

DOI :10.3880/j.issn.1006-7647.2009.01.003

最小二乘支持向量机参数反演方法及其应用

韩立炜 李宗坤 王建有

(郑州大学水利与环境学院 ,河南 郑州 450002)

摘要 :针对最小二乘支持向量机(LS-SVM)参数不易确定的问题 ,利用遗传神经网络模拟 LS-SVM 计算结果与参数之间的关系 ,提出了一种基于遗传神经网络(GA-BP)的参数选择方法 ,该方法利用正交分解法构建训练参数组 ,并将参数组代入最小二乘支持向量机以获得计算输出值 ,然后将计算输出值与训练参数组代入遗传神经网络进行训练并获得合适的 LS-SVM 参数。最后以土石坝渗流分析为例进行验证 ,结果表明该方法对优化选择最小二乘支持向量机参数十分有效 ,预测精度可达 10^{-4} 。

关键词 :最小二乘支持向量机 ;参数反演 ;遗传神经网络

中图分类号 :TP311 文献标识码 :A 文章编号 :1006-7647(2009)01-0008-04

Study and application of inverse method for determining parameters in Least squares support vector machines//HAN Li-wei LI Zong-kun ,WANG Jian-you (Zhengzhou University , School of Water Conser vancy and Environment Engineering , Zhengzhou 450002 , China)

Abstract :The parameters in Least Squares Support Vector Machines (LS-SVM) are not easy to determine. A new parameter-selecting method that simulates the relation of LS-SVM parameters with their results through a genetic neural network is put forward. The training parameter group is constructed with proper orthogonal decomposition. The resulting output is obtained by putting the training parameter group into LS-SVM. By substitution of the output and the train parameters set into the genetic neural network ,the most suitable parameters of LS-SVM are determined. Finally , the example of earth dam seepage is considered. Simulation results show that this method can yield optimized parameters and that the prediction accuracy is over 10^{-4} .

Key words :Least squares support vector machines ;back analysis ;genetic neural network

Vapnik^[1]在 1995 年提出了一种新型统计学习方法——支持向量机(support vector machines ,SVM) ,支持向量机基于结构风险最小化原理 ,具有完备的统计学理论基础和出色的学习性能 ,并在优化控制、模式识别、港口吞吐量^[2]、小流域泥石流流输沙量预测^[3]等领域获得了成功的应用。近年 ,Suykens 等^[4]提出了最小二乘支持向量机方法(least squares support vector machines ,LS-SVM) ,在非线性估计及模式识别中得到广泛应用。和其他学习算法一样 ,其性能依赖于学习机的参数。然而到目前为止 ,还没有指导 LS-SVM 参数选择的好方法。如何确定最优参数 ,一直是提高 LS-SVM 学习和泛化能力的主要研究问题之一。

目前 ,主要是通过试验来选择支持向量机的参数。例如 ,对于标定好的数据集 ,常用的方法是根据经验大致设定学习机器的参数 ,从数据集中随机取出一半进行训练 ,留下另一半进行测试 ,以检验该学习机器的推广能力。为了得到一个较优的参数 ,需

要多次改变参数 ,并进行训练和测试 ,最后根据历次测试的结果确定学习机器的参数^[5]。文献[6]采用模型训练的交互验证法结合对参数的测试来进行模型设计 ,应用交叉验证方法选择参数能避免支持向量机的过学习 ;文献[7]采用留一法与网格搜索法来确定参数 ,留一法能有效避免过拟合现象 ,网格搜索法可以安全地搜索到最优参数组合 ;文献[8]采用三步搜索技术来确定参数 ;文献[9]采用最小拟合差方法来确定参数。但是上述方法普遍存在训练与测试次数较多、花费时间较长的问题。针对这一问题 ,笔者提出了将 LS-SVM 与遗传神经网络(GA-BP)算法结合的参数反演方法。

1 最小二乘支持向量机

Suykens 等^[4]提出最小二乘支持向量机 ,优化指标采用平方项 ,只有等式约束 ,优化目标是最小化 $\frac{1}{2} \| \omega \|^2$,其表示为

$$\min_{\omega, b, \xi} \frac{1}{2} \|\omega\|^2 + \frac{1}{2} \gamma \sum_{i=1}^l \xi_i^2 \quad (i = 1, \dots, l) \quad (1)$$

$$\text{s.t.} \quad y_i [\omega^T \phi(X_i) + b] = 1 - \xi_i$$

式中: ω 为权值向量; γ 为正则化参数; ξ_i 为误差变量; b 为偏差量。

可得到线性方程组:

$$\begin{bmatrix} 0 & y^T \\ y & Q + \gamma^{-1}I \end{bmatrix}_{(l+1) \times (l+1)} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ e \end{bmatrix} \quad (2)$$

式中: $Q = \phi(X_i)\phi(X_j)$; $y = [y_1, \dots, y_n]$; $a = [a_1, \dots, a_n]^T$; I 为单位矩阵; e 为元素为 1 的向量。

本文核函数采用径向基核函数:

$$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \quad (3)$$

式中: σ^2 为径向基核参数; x, x_i 为支持向量内积运算中的项。在没有关于问题的先验知识时, 由这种核函数训练而成的模型具有比基于其他核函数的模型更好的总体性能。

a 和 b 可通过最小二乘法解得, 应用 LS-SVM 对非线性函数回归的结果为

$$f(x) = \sum_{i=1}^l a_i K(x, x_i) + b \quad (4)$$

最小二乘支持向量机的最大优点就是能显著提高标准支持向量机的训练速度, 简化计算。

由 LS-SVM 算法推导可知, 对于采用径向基核的 LS-SVM 的主要参数是 γ 和 σ^2 , γ 是权衡拟合曲线光滑度并使拟合误差最小化的参数, γ 值越大, 训练数据样本点拟合值越符合实际值, γ 减小会降低模型的复杂性; σ^2 增大会使拟合曲线更光滑。这 2 个参数在很大程度上决定了 LS-SVM 的学习能力和预测能力。

2 参数确定方法

在岩土工程模型参数的反分析中, 神经网络被用来建立待反演参数与相应计算位移之间的映射关系, 从而代替耗时的有限元计算以提高计算效率。仿效其原理, 假设 LS-SVM 计算过程为岩土工程计算的有限元计算, LS-SVM 参数为岩土本构模型中的待反演参数, LS-SVM 计算结果为有限元的计算位移, 则可利用 GA-BP 来模拟 LS-SVM 参数与结果之间的映射关系。

神经网络在进行反演时多采用 BP 网络, BP 算法的优点是寻优具有精确性, 但同时具有易陷入局部极小、收敛速度慢和引起振荡效应等缺点。遗传算法具有很强的宏观搜索能力, 且能以较大的概率找到全局最优解, 所以用它来完成前期的搜索能较好地克服 BP 算法的缺点。笔者将二者结合起来, 形成一种混合训练算法——GA-BP 算法, 达到优化

网络的目的。

在 matlab 中编制了反演支持向量机 γ 和 σ^2 的程序, 参数反演流程如图 1 所示。①选择若干组库水位、降雨与时效作为预测样本, 用于测试使用 GA-BP 反演出参数后 LS-SVM 的预测能力, 其他资料作为 LS-SVM 的训练样本。②选定 γ 与 σ^2 的大致范围, 利用正交试验原理产生 n 组 γ 与 σ^2 , 将每一组参数与训练样本一起代入 LS-SVM 进行计算, 得到 n 组测压管水位拟合输出值。③构造以 LS-SVM 拟合输出值为输入节点、以 LS-SVM 参数为输出节点的 GA-BP, 并进行训练。输入节点数为参数组的数量 n , 输出节点数为 2。④将实际输出值代入训练好的 GA-BP, 求得 LS-SVM 的 γ 和 σ^2 。⑤将 γ 和 σ^2 代入 LS-SVM 模型, 利用预测样本进行预测。

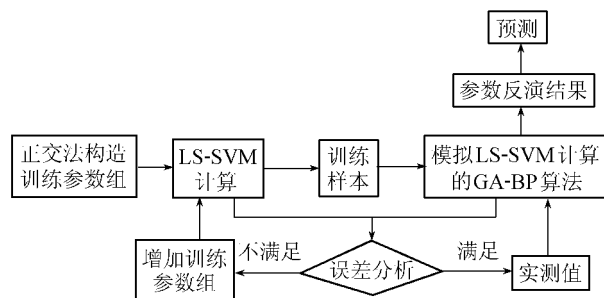


图 1 反演流程

3 工程实例

影响土石坝测压管水位的主要因素可归纳为库水位分量、降雨分量和时效分量三部分。笔者选用某水库 S3 测压管 1998 ~ 1999 年共 83 组测压管水位, 其中第 1 ~ 76 时步的 76 个测压管监测值作为训练样本(表 1), 第 77 ~ 83 时步的 7 个测压管监测值作为预测样本, 构建 LS-SVM 模型, 对训练样本进行学习, 在此基础上对训练样本进行拟合并对后 7 个预测样本进行预测。其中, LS-SVM 模型中的参数采用 GA-BP 算法进行拟合。

3.1 参数反演

选取时步 77, 78 作为反演样本, γ 选取 1, 21, 41, 61, 81, 101 共 6 种水平, σ^2 选取 0.1, 20.1, 40.1, 60.1, 80.1, 100.1, 依据正交试验设计原理, 给出 36 组不同的 γ 和 σ^2 组合试验, 对每一组试验, 利用 LS-SVM 计算拟合测压管水位值(限于篇幅未列出)。用其中的前 30 个样本训练网络, 用后 6 个样本测试网络。获得的模型作为映射进行参数反分析, 参数反演的搜索区间: γ 为 1 ~ 101, σ^2 为 0.1 ~ 100.1。

选取时步 77, 78 的实际水位 275.41 m 和 275.40 m 分别代入网络反演得到最小二乘支持向量机参数值, 取均值作为最终参数值, 为 $\gamma = 53.58$, $\sigma^2 = 45.01$, 代入 LS-SVM 模型中进行拟合与预测, 并

与BP神经网络的拟合与预测值进行比较 ,见表 1。

3.2 算法对比

定义拟合差函数 $s = \sum_{i=1}^n (Y'_i - Y_i)^2 / (n - 1)$ 作为衡量模型精度的指标 ,式中 Y'_i 为拟合时第 i 个训练样本的拟合值 , Y_i 为拟合时第 i 个训练样本的实测值^[8]。

当采用拟合差最小方法计算参数时 ,拟合差随 γ 的增大、 σ^2 的减小而减小。当 $\gamma > 20$ 时 ,拟合差随 γ 的变化趋缓 ,随 σ^2 的变化亦趋缓 ;当拟合差取最小为 0.00001 时 , $\gamma = 100.71$, $\sigma^2 = 0.11$ 。

分别将 $\gamma = 100.71$, $\sigma^2 = 0.11$ 及 $\gamma = 53.58$, $\sigma^2 = 45.01$ 代入 LS-SVM 进行拟合并预测 ,拟合值、预测值及相应的误差见表 1 和表 2 ,对比图如图 2 所示 (图中未连线的方框表示相应的预测值)。

由表 1 可知 ,LS-SVM 的拟合精度比 BP 神经网络的拟合精度要高。对比表 1、表 2 可知 :当 $\gamma = 100.71$, $\sigma^2 = 0.11$ 时 ,学习样本的拟合精度很高 ,总拟合差仅为 1.03×10^{-5} ,但预测值精度却不高 ,随着时步的增加 ,其预测精度越来越差 ,在第 83 时步的时候其绝对误差最大 ,达 0.534 3 m ;当 $\gamma = 53.58$, $\sigma^2 = 45.01$ 时 ,学习样本的拟合精度虽不如上一组参数 ,拟合差为 0.001 4 ,这对于渗流监测来说已足够精确。然而其外推预测精度较高 ,从表 2 可知其拟

合差为 0.0002。总之 ,选择 GA-BP 反演参数的 LS-SVM 模型比神经网络模型的拟合精度与预测精度都要高。

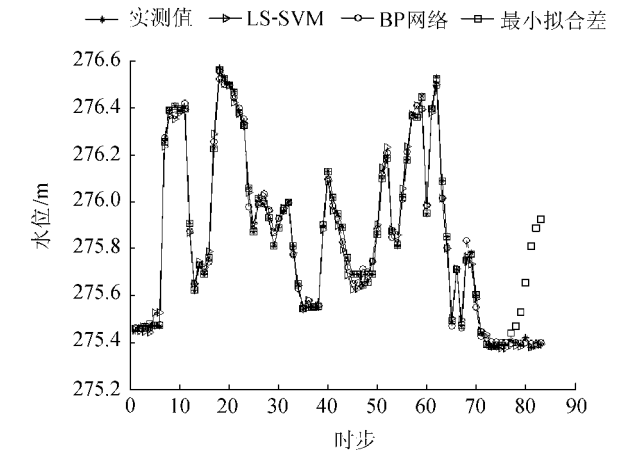


图 2 不同参数、方法拟合及预测值

分析其原因 ,主要是由于观测误差和模型误差的存在 ,利用拟合差最小获得参数的方法无法保证搜索到全局最优解。早在 1963 年 Tikhonov^[10]就证明 ,如果正问题是线性的 ,那么 ,反问题的解存在唯一并且连续地依赖于观测数据(稳定)。但是 ,当考虑到观测数据的观测误差时 ,即使正问题是线性的 ,反问题的目标函数是非凸的 ,反问题的解也不是唯一的。观测误差越大 ,目标函数的局部极小值数目

表 1 LS-SVM 不同参数对学习样本拟合的对照

时步	实测值	LS-SVM 拟合				神经网络拟合	
		$\gamma = 100.71 \sigma^2 = 0.11$		$\gamma = 53.58 \sigma^2 = 45.01$		拟合值	绝对误差
		拟合值	绝对误差	拟合值	绝对误差		
1	275.45	275.46	- 0.01	275.45	0.00	275.45	0.00
2	275.47	275.46	0.01	275.45	0.02	275.42	0.05
3	275.47	275.47	0.00	275.45	0.02	275.42	0.05
4	275.47	275.48	- 0.01	275.44	0.03	275.42	0.05
71	275.44	275.44	0.00	275.45	- 0.01	275.42	0.02
72	275.39	275.39	0.00	275.43	- 0.04	275.41	- 0.02
73	275.38	275.38	0.00	275.39	- 0.01	275.37	0.01
74	275.38	275.38	0.00	275.39	- 0.01	275.39	- 0.01
75	275.39	275.39	0.00	275.38	0.01	275.38	0.01
76	275.39	275.40	- 0.01	275.38	0.01	275.39	0.00

注 LS-SVM 在 $\gamma = 100.71 \sigma^2 = 0.11$ 和 $\gamma = 53.58 \sigma^2 = 45.01$ 时拟合差分别为 1.03×10^{-5} 和 0.001 4 ,神经网络拟合差为 0.0025。

表 2 LS-SVM 不同参数对检验样本预测的对照

时步	实测值	LS-SVM 拟合				神经网络拟合	
		$\gamma = 100.71 \sigma^2 = 0.11$		$\gamma = 53.58 \sigma^2 = 45.01$		预测值	绝对误差
		预测值	绝对误差	预测值	绝对误差		
77	275.41	275.44	0.03	275.39	- 0.02	275.43	0.02
78	275.40	275.47	0.07	275.39	- 0.01	275.44	0.04
79	275.39	275.53	0.14	275.38	- 0.01	275.42	0.03
80	275.42	275.65	0.23	275.40	- 0.03	275.42	0.00
81	275.38	275.81	0.43	275.38	0.00	275.41	0.03
82	275.40	275.89	0.49	275.39	- 0.01	275.41	0.01
83	275.39	275.92	0.53	275.39	0.00	275.40	0.01

注 LS-SVM 在 $\gamma = 100.71 \sigma^2 = 0.11$ 和 $\gamma = 53.58 \sigma^2 = 45.01$ 时拟合差分别为 0.1128 和 0.0002 ,神经网络拟合差为 0.0007。

越多。而 GA-BP 算法可以搜索到全局最优解,例如该算例比较结果中,拟合误差最小时,预测精度并不高,而采用反演方法计算的结果,虽然拟合误差不是最小,但预测精度却很高。

4 结 语

通过引入 GA-BP 算法来反演 LS-SVM 的 2 个重要参数,理论分析和仿真试验表明了这种方法的有效性,可以优化选择支持向量机参数,从而获得较好的模型结果,有效解决了人为选择支持向量机参数的随机性和盲目性,提高了预测的精度。

由于 LS-SVM 具有良好外推能力,不必知道太多的数据即可建模,而且在学习样本数量很多时又不易陷入“维数灾难”,能保证解收敛于全局最优,预测结果更接近于真实值,故可有效地根据土石坝原型观测资料分析大坝中存在的渗流问题。

参考文献:

[1] VAPNIK V N. The nature of statistical learning theory[M]. NY : Springer-Verlag ,1995.
[2] 武晓,宗蓓华.基于支持向量机回归的港口吞吐量非线性

性组合预测[J].河海大学学报:自然科学版,2005,33(4):482-484.
[3] 王涛,刘兴年,黄尔.基于 SVM 方法的小流域泥石流输沙量预测[J].水利水电科技进展,2008,28(2):1-3.
[4] SUYKENS J A K, VANDEWALLE J. Least squares support vector machine classifier[J]. Neural Processing Letter,1999,9(3):293-300.
[5] 赵洪波.基于遗传算法的进化支持向量机研究[J].绍兴文理学院学报,2004,24(9):25-27.
[6] 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.
[7] 余志雄,周创兵,李俊平,等.基于 V-SVR 算法的边坡稳定性预测[J].岩石力学与工程学报,2005,24(14):2468-2475.
[8] 郭辉,刘贺平,王玲.最小二乘支持向量机参数选择方法及其应用研究[J].系统仿真学报,2006,18(7):2033-2036.
[9] 刘开云,乔春生,滕文彦.边坡位移非线性时间序列采用支持向量机算法的智能建模与预测研究[J].岩土工程学报,2004,26(1):57-61.
[10] TIKHONOV A N. Solutions of ill-posed problems[M]. New York : John Wiley & Sons ,1977:47-87.

(收稿日期:2008-08-07 编辑:方宇彤)

(上接第 7 页)

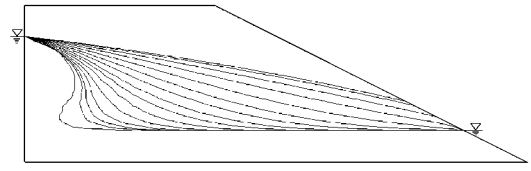


图 5 水位骤升时坡体内浸润线变化过程

在求解过程中每步计算最多迭代 3 次,有限元计算就满足误差收敛条件 $\epsilon \leq 10^{-12}$,计算收敛速度快。由图 4、图 5 可知,2 种工况的初期坡体内的孔隙负压较大,且初期自由面变化较快,随时间增长变化速度逐渐减小,直到趋于稳定,计算结果符合工程实际规律。

4 结 语

推导了饱和-非饱和非稳定渗流的整体平衡方程,对迭代求解过程中的数值弥散现象与参数拟合不收敛现象,提出了一种改进方法。从算例可以看出,文中所提的方法有效消除了有限元求解过程中的振荡现象,并加快了迭代收敛速度,可以推广到实际工程中进行应用。

参考文献:

[1] 毛昶熙.渗流计算分析与控制[M].北京:水利水电出版

社,1990.
[2] 吴梦喜.饱和-非饱和土体非稳定渗流数值分析[J].水利学报,1999(12):38-42.
[3] 高骥,雷光耀,张锁春.坝坝饱和-非饱和渗流的数值分析[J].岩土工程学报,1988,10(6):28-37.
[4] 杨代泉,沈珠江.非饱和土一维广义固结的数值计算[J].水利水运科学研究,1991(4):357-385.
[5] 岑建,詹美礼.水工渗流有限元网格剖分的控制断面节点对连法[J].水利水电科技进展,2006,26(5):48-50.
[6] 朱伟,秦建设,高玉峰.求解非饱和土堤初始浸润线的一种解析法[J].河海大学学报:自然科学版,2003,31(3):314-317.
[7] 刘晓青.土体-孔隙体静动相互作用有限元解法的研究[D].南京:河海大学,1997.
[8] 鲁子林.水库群调度网络分析法[J].华东水利学院学报,1983,11(2):35-48.
[9] 彭华,陈胜宏.饱和-非饱和岩土非稳定渗流有限元分析研究[J].水动力学研究与进展,2002,17(2):253-259.
[10] 彭华,陈胜宏.饱和非饱和非稳定渗流有限元加速技术[J].武汉大学学报:工学版,2001,34(3):9-12.
[11] LANE P A, GRIFFITHS D V. Assessment of stability of slopes under drawdown conditions[J]. Journal of Geotechnical and Geoenvironmental Engineering, ASCE, 2000,126(5):443-450.

(收稿日期:2008-07-22 编辑:方宇彤)