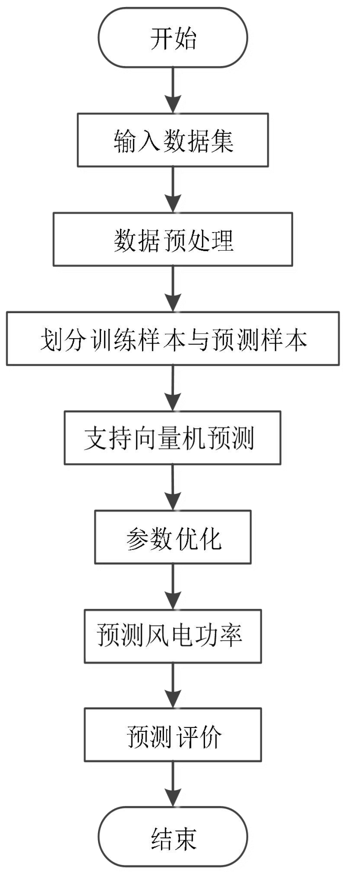

图 1 预测流程图

惩罚系数γ和径向基核函数中核参数σ的取值，直接决定 SVM 预测的性能，

因此采取参数寻优方法优化支持向量机功率预测的效果。通常采用网格法找出惩罚系数和核参数，在模型上用交叉验证的方式验证参数的预测误差，找出误差最小的那对参数，即为最优解。