

文章编号: 1000 5013(2010)05-0562-04

# 建设工程项目工序的 LS-SVM 工期预测模型

祁神军<sup>1,2</sup>, 张云波<sup>1</sup>, 烈云<sup>2</sup>,

(1. 华侨大学 土木工程学院, 福建 泉州 362021;

2. 华中科技大学 土木工程与力学学院 湖北 武汉 430074)

**摘要:** 鉴于传统工期预测的模糊性和随机性, 分析影响工程项目工期的因素及参数的获取方式. 采用最小二乘支持向量机(LS-SVM)构建建设工程项目工序工期的预测模型, 并用工程实例论证方法的有效性. 结果表明, 对类似工程或者同一工程项目的类似工序的进度执行状况进行学习, 采用 LS-SVM 的工期预测模型预测即将开展的工程项目的工序工期, 符合实际工期控制的要求. 与基于 BP 神经网络工期预测模型对比分析, LS-SVM 的工期预测模型的预测误差更小, 平均训练时间更短, 网络总误差更小.

**关键词:** 建设工程项目; 最小二乘向量机; 工序; 工期预测

**中图分类号:** TU 722

**文献标识码:** A

进度计划是建设企业集团工程项目进度控制的指挥棒, 而各工序的建造工期精确程度直接影响整个进度控制. 人为因素在工期估算中会产生随机性和模糊性<sup>[1-2]</sup>, 而工程条件的多重复杂性也会导致工期的各种可能性. 1987 年, Thomas 和 Yiakoumis 提出了影响高速公路项目的生产建造效率的因素, 包括环境因素、施工现场因素、管理因素及中断因素, 并建立了相应的劳动生产率和产出因素之间关系的模型<sup>[3]</sup>. 1997 年, Thomas 等<sup>[4]</sup>将中断因素分为资源、返工和管理 3 类. 2001 年, Ovararinand 等<sup>[5]</sup>指出, 影响建筑工程效率的 13%~35% 是工作面到材料仓库之间的距离不够引起的. Botero 等通过 10 个案例的研究<sup>[6]</sup>, 指出影响进度的因素一般包括 10 种因素, 如工作开展之前的准备、人力、进度安排的精确性、材料、进度执行过程中的协调、返工、机械设备、天气、设计, 以及其他因素. 最小二乘支持向量机<sup>[7]</sup> (Least Square SVM, LS-SVM) 是经典的支持向量机的发展和改进. 它采用等式约束代替不等式约束, 以平方项作为优化指标, 将二次规划问题转化为用最小二乘法求解线性方程组, 降低了计算的复杂性, 大大提高了求解速度. 本文考虑了工程项目工期的随机性和模糊性, 并根据工程实践提出诸要素的参数获取方式, 构建了基于支持向量机(LS-SVM)的工期预测模型<sup>[8]</sup>.

## 1 影响工序工期的因素及参数获取

### 1.1 影响因素分析

不同工程项目, 影响工期的因素是不同的. 考虑到大型工程项目建造过程的共性, 以及工序工期影响因素数据的可获取性, 综合国内外学者对工期影响因素的研究, 设定影响工序工期( $t$ )主要有如下 7 种因素: 建筑工程工序本身的工程量( $Q$ ), 在该工序可分配的人员( $P$ )、材料( $M$ )及设备( $E$ ), 可供该工序进行的工作面( $S$ ), 该工序的施工水文地质条件( $C$ ), 以及项目施工管理水平. 施工水文地质条件还包括施工阶段的平均气温、施工阶段冻土深度、施工阶段的降雨量、施工阶段的风速等<sup>[9]</sup>.

### 1.2 参数获取<sup>[10]</sup>

根据工程项目实践及参数的可获取性、适用性等原则, 提出参数获取的方式, 如表 1 所示.

(1)  $Q$  可从工程量清单预算中直接获取.

收稿日期: 2009 10 23

通信作者: 祁神军 (1982), 男, 讲师, 主要从事工程项目管理的研究. E-mail: qisj972@163.com.

基金项目: 国务院侨办科研基金资助项目(08QZR06); 华侨大学高层次人才科研启动项目(07BS404)

- (2) 当  $j$  工序仅有 1 种材料时,  $M = \frac{m_{i,j}}{Q_j}$ ; 当  $j$  工序存在多种材料时,  $M = \sum \frac{m_{i,j}p(m)_{i,j}}{Q_j}$ . 其中:  $m_{i,j}, p(m)_{i,j}$  分别表示  $i$  材料在  $j$  工序中的预算量和预算单价.
- (3) 当  $j$  工序仅有 1 种设备时,  $E = \frac{e_{k,j}}{Q_j}$ ; 而当  $j$  工序存在多种设备时,  $E = \sum \frac{e_{k,j}p(e)_{k,j}}{Q_j}$ . 其中:  $e_{k,j}, p(e)_{k,j}$  分别表示  $k$  设备在  $j$  工序中预算量和预算单价.
- (4)  $S$  是该工序可供工作的工作面与工程量之比.
- (5)  $C$  取 1, 2, 3, 4 共 4 个等级.

表 1 工程项目工期预测参数的取值  
Tab.1 Parameter values of project activity duration

参数名称	参数指标	计算方法	取值类型
$Q$	按工程量清单规则计算的工程量	实际取值	数值型
$P$	该工序可以调用的人员强度	实际取值	数值型
$M$	该工序建造材料的调用强度	加权取值	数值型
$E$	该工序建造设备的调用强度	加权取值	数值型
$S$	可供该工序的工作面强度	平均取值	数值型
$C$	定性指数	等级取值	数值型

2 模型的建立

设  $l$  组待训练类似工序样本, 即  $(x_1, t_1), \dots, (x_l, t_l) \in R^6 \times R, x_i = (Q_i, P_i, E_i, M_i, S_i, C_i)$ .  
(1) 若因素  $Q, P, E, M, S, C$  与工期  $t$  构成线性关系, 则设其线性回归函数为

$$f(x) = \tilde{\omega}x + b.$$

引入结构风险函数, 将回归问题转化为二次优化问题, 有

$$\min \frac{1}{2} \|v\|^2 + g \frac{1}{2} \sum_{i=1}^l x_i^2,$$

其约束条件为  $t_i = \tilde{\omega}^T x_i + b + \xi_i, i = 1, 2, \dots, l$ . 定义拉格朗日函数为

$$L = \min \frac{1}{2} \|\tilde{\omega}\|^2 + \gamma \frac{1}{2} \sum_{i=1}^l \xi_i^2 - \sum_{i=1}^l \alpha_i (\tilde{\omega}^T x_i + b + \xi_i - y_i),$$

根据非线性最优规划问题的 KKT( Karush- Kuhn- Tucker) 条件, 可得

$$\begin{cases} \frac{\partial L}{\partial \tilde{\omega}} = 0 \rightarrow \tilde{\omega} = \sum_{i=1}^l \alpha_i x_i, \\ \frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^l \alpha_i = 0, \\ \frac{\partial L}{\partial \xi_i} = 0 \rightarrow \alpha_i = \gamma \xi_i, \\ \frac{\partial L}{\partial \alpha} = 0 \rightarrow \tilde{\omega}^T x_i + b + \xi_i - y_i = 0, \quad i = 1, 2, \dots, l \end{cases}$$

最后的解为

$$\begin{bmatrix} 0 & I^T \\ I & x^T x + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix}.$$

(2) 若因素  $Q, P, E, M, S, C$  与工期  $t$  构成非线性关系, 则使用相同的非线性映射  $\phi$ . 然后, 把数据映射到一个高维特征空间并进行线性回归.

关键的问题是核韩式  $K(x, y)$  的选取, 使得  $K(x_i, y_i) = \phi(x_i)^T \phi(x_j)$ , 其非线性回归函数的解为

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

因此, 只需要通过训练样本求解出参数  $a, b$ , 即可对工程项目的工期进行预测<sup>[11]</sup>.

3 算例分析

3.1 样本来源

某一建设企业集团承建了某市地铁的一个标段, 沿线需要建造若干灌注桩. 根据建造完成的 9 个灌注桩的影响因素  $Q, P, M, E, C, S$  和工期  $t$  等的统计数据, 预测即将展开的第 10, 11 个灌注桩的工期. 样本统计数据和待测样本的基础数据, 如表 2 所示.

表 2 样本统计数据和待测样本的基础数据

Tab. 2 Statistical data of learned samples and basic data of forecasted samples

序号	$Q$	$P$	$M$	$E$	$C$	$S$	$t$
1	22. 20	10. 0	47. 60	0. 49	2	50. 24	6. 5
2	20. 20	11. 0	50. 60	0. 52	2	54. 24	6. 0
3	32. 20	16. 0	70. 20	0. 72	1	54. 24	5. 0
4	25. 20	8. 0	39. 20	0. 41	2	40. 24	6. 0
5	21. 40	9. 5	47. 06	0. 50	3	49. 20	9. 0
6	47. 08	21. 2	103. 50	1. 10	4	105. 00	17. 0
7	50. 30	22. 7	110. 86	1. 20	4	110. 00	18. 0
8	21. 40	11. 0	51. 20	0. 56	3	53. 45	6. 5
9	100. 00	50. 0	200. 00	2. 12	3	200. 00	16. 5
10	22. 50	12. 0	50. 00	1. 12	3	50. 00	待测
11	30. 50	10. 0	30. 00	2. 12	2	26. 00	待测

3.2 训练仿真

经过对样本的训练, 在网络总误差为 0. 011,  $b$  为- 0. 062 1,  $\alpha=[51. 101\ 5, -32. 167\ 2, -1. 078\ 1, -20. 687\ 6, 18. 633\ 5, 0. 551\ 3, 1. 295\ 0, -18. 963\ 9, 1. 315\ 5]$  的情况下, 构建训练模型.

在试验中, 同基于神经网络的工期预测方法进行了对比, 结果如表 3 所示. 表 3 中:  $t_r, t_c$  分别为实际工期值和预测工期值. BP 神经网络预测模型的输入层设置 6 个神经元, 且采用 Tansig 转换函数; 输出层节点数为 1, 采用 Purelin 函数, 且选择 Traingdx 函数进行学习和训练.

表 3 BP 神经网络与最小二乘支持向量机的仿真结果对比

Tab. 3 Comparison between the simulation results by BP model and LS-SVM model

$t_r/d$	$t_c/d$		绝对误差		相对误差	
	LS-SVM	BP 神经网络	LS-SVM	BP 神经网络	LS-SVM	BP 神经网络
6. 5	6. 447 5	6. 369 6	- 0. 052 5	- 0. 130 4	- 0. 814 3	- 2. 006 2
6. 0	6. 033 1	6. 089 6	0. 033 1	0. 089 6	0. 548 6	1. 493 3
5. 0	5. 001 1	4. 999 9	0. 001 1	- 0. 000 1	0. 022 0	- 0. 002 0
6. 0	6. 021 3	6. 041 1	0. 021 3	0. 041 1	0. 353 7	0. 685 0
9. 0	8. 980 8	8. 998 2	- 0. 019 2	- 0. 001 8	- 0. 213 8	- 0. 020 0
17. 0	16. 999 4	16. 998 2	- 0. 000 6	- 0. 001 8	- 0. 003 5	- 0. 010 6
18. 0	17. 998 7	18. 004 2	- 0. 001 3	0. 004 2	- 0. 007 2	0. 023 3
6. 5	6. 519 5	6. 499 9	0. 019 5	- 0. 000 1	0. 299 1	- 0. 001 5
16. 5	16. 498 6	16. 498 1	- 0. 001 4	- 0. 001 9	- 0. 008 5	- 0. 011 5

通过表 3 可得, BP 神经网络与 LS-SVM 模型仿真的绝对误差最大值分别为- 0. 130 4 和 0. 019 5. 其训练网络总误差分别为 0. 025 9, 0. 011 0; 平均训练时间分别为 82. 031 0, 0. 453 0 s; 平均训练误差率分别为  $1. 989\ 0 \times 10^{-4}$ ,  $2. 085\ 5 \times 10^{-4}$ .

由此可验证基于结构风险的最小化. 折衷考虑经验风险, 置信区间的 LS-SVM 的预测模型对工程项目工序工期进行预测时, 可能有更高预测精度和更快的训练速度<sup>[12]</sup>.

3.3 工期预测

对待测工序的工期进行预测, 第 10, 11 个灌注桩的预测工期分别为 3. 30, 9. 56 d, 实际工期分别为 3. 5, 9. 0 d, 预测工期与实际工期的偏差分别为- 3. 32%, 3. 90%. 导致误差结果的主要原因是, 在组织施工方面出现了窝工. 从大型工程项目施工的角度来讲, 绝对工期偏差在半天之内是允许的.

## 4 结束语

考虑工程项目工期的随机性和模糊性, 根据工程实践提出影响工期诸要素的参数获取方式, 构建基于支持向量机的工期预测模型. 该模型能准确地预测类似工程项目的工序工期, 有效编制工程项目的进度计划, 使工程项目建造过程稳定均衡、工作流持续、建造周期缩短、建造成本降低、产品质量和客户满意度提高. 在大型工程项目工序工期预测中, 该方法具有一定的指导意义和价值.

### 参考文献:

- [1] 陆歆弘. 模糊假言推理确定施工工期[J]. 基建优化, 1999, 20(5): 18-21.
- [2] 何飞, 张贵平. 工程施工工期概率: 模糊估算法[J]. 陕西建筑, 2000(2): 42-45.
- [3] SANDERS S R, THOMAS H R. Factors affecting masonry-labor productivity[J]. J Constr Engrg and Mgmt, ASCE, 1991, 117(4): 626-644.
- [4] THOMAS H R, MATHEWS C T, WARD J G. Learning curve models of construction productivity[J]. J Constr Engrg and Mgmt, ASCE, 1986, 112(2): 245-258.
- [5] OVARARIN N, POPESCU C M. Field factors affecting masonry productivity[C] // Proc of the 45th AACE: International Transaction Pittsburgh: [s. n.], 2001: 91-100.
- [6] BOTERO L F, ALVAREZ M E. Last planner: An advance in planning and controlling construction projects: Case study in the city of Medellin[C] // Proc of the 4th Brazilian Symposium on Construction Management and Economics Porto Alegre: [s. n.], 2005: 1-9.
- [7] SUVKENS J A K, VUNDEWULLE J. LeusL squares support vector machine classifiers[J]. Neural Processing Letter, 1999, 9(3): 293-300.
- [8] SALEM O, SOLOMON J, GENAIDY ETAL A. Site implementation and assessment of lean construction techniques [J]. Lean Construction Journal, 2005, 3(2): 1-21.
- [9] 张云波, 胡云昌. 人工神经网络的建设工期定额地域分类[J]. 华侨大学学报: 自然科学版, 2004, 25(3): 270-274
- [10] 祁神军, 丁烈云, 骆汉宾. 大型工程项目工序工期精准预测方法研究[J]. 重庆建筑大学学报, 2007, 29(6): 141-144.
- [11] 宋志宇, 李俊杰. 最小二乘支持向量机在大坝变形预测中的应用[J]. 水电能源科学, 2006, 24(6): 49-52.
- [12] 辛治运, 顾明. 基于最小二乘支持向量机的复杂金融时间序列预测[J]. 清华大学学报: 自然科学版, 2008, 48(7): 1147-1149.

## Forecast Model of Activity Duration Based on LS-SVM in Construction Engineering Project

QI Shen-jun<sup>1,2</sup>, ZHANG Yun-bo<sup>1</sup>, DING Lie-yun<sup>2</sup>

(1. College of Civil Engineering, Huaqiao University, Quanzhou 362021, China;

2. College of Civil Engineering and Mechanics, Huazhong University of Science and Technology, Wuhan 430074, China)

**Abstract:** Referring to fuzziness and randomness for activity duration forecast of construction engineering project by traditional ways, the influence factors of activity duration are analyzed, and parameter calculation is also proposed. A forecast model of activity duration based on the least square support vector machine (LS-SVM) is set up, and the analysis of a subway case confirms validity of this model. The model is trained by the schedule execution situation of the activities in other similar projects or the similar activities in the same project, the activity duration simulated by the model conforms with the request of schedule controlling. In the forecast model of activity duration based on LS-SVM, the prediction and network total errors is less, training time is shorter than the ones in the forecast model based on BP.

**Keywords:** construction engineering project; least square support vector machine; activity; activity duration forecast

(责任编辑: 钱筠 英文审校: 方德平)