

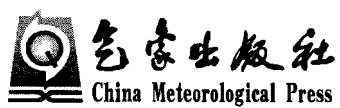
王定成 ◎ 著

支持向量机建模 预测与控制



支持向量机建模预测与控制

王定成 著



内容简介

统计学习理论在研究小样本情况下表现出了极好的性能,支持向量机和支持向量机回归是统计学习理论的重要组成部分,具有较好的推广能力和非线性建模特性。由于现有非线性建模与智能控制方法存在诸如模型结构的选取过分依赖经验、局部极小点等问题,因而,本书采用支持向量机和支持向量机回归进行建模预测与控制的研究。但基于二次规划的支持向量机和支持向量机回归不适合在线建模与实时控制,因此,本书在介绍支持向量机的基本理论基础上进行在线支持向量机回归算法的研究。主要包括:支持向量机与回归的理论与方法、支持向量机灾害天气预测、支持向量机回归非线性建模方法、支持向量机预测控制、支持向量机内模控制、在线稀疏最小二乘支持向量机回归模型与内模控制。

本书适合高等院校相关专业高年级学生、研究生使用,也可作为教师和科研人员的参考书。

图书在版编目(CIP)数据

支持向量机建模预测与控制/王定成著. —北京:气象出版社,2009.12

ISBN 978-7-5029-4918-1

I. ①支… II. ①王… III. ①向量计算机-系统建模-研究 IV. ①TP338

中国版本图书馆 CIP 数据核字(2009)第 242854 号

Zhichixiangliangji Jianmo Yuce yu Kongzhi

支持向量机建模预测与控制

王定成 著

出版发行:气象出版社

地 址:北京市海淀区中关村南大街 46 号

总 编 室:010-68407112

网 址:<http://www.cmp.cma.gov.cn>

策划编辑:李太宇

责任编辑:林雨晨

封面设计:博雅思企划

印 刷:北京中新伟业印刷有限公司

开 本:787mm×1092mm 1/16

字 数:300 千字

版 次:2009 年 12 月第 1 版

定 价:25.00 元

邮政编码:100081

发 行 部:010-68409198

E-mail: qxcbs@263.net

终 审:朱文琴

责任技编:都 平

印 张:10.25

印 次:2009 年 12 月第 1 次印刷

本书如存在文字不清、漏印以及缺页、倒页、脱页等,请与本社发行部联系调换

前 言

支持向量机是统计学习理论的重要组成部分。统计学习理论产生于 20 世纪 70 年代,是专门研究小样本的机器学习理论,与传统的基于无限样本假设的机器学习理论相比,更符合实际系统有限样本的情况。

由于实际大多系统的模型都具有非线性,因而,非线性模型可以更精确地表达真实系统的模型。对于非线性系统来说,由于没有统一的建模方法,系统数学模型的建立是非常困难的,有时甚至是不可能的。黑匣子模型不依赖于系统的精确数学模型,只根据系统的输入和输出就可以确定系统的模型,是非线性系统建模中经常采用的方法,由于现有的一些非线性建模方法存在诸如模型结构的选取过分依赖于经验、局部极小点等问题,因而,采用支持向量机和支

持向量机回归进行建模预测的尝试是有意义的。

在非线形预测控制中,预测控制和内模控制具有较好的稳定性和鲁棒性,但这些方法都属于基于模型的控制方法,模型的准确与否会直接影响控制的精度和稳定性。对于复杂的非线性系统来说,支持向量机回归提供了较好的建模方法,不仅不需要考虑系统的数学模型,而且相比较其他智能方法,具有更好的理论基础。在对不确定性非线性系统来说,传统的基于二次规划的方法,需要进行改进,即需要在线迭代的在线训练方法。本书研究了非线性系统的预测控制和内模控制,并研究了支持向量机回归在线训练的方法。将所研究的预测控制采用仿真模型应用于复杂的温室环境控制。

数值预报是气象预报的重要方法之一,多元分析与时间序列分析是数值预报较常用的方法。然而常用的方法,要么是线性回归,要么即便是非线性回归方法,所采用的模型非线性映射功能比较简单,也不能完全反映非线性规律;要么就是模型结构比较复杂,推广性差等不足。支持向量机回归具有较好的理论基础,对样本数据拟合采用的是结构风险最小化,具有很好的推广性,将这个方

法引入到气象预测是非常有益的尝试。因此本书在第 2 章介绍了一些支持向量机与支持向量机回归的基本原理及训练方法,同时在其他章节(如第 6 章),提供了一些快速训练算法,以期给气象领域的科研工作者一些参考。

我国地处欧亚大陆东南部、面临世界第一大洋——太平洋,在带来季风明显、气候条件优越的情况下,也使气象灾害比较频繁,因此进行气象灾害预测是非常重要的。简单的灾害预测就是判断未来某一年灾害或不是灾害,因此可以看成模式识别的分类问题。比较复杂的气象问题,如灾害的分级就可以看成是模式识别的多分类问题。因而灾害预测问题也就可以变成支持向量机的分类问题。本书通过凉夏和异常高温的例子实现了支持向量机的灾害预测。

本书的内容力求简单易懂,尽可能地不作理论上的阐述,以求读者能更快地了解支持向量

机进而使用支持向量机进行非线性建模预测与控制的工作。

本书的大部分工作是在博士研究生和博士后工作期间的研究成果,因此,在此特别感谢中国科学院合肥智能所方庭健研究员、中国农业大学汪懋华教授和南京航空航天大学姜斌教授的指导。南京信息工程大学葛群芳女士校对了全文,在此表示感谢。

本书的研究工作得到中国博士后科学基金(2004035375、20070411044)、江苏省博士后科研资助计划(070101413)、江苏省高等学校自然科学研究项目和南京信息工程大学科研基金项目的资助。在此深表感谢。

本书适合于高等院校高年级学生、研究生、教师和科研院所科研人员阅读。由于作者水平有限,书中难免有不妥之处,敬请读者批评指正。

作者

2009年11月

目 录

前言

第 1 章 绪论	(1)
1.1 概述	(1)
1.2 支持向量机的发展	(1)
1.3 SVMR、建模与控制的研究现状与存在的问题.....	(3)
1.4 智能控制研究的现状与面临的挑战	(7)
1.5 研究的动机与目标	(9)
1.6 研究的内容	(9)
第 2 章 支持向量机与支持向量机回归	(12)
2.1 概述.....	(12)
2.2 SVM 的基础	(12)
2.3 支持向量机.....	(16)
2.4 支持向量机回归.....	(20)
2.5 模型选择.....	(27)
2.6 支持向量机种类.....	(30)
2.7 本章小结.....	(36)
第 3 章 基于 SVM 的灾害天气预测方法	(38)
3.1 支持向量机灾害天气预测意义.....	(38)
3.2 支持向量机灾害天气建模与预测方法.....	(38)
3.3 支持向量机灾害天气预测实验.....	(39)
3.4 结论.....	(43)
3.5 本章小结.....	(43)
第 4 章 支持向量机回归非线性建模方法及应用	(44)
4.1 非线性建模理论研究现状与意义.....	(44)
4.2 系统过程建模的一般原理.....	(45)
4.3 支持向量机回归建模.....	(47)
4.4 结论.....	(53)
4.5 本章小结.....	(53)
第 5 章 支持向量机回归内模控制	(54)
5.1 概述.....	(54)
5.2 内模控制的基本原理.....	(54)

5.3	非线性内模控制的研究现状	(56)
5.4	支持向量机回归内模控制	(57)
5.5	SVMR 内模控制的仿真与分析	(60)
5.6	结论	(69)
5.7	本章小结	(69)
第 6 章	支持向量机回归预测控制	(70)
6.1	概述	(70)
6.2	预测控制的原理	(71)
6.3	支持向量机回归预测控制	(72)
6.4	SVMR 预测控制的实验与分析	(78)
6.5	结论	(94)
6.6	本章小结	(95)
第 7 章	在线稀疏最小二乘支持向量机回归	(96)
7.1	概述	(96)
7.2	最小二乘支持向量机	(96)
7.3	在线稀疏最小二乘支持向量机	(97)
7.4	在线稀疏最小二乘支持向量机收敛性分析	(100)
7.5	仿真实验及分析	(103)
7.6	结论	(107)
7.7	本章小结	(107)
第 8 章	基于 OS-LSSVM 的内模控制方法	(108)
8.1	概述	(108)
8.2	内模控制的基本原理	(108)
8.3	非线性内模控制的研究现状	(110)
8.4	OS-LSSVM 回归内模控制	(111)
8.5	OS-LSSVM 内模控制的实验与分析	(113)
8.6	结论	(127)
8.7	本章小结	(128)
第 9 章	数字温室环境 OS-LSSVM 建模与基于 SP 的控制方法研究	(129)
9.1	概述	(129)
9.2	温室智能控制的研究现状与存在问题	(130)
9.3	温室环境与作物生长模型	(131)
9.4	温室环境 OS-LSSVM 建模的研究	(132)
9.5	基于 SPA 的温室环境 OS-LSSVMR 控制方法	(138)
9.6	结论	(142)
9.7	本章小结	(143)
参考文献		(144)
后记		(155)

第1章 绪论

1.1 概述

机器学习不仅要学习得到规律,从得到的规律中较好地解释已知的实例,而且更重要的是对未来的现象或无法观测的现象做出正确的预测和判断,也即具有推广能力。统计学在解决机器学习问题中起着基础性的作用。然而,传统统计学研究的主要是渐近理论,即当样本趋向于无穷多时的统计性质,而实际问题中的样本数目是有限的,因而与传统统计学的这一性质相矛盾(张学工,2000)。统计学习理论,产生于20世纪70年代,研究的是基于有限样本情况下的机器学习问题,因而与实际问题相一致,可以较好地解决实际的学习问题。支持向量机(support vector machines, SVM),是在这一理论框架下产生的一种新的通用机器学习方法,产生于20世纪90年代,从其产生到现在的10多年里,其理论和应用都得到了很快的发展,取得了一些重要的成果,显示了这种理论与方法的优越性。

1.2 支持向量机的发展

1.2.1 统计学习理论的发展

统计学习理论是机器学习的一类,机器学习理论产生于20世纪60年代,其发展历程可以分为四个阶段(Vapnik,2000):

第一阶段主要是在20世纪50年代中叶到60年代,主要成果是学习机器的创立。如,1957年,Rosenblatt提出的感知器模型,标志着对学习过程进行数学研究的真正开始,为统计学习的产生提供了可能性。

第二阶段主要是在20世纪60—70年代,是统计学习理论基础创立阶段。如:1960年,Solomonoff提出的算法复杂度的思想(Solomonoff,1960);1962年,Tikhnov等发现解决不适定问题的正则化理论方法(Tikhnov,1963; Ivannov,1962);1957年 Rosenblatt,1962年 Parzen等提出了几种密度估计的非参数统计学方法;对于支持向量机来说,更为重要的是,1968年 Vapnik和 Chervonenkis提出了统计学习理论的核心概念 VC熵和 VC维的概念(Vapnik等,1968)。1982年,V. Vapnik进一步提出了具有划时代意义的结构风险最小化原理,为SVM的研究奠定了直接的、坚实的理论基础(Vapnik,1982)。

第三阶段是 20 世纪 80 年代,主要是神经网络的建立。1986 年,Le Cun (Le Cun, 1986)、Rumelhart 等独立地提出了后向传播(back propagation, BP)的神经网络思想(Rumelhart 等, 1986)。

第四阶段主要是在 20 世纪 90 年代神经网络替代方法的建立。其主要成果为:

(1)神经网络改进方法的研究,例如径向基函数模型的研究(Powell, 1992);

(2)支持向量机的出现,1992 年, B. Boser, I. Guyon 和 V. Vapnik 提出了最优边界分类器,由此形成了 SVM 的雏形(Boser 等, 1992)。

1.2.2 支持向量机的发展

支持向量机出现以后,其理论的发展和应用非常快,其主要的成果有:在 B. Boser 等提出最优边界分类器后, Cortes 和 Vapnik(1995)进一步探讨了非线性最优边界的分类问题。1995 年, Vapnik 出版“*The Nature of Statistical Learning Theory*”(统计学习理论的本质),系统地阐述了统计学习理论及 SVM 的概念和分类方法(Vapnik, 1995)。1997 年, Vapnik 等发表的论文“*Support Vector Method for Function Approximation, Regression Estimation, and Signal Processing*”(函数逼近、回归估计和信号处理的支持向量法)详细介绍了基于 SVM 方法的回归算法和信号处理方法(Vapnik 等, 1997)。1997 年, Müller 等采用 SVMR(support vector machines regression, 支持向量机回归)进行时间系列建模的研究(Müller 等, 1997),拓宽了支持向量机的研究领域,此后许多学者进行了这方面的研究(Mukherjee 等, 1997; Frontzek 等, 2001; Gunn 等, 1999; Brown 等, 1998; Drezet 等, 1998; Kecman 等, 2001)。1998 年, Smola 在博士论文中系统地研究了 SVM 的学习机理及在分类中的应用,为进一步完善 SVM 非线性算法做出了重要的贡献(Smola, 1998)。2001 年 Suykens 等采用了 SVMR 进行优化控制的研究(Suykens 等, 2001),使支持向量机的研究向控制领域发展,不仅开创了智能控制的新方向,而且进一步拓宽了支持向量机的研究领域。

近几年出现了许多发展和改进的 SVM 算法,如张学工的 CSVM(Zhang, 1999), Schölkopf (1999)的 ν -SVM, Joachims(1999)提出的 SVM^{light}, Hsu(1999)等提出的 B SVM, 这些算法主要是对支持向量机训练算法的改进。SVM 正日益引起广泛的关注,尤其值得一提的是 IEEE 于 1999 年在《IEEE Transactions on Neural Networks》出版了专辑(见《IEEE Transactions on Neural Networks》[1999, 10(5)]),这对于推动 SVM 的发展起了重要的作用。

1.2.3 国内统计学习理论与支持向量机的研究概况

早在 20 世纪 80 年代末期,边肇祺等(1988)就注意到统计学习理论的基础成果。进入 90 年代末期,一些专家学者已经认识到这个领域的重要性,越来越多的学者、研究人员和科研机构正积极地从事统计学习理论方面的研究。如清华大学闻芳等(1999)将 SVM 用于剪接位点的识别;清华大学卢增祥等(1999)提出交互支持向量机学习算法并将起其应用于文本信息过滤问题的研究(鲁小兵等, 1996);中国科学技术大学陶卿等(2000)发表了关于支持向量机的综述文章,清华大学的张学工(2000)发表了关于统计学习理论和支持向量机的论文;安徽大学的张铃(2001)研究了 SVM 理论与神经网络规划算法的关系,并指出基于 SVM 的算法与其 1994 年提出的神经网络的基于规划的算法是等价的。中国科学技术大学范劲松等(Fan 等, 2000; 2002)将统计学习理论应用于遗传算法的研究。从文献检索获悉,2003 年以后的几年里

国内相关的文献比以前要多得多,说明了统计学习理论和支持向量机的研究在国内受到重视。

1.3 SVMR、建模与控制的研究现状与存在的问题

SVMR(支持向量机回归)是支持向量机在实函数域的研究内容。近些年来,有关 SVMR 的新理论和新方法不断涌现,其应用领域也在不断扩大,从早期的时间序列建模到学习前馈控制(learning feed-forward control)、优化控制等方面的研究,充分显示了这种算法的优越性。

1.3.1 支持向量机回归学习算法的研究

Smola 等(2001)和 Scholkopf 等(2000)介绍了许多对偶问题的解法及实现。但传统的诸如内点法等二次规划算法不适合解决大数据量的问题,这是因为:首先,这些算法要求矩阵的计算和存储都在内存里进行,这对于大数据量的问题需要特别巨大的内存;其次,这些算法需要复杂的矩阵运算。因此,对大规模数据集的 SVMR 学习的算法必须进行改进。

块(chunking)算法是由 Cortes(1982)和 Vapnik(1995)提出的一种改进的学习算法。这种学习算法采用了并行依次更新 α_i 的方法,将优化方法的大型二次规划(quadratic programming)问题分解为一系列小规模二次规划问题来解决大数据量 SVMR 的学习问题。首先选择一个数据子集或工作集,并在这个数据子集解二次规划问题,得出的支持向量和当前不满足 KKT 条件(有关 KKT 条件的内容见 2.3.1 节)的 M 个数据组成新的数据子集或工作集,并在这个数据子集解二次规划问题,这个过程循环往复,直到所有样本满足优化条件为止。

块优化算法虽然对传统的 SVMR 算法有了一定的改进,但是对于规模巨大的数据集或建模数据不具有稀疏性的数据仍然不适用。为此 Osuna 等提出了分解算法(decomposition)(Liao 等,2001;Osuna 等,1997;Laskov,2000)。这种算法采用固定大小的数据子集,在解子集二次规划问题后,先从子集中移走一个样本,再加入一个不满足 KKT 条件的样本,并在这个数据子集进行优化,这个过程循环往复,直到所有样本都满足优化条件。

虽然分解法对块算法在大数据集和不具有稀疏性数据的学习方法进行了改进,但这种方法每次只优化一个数据,因而存在速度和效率问题,这对于大数据集的问题同样也是不适合的,而且固定子集大小是否能够包含所有支持向量也是一个要考虑的问题。Platt 等提出了 SMO(sequential minimal optimization,序贯最小优化)学习算法(Platt,1999;Hevade 等,2000)。这种方法每次重复过程只能优化两个拉格朗日(Lagrange)因子,由于只有两个参数得到优化而其余参数保持不变,因而优化可以不用二次规划方法而采用分析方法。因此该方法不仅适合大数据集的学习而且提高了运行的速度,因此,SMO 算法被认为是目前 SVMR 训练算法中最好的方法(Campbell,2000)。

1.3.2 支持向量机回归模型的选择

支持向量机回归模型的选择包括核函数、损失函数的选择以及与之相关的参数的选择。

1.3.2.1 SVMR 核函数的选择

对于非线性模型来说,回归模型中经常采用 RBF 核函数。以下就以 RBF 为例来说明如何选择核参数。对于 RBF 核来讲,需要选择的参数有:核的宽度 σ 和容量 C 。通常有如下几种方法:交叉验证法(cross-validation)、从统计学习理论导出 VC 维的界和贝叶斯(Bayes)法等。

交叉验证法:最简单的交叉验证方法是定义一个训练集、确认集和测试集,可以选择几组不同的 C 和 σ ,从训练集中的训练数据推导出支持向量系数 σ ,选择其中使确认集中数据错误最小的那一组 C 和 σ 作为模型的参数。测试集中的数据是完全独立的数据,在模型辨识的过程中,不参与模型的训练与确认,只是在训练完毕后用来测试模型。交叉验证法是较好的方法,对于大数据集倾向于采用 10 次交叉验证法。

统计学习理论的 VC 维学习方法:采用统计学习理论的方法导出模型推广错误的界(见(2-2-1)式),这个界用 VC 维来表示,用统计学习理论选择的核和容量 C 可以使 VC 维的上界最小,从而就可以确定模型的参数。采用这种方法虽然可以使 VC 维上界最小,然而这种方法需要在非线性空间计算超球半径(Steve,1998;Chapelle 等,2000)。

贝叶斯学习方法:贝叶斯学习方法被成功地应用于经典神经网络的训练与理解。在 LS-SVM 中的应用被发展为三个层次的推理:第一个层次的推理是考虑 w 的概率的分布(在一个潜在的无限维空间),确定正则项和最小二乘损失函数的可能性;第二个层次的推理是容量 C 的推导;第三个层次的推理是获得核的 σ 。

1.3.2.2 损失函数的选择

支持向量机回归中的损失函数(详见 2.4.1)常用的有二次损失函数(最小二乘损失函数)、最小模损失函数(拉普拉斯函数)、Huber 损失函数和 ϵ 不敏感损失函数等。Huber 认为在数据分布未知的情况下,Huber 损失函数具有最佳特性的鲁棒回归函数。如果已知噪声密度是一个对称函数,则最好的对回归的最小最大的估计(在最坏的噪声情况下对模型最好的估计)的损失函数为最小模损失函数。 ϵ 不敏感损失函数具有稀疏性,其他的损失函数则不具有稀疏性(Vapnik,2000;Hevade 等,2000)。

1.3.3 支持向量机回归建模的研究

采用 SVMR 进行时间序列建模的研究比较多(Müller 等,1997;Mukherjee 等,1997;Frontzek 等,2001;Gunn 等,1999;Brown 等,1998;Drezet 等,1998;Kecman 等,2001),除了时间序列建模以外,还有系统辨识、经验数据建模(Brown 等,1998;Kecman 等,2001)等建模方法。这部分简单介绍时间序列和系统辨识的建模方法研究。

1.3.3.1 非线性时间序列的建模

Müller 等(1997)对有噪声的时间序列进行了仿真实验,并与 RBF 网络仿真的结果进行了比较,结果表明,对于随机白噪声 RBF 与 SVMR 取得了相似的结果,但对于正规噪声则 SVMR 好于 RBF 网络。Mukherjee 等(1997)采用了三个时间序列进行了仿真实验并与多项式、比例、局部多项式、RBF 和神经网络等算法进行了比较,结果表明与其他方法相比 SVMR 具有很好的特性。Frontzek 等(2001)应用 SVMR 时间序列预测的原理,采用不同的核建立了单个神经元的动力学模型并对其进行了测试,与传统的 RBF 神经网络相比,高斯核的 SVMR 算法具有学习速度快和极好的一步预测效果的优点。

1.3.3.2 系统辨识

Drezet 等(1998)将 SVMR 应用于离散的线性和非线性系统辨识的研究。对于线性系统辨识其基本原理是将线性系统的系数矩阵 (a, b) 等于优化结构风险泛函所得权矩阵 (w) ,其中的 SVMR 采用线性 SVMR;对于非线性系统辨识其基本思想是将高维线性空间对应的非线性

性项的权值作为对非线性系统所辨识的参数,适合的核函数有高斯核、双曲正切核和多项式核。仿真结果表明,对于线性系统来说,当所估计的模型的阶数与实际模型阶数一致时,辨识的精度就高,反之辨识的精度就降低。对于非线性系统来说,辨识的精度与核函数的复杂性有关,精度随着核函数的复杂性的增加而降低。

1.3.3.3 支持向量机非线性过程建模的研究

应用 SVMR 进行非线性过程建模是近几年国内比较热门的研究。其中贾新春等(2003)采用基于 Mackey 证据框架的 v -SVR 进行系统辨识;冯瑞等(2003)采用用模糊支持向量分类算法 F-SVC 对输入数据进行预处理,得到多模型模糊隶属度;用模糊支持回归算法 F-SVR 建立多模型 MM 估计器,并应用中和滴定过程进行建模。李丽娜等(2003)采用 SVM 将支持向量机应用到典型的时变、非线性工业过程的连续搅拌反应釜的辨识;王定成等(2004a)从温室环境机理模型出发,分析了温室环境的特点,并采用在线 SVMR 进行温室环境的建模;阎威武等(2004a)采用最小二乘 SVMR 利用滚动时间窗的数据进行建模并将其应用于轻柴油凝固点的预测;崔万照等(2004)采用基于小波核的最小二乘 SVMR 进行系统辨识;阎威武等(2004b)采用贝叶斯证据框架下的最小二乘支持向量机进行建模。这些研究都取得了较好的效果。

1.3.3.4 在线训练算法的研究

在线训练一个突出的特点是支持向量机回归的学习不是一次离线进行的,而是数据逐一加入,不断进行的优化过程。因此,这种学习方法更要考虑训练的速度。增量学习算法是研究比较多的一种在线学习算法。Syed 等(1999)介绍了一种简单的增量训练方法。这种方法是每次选一小批常规二次规划算法能处理的训练样本进行训练,丢弃训练结果中的非支持向量,而将支持向量留下来和新进的样本一起进行训练,直到所有训练样本用完为止。这种方法与 Vapnik(1982)的块算法和 Osuna 等(1997)的分解算法(decomposition)非常相似,但不同的是块算法和分解算法对以前的非支持向量重新判断是否为支持向量,实质上是针对所有样本的训练,而 Syed 的增量算法则完全丢弃了非支持向量,实质上只与增量样本有关。其优点是减少了训练的样本,加快了训练速度,缺点是有可能导致有用的支持向量的丢失,最终导致建模预测的不准确。Syed 在文献的 Australian(澳大利亚人)、Diabetes(糖尿病)、German(德国人)等 12 个数据库的实验也说明了这个问题。

一种改进的支持向量机回归训练算法的目的是防止支持向量机回归训练过程中有用支持向量的丢失。Cauwenberghs 等(2001)在支持向量机进行模式识别的研究中实现了这一目的;马勇等(2004)、Martin 等(2002)采用 Cauwenberghs 等(2001)的方法在支持向量机回归中实现了这一算法。算法主要是根据 KKT 条件将训练样本数据分成三个集合,对新增加的样本 (x_c, y_c) 调整其系数 β_c 使其适合 KKT 条件,同时调整其他训练样本的系数使其也满足 KKT 条件;或者对于存在样本集里的支持向量数据 (x_c, y_c) ,在满足 KKT 条件下调整其系数 β_c 使其成为非支持向量数据,从而可以从训练样本集数据中移走,而不影响训练的结果,与此同时,调整其他训练样本的系数使其也满足 KKT 条件。这种方法实质上还是针对所有的样本数据进行训练。

Frieß 等(1998)介绍的序贯分类的训练算法,Vijayakumar 等(1999)将算法改进,使其不断适合分类而且也适合回归方法。其基本思想是借助感知机中的 Adatron 算法的原理来改变拉格朗日系数,具体来说通过序贯加入的样本的预测误差来修改支持向量机回归样本的系数

α, α^* , 其本质上是一个爬山的寻优算法, 通过反复的修改序贯加入样本的系数, 使支持向量机回归最终收敛。这种方法形式比较简单, 每次更新的运算量小、需要的内存也不大。然而这种训练算法每次迭代只改变一个样本的系数, 而且每次训练都是所有样本都需要进行反复迭代。因此, 算法的复杂性与时间成正比, 对于在线训练的支持向量机回归运算来说, 由于随着时间的推移, 样本数据在不断增加, 用于在线训练还有一定的局限性。Engel 等(2002)在 Frieß 等(1998)的基础上提出一种 SOG-SVR(sparse online greedy support vector regression, 稀疏在线贪婪的支持向量回归法)。算法是这样的: 在第 t 步, 在观察第 $t-1$ 个样本后, 选择 m 个线性独立的基向量 $\{\phi(\tilde{x}_j)\}_{j=1}^m$ ($\tilde{x}_j$ 为在第 t 步时的字典里的样本)的字典, 在线训练过程中, 只改变字典里样本 $\tilde{x}_j$ 系数, 由于一般情况下, $t \gg m$, 这样就减少了训练和模型运算的时间。Kivinen 等(2004)将支持向量机系数的调整分两种情况: 新加入的样本的系数采用损失函数的梯度来进行调整, 以往的系数采用乘以系数 $1 - \eta\lambda$ ($\eta\lambda < 1, \lambda > 0$)。随着样本数量的增加, 对于旧的样本会被遗忘掉, 因此这实际上是一种近似的支持向量机训练方法, 虽然可以保证训练的实时性, 但模型的精度却不一定能得到保证。

1.3.4 支持向量机回归控制的研究

Suykens 等(2001)和 Kruif 等(2001)采用了 SVMR 进行控制算法的研究。Suykens 等采用最小二乘支持向量机进行优化控制的研究, 其基本思想是将 N 步优化控制的目标函数和 N 步预测误差作为损失函数, 约束条件除 SVMR 的约束条件外, 将系统的动力学方程和控制规律的表达式也作为约束条件。最小化风险泛函即可求得具有核函数的控制规律, 并通过曲线跟踪、倒立摆和球与光束的例子进一步说明了该优化控制方法的控制效果。

Kruif 等(2001)将 SVMR 用于学习前馈控制(LFFC)中的学习前馈结构(LFF), 并分析了 SVMR 作为 LFFC 中的 LFF 的可行性。在学习前馈控制中, 如果学习前馈是过程的逆, 那么当干扰出现时仍能获得很好的控制。采用 SVMR 进行离线学习, 并用线性马达系统进行了仿真实验, 与 B 样条神经网络相比, 具有极好的性能。

此外, Millerioux 等(2002)采用支持向量机回归来设计多中心非线性观测器。Andrea 等(2001)研究了塑料注射成型过程的支持向量机回归自适应控制过程的硬件实现。

此外, 国内采用 SVMR 进行预测控制的研究, 其主要机理是采用 SVMR 进行非线性对象的建模与预测。王定成等(2004a)采用 SVMR 进行内模控制的研究, 通过 SVMR 建立非线性系统的正向模型和设计逆模控制器, 仿真实验说明了 SVMR 内模控制具有较好的控制效果。张浩然等(2003), 王定成等(2004c), 王宇红等(2004), 刘斌等(2004), 叶美盈(2005), Zhong 等(2004)都是采用 SVMR 进行建模的支持向量机回归控制预测控制的研究。Kulkarni 等(2003)将混沌系统分解成线性部分和非线性部分, 用离线训练的 SVMR 逼近非线性部分, 预测非线性误差, 来补偿对于线性部分设计的反馈控制器的误差。

1.3.5 支持向量机回归软测量方法的研究

熊志化等(2004), 马勇等(2004), 陶卿等(2004), 冯瑞等(2004), 阎威武等(2003), 阎威武等(2004a)采用 SVMR 进行软测量的研究。熊志化等(2004)采用 SVM、核主元回归和核偏最小二乘法进行软测量的建模, 以及 SVM 与核主元回归、SVM 与核偏最小二乘法的混合软测量建模并将其应用于工业萘初馏塔酚油含氮量的软测量; 阎威武等(2003)采用最小二乘支持

向量机进行软测量建模并应用于轻柴油凝固点的预估;冯瑞等(2004)提出一种适合于局部学习的加权支持向量机移动建模方法并应用于 Box-Jenkins 煤气炉和重油催化裂化装置的软测量。研究结果表明 SVMR 在软测量方面取得了较好的结果。

1.3.6 支持向量机故障诊断的研究

故障诊断分为两种:一种是定性的方法,通常是采用分类的方法识别故障;另一种是定量的方法,通常采用评估系统模型预测输出与实际系统输出残差的方法,从而可以确定故障的程度。目前大多采用前者进行故障诊断(胡寿松等,2001;王定成等,2004b;肖健华等,2001;李凌均等,2002;He 等,2002;Tomas 等,2005;Bernardete 等,2005;Gianluca 等,2004),主要采用多分类的方法进行故障类型的确定。而后者研究的内容却较少。

1.3.7 支持向量机回归其他方面的研究

随着支持向量机研究的深入支持向量机的研究不仅在理论和应用方面得到发展,而且在硬件实现方面也进行了研究(Anguita 等,1998;1999;2000)。此外,支持向量机与其他软计算方法结合的研究,例如 Jeng 等将支持向量机与模糊逻辑结合的研究(Jeng 等,1999;Inoue 等;Lin 等,2002);Jayadeva 等将支持向量机与神经网络结合的研究(Jayadeva 等,2002;Anguita 等,2002);Strauss 等将支持向量机与小波结合的研究 (Strauss 等,2001;Millet-Roig 等,2000)。

1.3.8 存在的问题

虽然支持向量机回归以及其在建模与控制的研究方面取得了一定的成果,然而这些研究仍然存在一些有待深化研究的问题:

(1)在 SVMR 理论与算法的研究方面,SMO 算法虽然是大数据样本学习的较好算法,然而其训练是针对所有样本的训练,不适合不确定性温室环境在线建模的训练,因此寻找新的在线学习算法是 SVMR 研究必须解决的问题。

(2)支持向量机在回归建模的研究方面具有很好的特性,但却缺乏有效的在线建模与预测的 SVM,这方面的研究有待探索和深化。

(3)SVM 在控制领域的研究刚刚起步,目前尚没有提出有效的 SVMR 控制理论来解决非线性、时变、干扰、复杂的控制问题,因此提出新的 SVMR 控制理论和方法是需要重点解决的问题。

(4)统计学习理论在理论方面的研究比较突出,但在实际应用方面的研究却较少,扩大应用领域的研究是支持向量机回归的重要课题。

1.4 智能控制研究的现状与面临的挑战

1.4.1 控制理论发展概述

控制理论与量子理论、相对论一起被称为 20 世纪上半叶科学发展的三大飞跃,经过几十年的发展,其应用和影响已经深入到社会生活的各个方面,使人类大大突破了自身能力的局

限。控制理论的发展大体上分为经典控制理论、现代控制理论和智能控制理论三个阶段。如果把 1932 年奈奎斯特(H. Nyquist)关于反馈放大器稳定性的经典论文作为起点,经典控制理论主要是其后的 30 年发展起来的。经典控制理论主要处理单变量常系数线性系统的控制问题。系统的模式比较单纯,基本的分析和综合的方法是基于频域的和图解的方法。通过引入等价的线性化环节,并把相应的控制系统的概念和分析方法加以修正和拓展,经典控制理论也能有效地处理一些包含非线性环节的简单非线性控制系统(高为炳等,1994)。

现代控制理论产生于 20 世纪 60 年代前后,是对经典控制理论的精确化、数学化和理论化,克服了经典控制理论仅能分析 SISO 系统的局限性,是研究多变量线性控制系统,其中特别重要的是对控制系统本质的基本理论的建立,如可控性、可观测性、实现理论、典型型、分解理论等。这些理论促使非线性系统、最优控制、自适应控制、辨识与估计理论、卡尔曼滤波等发展为独立的学科分支(高为炳等,1994;陈翰馥,1994)。

1.4.2 智能控制理论发展概述

然而经典控制理论和现代控制理论本质上研究的都是线性时不变系统的控制问题,而实际系统都是非线性的,在控制系统中的线性系统实际上是对非线性系统的一种理想化或近似的描述。随着控制理论研究的深入和控制要求的提高,需要解决的控制问题是更接近实际的问题,即非线性和时变等特性更加突出的问题。智能控制是以无模型为特征的更接近人脑思维方式的一种控制理论,是以知识信息为基础进行学习和推理,用启发式方法来引导求解过程,适合含有复杂性、不确定性和模糊性系统的求解过程(陈翰馥,1994)。

智能控制产生于 20 世纪 60 年代,其间的研究主要是将人工智能用于控制系统中的学习和记忆,是智能控制的萌芽。70 年代是智能控制的诞生和形成期。80 年代,智能控制的研究进入了迅速发展时期。1984 年专家系统被引入到控制系统,出现了专家控制;与此同时,神经网络被引入到控制领域,出现了神经网络控制。90 年代,智能控制的种类进一步增多,出现了诸如仿人智能控制、遗传算法控制、小波控制等智能控制(盛万兴等,1999)。

智能控制中的专家系统控制是一种基于知识的控制方法,它作为人工智能在控制领域的较早尝试,在解决某些复杂系统的控制问题上取得了较满意的效果。但是这种基于知识的专家系统在知识获取、知识表达和推理方式上存在着严重的缺陷,例如知识的来源主要靠专家经验、知识“瓶颈”、知识“组合爆炸”等。这些缺陷限制了专家系统控制理论的发展和应用。

智能控制中的模糊控制和神经网络控制在一定程度上避开了这些问题。典型的模糊控制系统是由三模块结构构成的:模糊化、模糊规则合成推理和模糊判决。其核心是以自然语言表示的控制规则,易为人们所接受,且构造容易,鲁棒性强。模糊控制特别适用于那些难以建模或无法建模的对象。然而模糊系统的设计往往包含过多的经验成分,这使得设计的控制器的性能的优劣,取决于设计者的经验的优劣;由于模糊控制是先对控制变量论域离散、量化,再进行推理、决策的过程,由于对语言变量值的分档不可能太细,因此模糊控制的精度较差。模糊控制虽然具有良好的鲁棒性,但由于知识获取困难,使得模糊控制的适应能力有限(Warwick, 1998)。

神经网络控制是 20 世纪 90 年代比较活跃的智能控制,这是因为神经网络具有如下特性(焦李成,1993):

(1)对复杂不确定性问题的自适应和自学习能力(作为控制系统中的补偿环节、自适应环

节等);

(2)对任意非线性关系的表示能力(用于非线性系统的辨识、控制等);

(3)对网络的非线性动力学带来的快速优化计算能力(复杂控制问题的优化计算等)。

神经网络控制系统按照其在控制系统中的作用可分为三种控制系统(焦李成,1993):

(1)神经网络作为模型的各种控制系统;如内模控制、模型参考自适应控制、预测控制中充当对象的模型;

(2)充当各类控制器,如神经网络自适应控制器等;

(3)在控制中起优化计算作用,如神经网络优化控制器和预测控制等。

尽管神经网络具有强大的逼近非线性映射的能力,为神经网络的研究起过强烈的推动作用,但它却具有一些难以克服的缺陷。网络的学习算法是梯度下降法,它由此调节权值使目标函数达到极小,由于基本处理单元的非线性,导致了算法的可能不收敛,有时甚至连局部最小也达不到,它的缺陷具体体现在(何玉彬等,2000):

(1)存在局部极小问题,造成网络的局部收敛,影响系统的控制精度;

(2)学习速度慢、训练时间长限制了神经网络在实时控制中的作用;

(3)理想的训练样本提取困难,影响了网络的训练速度和训练质量;

(4)网络结构不易优化,特别是隐层节点数目的选取常常带有盲目性;

(5)尚未从理论上完全解决神经网络学习算法的收敛性和神经网络控制系统的稳定性。

1.5 研究的动机与目标

统计学习理论在研究小样本情况下表现出了极好的性能(Vapnik,2000),支持向量机是统计学习理论的重要组成部分,但 SVMR 算法是一个新的算法,同其他任何新的算法一样也需要不断的完善。同时,由于专家系统等智能控制理论存在一定的缺陷,传统的神经网络智能控制理论又存在诸如模型结构的选取、算法的收敛性、解的唯一性等问题,需要研究新的智能控制方法。SVMR 具有好的推广特性和非线性建模的能力,因而适合解决非线性、大时延以及不易建模系统的控制问题,然而现有 SVMR 的在线建模与实时控制中有一定的问题,因此本书进行在线 SVMR 算法及建模与控制的研究。

本书的研究目标旨在通过支持向量机回归理论的研究改善支持向量机回归训练的实时性和算法的稀疏性,以满足在线建模和实时控制的需要;建立支持向量机回归智能控制方法并将其应用于复杂系统的控制。

1.6 研究的内容

本书对支持向量机与支持向量机回归的理论基础及常见的训练方法以及常见的种类进行了介绍,在此基础之上,研究了支持向量回归算法的改进、支持向量机回归的智能控制算法(内模控制与预测控制)的研究,以及支持向量机的建模应用于重大灾害天气预测,将支持向量机回归建模应用于交通流的预测,将 OS-LSSVM 应用于非线性、时滞、不确定性环境的温室等的建模,将研究的 SVM 控制用于仿真的温室环境控制。具体内容安排如下。

第1章 绪论

简述统计学习理论、支持向量机和智能控制发展的过程。介绍支持向量机回归理论、建模及智能控制研究国内外现状和存在问题,揭示本书研究工作的背景。

第2章 支持向量机与支持向量机回归

本章是本书的支持向量机与支持向量机回归基础理论部分,部分内容是以后各章的理论基础,但更重要的出发点是给读者以相关方面较为全面的了解。首先系统地介绍了支持向量机的基础,即,学习机器的一般方法、复杂性、推广性能的有关理论以及关于统计学习理论的特点、研究内容,其目的是为了说明支持向量机与支持向量机回归所具有的特点;其次对支持向量机和支持向量机回归的基本概念和学习的机理进行了介绍;再其后对加快支持向量机训练速度和减小内存占用的常用的训练方法进行了介绍;最后介绍了支持向量机与支持向量机回归的模型选择及常见的几种类型及其学习的原理。

第3章 基于 SVM 的灾害天气预测方法

提高预报的准确率,对减少灾害带来的损失是非常有益的。灾害天气是小概率事件,常规统计方法建立的统计预报模型预报的准确率不够理想。基于统计学习理论的支持向量机进行灾害天气预测预报的研究,首先解决了样本数据的构成与核函数及参数选取的问题,然后将其应用于夏天异常高温和凉夏预测,取得了较好的效果。

第4章 支持向量机回归非线性建模方法及应用

利用 SVMR 具有很好的非线性建模的特点建立过程的非线性模型。采用仿真将支持向量机建模应用于具有时变和不确定环境的交通流量的建模与预测,取得了较好的效果。

第5章 SVMR 内模控制

分析内模控制的工作原理、具有的特点以及国内外研究的现状及存在的问题,针对非线性、不确定性环境不易建模的特点,提出支持向量机回归内模控制方法,这种控制方法采用 SVMR 建立内模控制的正向模型,对模型可逆并且唯一的非线性系统设计逆模控制器。对非线性系统模型和温室环境仿真模型的控制进行了实验并与神经网络内模控制的控制效果进行了比较,实验表明 SVMR 内模控制优于神经网络内模控制。

第6章 SVMR 预测控制

分析预测控制的工作原理、具有的特点以及国内外研究的现状及存在的问题,针对非线性、不确定性环境不易建模的特点,提出支持向量机回归预测控制理论,这种控制理论采用 SVMR 作为预测控制的预测模型部分,滚动优化部分采用两种非线性优化算法的研究,分别为:Levenberg-Marquardt 优化方法和遗传算法优化技术相应的预测控制分别为:SVMR-LM 预测控制和 SVMR-GA 预测控制。详细讨论了 SVMR-GA 预测控制的工作机理以及解决 GA 在实时控制中快速收敛算法的研究。对两种预测控制都进行了非线性仿真实验的研究,实验表明两种 SVMR 预测控制都具有很好的非线性控制特性。