

文章编号:1673-0062(2011)01-0018-05

支持向量机在混凝土强度预测中的应用研究

王 志,丁德馨*,张志军

(南华大学 核资源与核燃料工程学院,湖南 衡阳 421001)

摘 要:支持向量机是基于统计学习理论框架下的一种新的通用机器学习方法.文中提出了基于支持向量机的混凝土强度预测方法,并在 MATLAB 中编制了相应的支持向量机程序,建立了相应的混凝土强度预测模型.以实例数据为学习样本和测试样本讨论了基于支持向量机的混凝土强度预测方法及可行性.研究表明支持向量机可以较好地表达混凝土强度与其影响因素之间的非线性映射关系.用支持向量机来预测混凝土强度是可行的,它为预测混凝土强度提供了一种新的方法.

关键词:混凝土强度;支持向量机;预测;参数分析

中图分类号:TU27 **文献标识码:**B

Study of Support Vector Machine and Its Application in Concrete Strength Prediction

WANG Zhi, DING De-xin*, ZHANG Zhi-jun

(School of Nuclear Resources and Nuclear Fuel Engineering, University of South China,
Hengyang, Hunan 421001, China)

Abstract: SVM is a new machine learning method based on the statistical learning theory. This paper proposed forecasting methods of concrete strength based on SVM, and the corresponding SVM program and concrete strength prediction model has been established by MATLAB. The instance data for learning samples and test samples were used to discuss forecasting methods of concrete strength and feasibility based on SVM. The results indicate that it is able to express the nonlinear mapping relation between the strength of concrete and its influence factors better. Using SVM to predict the strength of concrete is feasible. It will provide a new method to predict concrete strength.

key words: concrete strength; support vector machine (SVM); prediction; parameter analysis

混凝土的抗压强度是混凝土结构设计的重要依据,按照传统的方法通常要对混凝土试件进行

收稿日期:2010-09-12

作者简介:王 志(1985-),男,湖南衡阳人,南华大学核资源与核燃料工程学院硕士研究生.主要研究方向:岩土工程.* 通讯作者.

标准养护后通过测试获得. 如果能在浇筑混凝土后数小时内得到其预期的强度对于提高施工的质量和进度具有实用价值. 因此许多试验和分析的方法被研究开发出来用于预测早期混凝土的强度. 支持向量机在对样本数据拟合与预测时有着非常高的精确度^[1-2]. 支持向量机是基于统计学习理论的一种新的学习方法, 最早 Vapnik 教授及其合作者提出^[3], 由于它的优良特性, 引起了许多研究者的兴趣. 本文采用支持向量机的函数拟合方法来预测混凝土的抗压强度, 应用 MATLAB^[4] 中的优化工具箱, 并在 MATLAB 中编制了相应支持向量机的程序. 利用编制的支持向量机程序, 通过对 30 个训练样本的学习和 10 个测试样本的预测, 讨论了运用支持向量机进行混凝土抗压强度预测的可行性和有效性.

1 支持向量机的基本理论

1.1 支持向量机的基本思想

支持向量机的基本思想是通过用内积函数定义的非线性变换将输入空间变换到一个高维空间, 在这个高维空间中寻找输入变量和输出变量之间的一种非线性关系, 其基本结构如图 1. 支持向量机有着严格的理论基础, 采用结构风险最小化原则, 具有很好的推广能力; 支持向量机算法是一个凸二次优化问题^[5-6], 保证找到的解是全局最优解; 能较好地解决小样本、非线性、高维数等实际问题, 因此支持向量机是当今研究的热点问题.

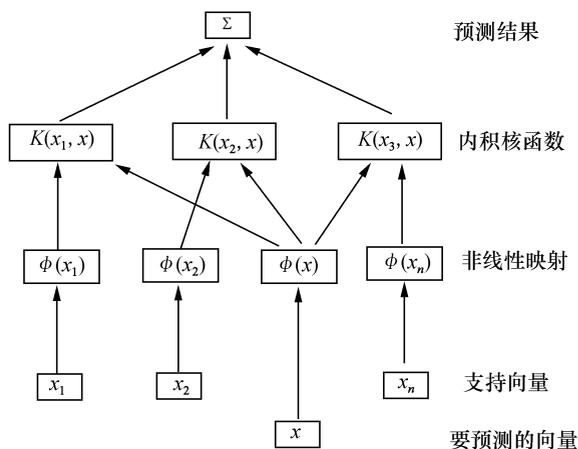


图 1 支持向量机结构示意图

Fig.1 Sketch of support vector machines

1.2 支持向量机方法简介

支持向量机函数拟合分为线性拟合和非线性

拟合^[7], 首先考虑用线性拟合函数 $f(x) = \omega x + b$ 拟合数据 $\{x_i, y_i\}, i = 1, 2, 3, \dots, n, x_i \in R^n, y_i \in R$ 的问题. 假设所有训练数据, 在 ε 精度下无误差地用线性函数拟合即

$$\begin{aligned} y_i - \omega x_i - b &\leq \varepsilon \\ \omega x_i + b - y_i &\leq \varepsilon \end{aligned} \quad (i = 1, 2, \dots, k) \quad (1)$$

优化目标是使 $\frac{1}{2} \|\omega\|^2$ 最小. 根据统计学习理论, 在这个优化目标下可取得较好的推广能力. 考虑到允许误差的情况, 引入松弛因子 $\xi_i \geq 0$ 和 $\xi_i^* \geq 0$, 则式(1)变为

$$\begin{aligned} y_i - \omega x_i - b &\leq \varepsilon + \xi_i \\ \omega x_i + b - y_i &\leq \varepsilon + \xi_i^* \end{aligned} \quad (i = 1, 2, \dots, k) \quad (2)$$

优化目标是使 $\frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^k (\xi_i + \xi_i^*)$ 最小, 其中常数 $C > 0$ 表示对超出误差 ε 的样本的惩罚度. 通常可采用优化方法得到其对偶问题.

其函数表达式为:

$$\begin{aligned} \max W(a, a^*) = &-\frac{1}{2} \sum_{i,j=1}^k (a_i - a_i^*)(a_j - a_j^*) \times \\ &(x_i x_j) + \sum_{i=1}^k y_i (a_i - a_i^*) - \varepsilon \sum_{i=1}^k (a_i + a_i^*) \end{aligned} \quad (3)$$

令:

$$\sum_{i=1}^k (a_i - a_i^*) = 0 \quad (4)$$

$$0 \leq a_i, a_i^* \leq c, i = 1, 2, \dots, n$$

由式(3)、式(4)可得到支持向量机拟合函数为:

$$\begin{aligned} f(x) &= \omega x + b \\ &= \sum_{i=1}^k (a_i - a_i^*) (x x_i) + b \end{aligned} \quad (5)$$

其中, a_i, a_i^* 将只有小部分不为 0, 它们对应的样本就是支持向量.

对于非线性问题, 可以通过非线性变换将原问题映射到某个高维特征空间中的线性问题进行求解. 在高维特征空间中, 线性问题中的内积运算可用核函数来代替, 即 $K(x_i, x_j) = \Phi(x_i) \Phi(x_j)$, 核函数可以用原空间中的函数来实现, 没有必要知道非线性变换的具体形式. 即式(3), 式(4), 式(5)可变为如下形式:

$$\max W(a, a^*) = -\frac{1}{2} \sum_{i,j=1}^k (a_i - a_i^*)(a_j - a_j^*) \times$$

$$K(x_j, x_j) + \sum_{i=1}^k y_i (a_i - a_i^*) - \varepsilon \sum_{i=1}^k (a_i + a_i^*) \quad (6)$$

令:

$$\sum_{i=1}^k (a_i - a_i^*) = 0 \quad (7)$$

$$0 \leq a_i, a_i^* \leq c, i = 1, 2, \dots, k$$

$$f(x) = \omega x + b$$

$$= \sum_{i=1}^k (a_i - a_i^*) K(x, x_i) + b \quad (8)$$

目前常用的核函数主要有:

多项式核函数:

$$K(x, y) = (x \cdot y + 1)^d, d \text{ 为阶数} \quad (9)$$

径向基核函数:

$$K(x, y) = \exp\{-\gamma \|x - y\|^2\} \quad (10)$$

Sigmoid 核函数:

$$K(x, y) = \tanh\{\omega(x \cdot y) + c\} \quad (11)$$

2 混凝土强度预测

2.1 混凝土强度影响因素

混凝土是一种由水泥、粗细骨料等加适量的水混合而成的人工材料。影响其强度的因素有水泥的强度等级、水灰比、水的用量、水泥用量、粗骨料的尺寸、砂的细度模数、砂率、骨灰比、坍落度、成型条件、养护条件和测试条件等。其中成型条件、养护条件和测试条件等3个因素在标准养护下符合规范要求可视为不变量。

2.2 混凝土强度预测模型

假设有如下混凝土强度的实测数据 (x_i, y_i) , $i = 1, 2, \dots, k$, $x_i \in R^n$ 为第 i 个样本的影响因素 $y_i \in R$ 为混凝土强度第 i 个实测值。建立混凝土强度的支持向量机模型即寻找函数关系:

$$y_i = f(x_i) \quad (12)$$

$$f: R^n \rightarrow R$$

式中: R^n 为影响混凝土强度的9个因素,即水泥的强度等级,水灰比,水的用量,水泥用量,粗骨料的尺寸,砂的细度模数,砂率,骨灰比,坍落度; R 为混凝土强度的指标。

根据支持向量机理论,混凝土强度预测模型的建立,也即寻求如下的表达式成立:

$$f(x) = \omega x + b = \sum_{i=1}^k (a_i - a_i^*) K(x, x_i) + b \quad (13)$$

式中: x 为要预测的混凝土强度的影响因素; x_i 为 k 个样本中第 i 个样本; $K(x, x_i)$ 为核函数,可在式(9),式(10),式(11)中选择。

2.3 学习样本和测试样本的选择

本文从文献[2]中收集了40个混凝土的实例,对于混凝土强度预测模型,随机抽取30个样本作为预测模型的学习样本,其余的10个样本用于测试建立的模型见表1,学习样本见表2。

表1 测试样本

Table 1 Test samples

序号	水泥强度 /MPa	水泥用量 /($\text{kg} \cdot \text{m}^{-3}$)	水灰比	骨灰比	用水量 /($\text{kg} \cdot \text{m}^{-3}$)	砂率/%	最大粒径 /mm	砂细模数	坍落度 /cm	28 d 强度 /MPa
1	48	426	0.46	2.78	194	33	31.5	1.9	2	34.5
2	48	464	0.41	2.56	190	32	31.5	1.9	2	35.4
3	48	454	0.42	2.54	192	32	31.5	1.9	2.3	35.9
4	59.3	274	0.62	4.405	170	38	40	2.6	4.5	28.2
5	48	431	0.46	2.682	199	35	40	1.9	3.5	34.2
6	48	256	0.76	5.1	195	33	40	1.9	5.1	15.1
7	48	260	0.77	4.64	200	38	40	1.9	7	15.6
8	48	266	0.78	4.46	208	38	40	1.9	5.2	14.6
9	48	254	0.79	4.84	200	37	40	1.9	5	14.7
10	59.3	346	0.51	3.604	176	33	40	2.6	4.5	35.6

2.4 模型选择及参数的确定

混凝土强度预测模型的建立,主要是选择相应的支持向量机参数^[8]:核函数和惩罚因子 C ,它们对预测结果的影响很大,它们的合理确定直接影响着模型的精度和推广能力。本文通过对各种

核函数进行测试,最终确定:预测模型的核函数为径向基核函数,最优参数为 $C = 4, \gamma = 0.0625$;即可得到混凝土强度的支持向量机预测模型。利用上面建立的模型对表1中的10个样本进行预测,预测结果见表3,测试样本的预测结果图见图2。

表2 学习样本
Table 2 Learning samples

序号	水泥强度 /MPa	水泥用量 /($\text{kg} \cdot \text{m}^{-3}$)	水灰比	骨灰比	用水量 /($\text{kg} \cdot \text{m}^{-3}$)	砂率/%	最大粒径 /mm	砂细模数	坍落度 /cm	28 d 强度 /MPa
1	48	382	0.61	2.99	235	36	31.5	1.9	6	25.9
2	48	330	0.65	3.42	214	39	31.5	1.9	4.5	23.5
3	59.3	279	0.61	4.265	170	39	40	2.6	5.5	27.4
4	48	335	0.58	3.75	195	32	40	1.9	6	24.2
5	48	339	0.57	3.79	193	31.5	40	1.9	5	28.4
6	48	375	0.62	3.07	231	36	31.5	1.9	4	27
7	48	374	0.61	3.3	196	31	40	1.9	7.5	28
8	59.3	315	0.65	3.543	205	40	31.5	1.9	1	23.3
9	59.3	321	0.7	3.315	225	42	31.5	1	3	30
10	48	507	0.47	2.36	236	29.5	31.5	1.9	4	32.6
11	48	261	0.77	4.68	202	37	40	1.9	5.5	14.8
12	48	266	0.8	4.73	213	35	40	1.1	6.5	14.9
13	48	260	0.79	4.707	206	36	40	1.9	5	17.5
14	48	265	0.79	4.713	209	36	40	1.9	4.4	19.3
15	48	298	0.68	4.034	204	36	40	1.9	8	20.1
16	48	260	0.75	4.692	194	37	40	1.9	6.5	19
17	59.3	400	0.48	2.948	193	34	31.5	2.6	0.5	38.5
18	48	291	0.69	4.254	200	35	40	1.9	6	19.4
19	48	361	0.54	3.576	195	30	40	1.9	5.7	26
20	48	323	0.58	3.901	189	34	40	1.9	5	27.4
21	48	322	0.62	3.835	200	34	40	1.9	5.5	23.6
22	48	364	0.55	3.481	200	31	40	1.9	7.1	24.7
23	59.3	400	0.49	2.78	195	37	31.5	1.9	0.5	45.2
24	48	265	0.79	4.713	209	36	40	1.9	4.4	19.3
25	48	298	0.68	4.034	204	36	40	1.9	8.0	20.1
26	48	260	0.75	4.692	194	37	40	1.9	6.5	19
27	59.3	400	0.48	2.948	193	34	31.5	2.6	0.5	38.5
28	48	291	0.69	4.254	200	35	40	1.9	6.0	19.4
29	48	361	0.54	3.576	195	30	40	1.9	5.7	26
30	48	323	0.58	3.901	189	34	40	1.9	5.0	27.4

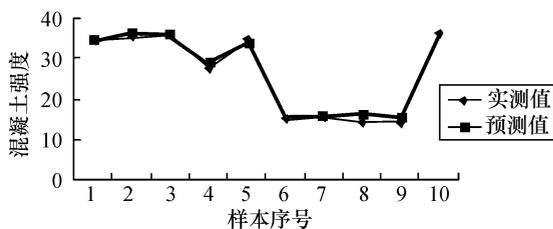


图2 测试样本的预测结果

Fig. 2 Predicted results of test samples

3 模型参数对预测结果的影响及成果分析

在模型的建立过程中,考虑到支持向量机参数对预测结果的影响,对大量的参数进行了测试,最终确定了本文中的参数。

根据已经确定的混凝土预测模型,对表1中的10个样本进行预测,得出相应的预测混凝土强度,如表3所示。从表3可看出除了样本8的相

对误差达到 11.75% 之外,其他样本的相对误差均较小,平均误差为 2.22%,这表明用 SVM 进行混凝土强度预测是合理可靠的。

表 3 支持向量机预测结果与实际值对比
Table 3 Comparison of SVM-predicted results with measured values

序号	实际值	预测值	绝对误差	相对误差%
1	34.5	34.1358	-0.3642	1.06
2	35.4	36.1191	0.7191	2.03
3	35.9	35.6606	-0.2394	0.67
4	28.2	28.4486	0.2486	0.88
5	34.2	33.9554	-0.2446	0.72
6	15.1	15.3332	0.2332	1.54
7	15.6	15.5873	-0.0127	0.08
8	14.6	16.3159	1.7159	11.75
9	14.7	15.1128	0.4128	2.81
10	35.6	35.3575	-0.2425	0.68
平均值			0.4431	2.22

4 结 论

1) 混凝土强度受很多因素的影响,并且它们之间的关系是高维的、高度非线性的、复杂的关系,用传统的建模方法很难处理,本文的支持向量机方法很好地处理了这种关系。

2) 利用支持向量机进行建模比较简单,不需

要对数据进行过多的预处理,并且由于支持向量机算法对应的为二次规划问题,找到的解为全局最优解。

3) 对于支持向量机,其拟合和预测效果关键是如何选择核函数及其参数、惩罚因子 C 。迄今为止尚无这方面的方法和结论,有待于进一步研究,使其具有普遍性,以便得以推广。

参考文献:

- [1] 武换娥,丁圣果,巩玉志,等. 支持向量回归机在混凝土强度预测中的应用研究[J]. 工业建筑,2007,37(增刊):991-995.
- [2] 王继宗,倪鸿光,何锦云,等. 混凝土强度预测和模拟的智能化方法[J]. 土木工程学报,2003,36(10):24-29.
- [3] Vapnik V. The nature of statistical learning theory[M]. New York:Springer Verlag,1995.
- [4] 管爱红,张红梅,杨铁军,等. MATLAB 基础与应用教程[M]. 北京:电子工业出版社,2009.
- [5] 邓乃扬,田英杰. 支持向量机——理论、算法与拓展[M]. 北京:科学出版社,2009.
- [6] 郭科,陈聆,魏友华. 最优化方法及其应用[M]. 北京:高等教育出版社,2010.
- [7] 赵洪波,冯夏庭. 支持向量机函数拟合在边坡稳定性估计中的应用[J]. 岩石力学与工程学报,2003,22(2):241-245.
- [8] Steve R Gunn. Support vector machines for classification and regression[M]. England:University of Southampton, 1998.